An Examination of the Empirical Derivatives of the Favorite-Longshot Bias in Racetrack Betting*

Russell S. Sobel

Department of Economics
PO Box 6025
West Virginia University
Morgantown, WV 26506
rsobel2@wvu.edu

S. Travis Raines

Department of Economics
CB #3305
University of North Carolina-Chapel Hill
Chapel Hill, NC 27599
traines@unc.edu

Draft Date: March 17, 2000

Abstract

Market efficiency dictates it equally profitable to bet on any racing participant, including the favorite or longshot. However, a well-documented anomaly is that racetrack bettors tend to overbet longshots and underbet favorites. We present and test two theoretical explanations for this favorite-longshot bias. The unparalleled richness of our data allows us to explore how the bias changes with several key variables. We find the most popular current explanation for the bias, the risk preference model, cannot explain the data as well as an information-based model, in which the bias depends on bet complexity and the information possessed by bettors.

JEL Classifications: G14, D84, D44

Keywords: pari-mutuel betting, favorite-longshot bias, market efficiency, wagering

* The authors would like to acknowledge a research grant from the College of Business and Economics at West Virginia University. In addition, we have benefited from comments and discussions with Bill Trumbull, Ron Balvers, Thomas Garrett, Raymond Sauer, Gary Wagner, Ashley Lorance, and William Grant. Ziyu Luo provided research assistance. We are also indebted to Alex Tucker, comptroller, and Carl Meyers, racing secretary, of Wheeling Downs Racetrack and Gaming Center for help in data collections. All errors remain our own.

An Examination of the Empirical Derivatives of the Favorite-Longshot Bias in Racetrack Betting

1. Introduction

Market efficiency requires that betting on racetrack favorites is equally profitable as betting on the longshot, or any other racing participant. However, a common anomaly in the empirical literature is to find a "favorite-longshot bias" in which the return to betting on favorites is higher than the return to betting on longshots.[1] Previous authors have offered many possible explanations for the bias, with the most popular being risk-loving behavior on the part of bettors and the presence of poorly informed casual bettors at the track.[2]

While market efficiency implies equality of *expected return* across bets, risk-loving bettors themselves are not indifferent between bets of equal expected return if the bets have different variance. The higher variance provided by a bet on the longshot makes it a more attractive bet to a risk-loving bettor. The end result is that the expected return on longshots must fall, and the expected return on favorite bets must rise, until they provide equal *expected utility*.

The favorite-longshot bias is simply the result of too much money being bet on longshots, and too little on favorites, relative to their true chances of winning. The risk-loving model can explain this phenomenon, but an alternative is that casual, or less-informed, bettors simply bet too evenly across the racing entrants. One can envision a casual bettor whimsically selecting which bet to place by the name of the horse or even its number. Even legitimate attempts to digest the statistics presented in the racing program will result in errors that would not be made by regular or serious bettors having more familiarity with racetrack betting. This lack of information, or lack of ability to digest the information provided, results in the longshot being bet too heavily (overbet) and the favorite being underbet relative to the true probabilities of win.

In this paper we first derive specific, testable empirical predictions from both the risk-preference and information models. We then test these predictions on data from two greyhound racing tracks. Previous literature has repeatedly demonstrated the presence of this favorite-longshot bias, but has never made a serious attempt to see whether the bias changes predictably with specific variables that could differentiate between these possible explanations for the bias. However, examining these "empirical derivatives" of the favorite-longshot bias has the highest potential for producing a true understanding of this anomaly. For each race, we collect data for 5 different bets (win, place, show, quinella, and perfecta). We differentiate races by the grade of the race (6 track-determined quality levels). Most importantly, we also have information about the total volume of betting for the day, the track attendance, and the day of the week from which we can infer the relative attendance of less-informed, casual bettors. This unparalleled set of supplementary variables enables us to directly test our hypotheses and to differentiate between the risk and information models as the cause of the bias.

Our results show that an information model in which poorly informed bettors simply bet too evenly across the racing entrants explains the favorite-longshot bias much better than a risk-preference model. As predicted by the information model, the favorite-longshot bias changes significantly with the level of attendance of casual bettors at the track, the amount of prior performance information available for the entrants, the number of entrants in the race, and across bets of different complexity. One of our more interesting results is that well-informed serious bettors, when faced with simple bets, tend to display an *opposite* favorite-longshot bias. This result provides a possible explanation for why several studies, in contrast to the rest of the literature, have found an opposite bias in racetrack wagering.[3] It also may explain why sports betting and other markets can display an opposite bias to the one normally found in racetrack wagering.

2. Some Analytics of the Risk and Information Models

In this section we develop two models that summarize the main theoretical and empirical implications of the risk and information theories. We focus our derivations on several key differences in the predictions of these models for empirical testing.

The Risk Preference Model

The risk model has received most of the attention in the literature explaining the favorite-longshot bias. Here, we replicate the risk preference model of previous authors, then extend the model to explicitly express the relationship in terms of objective and subjective probabilities for empirical testing.

Let π_i denote the true probability of winning for racing entrant *i*. This is called the "objective probability of win" and can be estimated as the actual percent of races in which the entrant is observed to win. Let ρ_i denote the individual bettor's (or racetrack betting market's) estimate of the entrant's probability of win, which is known as the "subjective probability of win." Grifith (1949) and McGlothlin (1956) were the first to formally note that the subjective probability of win is the percent of the total betting pool that is actually bet on entrant *i*. Track odds for entrant *i*, O_i, are defined in terms of the betting pool share (the subjective probability) for that entrant, such that:[4]

[1] $$O_i = \frac{(1-\rho_i)}{\rho_i}$$

The track payout per dollar if the bet wins, $\$_i$, is given by $\$_i = O_i + 1$ so through substitution:

[2] $$\$_i = \frac{1}{\rho_i}$$

Rosett (1965) finds for a general class of continuous functions, that the equilibrium set is confined to the relationship $R = \lambda\pi^{\delta}$, where *R* is the return on a dollar bet and δ is a measure of the bettor's risk preference, which is a function of the bettor's underlying preferences and initial level of wealth. Here, $0 < \delta < 1$ implies risk-loving behavior, with $\delta < 0$ for risk-averters.

