BETTING MARKET EFFICIENCY AT PREMIERE RACETRACKS

Marshall Gramm, Rhodes College

ABSTRACT

Accessibility to betting markets has increased dramatically with the simulcasting of races. Based on data from the two largest tracks in the United States, this paper is an empirical analysis of the favorite-longshot bias. Bets on races from these tracks are taken at other tracks, off-track betting facilities, casinos, and even by phone or online and are incorporated into the same mutuel pool. Despite all this, the results indicate betting market inefficiencies still exist. Betting markets are analyzed using three traditional methods (1) the grouping of horses by favorite position, (2) the grouping of horses by odds, and (3) a clustered tobit regression of net returns on odds. Results show that lower odds horses are underbet, consistent with previous studies. Incorporating market participation, however, reveals that as betting dollar volume increases the bias is reduced.

INTRODUCTION

Betting markets have been studied by social scientists as a perceived controlled repeated experiment of asset markets and behavior (see Sauer (1998) for an overview). These experiments are repeated numerous times daily around the world in parimutuel markets for betting on horse and dog racing. The unique aspect of parimutuel betting is that the track acts only as an intermediary or market maker who extracts a certain amount (15-20%, called the take) from the betting pool and then redistributes the rest to the holders of the winning tickets. Odds are displayed for each betting interest in a race.

In an efficient betting market with perfect information, the expected return on any unit bet would equal one minus the track take. This implies that the probability of a horse winning a race would be equivalent to the percentage of money bet on that horse. Previous studies have shown that racetrack betting markets are not efficient (see Thaler and Ziemba (1988) for an overview). Numerous empirical studies have found the existence of some form of a favorite-longshot bias. In a traditional favorite-longshot bias, the public systematically underbets favorites and overbets longshots causing favorites to earn a higher expected return than longshots. There have been few instances of a reverse favorite-longshot, where the public systematically overbets the favorite (see Busche and Hall (1988), Swindler and Shaw (1995)). Any form of a favorite-longshot bias is a violation of market efficiency. 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 Hall (2000b)).

This paper is an empirical replication using new data with analysis of the relationship between market efficiency and market size looking specifically at Saratoga (Saratoga Springs, NY) and Del Mar (Del Mar, CA), the two premiere tracks in the United States. Data includes all 79 racing days (with 8 to 11 races each) at the two tracks in 2001. The paper reveals the existence of a favorite-longshot bias for win bets, in particular a bias against horses with very short odds. This bias is consistent with many previous studies. As market size, measured by the size of the betting pool, increases, the favorite-longshot bias is reduced.

Data

The dataset consists of all races run at Saratoga and Del Mar in 2001. These are the two most highly publicized meets in the country. In 2001, attendance at Saratoga and Del Mar averaged 28,102 and 15,456, which is enormous in an age where fan support for horse racing is in decline. Races at Del Mar and Saratoga are simulcast to almost every racetrack across the country as well as to casinos in Las Vegas and into homes through the TVG network which serves bettors with phone or internet betting accounts. Mutuel pools are co-mingled, meaning that every dollar bet on a race, whether it is done at Del Mar or Saratoga, at any other racetrack, at an off track betting facility, at a casino, by phone, or online, goes into the same pool. Del Mar and Saratoga are the two most popular tracks for simulcast bettors who fatten up the mutuel pools betting millions of dollars ($11.7 and $15.6 million daily at Del Mar and Saratoga respectively). Trainers and owners point their horses for these meets because of the large purses (median purse size of $47,000 at Del Mar and $43,000 at Saratoga) and prestigious stakes races (51 graded stakes races of which 18 are grade I). The media focus on these thoroughbred meets is only exceeded by the Triple Crown and Breeders' Cup. There were 5,894 different betting interests in 709 races run at the two racetracks which meant that a bettor had an average of 8.3 different choices with which to back with a win bet. The average amount in the mutuel pool (win, place and show) for a given race was $440,000 with the largest being $2,879,751 on the 132nd running of the Travers Stakes at Saratoga. The smallest pool size was $138,000, which would be considered a large pool for most other tracks. Odds on horses ranged from 0.10 to 1 at the low end to 195.50 to 1. The track take at Saratoga was 15% and Del Mar was 15.43% and payouts were rounded down to the nearest nickel or dime (breakage). Race size varied from 3 betting interests to 12 betting interests. Horses that were coupled in betting were treated as one betting interest as they are in the mutuel pool. All data were collected from the 2002 Del Mar and Saratoga Players' Guide published by the Daily Racing Form.

