

An Economic Analysis of Pari-mutuel Race Competitiveness

Introduction

Individual bettors are interested in the expected return from their bets. That is, they are concerned with identifying and placing winning bets. Economists researching pari-mutual racing markets are concerned with efficiency. They investigate individuals' betting behavior to determine if it indicates markets are efficient. But racetrack officials are not primarily interested in the odds on an individual horse, the expected returns from a bet or the overall efficiency of markets at their racetrack. They are primarily interested in earning a profit from providing their product, i.e. the entertainment and gambling opportunities provided by horse races. In order to do so, they are interested in scheduling competitive horse races.

In terms both of entertainment and betting, races with horses that have very dispersed odds (i.e. strong favorite and longshot) are less appealing. Consider a race with Secretariat racing against a field of cart horses. Such a race would not create much interest for either spectators or bettors. Watching Secretariat finish lengths ahead of the field, as expected, would not create much more excitement than watching Secretariat complete a workout. And the probability of winning a bet on a cart horse would be extremely low, as would the return to a bet on Secretariat. To assure the interest of spectators and bettors, racetrack officials schedule races in an attempt to maximize interest in the race, and therefore handle (the amount of money bet on the race) and profits.

Background

The race card for any given day contains races scheduled by the Racing Secretary. The Racing Secretary's job is to "bring together horses and purse money in a way to encourage people to come to the racetrack and bet on those horses." (Smith, 1998) The way this is done has changed a bit over time. "In the old days, racing secretaries at Thoroughbred tracks designed handicap races in which they assigned weights to individual horses according to their talents. In a perfectly handicapped race, all horses would finish in a dead heat." (Smith, 1998) Today, racing secretaries try to meet the goal of prompting the closest possible finish by writing conditions for each race. Conditioned races have qualifications for each entry. They specify qualification like the age, number of starts, past success (number of wins or money won), or sex. Race conditions also include whether a race is a claiming or allowance race.

A claiming race is a race in which the horses entered are available for purchase at a specified price. Eligible individuals (not just anyone can claim a horse) drop a claim into the claiming box BEFORE the start of a race. If more than one individual claims a horse, the new owner is determined by drawing lots. At the start of the race, the person making the claim owns the horse, "win, lose, or dead." (Smith, 1998) The claiming price for each race is set by the Racing Secretary to approximate the value of the horses the race is designed to attract. A horse's value is determined by its racing ability. Claiming races are an attempt to equalize the quality of the field. No one would want to enter a horse with a value well above the claiming price (a "ringer") because they would lose ownership of the horse at the start of the race, for a price below its value. A horse with a value well below the claiming price of the race is very unlikely to win the race, so the owner stands little chance of receiving any purse money and no one would want to pay the high claiming price.

The term "allowance" as in an allowance race means that horses get a break (allowance) in the weight they must carry based on their past performance. The races are set for varying levels of competition, and usually limit entries to horses that haven't won a certain number of races in a certain length of time. For example, an allowance race might be held with the specifications below. These types of weight allowances are not relevant to harness races, since there are no jockeys.

Four year olds and upward which have not won two races other than maiden, claiming, or starter since October 2000 weight 122. Non-winners of two races since November 17, allowed 3 lbs. A race since November 17, 6 lbs.

Allowance races "usually attract better horses than claiming races because the horse is not for sale." (Smith, 1998) They are established to assure some degree of competitiveness between entrants.

Racetracks and their officials (particularly the racing secretary) are very concerned with the distribution of odds among horses in a race and keeping the dispersion of the odds from becoming too great.

This study considers a measure of the dispersion of odds in individual races to determine what factors significantly influence race competitiveness. The odds on horses in a race are determined by the amount of money bettors wager on each horse.

Thus they represent the competitiveness of the race *as perceived by bettors*. There is ample evidence that the perceptions of bettors with respect to the probability of a given horse winning a race are very reliable. "The odds generated by racetrack betting efficiently predict the order of finish, and provide reasonably good estimates of the probability of winning." That is, the "standard definitions of market efficiency are generally satisfied" by the market for pari-mutuel races. (Sauer, 1998) So a race with concentrated odds is likely to be a close, competitive race.