Notice that a bet pays $\$_i$ with probability π_i and nothing with probability $(1-\pi_i)$. The expected return, *R*, is thus $\$_i \pi_i$. Substitution produces: [5]

[3] $$\$_i = \lambda\pi_i^{\delta-1}$$

Combining equations [2] and [3], we can write the subjective probabilities in terms of the objective probabilities:[6]

[4] $$\rho_i = \psi\pi_i^{\beta}$$

Equation [4] allows us to show the favorite-longshot bias within the context of the risk model. If $0 < \beta < 1$ then a favorite-longshot bias exists with favorites underbet and longshots overbet. If $\beta > 1$ then an opposite favorite-longshot bias exists, with favorites overbet and longshots underbet. More formally, this result and its empirical implications are given in Proposition 1.

> **Proposition 1: (Risk Model)** Given full rationality and efficient markets, a preference for risk will produce a general relationship between the subjective betting probabilities (ρ) and objective win probabilities (π) of the form $\rho = \psi\pi^{\beta}$. Here, β is decreasing in the preference for risk, with $0 < \beta < 1$ for risk lovers and $\beta > 1$ implying risk aversion. The relationship can be expressed as: $\log(\rho) = \alpha + \beta \cdot \log(\pi)$, where $0 < \beta < 1$ implies a regular favorite-longshot bias, and $\beta > 1$ an opposite favorite-longshot bias.

The Information Model

Here, we develop an information model for risk-neutral bettors using Bayesian updating.[7] In this alternative model explaining the favorite-longshot bias, it is the information level of the bettor rather than risk aversion that determines the presence and extent of the bias.

Let us begin by assuming that prior to arriving at the track, each bettor has no knowledge about the racing entrants. In the absence of any information, the bettor must believe that all entrants are equally likely to win. Thus, the mean of the bettor's prior beliefs about the entrants probability of win, ρ_P, must be

[5] $$\rho_P = \frac{1}{N}$$

where N is the number of racing entrants. Once at the track, the bettor collects a set, k, of information signals regarding the racing entrants true probability of win. This information is generally provided to bettors in a racing program that is available for purchase. Regular or serious bettors will be able to utilize the information signals better than casual or infrequent bettors. Let ρ_I denote the mean probability implied by the information signals received by the bettor. Using Bayes' theorem, the mean of the bettor's posterior probability of win distribution, ρ, on which the bettor makes her bets, is given by:

[6] $$\rho = (1-h)\rho_P + h\rho_I$$

where *h* is a precision parameter that allocates the relative weight placed on the new information and the prior belief.[8]

Following Zellner (1987), the precision parameter $h = h(k, \sigma_P, \sigma_I, c)$ is a function of the information received, the variance of the priors and the information signals, and of the degrees of freedom (or complexity), c. Without assuming any specific form for h, the relationship with its arguments is intuitively obvious. $\frac{\partial h}{\partial k} > 0, \frac{\partial h}{\partial \sigma_I} < 0, \frac{\partial h}{\partial c} < 0$. This simply says that as a bettor becomes more information rich, she will choose to place less weight on her prior beliefs. However, if there is high variance in the new information, or the complexity of processing the information is great, then the bettor will place more weight on her prior beliefs.

Now we can build the information model with the subjective probability as a function of the true probability of win. Let the relationship between the entrant's true probability of win and the mean of the information signals conveyed to the bettor through the information be given by:

[7] $$\rho_I = \nu + \phi\pi$$

If the bettor's interpretation of the information signals is fully rational then $\phi = 1$ and $\nu = 0$. If bettors overreact to information, as in Thaler (1988), then $\phi > 1$.

Combining equations [5], [6], and [7] gives the posterior mean equation from which the bettor places her bets

[8] $$\rho = \alpha + \beta\pi$$

where $\alpha = \frac{(1-h)}{H} + h(\nu + \phi)$ and $\beta = h\phi$. Strong form market efficiency would be consistent with $\beta = 1$ and $\alpha = 0$. Once again, $0 < \beta < 1$ gives the traditional favorite-longshot bias, while $\beta > 1$ produces an opposite favorite-longshot bias, where favorites are overbet and longshots underbet. Since $\beta = h\phi$, this can occur if bettors overreact to information and the weight placed on the prior is small. This result and its empirical implications are given in Proposition 2.

> **Proposition 2: (Information Model)** The general relationship between the subjective betting probabilities (ρ) and objective win probabilities (π) from an information model with risk neutral agents is $\rho = \alpha + \beta\pi$. β is large if bettors overreact to information,

small if they possess little information, the variance of the information is large, or the more complex the relationship that must be estimated with a given information set. $0 < \beta < 1$ characterizes the favorite-longshot bias, while $\beta > 1$ implies an opposite favorite-longshot bias.

To summarize these points, let's look at several key differences between the risk and information models. Most noticeably, the risk model suggests a non-linear relationship between objective and subjective probabilities (or more precisely, a relationship that is linear in log form), while the information model predicts a simple linear relationship. Another key feature of these models comes from the nature of probabilities. Regardless of the type of bias, the probabilities over all entrants must sum to one. Thus, there must be at least one point, a "crossing point" on the zero to one interval where the subjective and objective probability lines intersect (where $\rho = \pi$). Above this probability entrants will either overbet or underbet with the opposite being true below that point. These two models provide different predictions about that crossing point. Substituting $\rho = \pi$ into the risk model produces a cross point at $\pi = \lambda^{1/(1-\beta)}$. The same substitution into the information model gives a cross point at $\pi = \frac{1}{N}$. So, the size of the racing pool *N*, (which is typically eight racing entrants in our data) determines the crossing point in the information model, whereas in the risk model the crossing point could occur at any value between 0 and 1 depending on the bettor's risk preference.

