CHAPTER 8

Unifying the Favorite-Longshot Bias with Other Market Anomalies

Russell S. Sobel
Department of Economics, West Virginia University, USA.

Matt E. Ryan
Department of Economics, West Virginia University, USA.

1. Introduction 138
2. Biases Found in the Previous Literature 138
3. What Causes the Favorite-Longshot Bias at the Racetrack? 140
   3.1. The Casual Bettor 143
   3.2. The Serious or Regular Bettor 144
   3.3. The Arbitrageur 148
4. Is it Risk or Information? 150
5. Can the Model Explain the Biases in Other Markets? 156
6. Conclusion 157
References 158
1. INTRODUCTION

The favorite-longshot bias in racetrack wagering is perhaps the most well-documented anomaly in the literature on betting market efficiency. Dozens of studies have shown a consistent bias in which the expected return from betting on favorites exceeds the expected return from betting on longshots. This chapter focuses on the bigger picture with respect to this anomaly, namely to what extent it can be viewed as a bias that can be integrated with the differing biases found in other financial and betting markets—biases that are often in the opposite direction.

This chapter begins by reviewing the empirical findings on how the favorite-longshot bias differs across different types of bets, and different betting markets. The chapter then proceeds by discussing the different types of bettors, and also the different possible theories that have been presented to explain this bias, and whether they can be viewed in a broader context to provide a consistent explanation for these many biases. Finally, this chapter discusses the issue of arbitrage, and specifically why arbitrage may not eliminate these differential returns across bets.

2. BIASES FOUND IN THE PREVIOUS LITERATURE

Economists have spent considerable effort exploring the efficiency of market outcomes. Betting markets, because of their unique characteristics, provide an interesting case in which to conduct empirical studies of efficiency. At the top of the list of favorable characteristics is the outright abundance of data available—a facet that makes analyzing the stock market fruitful as well. However, Thaler and Ziemba (1988) note distinct advantages that betting markets have over financial markets:

The advantage of wagering markets is that each asset (bet) has a well-defined termination point at which its value becomes certain. The absence of this property is one of the factors that have made it so difficult to test for rationality in the stock market. Since a stock is infinitely lived, its value today depends both on the present value of future cash flows and on the price someone will pay for the security tomorrow. Indeed, one can argue that wagering markets have a better chance of being efficient because the conditions (quick, repeated feedback) are those which usually facilitate learning.

Despite the feedback and large-scale repeated-play aspect of betting markets, the search for market efficiency has led to the identification of an unexpected, and persistent, anomaly known as the favorite-longshot bias. In practice, weak-form market efficiency would imply that betting on racetrack favorites should be equally profitable as betting on longshots. Dozens of empirical studies of racetrack betting, however, have consistently found that the return to betting on favorites is higher than the return to betting on longshots. While Thaler and Ziemba (1988) provide perhaps the best overview of this body of literature, some of the more influential individual studies include Ali
Russell S. Sobel and Matt E. Ryan

(1977, 1979), Asch et al. (1982, 1984, 1986), Busche and Hall (1988), Figlewski (1979), Hausch et al. (1981), Losey and Talbott (1980), Snyder (1978), Swidler and Shaw (1995), Tuckwell (1983), and Ziemba and Hausch (1986). In some particular instances, such as discrepancies in win bets versus place and show bets (Hausch et al., 1981) as well as in exacta markets (Ziemba and Hausch, 1986), the return differential has actually been strong enough to produce deviations from weak-form market efficiency as well, in the form of profitable (positive expected value) wagers.

This unexpected bias identified in betting markets, while present in the vast majority of studies, does not uniformly hold true. Several notable studies have, in fact, found an exactly opposite bias, with favorites being overbet and longshots under-bet, an opposite favorite-longshot bias. Swidler and Shaw (1995) find this opposite bias for a smaller Class II racetrack in Texas, and Busche and Hall (1988) find this opposite bias at a racetrack in Hong Kong that had much higher betting volumes than the US racetracks used in other studies. More recently, Sobel and Raines (2003) have identified a situation in which the bias changes from the regular favorite-longshot bias to an opposite bias, under certain predictable conditions. Gramm and Owens (2005) find that the favorite-longshot bias diminishes for races with larger betting pools and more entrants. While these studies might appear at odds with the other literature on racetrack wagering, their results are actually more consistent with findings from studies of efficiency in markets other than racetrack wagering.