There is one anomaly regarding the efficiency of race odds in predicting race competitiveness, known as the "favorite-longshot bias." "There are empirical irregularities that are inconsistent with generic notions of efficiency." (Sauer, 1998) The favorite-longshot bias refers to a situation, found empirically in a variety of (though not all) studies, where the expected return to betting favorites is higher than the expected return to betting longshots. The bias is hypothesized to result from the consumption benefits from betting (i.e. the utility derived from entertainment, as opposed to gambling, aspect of racing) and from the differential returns earned by informed (versus uninformed) market participants. (Sauer, 1998)

Previous studies of wagering markets have focused on market efficiency and what economists can learn about financial markets from the "laboratory" and data that racetracks provide. To do so, previous studies have used the odds on individual horses as they relate to the horses finish in a race. This study is unique in that it looks at the aggregate odds on individual races, that is, a measure of the concentration/dispersion of odds in a race. This measure is referred to as the competition index.

Calculation of the Competition Index

The competition index is a measure of the allocation of money bet on a race and therefore the concentration or dispersion of the odds on horses in a race. This index is patterned after the Herfindahl Index, which uses sales percentages in an industry to measure market concentration.

The Herfindahl Index is given as

$$H = \sum_{i=1}^M X_i^2$$

where each X_i is firm i 's percent of total sales and m is the number of firms in the industry. This index takes into account the size distribution of firms in an industry by squaring the market shares to give greater weight to large firms. In the case of a perfectly competitive industry, with an infinite number of small firms, the Herfindahl Index would approach 0. In a monopoly industry, one firm has 100% of industry sales and therefore the Herfindahl Index equals 100^2 , or 10,000. The Herfindahl index is used by the U.S. Justice Department to determine guidelines for evaluating proposed mergers.

The competitiveness index uses the same approach as the Herfindahl Index to evaluate the competitiveness of a race (as determined by bettors). It takes into account the distribution of bets placed on a race by squaring the percent bet on each horse to give greater weight to favorites.

The competition index is given as

$$C = \sum_{i=1}^m X_i^2$$

where each X_i is the percent of the total amount bet that is placed on horse "i" and m is the number of horses in a race. In a "perfectly competitive" race, with an infinite number of horses, each with a small amount bet on them, the competitiveness index would approach 0. A more realistic example would be a race with 10 horses where 10% of the total amount bet is placed on each horse. In this example, the competitiveness index equals .1. In a race where the outcome is certain (according to bettors), all money would be bet on the favorite horse and the index would equal 1² or 1. The competitiveness index measures the concentration/dispersion of odds on a race, which is part of the determination of the expected return on a bet (both the chances of winning and the payoff in the event of a winning bet).

The percent of the total amount bet on a race that is bet on a given horse is used to calculate the odds of a bet on that horse. The share of the win pool bet on each horse is related to the odds on a horse as given in the equation below.

$$X_i = 1 - [Y_i / (Y_i + 1)]$$

Where;

X_i = the share of the win pool bet on the ith horse

Y_i = the odds on the ith horse, given as "Y to 1" (for example if the odds are 2 -1, $Y = 2$)

The example below shows how the distribution of the amount bet on a race affects the odds on the race and the competition index. The example uses the win pool (bets on each horse to win the race), reports all odds as X -1 (which may not be how the odds would be reported at a track, for example the track may report odds as 1 - 9), and assumes no takeout or breakage. In reality, the takeout (amount of the pool that is not distributed to bettors but is used to pay the combined profit of the track and the state) and breakage (amount of the pool that is not distributed to bettors due to rounding off payoffs to the lowest dime) are taken out of the betting pool before odds are determined.

	Race 1		Race 2	
	Win Pool	Odds	Win Pool	Odds
Horse 1	1000	4 -1	3000	.67 -1
Horse 2	1000	4 -1	1500	2.3 -1
Horse 3	500	9 -1	200	24 -1
Horse 4	1500	2.3-1	0	
Horse 5	500	9 -1	100	49 -1
Horse 6	500	9 -1	200	24 -1
Total	5000		5000	
Index	0.20		.45	

The example shows two races with the same total amount bet in the win pool (\$5000), but each race has a very different distribution of amounts bet on each of the six horses in the race. Race 1 has money bet on

each horse and no overwhelming favorite in the race (according to the expectations of the bettors as revealed by their bets). Race 2 has a clear favorite, with at least double the amount bet to win as compared to any other horse. Race 2 also has a definite long shot, with no one willing to bet on horse #4. The competition index shows the difference in the dispersion of money bet on the two races (and therefore odds on the horses in each race).

The Competition Index in Reality

Theoretical examples illustrate how the competition index is calculated, but do not give insight into the dispersion of odds in real races. Statistics describing the value of the competition index for 282 races held at a harness track and 195 races held at a Thoroughbred track during July 2000 are reported in the table below.