In an applied example, the information model predicts that the betting-market bias on a racing entrant of *given* true probability of win should change with the number of other contestants in the race. The risk model predicts that it would remain unchanged. For example, if two racing entrants both had a 20 percent true probability of win, but one was in a race with 6 entrants and the other in a race with 4 entrants, the information model predicts the bias would be different across the two races, while the risk model predicts the bias would be the same.

We also could extend these models to examine different types of bettors when faced with situations involving the same objective probabilities. In particular, it is a rather simple extension to show that when comparing the subjective betting probabilities of two groups of bettors, the relationship will be linear if caused by an information differential, but nonlinear if caused by a risk preference differential. With these empirical propositions, we now analyze the data to determine the merit of the risk and information models.

3. Data Description

The data analyzed in this paper consists of 2,799 races containing 22,392 racing entrants from two greyhound racing facilities in West Virginia, the Tri-State Greyhound Racing Park and Wheeling Downs.[9] The distance between these two tracks is sufficient to ensure virtually no overlap in the customer base. Because we use data from wagering at greyhound dog racing, instead of thoroughbred and harness horse racing, it is worth discussing the similarities and differences between dog and horse wagering.[10] The most important aspect, the structure of betting, is identical. The exact types of bets, determination of odds and payouts, information provided to the bettors, and legally mandated track-take percentages, are identical. One key difference is that because the cost of providing a dog race is lower, the sport can be profitably run with significantly lower attendance. The attendance at the tracks used in previous studies has generally been around 15,000 per night, while at these dog-racing tracks the average attendance is approximately 1,300. Betting data also suggest that serious bettors more heavily dominate the dog tracks studied here than the horse racing tracks used in most previous studies. At both of the tracks in this study, exotic, multiple-dog bets capture more than half of all bets, and as a percent of all bets are about 4 to 5 times higher than at an average horseracing track. Thus, in a relative sense, these tracks are more heavily dominated by well-informed regular bettors than the average horseracing track.

For each race in the sample data on 5 bets (win, place, show, quinella, and perfecta) were collected, along with several supplementary variables for each race. First, each race is graded by the quality of the dogs running in the race. The main grades are, from best to worst: AA, A, B, C, D, and M. Significant differences exist in the probabilities of win, the average quality level of the dogs, and the amount of previous performance information available to bettors across these different grades. Data is also collected regarding the total volume of betting for the day, the attendance, and the day of the week. At both tracks, races are offered on Monday evening, Wednesday evening, Thursday evening, Friday evening, Saturday afternoon, Saturday evening, and Sunday afternoon. In addition, Wheeling Downs also offers a Wednesday afternoon race.[11] The most important difference between days is that attendance is 2 to 3 times larger on weekends than during weeknights. Because a substantial portion of this analysis will rest on the difference in attendance between weekday nights and weekends, the next section explores this difference in more detail.

4. Weekdays versus Weekends: The Casual Bettor Effect

Less-informed, or what we call "casual," bettors tend to bet significantly less per person and also less heavily in the exotic betting markets than do regular, or "serious," bettors.[12] Thus, differences in betting volumes per person and the types of bets placed allows some insight into the relative attendance of these two groups across different nights at these tracks. We had expected that Friday and Saturday night races would attract a much larger share of casual bettors relative to weekday nights. The data in Table 1 strongly support that conjecture.

[Table 1 about here]

At both tracks, average betting per person per race is much higher on weeknights suggesting a higher proportion of serious bettors. For every dollar spent by the average weekend attendee, the average weekday attendee spent $1.10 at Wheeling Downs and $1.18 at Tri-State. Examining the extreme day comparisons show betting is almost 50% higher per person per race during some weeknights versus the busiest weekends.[13] Besides betting a larger sum of money on average, serious bettors also tend to concentrate more heavily on the exotic bets involving multiple dogs. The data show this as well, because as a percentage of all bets, the most exotic bet, the superfecta, is 31% higher on weeknights at Wheeling downs. The data in Table 1 support the notion that weeknight attendance is composed of a relatively higher proportion of serious bettors than weekends. The large influx of casual bettors can also be seen in the attendance numbers which show average attendance almost doubling on weekends.

Since different days contain different proportions of serious and casual bettors, the next step is to see whether betting on favorites (and longshots) differs significantly between days. It is worth noting at the outset that there is no reason to expect that the underlying probabilities for the dogs differ across nights. At both tracks the dogs cycle through the nights, usually having two to four days between appearances. Thus, a dog racing on Saturday night will generally race again on Wednesday evening, then next on Friday evening, then Monday evening, and so on.[14] Because there is no difference in the objective win probabilities, differences in subjective betting probabilities across nights truly reflect differences in the behavior of casual versus serious bettors when faced with the same situations, often involving the exact same dogs.

[Figure 1 about here]

Figure 1 shows the percentage of win bets placed on the favorite relative to the longshot by racing day. The graph is arranged by average nightly attendance and data are shown individually for each track. Across both tracks the same pattern is clear. As attendance rises, the share of bets placed on the favorite falls while the share placed on the longshot rises. On very low attendance weekday nights, eight times more is bet on the favorite to win than the longshot, compared to only five times as much on weekends. Figure 1 only shows the ratio of favorite to longshot betting, however this trend does occur individually for both the favorite and longshot. For example, on Saturday nights, favorites receive an average of 26.3% of win bets, yet on Wednesday nights they capture 29.2% of the bets. For longshots, the opposite occurs. Saturday evenings find an average of 5.4% of all win bets made on the longshots, while only 3.7% of all bets are for the longshot to win on Wednesday nights. Table 2 shows win bets for favorites and longshots over all possible days and other key statistics.

[Table 2 about here]

The data in Table 2 summarize the above findings. On the three weekday evenings bettors wager a much larger share on the favorite than do bettors on the weekends. Similarly, weekend betters place relatively more bets on the longshot.