Studies of Las Vegas sports wagering have concentrated on betting on National Football League (NFL), National Basketball Association (NBA), National Hockey League (NHL), and Major League Baseball (MLB) games. In these markets, deviations from market efficiency have been found in about three-fourths of the studies, and in the vast majority of cases where it existed, it was an opposite favorite-longshot bias. In the terminology of this literature, bettors tend to overvalue favorites. Gandar et al. (1988) show that a strategy of betting on NFL longshots produces a return greater than betting on favorites for point spread betting on NFL games. Kochman and Badarinathi (1992) found through a simple analysis of role (underdog or favorite), location of the game (home or away), and month in which the game was played, that a rate of success above break-even could be achieved for wagering on NFL games between 1986 and 1990. Zuber et al. (1985) found an exploitable inefficiency in NFL point spread betting market during the 1983 regular season, while Lacey (1990) finds some profitable opportunities in certain betting rules in NFL point spreads from 1984 to 1986. In professional baseball, Woodland and Woodland (1994) find an opposite favorite-longshot bias in baseball betting against the odds, which they confirm still exists after including 10 additional years of data in their updated study, Woodland and Woodland (2003). In their study of NHL betting markets, Woodland and Woodland (2001) find a strong opposite favorite-longshot bias in baseball betting against the odds, which is confirmed in an updated and corrected version of this study by Gandar et al. (2004). Similar to the case of the regular favorite-longshot bias, these opposite favorite-longshot bias findings in professional sports betting markets are not always found. Sauer (1988), for example, finds efficiency in NFL over/under betting, while Johnson and Pawlukiewicz (1992) find efficiency in over/under betting in the NBA.

In the stock market, there are several findings related to the studies of betting summarized above. The first is the finding that portfolios of recent loser stocks seem to
outperform portfolios of recent winner stocks. To the extent that recent winner stocks are comparable to racetrack favorites, this is an opposite favorite-longshot bias. Several of the articles demonstrating (and providing explanations for) this bias in the stock market include De Bondt and Thaler (1985, 1987, 1990), Chan (1988), Bremer and Sweeney (1991), Brown and Van Harlow (1988), Lehmann (1990), and Howe (1986). The best overall summaries of the literature on stock market efficiency are De Bondt and Thaler (1989) and Thaler (1992).

More closely related is the finding of De Bondt and Thaler (1990) that professional stock market forecasters’ predictions about company earnings tend to have an opposite favorite-longshot bias. They regress actual company earnings on forecasted earnings and find a slope coefficient significantly less than one. This implies that the companies with the highest forecasted earnings tended to be over-estimates of true earnings, while the companies with the lowest forecasted earnings tended to be underestimates of true earnings. In studies of betting behavior, this type of bias would result in favorites being overbet and longshots being under-bet. This phenomenon also manifests itself in variance comparisons, as the variance in the predicted values is significantly higher than the variance in the true-underlying values.

Perhaps most ironically, while economists have struggled to find an explanation for these biases, a similar opposite favorite-longshot bias has been found to exist in professional economists’ own economic forecasts in the Livingston survey by Ahlers and Lakonishok (1983).

The link between deviations from market efficiency in these many markets has not been clearly made in the previous literature. They have been largely viewed as separate anomalies, each with its own potential explanation. However, the key to unlocking the mystery of why human behavior results in departures from market efficiency may lie precisely in developing a unified theory that simultaneously explains the phenomenon occurring in these different markets. In fact, with so-called anomalies, or departures from efficiency, appearing in every strand of economic research from international finance to labor economics, a unified theory is essential to our general understanding of how markets work. This chapter attempts to provide a starting point for that unified framework.