Table 1 | Competition Index Values

	Harness Track	Thoroughbred Track
Mean	.382	.321
Standard Deviation	.112	.092
Median	.357	.304
Minimum	.170	.154
Maximum	.910	.737
N	283	195

The high/low values from the actual races are higher/lower than the values in the hypothetical example and show considerable variation in the dispersion of odds in these races. Harness races showed a higher mean and standard deviation than the Thoroughbred races, perhaps because of the difference in the options available for setting conditions (e.g. weight allowances) or the number of races and the number of racehorses from which tracks can draw competitors.

A bettor can select from a range of different situation with regard to the level of competitiveness perceived by bettors as a whole. Bettors can select from high return/low probability or low return/high probability bets in a race with a high competition index (and a likely winner). Or they can choose to bet on a race with a low competition index and experience the excitement cheering for their selection in what is expected to be a close race.

The Model

The equation below is used to determine what factors significantly affect the competitiveness index of a race. The equation is estimated for a harness track and a Thoroughbred track, using data from races held during July, 2000.

Harness Index = f (#starters, gait, 2YrOld, 3YrOld, claiming, fastclass, purse, %Acct)

Tbred Index = f (#starters, 2YrOld, 3YrOld, claiming, allowance, dist, turf, purse, %Acct)

The definitions of variables are given in Tables 1 and 2.

Table 1 | Definition of Harness Variables

Index Competition Index (see discussion in text)
#Starters (-) Number of horses starting in the race
Gait (?) = 1 if the race is for pacers, = 0 if the race is for trotters
2YrOld (+) = 1 if the race is restricted to 2 year olds, = 0 otherwise
3YrOld (+) = 1 if the race is restricted to 3 year olds, = 0 otherwise
Claiming (-) = 1 if the race is a claiming race, = 0 otherwise
Fastclass (-) = 1 if the race is an invitational or stakes race, = 0 otherwise
Purse (-) The size of the purse awarded to race winners
%Acct (+) Percent of race handle bet via an account

Table 2 | Definition of Thoroughbred Variables

Index Competition Index (see discussion in text)
#Starters (-) Number of horses starting in the race
2YrOld (+) = 1 if the race is restricted to 2 year olds, = 0 otherwise
3YrOld (+) = 1 if the race is restricted to 3 year olds, = 0 otherwise
Claiming (-) = 1 if the race is a claiming race, = 0 otherwise
Allowance (-) = 1 if the race is an allowance race, = 0 otherwise
Distance (-) = 1 if the race is a sprint race, = 0 otherwise
Surface (-) = 1 if the race is on turf, = 0 if the race is on dirt
Purse (-) The size of the purse awarded to race winners
%Acct (+) Percent of the race handle bet via account

The number of starters is included due to the connection between the allocation of the betting pool and the odds. Assuming there is a bet on every entrant in a race (presumably every entrant stands some chance of winning the race in some bettor's eyes, e.g. the owner), more entrants mean that the maximum percent of the pool that can go to the favorite is lower. It also means added information is needed to handicap the race. Therefore, as the number of starters rises, the competition index should fall.

In harness races, there are two distinct gaits at which races are held. There are pacing races and trotting races. If there is any consistent difference in the competitiveness of the two types of races, the gait variable will have a significant effect on index.

Younger horses will have less racing experience. Racehorses usually begin their racing career at age 2 or 3. There are fewer young horses racing, especially early in the season, and therefore the conditions for their races must be written broad enough to attract enough entrants. The month of July is early in the racing season for horses beginning their career. The wide range of experience and ability resulting from the conditions written for early season races for young horses can lead to a wider dispersion of odds and therefore a higher index.

Claiming races for both Thoroughbred and harness races, and allowance races for Thoroughbreds are designed to make races more competitive. Therefore, the claiming race variable is expected to have a significant negative effect on the index, if claiming races achieve their purpose.

Stakes races are races that offer high purses in order to attract high quality horses. If races with top purses are more competitive, then purse should have a negative effect on index.

Premier races, such as stakes and invitational harness races (races designated as "Fastclass") and Thoroughbred turf races attract the highest quality entrants and are expected to be very competitive.

Existence of the favorite-longshot bias will cause the index to be lower, as bettors underbet the favorite (driving the odds up) and overbet longshots (driving the odds down). That is, the race odds will be less dispersed with a favorite-longshot bias. Research indicates the favorite-longshot bias depends on the consumption benefits of attending races as well as the number of informed bettors. (Sauer, 1998, Piron and Smith, 1995) The %Acct variable increases as more of the betting pool is contributed by account bettors. Account bettors place their bets via telephone or computer. They do not receive the consumption benefits from the race. In addition, bettors holding and using accounts are presumed to have greater information and spend more time handicapping races. As a greater percent of the race handle comes from these bettors, the favorite-longshot bias is less likely to exist. Ray (2001) found that higher live betting handles are associated with lower competition indices.