[Figure 2 about here]

Figure 2 shows a plot of the weekday subjective betting percentages against the weekend subjective betting percentages for all racing entrants (not just the favorite and longshot). Looking at the data, it appears to be a near perfect linear relationship. If this is true, this suggests that differences in risk preference cannot be responsible for the favorite-longshot bias. While the linearity does not prove the information model, it does not lead us to reject it as a possible explanation of the favorite-longshot bias. If the information model is the true explanation, then we should expect that, since casual bettors with little information bet too evenly across all dogs, that the betting percentages should be a perfect linear rotation of the true probabilities toward the diffuse prior probability of 12.5%.[15] Regressing weekend betting percentages (Γ) as a function of weekday betting percentages (Λ) yields the following (standard errors in parenthesis):

Double-Log Risk Model:

$$\ln(\Gamma) = -0.331 + 0.831 \cdot \ln(\Lambda)$$

(0.024) (0.010)

Linear Information Model:

$$\Gamma = +0.017 + 0.868 \cdot \Lambda$$

(0.001) (0.009)

Figure 2 also shows these estimated equations. It is quite visible that the linear model provides a better fit of the data. The linear model's predictions of extreme values are much closer than the risk preference model's.[16] It is also important to note that the point at which the linear model would intersect a 45-degree line shows the probability where the rotation takes place. This probability is very near the diffuse prior of 12.5 percent.[17] Since the risk model allows for this rotation to occur anywhere, this finding does not damage the risk model, but it does provide evidence in favor of the information model.

To summarize this section, the data suggest that the presence of casual, less-informed bettors is higher on weekends than on weeknights. When looking at betting across these nights, betting on the favorite falls while betting on the longshot increases relative to weeknights. The evidence suggests that this is not due to a higher preference for risk in these bettors. We note, however, that the weekend betting percentages are a perfect linear rotation of the weeknight betting percentages around a diffuse, less-informed, prior. Casual bettors simply tend to bet more evenly across dogs than serious bettors, so any dog with a higher than average probability is less heavily bet by casual bettors, while any dog with less than average probability of win is more heavily bet.

5. Time of Day Effects: Risk and Bettor Wealth

A few previous studies have suggested that the favorite-longshot bias becomes more severe as the racing day progresses. This conflicts with most standard models of risk behavior. Traditional models suggest that as wealth declines throughout the day (because of the track take), agents should become less risk-loving, not more so as the previous studies have found. Thaler (1992) suggests that one explanation is Kahneman and Tversky's (1984) model of mental accounting, where bettors attempt to break even for the day by betting more heavily on the longshot.

To explore further evidence either for or against the risk theory, we first grouped the data into the first thirteen races and the final two to contrast betting behavior as has been done in previous studies. An immediate problem became apparent after viewing the data, making this grouping simply inadequate. As was mentioned earlier, races vary in grade from AA to M. The data suggests that favorites win a higher proportion of AA races than any other grade. In fact, the proportion of races that a favorite wins declines as the grade of the race becomes lower. This is relevant because in almost 90% of the days at both tracks the final two races were grade AA. Because favorites win a larger share of grade AA races, and the final race tends to be grade AA, it is natural to expect betting on the favorite to be higher in the final race irrespective of risk, wealth, or mental accounting. No previous authors have attempted to control for this important factor. To adjust for these grade effects, we performed a simple race level regression controlling for race number (a constant was omitted) and grade (AA was the excluded group). Figures 3a and 3b show the raw race averages and the averages after controlling for grade for the favorite and longshot, respectively for each and every race of the day.

[Figures 3a and 3b about here]

Both figures clearly show the importance of controlling for grade effects in the last few races. The raw data show that betting on the favorite rises significantly in the last race, which disappears after controlling for the fact that it is a grade AA race. Looking at the corrected data, betting on the favorite remains fairly constant throughout the racing day. The notable exception is in the final two races, where betting on the longshot reverses its slight downward trend and rises up again, especially in the last race of the day. This last race effect appears to be isolated to the longshot, with no similar last race effect happening for the favorite. However, again this evidence goes against the idea that risk preferences are a significant determinant of betting behavior. We know that, due to the track take, the average bettor's wealth declines throughout the evening. Yet, the patterns exhibited in Figure 3a and 3b do not show people moving toward more risk adverse bets, as traditional risk theory would indicate. In fact, the upward trend in longshot betting is supportive of the idea of mental accounting.[18]

One plausible explanation of the general trend over the last half of the evening is that attendance falls significantly through the evening. Track officials estimate that at most 60 percent of the peak attendance remains for the final race. It is reasonable to believe that casual bettors make up a disproportionately large number of those who leave early, implying that the later races are dominated by serious bettors that bet more heavily on the favorites. Thus, the pattern through the course of the racing day is more supportive of the information theory (casual/serious bettor differences) than of the risk theory.

6. Comparing Betting to Dog Performance: Rates of Returns to Bets

To this point, we have concentrated solely on comparisons of betting behavior, that is subjective probabilities, independent of the true, objective probabilities of win. We have been able to develop conclusions about the difference in the betting behavior of casual and serious bettors. To this point, the raw data across days and throughout the races of a day are not supportive of the risk model, but instead are consistent with the predictions of the information model. However, the subjective betting probabilities must be compared with the objective probabilities to know whether or not favorites (or longshots) are overbet or underbet relative to the standard of market efficiency.

The objective probabilities of win were computed individually for several subsets of the data and we performed tests to determine whether these objective probabilities differed across groups. Races were divided into groups corresponding to the grade of the race. The objective probabilities differed significantly across grades, so we will maintain this division throughout, where possible. No significant differences were found among other subsets based upon day of the week, weekday verses weekend, and race number (controlling for grade). From this, we gathered that objective probabilities of win do not change across these subsets, as should be expected since dogs rotate through the racing days and race numbers. Also, the objective probabilities did not significantly differ between the two tracks, suggesting that there will be no loss of generality from pooling the data.