Empirical Tests for Market Efficiency

The favorite-longshot bias can be studied by breaking the data into similar groups and calculating the subjective probability, what the general public feels the horses chances are as revealed by the odds, and the objective probability, the actual percentage of winners in the group. The subjective probability is then compared to the objective probability and a significant difference indicates mispricing and market inefficiency. The total amount bet on all horses can be expressed as W with w

denoting the amount bet 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 is 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). The subjective probability is the fraction of money bet on a horse to win $\frac{w_i}{W} = \frac{1-t}{Odds_i + 1}$. 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 binomial statistic. For a sample of n horses a z-statistic can be computed and $z = (\psi - \zeta)\sqrt{n/\zeta(1-\zeta)}$ where ψ is the subjective probability and ζ is the objective probability (see Ali (1977) and Busche and Walls (2000b)). Since the subjective probability or bet fraction is not a random variable, but set for each grouping, independence is not violated. Z-scores with an absolute value of 1.96 or greater indicate a statistically significant difference between the objective and subjective probability at the 5% level. The return on a win bet is equal to -1 if the horse loses and the odds if the horse wins. The return on a unit bet to win on all 5,894 betting interests would have been -20.51%.

One method of grouping involves ranking the horses in each race from most favorite (lowest odds) to least favorite (highest odds). The horses are divided into eight groups by their favorite position in the race from 1 (most favorite, lowest odds) to 8-12 (least favorites, odds rankings of 8th and above). The 8th through 12th favorites were combined because of the small number of observations. The results are summarized in Table 1.

Table 1:
Data Grouped by Favorite Position

Favorite	Runners	Winners	Objective Probability	Subjective Probability	Standard Deviation	z-stat	Win Return
1	709	245	34.56%	34.43%	0.1022	-0.07	-18.00%
2	716	147	20.53%	20.37%	0.0474	-0.11	-16.39%
3	709	105	14.81%	14.89%	0.0368	0.06	-17.00%
4	705	68	9.65%	10.65%	0.0278	0.90	-22.78%
5	698	48	6.88%	7.65%	0.0224	0.81	-23.01%
6	656	38	5.79%	5.61%	0.0190	-0.20	-19.73%
7	571	36	6.30%	4.14%	0.0158	-2.13	8.50%
8-12	1130	22	1.95%	2.48%	0.0122	1.29	-39.07%

Difference 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 objective probability is slightly larger than the subjective probability in the first two favorite positions but the difference is not significant. Interestingly enough the 7th favorite had an objective probability of 6.30% which was significantly larger than the subjective probability of 4.14% (almost 20-1). The return on a unit win bet on the 7th favorite was 8.50%, the only return that was not a double-digit negative number. The lowest odds ranking category (8th through 12th favorites) had a z-statistic of 1.29 (significant at 20%) which is some evidence of the overbetting of extreme longshots. However, coupling the underbetting of the 7th favorite with the overbetting of the 8th-12th favorites might lead us to dismiss the longshot bias.

Another method of grouping betting interests is by odds ranges. Using the example of Busche and Walls (2000a), odds ranges are selected to set equal the total amount of money bet on each grouping of horses. The data are divided into ten odds groupings and the midpoints of each group are reported along with the rest of the results in Table 2.

Table 2
Data Grouped by Parimutuel Odds

Odds Midpoint	Runners	Winners	Objective Probability	Subjective Probability	Standard Deviation	z-stat	Win Return
0.75	132	82	62.12%	51.38%	0.0694	-2.54	2.54%
1.25	190	71	37.37%	37.70%	0.0263	0.09	-16.76%
1.75	226	65	28.76%	30.33%	0.0160	0.52	-19.62%
2.25	293	69	23.55%	25.43%	0.0125	0.76	-21.64%
3.00	324	63	19.44%	21.25%	0.0118	0.82	-22.76%
3.75	394	74	18.78%	17.79%	0.0093	-0.50	-10.23%
5.00	498	72	14.46%	14.43%	0.0097	-0.02	-16.43%
6.75	630	69	10.95%	11.05%	0.0092	0.08	-14.97%
10.00	905	68	7.51%	7.87%	0.0095	0.41	-18.83%
30.00	2302	76	3.30%	3.40%	0.0154	0.26	-26.60%

The evidence of a favorite-longshot bias is much clearer using this method of grouping. Heavy favorites were underbet, as seen by a large disparity between objective and subjective probability (10.74%) with a difference significant at the 1% level (z = -2.54). The return on win unit bets on these heavy favorites were positive, violating weak form market efficiency as defined by Thaler and Ziemba (1998). No other odds grouping had significant differences between favorites and longshots and all had negative double-digit returns.