A significant, positive coefficient for the %Acct variable is evidence that the favorite-longshot bias may be decreasing the index when there is a greater percentage of live bettors determining odds.

The Thoroughbred equation also includes variables to account for race distance, track surface, and for allowance races. Harness races were all mile long races in this data set (as is generally the case for harness races) and the track had the same surface. The Thoroughbred races in the data set were run at varying distances and on either turf or dirt tracks. If the type of track or distance systematically affects handicapping, it could be shown by a significant effect on the index.

The Thoroughbred track runs allowance races, an additional class of race not applicable to harness races. In these races, racing secretaries use conditions and weight allowances to design competitive races. If the racing secretary is successful, the allowance races will have significantly reduced index values.

Data

The data used to estimate the harness equation are for all races held at a U.S. harness track during July 2000. Data for all variable except %Acct come form the United States Trotting Association (the organization that oversees U.S. harness racing). The %Acct variable was constructed from proprietary data from the harness track.

The data used to estimate the Thoroughbred equation are for all races held at a U.S. Thoroughbred track during July 2000. Data for all variable except %Acct come form the United States Jockey Club (the organization that oversees thoroughbred racing). The %Acct variable was constructed from proprietary data from the Thoroughbred track.

Results

The means and standard deviations of the variables are reported in Table 3.

Table 3 | Descriptive Statistics

Harness Variables	Mean	Standard Deviation
Index	.382	.112
#Starters	8.133	.951
Gait	.362	.481
FillyRest	.220	.415
2YrOld	.131	.338
3YrOld	.018	.132
Claiming	.418	.494
FastClass	.078	.269
Purse	6403	7212
%Acct	.447	.077
Thoroughbred Variables	Mean	Standard Deviation
Index	.321	.092
#Starters	7.88	1.834
2YrOld	.036	.187
3YrOld	.082	.275
Claiming	.564	.497
Allowance	.174	.380
Distance	.431	.496
Surface	.200	.401
Purse	15507	16610
%Account	.226	.043

The results for the harness equation are reported in Table 4. The results for the Thoroughbred equation are reported in Table 5. The F-statistic for both equations is significant at the 99% level. The R² values are acceptable for cross-sectional data of this type, but the R² indicates the Thoroughbred equation explained more of the variation in the competition index than the Harness equation.

Table 4 | Harness Equation**Variable Coefficient t-Statistic**

	Constant	.48	6.72***
	#Starters	-.017	-2.34**
Gait		-.041	-2.67***
	2YrOld	.072	3.18***
	3YrOld	.185	3.59***
	Claiming	-0.21	-1.28 ⁺
	FastClass	-.007	-.26
	Purse	2.65E-07	.25
	%Account	.125	1.64*
	F-Statistic		5.68***
	Adj R ²		.15
N		283	

*** Significant at the .01 level

** Significant at the .05 level

† Significant at the .10 level

* Significant at the .10 level, 1-tailed test

Table 5 | Thoroughbred Equation**Variable Coefficient t-Statistic**

Constant	.511	13.08***	
#Starters	-.025	-7.84***	
	2YrOld	.194	6.18***
	3YrOld	.010	.48
	Claiming	-0.024	-1.67*
	Allowance	-.030	-1.70*
	Distance	.010	.91
	Surface	.011	.78
	Purse	-2.6E-07	.45
			%Account .084 .65
			F-Statistic 11.99***

*** Significant at the .01 level

* Significant at the .10 level

The number of starters had a significant, negative effect on the competition index, as expected for both harness and Thoroughbred races. More entrants decreased the percent of the win pool bet on the favorite horses.

Ceteris paribus, the competition index for races held for young horses were significantly larger. Races for younger horses had stronger favorites and longer longshots. The competition index rose for 2 year-olds' races (both Thoroughbred and harness) and for 3 year-olds' harness races. The fact that 3 year-old harness horse's races had significantly higher competition indices and 3 year-old Thoroughbred's did not may be explained by the difference in information available to those betting on 3 year-old races. Standardbred horses (e.g. horses racing in harness races) tend to have longer careers and race more often. Thus, a three-year-old Standardbred has raced much more often and has more information available for bettors.

Claiming races, both harness and Thoroughbred, had significantly lower competition indices, ceteris paribus. This shows that Race Secretary's attempts to create competitive races through claiming races is successful. Further, Thoroughbred allowance races, set up by Race Secretaries to create competitive races are also effective for decreasing the competition index.