Table 3 shows the subjective (betting) probabilities versus the objective probabilities in the win market for all of the pooled races, and separately for weekdays and weekends. Dogs are ranked from first favorite down to the eighth favorite (which is the longshot). The test statistic for the null hypothesis that the subjective probability is the same as the objective probability is shown in the table. Additionally, we provide the realized rates of return for bets on each dog.

[Table 3 about here]

Once again the data in the table is consistent with the ideas presented this far. Looking at the data, there is an opposite favorite-longshot bias, one that is very pronounced on weekday evenings. Betting on the longshot to win produces positive rates of return in both weekdays and weekends, though the return is much greater on the weekdays. The rate of return from betting on the favorite is profoundly negative, reaching levels of negative 34 percent, much larger than the expected loss implied by the track take of 17 percent. On weekends, the rate of return to favorites is still negative, though not nearly as much.

[Figure 4 about here]

Now we explore how this bias changes with respect to the grade (quality level) of the race. Figure 4 shows the favorite's probability of win in each racing grade both for the pooled data and each track individually. The win probabilities of the favorite across racing grades follow an identical pattern at the two tracks. It is important to note that these grade-level differences are not just random fluctuations across grades, but the true probability differentials across grades that exist at both tracks. With the true probabilities differing across grades, so do the optimal betting percentages. Figures 5a and 5b show the average subjective and objective win probabilities by racing grade for the favorite and longshot respectively.

[Figures 5a and 5b about here]

When comparing the differences in the subjective probabilities across weekday nights and weekends, it appears as if the betting line is almost shifted in a parallel fashion. While not shown, when the data are further subdivided by individual days, the betting lines form almost a perfect continuum of parallel lines. The only divergence from this is in grade D on the favorite. When calculated for each track separately, however, the Wheeling Downs data for grade D appears shifted just like in the other grades (shown by the dashed line). The data for the Tri-State track that pull the weekend betting on grade D upward, implying that this is probably due to randomness in the data.

The most important comparison is between the subjective and objective probabilities. In the highest grade, AA, there is an opposite favorite-longshot bias on both weekends and weekdays. In grade M, which is the entry-level race for new dogs, a regular favorite-longshot bias exists on both nights, but most significantly on weekends. The very critical points made by these data are that in some grades an opposite bias exists, while in others the standard bias exists. Had all of our races been of the lower grades D and M, our overall data presented earlier would show a strongly significant underbetting of favorites and overbetting of longshots, just as most previous horseracing studies have found. It is in these markets with the *least* prior information available to bettors that the bias is most heavily skewed toward a regular favorite-longshot bias. That is, longshots tend to be more overbet (and favorites less underbet) when less good information is available to the bettors. In the top AA grade, where there is good, reliable past performance information and consistently performing dogs are present, *the same bettors, on the same evenings*, tend to bet too heavily on the favorite, producing an opposite favorite-longshot bias.

To summarize this section, we have shown that a significant opposite favorite-longshot bias exists in overall win betting at these tracks. Here, the return to betting on longshots is greater than the return to betting on favorites. This bias becomes much more pronounced on weekday nights that are dominated by serious bettors. However, this opposite bias can disappear, and in fact reverse, depending on the race grade, specifically grades D and M, the races with the least good information available to bettors. It is races in the highest grades with the best information that tend to contain the opposite bias that is pulling the pooled results presented earlier.

7. Evidence from Other Bets

Given our results from the win market, we briefly explore some other bets to see if our predictions for the favorite-longshot bias hold. The other bets in this section are the place, show, quinella, and perfecta bets. According to Thaler (1992) bettors have problems in estimating the complex relationships for these more exotic bets. As a result, if the information model is correct, we should see that the biases turn toward the typical favorite-longshot bias, as the bets get more complex, because as a bet becomes more complex even serious bettors become relatively uninformed. That is, higher bet complexity is analagous to having less informed agents present.

[Table 4 about here]

The place bet is a bet on a single dog that pays if the dog comes in either first or second. The show bet is similar, except it pays if the dog finishes first, second, or third. Table 4 shows the objective and subjective, as well as the rate of return for the place and show markets. Overall, these markets tend to show more of a traditional favorite-longshot bias than the win market, especially the show market. It is easy to notice that the average rate of return in the place and show market is higher than the average return in the win market. Asch, Malkiel, and Quandt (1984, 1986) claim this comes from the mandated minimum payout of $2.20 on a $2 bet. This constraint is frequently binding in the place and show markets.

Finally, to check the validity of the information model, we examine these more complex bets to see if the properties of the information model hold true. Table 5 shows the results of the simple regressions for the win, place, show, quinella, and perfecta markets. Quinella and perfecta bets are bets over a combination of two dogs. For the quinella, a bettor selects two dogs. If they finish first and second, in either order, the bet pays. For perfecta, sometimes called exacta, the bettor selects two dogs, but they must finish in the exact order selected.[19] Remember, a slope coefficient being greater than one implies an opposite favorite-longshot bias, while less than one is indicative of the traditional favorite-longshot bias. A slope of one would be consistent with efficiency.

[Table 5 about here]

From Table 5, we can see that a significant regular favorite-longshot bias is present for the quinella and perfecta bets. Table 5 clearly shows that the slope coefficients move in favor of the regular favorite-longshot bias (the coefficient declines in value) as the complexity of the bet increases. This is precisely the prediction of the information model.

We also see that the rotation points are almost identical to what would be predicted by the information model. This suggests that the cause of the favorite-longshot bias is that bettors are simply betting too evenly across dogs as the bets become more complex. As the bets become more complex, dogs that have above average probabilities will become increasingly underbet, while those with below average probabilities will become more overbet.