3. WHAT CAUSES THE FAVORITE-LONGSHOT BIAS AT THE RACETRACK?

Previous authors have attempted to provide theoretical explanations for the existence of the specific bias found in their study. Some of the explanations for the regular favorite-longshot bias in racetrack betting include risk-loving behavior on the part of bettors, poorly informed casual bettors who bet too evenly across all entrants, bettors getting direct utility from betting on longshots, a systematic tendency of individuals to over-estimate the chances of low probability outcomes, and bettors having preferences over the skewness of the payouts. Ali (1977) suggests that this bias could simply be a general feature of any odds-based parimutuel system. In contrast to many of the utility-based theories grounded in the idea of a risk-loving bettor, Golec and Tamarkin
Russell S. Sobel and Matt E. Ryan

(1998) conclude that after controlling for skewness, the data are also consistent with risk-aversion. Theories explaining the favorite-longshot bias cover a broad spectrum, to say the least.

Despite their range, previous theories for the favorite-longshot bias can be generally grouped into two categories: risk-preference theories and information-perception theories. The first group of theories attributes the bias to a preference for risk among gamblers. Utility curves can be constructed so as to generate the desired effect; Friedman and Savage (1948) posited that consumers have a globally risk-averse utility function, yet over the relevant range concerning gambling and betting, the utility function becomes locally risk-loving. Weitzman (1965) introduced the concept of the “representative bettor,” along with this representative bettor behaving as an economic actor. Seeing as any rational actor would take the higher expected return over the lower one, and thus equilibrate the market of odds via expected returns, the representative bettor must be logically risk-loving in shifting his bets disproportionately (in profit terms) toward longshots. Starting with the conclusion that bettors are risk-loving, Quandt (1986) shows that the favorite-longshot bias must exist in order to clear the market. Because of the additional utility derived from taking high-risk (high variance) bets, the payoff of such a bet must necessarily be smaller for the high-risk longshot bet than the low-risk favorite bet.

The second group of theories, the information-perception theories, places the activities of the bettor as a reactionary to new information at the time of betting, as opposed to being naturally predisposed to take riskier bets. This literature generally assumes that bettors do not perfectly absorb information. Snowberg and Wolfers (2004), for example, point to “studies by cognitive psychologists that suggest that people are systematically poor at discerning between small and tiny probabilities (and hence they will price each similarly). Further, certain events are strongly preferred to extremely likely events, leading to even very likely events to be under-priced.” Ultimately, it is the representative bettor’s inability to process information correctly that leads to a favorite-longshot bias in the information-perception theories.

One common aspect in these explanations is the utilization of a representative bettor model in which all bettors are assumed to be acting in one particular manner. However, Sobel and Raines (2003) unambiguously show that this is not the case. The authors analyze nearly 2,800 races at two dog tracks in West Virginia.1 By analyzing attendance, and the breakdown of types and sizes of bets placed on different days of the week at the same tracks, and on the exact same set of racing entrants, they find that some days, in

---

1There exist some differences between betting on dog racing as compared to more traditional horse racing. While the structure of the betting, types of bets, track-take percentages and information provided to bettors are all identical to horse racing, dog racing tends to exhibit a far higher percentage of serious bettors due to the fact that the incidence of exotic bets is four to five times higher than at horse tracks. For further explanation of serious bettors, see Section 3.1. In addition, dog tracks are much less costly to run than horse tracks. The typical attendance found in studies of horse racing is generally around 15,000 per night; the average dog track attendance in the Sobel and Raines dataset is approximately 1,300. The data exclude races in which entrants were scratched from the race, so that all races examined include eight entrants. However, races with fewer entrants are analyzed separately in the original paper and also provide evidence in favor of their hypothesis. In addition, races in which there were ties in the finish were also excluded to avoid problems in computing objective probabilities.
particular weeknights, have a substantially larger percentage of serious bettors, while the weekends have a higher portion of casual bettors. This observation follows from the fact that serious bettors are characterized by larger average wagers and a higher incidence of exotic bets. Casual bettors, on the other hand, tend to wager less, and wager on simpler bets.\(^2\) Table 1 shows the breakdown of favorite and longshot wagers during the weekdays and the weekends in the Sobel and Raines dataset, along with data on average wager size and exotic bet frequency.\(^3\)