The distance and surface associated with Thoroughbred races did not significantly influence the distribution of odds in the race, as measured by the competition index. However, the gait variable for harness races was negative and significant. This means that trotting races were more competitive, while pacing races had more dispersed odds (higher competition indices). This may be explained by the larger number of pacing races. Bettors have more experience and information betting on pacing races. Also, pacers tend to be faster and more consistent than trotters, which may lead bettors to have more confidence about predicting the winner of pacing races. Breaking gait, a significant problem in a harness race, happens more often in trotting races than in pacing races. Therefore pacing races are easier to handicap (i.e. they are less often affected by breaks in gait). Finally, while trotting races make up about 36% of the races in this sample (but less than 20% of U.S. harness races), some prestigious harness races (such as the Hambletonian) are trotting races. Therefore trotting races can have very competitive harness horses.

In neither type of racing did offering higher purses (ceteris paribus) lead to more competitive races. Holding constant the purse for a race, Race Secretary efforts were what led to lower competition indices.

Finally, the %Acct variable was significant for the harness equation, but not for the Thoroughbred equation. For harness races, when a higher percent of the race handle was from account bettors (the more informed bettors who do not receive consumption benefits from attending races), the competition index was significantly higher. This means that there was a stronger favorite and longer longshot(s). This is evidence that live bettors on harness races may in fact contribute to a favorite-longshot bias by overbetting longshots and underbetting favorites, lowering the competition index of races.

The difference between the results for harness and Thoroughbred %Acct variables is likely due to the differences in the size of the two parts of the horse racing industry. The odds for Thoroughbred races may come from a more efficient market due to the increased availability of information and the larger number of potential entrants, races, and bettors.

The average competition index is lower for Thoroughbred races, as is the standard deviation. Race Secretaries at Thoroughbred tracks have more horses from which to fill a given race and can therefore write more specific race conditions. The larger number of races and bettors leads to a more informed market that is less susceptible to a favorite-longshot bias. The larger adj R² for the Thoroughbred

equation shows that the variation in the competition index is easier to explain using race characteristics (rather than horse-specific information) for Thoroughbred races than harness races.

Conclusions

The competition index, modeled after the Hirfindahl Index, is useful for measuring the competitiveness of individual horse races and evaluating the effect of race characteristics and conditions on competitiveness. The range for the competition index for a sample of 478 races held at two tracks during July 2001 (both harness and flat races) was from .17 to .91. The average competition index was lower, and less variable for Thoroughbred races.

While much of the difference in the competition index from race to race is due to differences in the individual characteristics of horses entered in the race, the model presented had a significant f statistic for both harness and thoroughbred races. The equation estimated to explain the variation in the competition indices of different races had more explanatory power for Thoroughbred races (an R² of .34 versus .15), indicating that race characteristics and conditions explain more of the variation in competitiveness of thoroughbred races. This is perhaps due to the larger number of Thoroughbred horses that enable Racing Secretaries to better write conditions to create competitive races.

These results indicate several ways that the race competitiveness can be increased. Increasing the number of horses per race is a direct way to reduce the competition index. Since high purses did not significantly affect the competition index (*ceteris paribus*), it may be possible to change the purse structure for races to increase the number of horses running in claiming races.

Increasing the number of races for older, experienced horses will lead to more races with higher competition indices, *ceteris paribus*. Again it may be possible to create more races that are more competitive by changing the way conditions are written and purses distributed so that more horses continue to race and have longer careers. This may also help to create more "celebrities" in horse racing by having competitors for enthusiasts to follow over a longer period of time, i.e. creating equine "Cal Ripken's."

The results also show that Race Secretary's efforts to increase competitiveness by writing conditions for claiming and allowance races are successful in significantly lowering the competition index. However, increasing purses and offering stakes and other top level races does not, by itself, lead to increased competitiveness. A race card with more claiming races with the right conditions (e.g. older horses, pacing races) and more entrants per race seems to be the way to have the greatest effect on competitiveness.

Finally, there is evidence that an increase in the number of live bettors (as compared to account bettors) on harness races decreases the competition index. This may be the result of a favorite-longshot bias in the betting market.

Bibliography

Piron, R. and Smith, L.R. (1995) Testing Riskloving in an Experimental Racetrack, *Journal of Economic Behavior and Organization*, 27, 465-74.

Ray, M. (2001) " Taking Handle into Account: An Economic Analysis of Account Betting." University of Arizona RTIP Working Paper. To be presented at the 2001 Symposium on Racing, Tucson, AZ, December, 2001.

Sauer, R (1998) The Economics of Wagering Markets, *Journal of Economic Literature*, 36, 2021-64.

Smith, S. (1998) *The Complete Idiots Guide to Betting on Horses*. New York: Simon and Schuster Macmillan.