A third method used in analyzing the data for the existence of a favorite-longshot bias is the use of regression analysis based on methods suggested by Vaughan Williams and Paton (1998b). Dividing horses into groups based on favorite position or odds categories results in measurement error bias which can be eliminated by using information on individuals betting interests, their respective odds, and actual returns (-1 for a non-winner and the odds for a winner). The equation estimated is

$$NR_{ij} = \beta_0 + \beta_1 Odds_{ij} + \varepsilon$$

where NR_{ij} is the actual net return to a unit win bet on the *i*th horse in the *j*th race and $Odds_{ij}$ are the odds on the *i*th horse in the *j*th race. Because the data is censored at -1, a Tobit regression is preferred. Since horses within races are interdependent, observations are clustered within races and assumed independent across races. If $\beta_1 < 0$ then a traditional favorite longshot-bias exists where the returns are greater for lower odds horses. Results from the clustered tobit regression are included in Table 3.

Table 3
Clustered Tobit Regressions

	Win Betting	
	Coef.	**Coef.**
Odds	-0.461	-0.481
Z-score	-13.36	-13.81
Slope	-0.0409	-0.0421
Odds*Z		0.085
Z-score		2.79
Slope		0.0074
Constant	-13.255	-13.155
Z-score	-11.26	-11.29
Observations	5,894	5,878
Clusters	709	707
Log Likelihood	-4097.09	-4083.08

Regression results indicate that a traditional favorite-longshot bias exists. The coefficient on Odds is statistically significant (z = -13.81) and the marginal effect is -0.0409. This can be interpreted as a $1 increase in the odds decreasing returns from a unit win bet by 4.09¢. This is lower than the bias found in previous studies. Gander, Zuber, and Johnson (2001) had a -0.0467 slope coefficient for races in New Zealand in the mid-1990s and Vaughan Williams and Paton (1998b) found a slope coefficient of -0.0636 for races from the UK in 1992. The intercept is negative and significant which is expected since net returns are diminished by the track take.

Market Size and Efficiency

To determine the relationship between market size and efficiency, assume all bettors can be classified as either professional players or recreational players and that the amount they bet on each race is *P* or *R*, respectively. The total amount bet is $W_j = P_j + R_j$ for the *j*th race. The total amount an individual professional player bets is based on whether she is informed or not informed about race *j*. An informed bettor is able to estimate the true (objective) probabilities for horse *i*. The individual professional bets on the race if she is informed and passes if uninformed. If *I* denotes the fraction of professional bettors who are informed then the total amount professionals bet on race *j* is $P = f(I)$ where $f' > 0$. For a given race as the number of informed bettors increases, *P* increases, *W* increases, the ratio $\frac{P}{W}$ increases, and the odds on the tote board approach the true odds. As $\frac{P}{W}$ approaches 1, the favorite-longshot bias disappears. Recreational players bet similar amounts on all races. Since $W_j = P_j + R_j$, the mutuel pool increases as information increases which in turn increases market efficiency.

The influence of market size on efficiency is estimated by employing data on the total dollar size of the mutuel pool into the regression model. By analyzing the effects of changes in pool size over the course of a day at the track we can identify races in which bettors are more informed. Busche and Walls (2000b) found that betting markets were inefficient only at tracks with low betting volume but did not examine variation in pool size within tracks. Vaughan Williams and Paton (1998b) controlled for information by separating higher graded races and found that for these races betting markets are efficient (coefficients were insignificant).