This data is very damaging for the risk model. The least risky market (the show bet) and the most risky market (perfecta) both demonstrate the usually favorite-longshot bias, while the mid-level risk markets have either an opposite bias or none at all. Also, the risk model predicts a non-linear relationship between the objective and subjective probabilities. Yet, this relationship is strongly rejected in favor of the linear model.[20]

[Figures 6a and 6b here]

Perhaps the most striking visual demonstration of the rejection of the risk model to explain the data comes from the quinella and perfecta bets. Figure 6a shows the rates of return from each of the 28 possible quinella combinations. A significant regular favorite-longshot bias is present for the quinella bet, and the perfecta bet has a pattern identical to the one for the quinella bet. The relationship in the figure is strikingly similar to the figures presented in previous papers.[21] While the relationship between objective probability and return is quite nonlinear, when the data are properly viewed as subjective betting probabilities versus objective betting probabilities the picture changes dramatically, as is shown in Figure 6b. Again, the relationship in the perfecta market was identical, so only the quinella is shown. The relationship is almost perfectly linear, in rejection of the risk explanation that implies a nonlinear logarithmic relationship between objective and subjective probabilities. For comparison, the fitted regression lines for the risk and information models are shown in the figures. The linear relationship between objective and subjective probabilities from the estimated information model regression is much closer to the data points than the estimated double-log risk model. Just as for the win data shown earlier, the difference is most noticeable in the tails. The nonlinear relationship between the rate of return and objective probabilities shown in Figure 6a is also more closely approximated by the information model. It is important to note that while the information model predicts a linear relationship between objective and subjective probabilities, that this implies a nonlinear relationship between the probabilities and the rate of return. Even in the rate of return data, the information model's nonlinear pattern fits the data better than the nonlinear pattern produced by the risk model, especially for the upper tail where the model's predictions differ most. As might be expected from the raw data graph, a Box-Cox test strongly rejects the double-log model in favor of the linear model for this data (at a 1% level of significance).

One final test of the information model involves data that was originally excluded from our analysis. Dogs can become disqualified from racing, leaving only seven dogs to race. Previous studies have simply pooled all races, regardless of the number of entries together. However if the information model holds true, the diffuse prior probability for each dog in a seven-dog race should be 1/7, and thus the bias should differ from the races with eight dogs However, one problem that arises is there were not enough races with only seven dogs to gain precise estimates of the objective win probabilities or to perform the analysis by grade. Thus, we proceed as we did in the early sections of this paper for the win market by comparing weekday night and weekend subjective betting probabilities to gain evidence of the rotation point.

[Table 6 about here]

Table 6 shows these estimates for both races with eight and seven dogs using the above technique. While the relatively small number of races with seven dogs limits precision, the estimates still support the information model. Again, when only seven dogs race, casual bettors bet too evenly across all dogs. Any dog whose probability of win is less than the diffuse prior is bet on more heavily by casual bettors. This is the strongest evidence yet against the risk preference model, as it would predict that a bet with a given true win probability would be bet on equally regardless of the number of other dogs in the race. However, as the number of dogs racing declines, the rotation point does change, which is a prediction of the information model but *not* the risk model. It, in fact, is a rejection of the risk model because that model would suggest the rotation point remain unchanged. In a race with 8 entrants, a dog with a 13.5% true probability of win has an *above* average probability of win, and is thus underbet by less-informed bettors. In a race with 7 dogs however, a dog with a 13.5 percent true probability of win has a *below* average probability of win and is thus overbet by less informed bettors. Only the information model has a prediction consistent with this data. It also points to a problem in previous studies that typically pool racing data across races with different numbers of entrants.

In summary, this section has shown that the favorite-longshot bias differs across different bets present at the track, even though they are offered at the same time and bettors may participate in any bet. In particular, across the place and show bets, just as for the win bet, the higher relative attendance of less-informed casual bettors on weekends results in more even betting across the dogs, causing the bias to move more toward a regular favorite-longshot bias. This holds despite the fact that the original biases were different across the bets. In addition, the bias shows a predictable relationship with the complexity of the bet. As bets become more complex, the bias moves more toward a regular favorite-longshot bias. Following the information model, this is simply the result of the increasing complexity of the bet essentially makes bettors less informed. The lower degree of information relative to the level necessary to accurately judge the true probabilities results in more even betting across the dogs. Finally, this section has provided further evidence from races where one dog was eliminated from the race, leaving only seven dogs instead of eight to race. The effect of this on betting percentages was precisely in line with the information model's prediction that the rotation would be around the prior of 1/7 th instead of 1/8 th.

8. Conclusions

In an attempt to better understand the causes and nature of the favorite-longshot bias, this paper has effectively demonstrated several "empirical derivatives" of the bias. The bias changes in accordance with the predictions of an information-based model, but not a risk-based model. As the proportion of causal (less-informed) bettors at the track expands, the share of the betting pool bet on the longshot rises, moving the market further in the direction of a favorite-longshot bias. In addition, on any given night, more complex bets tend to demonstrate more of a favorite-longshot bias because higher bet complexity functions similar to having less informed bettors.

The fact that the relationship between the subjective probabilities of bettors and the true probabilities for the racing entrants is linear, rather than nonlinear, forms a strong argument for discounting the risk model in favor of an information model. In addition, the critical "crossing point" probability at which entrants on one side are overbet while on the other side are underbet changes with the number of bets (or entrants). This is consistent with the betting being rotated around a diffuse prior of 1/N, not with a risk-based model of utility. Essentially, casual bettors bet too evenly across the possible outcomes thus underbetting the favorite and overbetting the longshot. This same effect is present on very complex bets where it is hard for bettors to understand critical relationships.

One of our interesting findings is that on nights dominated by more serious (better-informed) bettors, there is an opposite favorite-longshot bias on the simpler bets, particularly in the highest grade AA races. This is consistent with the idea that well-informed bettors may overreact to information, a theory proposed in Thaler's research. If so, our findings may be a critical link between the literature on racetrack wagering and the few studies that have found opposite biases in racetrack betting, as well as the literature on sports betting (NFL, NBA, etc.) and stock market forecasting where an opposite bias generally dominates.