As can be seen in Table 1, the average bet per person per race is 10% higher on weekdays than on weekends, and the proportion of exotic bets is over 30% greater. Both of these outcomes are consistent with a higher proportion of casual bettors at the track on weekends than on weeknights. The result of this difference in the favorite and longshot betting patterns is also clearly visible. On the average weekday night, for every $1 bet on the longshot, there is $7.50 bet on the favorite. On the weekends, however, this falls substantially to $5.20 bet on the favorite for every $1 bet on the longshot. This discrepancy cannot be explained by a difference in the characteristics of the races themselves, because at the two tracks studied the exact same racing entrants are used across days. This visible difference, then, is entirely due to a difference in the types of bettors at the track—a difference that is not arbitraged away.

Figure 1 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). Each point is also for a specific grade of race. The observation in the upper right hand corner, for example, is the average betting on the first favorite in grade AA races on weekdays plotted against the same value for weekends. There is a clear rotation

\(^2\)Casual bettors also tend to leave the track earlier, allowing later evening races to actually have a different proportion in the bettor mix than earlier races.

\(^3\)At both tracks studied, races are offered on Monday evening, Wednesday evening, Thursday evening, Friday evening, Saturday afternoon, Saturday evening, and Sunday afternoon. In addition, one track, Wheeling Downs, also offered a Wednesday afternoon race. The weekdays subsample includes all races on worknights (Monday, Wednesday, and Thursday nights), while the weekend subsample includes Friday evening, and both races on Saturday. Wednesday afternoon and Sunday races are not included in either subsample, but are included in all data showing combined numbers for all races. Wednesday afternoons are special events at which large tour groups are present, appearing more like a weekend. Interestingly, Sunday at one track appears more typical of a weekend, while at the other track more typical of a weekday night, in terms of average attendance and betting patterns. For these reasons it was included in neither subgroup.
of the bets, showing an outcome on weekends that might best be described as the weekday bets, just with an added proportion of uniform bets across all racing entrants and grades. Statistical testing shows clearly that the objective probabilities across the two tracks, different grades, and different racing conditions are identical. This difference in betting percentages occurs despite the exact same probability structure facing the bettors.

Given the clear presence of heterogeneous bettors, and the potential for heterogeneous bettor models to explain the favorite-longshot bias, we now turn to an explicit discussion of the different types of bettors generally participating in betting markets. We identify three such categories: the casual bettor, the serious or regular bettor, and the arbitrageur.

3.1. The Casual Bettor

The casual bettor does not attend the track frequently, and is thus less able to use information to make an informed bet.⁴ Information used for a wealth-maximizing bet can

---

⁴Piron and Smith (1995) and Sobel and Raines (2003) are the only studies that attempt to explicitly account for casual bettors.
come from a variety of sources, including previous race performance (and quality of opponents), track conditions and locale, as well as initial race odds and preliminary wagering. The casual bettor can simply be viewed as being able to only fully integrate some of this information. In a Bayesian framework, the casual bettor simply has more weight on a uniform prior in the construction of the posterior probabilities. In many cases these bettors are at the track accompanying friends or family and have other non-wealth-maximizing objectives, or could even be deriving utility from the act of gambling itself. Casual bettors will make bets in a more arbitrary fashion, and are also more likely to choose simpler and less risky types of bets. Arbitrary bets could be based on any predisposed preference—the race position or number of the entrant, the color the entrant is wearing, or even the entrant’s name or appearance. By definition, these bets will tend to be more evenly distributed across all participants than the true probabilities would suggest. Thus, the actions of this group of bettors results in a pool of bets that are placed too heavily on longshots and too lightly on favorites, which by itself can generate a regular favorite-longshot bias in which the return to betting on the favorites is higher than the return to betting on longshots. These casual bettors will overbet any racing entrant with an objective probability lower than the average probability of 1/N (where N is the number of racing entrants), and will under-bet any racing entrant with an objective probability higher than average. The result is a linear rotation of the betting proportions around 1/N. In a game theoretic approach, making money by playing against these causal bettors is the focus of our second group, which we term serious or regular bettors.