To test for the effects of increased bettor participation (as measured by mutuel pool size) an interaction term is added to the clustered tobit regression. The interaction term is the odds multiplied by the mutuel pool size that has been standardized based upon that specific day's mean and standard deviation, $Z_j = \frac{W_j - \overline{W}}{sd(W)}$. There are 79 days of racing (36 at Saratoga and 43 at Del Mar); consequently there are 79 unique means and standard deviations of mutuel pool size (results were similar when Del Mar and Saratoga were analyzed separately). Standardizing pool size simplifies interpretation. The equation estimated is

$$NR_{ij} = \beta_0 + \beta_1 Odds_{ij} + \beta_2 Odds_{ij} * Z_j + \varepsilon \, .$$

If increasing pool size increases market efficiency by reducing the favorite-longshot bias then $\beta_2 > 0$ (given that $\beta_1 < 0$). The outcomes summarized in Table 3 support this result for win bets. The coefficient on the interaction term is positive and significant (z = 2.79). A $1 increase in the odds still reduces returns by more than 4.2¢, but if the pool size is one standard deviation above the mean, the marginal effect is reduced to 3½ ¢. The partial effect of odds on return is now dependent on pool size with $\frac{\Delta NR}{\Delta Odds} = -0.0421 + 0.0074Z$. Therefore, as market participation increases the favorite-longshot bias is reduced and the market becomes more efficient. Relating the result back to professional and recreational bettors, the more bettors who are informed about a particular race, the more that is bet, and the closer odds move to their true value.

Conclusion

This paper suggests that the favorite-longshot bias is still prevalent at the two largest tracks in the country despite the fact that betting markets are more accessible than ever. Races are simulcast across the country and people can participate at their local track, casino, off-track betting facility (OTB), online, or by phone. All races run during the 2001 meet at Del Mar and Saratoga make up the data set which was analyzed for betting market inefficiencies using three traditional methods (1) the grouping of horses by favorite position, (2) the grouping of horses by odds, and (3) a clustered tobit regression of net returns on odds. Grouping the horses by favorite position reveals the oddity of 7th favorites being underbet and a positive return offered for a unit win bet. Grouping horses by odds indicates that heavy favorites are underbet and yield a positive return. The clustered tobit regression model shows the existence of a traditional favorite-longshot bias (underbet favorites and overbet longshots). The relationship between betting market participation and betting market efficiency was analyzed by incorporating the volume of betting in the regression model. The results show that an increase in the mutuel pool size reduces the favorite longshot bias.

REFERENCES

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

Asch, Peter; Malkiel, Burton G; Quandt, Richard E. (1982) Racetrack Betting and Informed Behavior, *Journal of Financial Economics*, **10**, 187-94.

Blood Horse Magazine. September 15, 2001

Blood Horse Magazine. September 7, 2002

Busche, Kelly; Hall, Christopher D. (1988) An Exception to the Risk Preference Anomaly, *Journal of Business*, **61**, 337-46.

Busche, Kelly; Walls, W David. (2000a) Probability and Utility Estimates for Asian Horse Track Gamblers, *Singapore Economic Review*, **45**, 59-71.

Busche, Kelly; Walls, W David. (2000b) Decision Cost and Betting Market Efficiency, *Rationality and Society*, **12**, 477-492.

Daily Racing Form. Del Mar 2002 Players' Guide

Daily Racing Form. Saratoga 2002 Players' Guide

Gandar, John M; Zuber, Richard A; Johnson, R Stafford. (2001) Searching for the Favourite-Longshot Bias Down Under: An Examination of the New Zealand Pari-mutuel Betting Market, *Applied Economics*, **33**, 1621-29.

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

Hausch, Donald B; Ziemba, William T; Rubinstein, Mark. (1981) Efficiency of the Market for Racetrack Betting, *Management Science*, **27**, 1435-52.

Hurley, William; McDonough, Lawrence. (1995) A Note on the Hayek Hypothesis and the Favorite-Longshot Bias in Parimutuel Betting, *American Economic Review*, **85**, 949-55.

Hurley, William; McDonough, Lawrence. (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, Raymond D. (1998) The Economics of Wagering Markets, *Journal of Economic Literature*, **36**, 2021-2064.

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, Dek; Farmer, Amy. (1996) Optimal Betting and Efficiency in Parimutuel Betting Markets with Information Costs, *Economic Journal*, **106**, 846-68.

Thaler, Richard H, Ziemba, William T. (1988) Anomalies. Parimutuel Betting Markets: Racetracks and Lotteries, *Journal of Economic Perspectives*, **2**, 161-174.

Vaughan Williams, Leighton; Paton, David. (1998a) Do Betting Cost Explain Betting Biases? *Applied Economics Letters*, **5**, 333-35.

Vaughan Williams, Leighton; Paton, David. (1998b) Why are some Favorite-Longshot Biases Positive and some Negative? *Applied Economics*, **30**, 1505-1510.