Opportunities to confirm this hypothesis also exist in several other markets. The literature shows that these types of biases not only exist in the markets for sports wagering and the stock market, but also applications to the winner's curse and the market for lemons. In addition, we see far-reaching implications to models where consumers face imperfect information about quality or price differences across firms, and for public economics' models of rent seeking, voter information, and electoral competition.

References

Ahlers, David and Lakonishok, Josef. "A Study of Economists' Consensus Forecasts." *Management Science*, October 1983, 29, 1113-25.

Ali, Muktar M. "Probability and Utility Estimates for Racetrack Bettors." *Journal of Political Economy*, 1977, 85, 803-15.

Ali, Muktar M. "Some evidence on the Efficiency of a Speculative Market." *Econometrica*, 1979, 47, 387-92.

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

Asch, Peter, Malkiel, Burton G. and Quandt, Richard E. "Market Efficiency in Racetrack Betting." *Journal of Business*, April 1984, 57, 165-75.

Asch, Peter, Malkiel, Burton G. and Quandt, Richard E. "Market Efficiency in Racetrack Betting: Further Evidence and a Correction." *Journal of Business*, January 1986, 59, 157-60.

Asch, Peter, and Quandt, Richard E. "Efficiency and Profitability in Exotic Bets." *Economica*, August 1987, 54, 289-98.

Busche, Kelly and Hall, Christopher D. "An Exception to the Risk Preference Anomaly." *Journal of Business*, 1988, 61(3), pp. 337-46.

De Bondt, Werner F.M. and Thaler, Richard H. "Do Security Analysts Overreact?" *American Economic Review Papers and Proceedings*, May 1990, 80(2), pp. 52-7.

Figlewski, S. "Subjective Information and Market Efficiency in a Betting Model." *Journal of Political Economy*, 1979, 87, pp. 75-88.

Gandar, John, Zuber, Richard, O'Brien, Thomas and Russo, Ben. "Testing Rationality in the Point Spread Betting Market." *Journal of Finance*, September 1988, 43(4), pp. 995-1008.

Golec, Joseph, and Tamarkin, Maurry. "Bettors Love Skewness, Not Risk, at the Horse Track." *Journal of Political Economy*, February 1998, 106(1), pp. 205-25.

Grifith, R. M. "Odds Adjustments by American Horse-Race Bettors." *American Journal of Psychology*, April 1949, 62, pp. 290-4.

Harville, David A. "Assigning Probabilities to the Outcome of Multi-Entry Competitions." *Journal of the American Statistical Association*, 1973, 68, pp. 312-6.

Hausch, Donald B., Ziemba, William T. and Rubinstein, Mark. "Efficiency of the Market for Racetrack Betting." *Management Science*, December 1981, 27(12), pp. 1435-52.

Hurley, William and McDonough, Lawrence. "A Note on the Hayek Hypothesis and the Favorite-Longshot Bias in Parimutuel Betting." *American Economic Review*, September 1995, 85(4), pp. 949-55.

Kahneman, Daniel and Tversky, Amos. "Choices, Values, and Frames." *American Psychologist*, 1984, 39, pp. 341-50.

Losey, Robert L. and Talbott, John C. "Back on Track with the Efficient Markets Hypothesis." *Journal of Finance*, September 1980, 35, pp. 1039-43.

McGlothlin, W. H. "Stability of Choices among Uncertain Alternatives." *American Journal of Psychology*, December 1956, 69, pp. 604-15.

Rosett, R. N. "Gambling and Rationality." *Journal of Political Economy*, December 1965, 73(6), pp. 595-607.

Sauer, Raymond D. "The Economics of Wagering Markets." *Journal of Economic Literature*, December 1998, 36(4), pp. 2021-64.

Shin, H. S. “Optimal Betting Odds Against Insider Traders.” *The Economic Journal*, September 1991, 101(3), pp. 1179-85.

Snyder, Wayne W. "Horse Racing: Testing the Efficient Markets Model." *Journal of Finance*, September 1978, 33(4), pp. 1109-18.

Swidler, Steve and Shaw, Ron. "Racetrack Wagering and the 'Uninformed Bettor': A Study of Market Efficiency." *The Quarterly Review of Economics and Finance*, Fall 1995, 35(3), pp. 305-14.

Terrell, Dek. “Biases in Assessments of Probabilities: New Evidence from Greyhound Races.” *Journal of Risk and Uncertainty*, November 1998, pp. 158-66.

Terrell, Dek and Farmer, Amy. “Optimal Betting and Efficiency in Parimutuel Betting Markets with Information Costs.” *The Economic Journal*, July 1996, 106(2), pp. 846-68.

Thaler, Richard H. and Ziemba, William T. "Anomalies: Parimutuel Betting markets: Racetracks and Lotteries." *Journal of Economic Perspectives*, Spring 1988, 2(2), pp. 161-74

Thaler, Richard H. *The Winner's Curse: Paradoxes and Anomalies of Economic Life*. Princeton, NJ: Princeton University Press, 1992.

Tuckwell, Roger H. "The Thoroughbred Gambling Market: Efficiency, Equity, and Related Issues." *Australian Economic Papers*, June 1983, 22, pp. 106-18.

Weitzman, M. "Utility Analysis and Group Behavior: An Empirical Study." *Journal of Political Economy*, February 1965, 73(1), pp. 18-26.

Williams, L. V. and Paton, David. “Why are some Favourite-Longshot Biases Positive and Others Negative?” *Applied Economics*, July 1998, 30(3), pp. 1505-10.

Woodland, Linda M. and Woodland, Bill M. "Market Efficiency and the Favorite-Longshot Bias: The Baseball Betting Market." *Journal of Finance*, March 1994, 49(1), pp. 269-79.

Zellner, Arnold. *An Introduction to Bayesian Inference in Econometrics*. Malabar, FL: Robert E. Drieger Publishing, 1987.