3.2. The Serious or Regular Bettor

The serious or regular bettor is the type of person who attends the track regularly. He or she is an informed participant in the betting market. He or she places bets with a strong understanding of the participants in the races, and can digest the vast amount of information provided in the racing forms. He or she believes he or she has a knack for picking the entrant most likely to win based on past performance. He or she will consistently pick the winning entrants or combinations more frequently than the casual bettors at the track. This type of bettor gets utility from the action of picking the winning entrant, or winning money, more often through the betting night than other serious and casual bettors in his or her social group at the track. Making money by playing against causal bettors who cannot accurately identify the likely winning entrants is the goal of serious or regular bettors.

This type of bettor is likely to place bets to win, place, or show on who he or she thinks is likely to win the race, and to bet on the likely combinations of the best entrants in the multi-entrant, exotic bets. By doing so, the actions of the serious or regular bettor generally lead to an opposite favorite-longshot bias in which the favorites are overbet and yield a lower return. These bets are often supplemented with much riskier exotic bets, including picking the top two or three finishers. These bets are often placed with a wheel or box option, in which the bettor picks his or her favorite one or two entrants, and then buys a single bet on every possible combination of those two entrants near the
top, with all other entrants in the remaining places. Almost by definition, these types of bets cause a favorite-longshot bias, because they place a single bet on each of many possible combinations, not all of which are equally likely to occur.

A better term for these types of bettors may be winner pickers, as this is their betting strategy’s main focus. These bettors examine and understand the odds system, and frequently utilize the displayed odds prior to the race as a supplement to their own unique information to further help their ability to pick the entrants most likely to win. These serious or regular bettors sometimes have inside information not available to all bettors—for example, they could be closely associated with the owners or trainers. Further, they may think their analytical abilities based on their experience allow them to pick the true winners better than the other bettors against which they compete. Again, they use this information to attempt to pick the winning entrant, and often wait to place their bets until near the close of the betting period so as to not reveal their actions (through the displayed odds) to other bettors. Gandar et al. (2001) show that bets placed closer to the end of betting more closely correspond to the true objective probabilities than early bets.

These serious bettors, however, as is more often assumed, are not arbitrageurs. Serious bettors and arbitrageurs are two distinctly different groups. Sobel and Raines consistently find in their data that the number of clearly visible arbitrage opportunities present at the track is actually higher on days dominated by more serious or regular bettors.5

Table 2 shows the estimates of how the favorite-longshot bias in the win betting market differs between weekdays and weekends.6 On the weekday races, dominated by a larger proportion of serious bettors, there is a strong opposite favorite-longshot bias that weakens as the betting pool rotates more heavily toward longshot betting on weekends. Had the effect been slightly larger on weekends, the bias would turn into a regular favorite-longshot bias on weekends. On weekdays, despite the presence of a large proportion of more serious bettors, there is a clearly arbitragable positive rate of return on longshot betting in the win market. One important implication of this finding is that the biases are not the result of arbitrage equilibriums, but rather are allowed to persist because of its absence.7

To explore for any time-of-day effects, the data were grouped into the first 13 races and the final two races of the day. An immediate problem is that races vary in grade from AA to M. The data suggest 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