Zuber, Richard A., Gandar, John M. and Bowers, Benny D. "Beating the Spread: Testing the Efficiency of the Gambling Market for National Football League Games." *Journal of Political Economy*, August 1985, 93(4), pp. 800-6.

Endnotes

[1] See Sauer (1998) and Thaler and Ziemba (1988) for excellent overviews of this literature. Some of the most heavily cited studies include Ali (1977, 1979); Asch, Malkiel, and Quandt (1982, 1984, 1986); Busche and Hall (1988); Figlewski (1979); Hauch, Ziemba, Rubinstien (1981); Losey and Talbot (1980); Swidler and Shaw (1995); Snyder (1978); and Tuckwell (1983).

[2] For examples of possible explanations, see Rosett (1965), Weitzman (1965), Ali (1977), and Golec and Tamarkin (1998). The variety of explanations includes bettors getting direct utility from betting on longshots, a systematic tendency for people to overestimate the chances of low probability outcomes, bettors having preference over the skewness of payouts, or a model that suggests bias is inherent in any odds-based parimutuel system.

[3] Swindler and Shaw (1995) and Busche and Hall (1988), as well as several studies on professional sports betting, have found an exactly *opposite* favorite-longshot bias, with favorites being overbet and longshots underbet. These racetracks arguably consisted of a much higher proportion of serious bettors than the average racetrack. One was a smaller "Class II" racetrack in Texas, and the other a racetrack in Hong Kong that had much higher betting volumes than the racetracks used in these other studies. Other markets with biases analogous to an opposite favorite-longshot bias also appear to involve well-informed participants dealing with more complex issues, such as betting on NFL, NBA and MLB games, and survey forecasts of economists and stock market earnings forecasters. See, for example, Gandar, Zuber, O'Brien and Russo (1988), Sauer (1988), Woodland and Woodland (1994), De Bondt and Thaler (1990), and Ahlers and Lakonishok (1983).

[4] To simplify the derivation, we have ignored the track take, *t*. Including the track take, the true track odds are $O_i = \frac{(1-p')}{p'}$ where $p' = p \cdot (1-t)$. For all bets in this study, all bets are \$2 and the necessary adjustments have been made for calculations.

[5] An alternative method of deriving this relationship between the objective and subjective probabilities is to begin with the equilibrium condition of Ali (1977). Here, the equilibrium set is of the form $U(1+O_i) = \frac{\omega}{\pi_i}$, where ω is the favorite's subjective probability of win and $U(\cdot)$ is the bettor's utility function. Combining these models allows us to form the basis for our risk model, that $U(1+O_i)=\omega/\pi_i$. Ali estimates his equation by imposing a general form on the utility function of $U=\alpha Y^{\gamma}$ where Y is the bettor's wealth. The parameter γ measures the risk preference of the bettor, where $\gamma>1$ implies risk-loving, $\gamma<1$ implies risk-aversion, and $\gamma=1$ implies risk-neutrality. His empirical estimates are $\alpha=0.28$ and $\gamma=1.18$ for his racetrack sample. Continuing by substitution of the utility function gives, $\alpha(1+O_i)^{\gamma}=\omega/\pi_i$ and by substitution for the odds yields $\alpha(1/\rho_i)^{\gamma}=\omega/\pi_i$. Through simplification this may be rewritten as $\rho_i^{\gamma}=(\alpha/\omega)*\pi_i$. Allowing $\lambda=(\alpha/\omega)^{-\gamma}$ and $\beta=-\gamma$ produces the general form of the equation, $\rho_i=\lambda\pi_i^{\beta}$. Despite the differing starting points, Ali and Rosett's models both simplify to the same general relationship between objective and subjective probabilities.

[6] Where $\psi = \lambda^{-1}$ and $\beta = 1 - \delta$.

[7] Other information models for racetrack betting have been derived by Williams and Paton (1998), Shin (1991), and Terrell and Farmer (1996). These models differ from ours because they build and test their models for fixed odds, as opposed to bet-dependent subjective odds.

[8] See Zellner (1987) pages 13-15 for a derivation and explanation of the Bayesian information model and precision parameter.

[9] These data cover the period from March 1, 1996 to July 14, 1997 for Wheeling Downs and October 19, 1996 to June 26, 1997 for Tri-State.

[10] Terrell (1998) also considers Greyhound Racing and finds support for several biases that were originally found by other authors in horseracing.

[11] Tri-state also offered one Monday afternoon race for Memorial Day.

[12] More "serious" betters also tend to bet more heavily on bets with less visible odds, see Asch and Quandt (1987).

[13] This result is large despite an inherent bias in the raw data in the opposite direction. Betting per person is understated for weekday nights because attendance is measured as the number of persons who enter the track and many bettors leave early on weeknights. Track officials estimated that in the final one-fourth of a weekday night's races attendance is probably between 50 and 60% percent of the peak level, implying that the data understate the difference in per person betting by about 20 percent.

[14] In addition, empirical testing confirms that there is no statistically significant difference in the objective or true win probabilities, nor the average racing time of the winning dog, across nights or across the two tracks.

[15] For races with 8 dogs. The true diffuse prior is 1/N, as discussed earlier.

[16] In addition, the Box-Cox test rejected the double-log specification at a 1% level of significance.

[17] This is derived from the intersection of a 45-degree line, which gives a probability of:

$$\frac{0.017}{1 - 0.868} \approx 12.9\%$$

[18] Supporting evidence for these conclusions are found in other bets (place and show), as they exhibit the same pattern.

[19] Perfecta betting data was only available for Wheeling Downs.

[20] The double log specifications of the models presented in the table all have lower R-squares, significantly heteroskedastic residuals, and a Box-Cox test rejected the log specification in all cases. The figures show the superiority of the linear relationship most evident in the tails.

[21] See Thaler and Ziemba (1988), figure 1, and Snyder (1978), figure 1.