5In their study this is weekday evenings when attendance is lower and when exotic betting is much heavier. This result also holds for “Superfecta Sundays” when these types of bettors dominate the track.
6In cases where two racing entrants tied for being the favorite (having identical odds and thus subjective probabilities), each was viewed as the first favorite with weight one-half and the second favorite with weight one-half.
7In Ali’s (1977) model, for example, the favorite-longshot bias is explained as a natural result of bettors making bets relative to the market odds, when those odds diverged from the bettor’s guess about the true win probabilities, thus making it a result of arbitrage.
<table>
<thead>
<tr>
<th>Favorite</th>
<th>All days (2,558 races)</th>
<th></th>
<th></th>
<th>Weekdays (952 races)</th>
<th></th>
<th></th>
<th>Weekends (1,066 races)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Objective probability of win ((\pi))</td>
<td>Subjective probability of win ((\rho))</td>
<td>Rate of return from bet</td>
<td>Objective probability of win ((\pi))</td>
<td>Subjective probability of win ((\rho))</td>
<td>Rate of return from bet</td>
<td>Objective probability of win ((\pi))</td>
<td>Subjective probability of win ((\rho))</td>
</tr>
<tr>
<td>1</td>
<td>27.64%</td>
<td>25.27%</td>
<td>−23.72%</td>
<td>28.33%</td>
<td>24.53%</td>
<td>−27.41%</td>
<td>27.36%</td>
<td>25.69%</td>
</tr>
<tr>
<td>2</td>
<td>18.45%</td>
<td>18.04%</td>
<td>−18.89%</td>
<td>18.82%</td>
<td>20.12%</td>
<td>−9.47%</td>
<td>18.23%</td>
<td>16.01%</td>
</tr>
<tr>
<td>3</td>
<td>14.26%</td>
<td>14.82%</td>
<td>−13.90%</td>
<td>14.22%</td>
<td>14.06%</td>
<td>−18.18%</td>
<td>14.25%</td>
<td>15.76%</td>
</tr>
<tr>
<td>4</td>
<td>11.55%</td>
<td>11.53%</td>
<td>−18.45%</td>
<td>11.48%</td>
<td>12.41%</td>
<td>−10.79%</td>
<td>11.57%</td>
<td>11.02%</td>
</tr>
<tr>
<td>5</td>
<td>9.49%</td>
<td>9.27%</td>
<td>−17.61%</td>
<td>9.33%</td>
<td>9.45%</td>
<td>−16.67%</td>
<td>9.58%</td>
<td>9.80%</td>
</tr>
<tr>
<td>6</td>
<td>7.78%</td>
<td>8.84%</td>
<td>−7.72%</td>
<td>7.60%</td>
<td>7.67%</td>
<td>−19.73%</td>
<td>7.85%</td>
<td>9.04%</td>
</tr>
<tr>
<td>7</td>
<td>6.23%</td>
<td>6.48%</td>
<td>−12.64%</td>
<td>5.96%</td>
<td>5.64%</td>
<td>−20.09%</td>
<td>6.36%</td>
<td>7.18%</td>
</tr>
<tr>
<td>8</td>
<td>4.61%</td>
<td>5.76%</td>
<td>17.48%</td>
<td>4.26%</td>
<td>6.13%</td>
<td>36.96%</td>
<td>4.80%</td>
<td>5.50%</td>
</tr>
</tbody>
</table>
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. 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 2 and 3 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.

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.8

8This is also evidence against a risk-preference theory of the favorite longshot bias. As the average bettor’s wealth declines throughout the evening (due to track-take), risk-preference would suggest they move to more risk-averse bets, but they don’t. One plausible explanation of the general trend over the last half of the evening is that attendance falls significantly through the evening, and it is disproportionately casual bettors who leave early, implying that the later races are dominated by serious bettors who 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.
3.3. The Arbitrageur

The arbitrageur is the typical type of agent modeled in economic theory. He or she takes into account all available information, and attempts only to maximize his or her return on investment with no other utility considerations, and risk neutral preferences. As in other markets, arbitrageurs are the agents that drive markets toward efficiency. Profit opportunities are quickly identified and acted on, eliminating any deviations from efficiency. Most important to this type of bettor is the information conveyed in the market odds at the track that constantly change to reflect betting intensities, and how these compare, for all entrants, to his or her best estimate of the true probabilities. The arbitrageur cares little about who might win the race, and more about which entrant has the highest expected return based on any divergence he or she sees.

The actions of arbitrageurs lead to market efficiency, in the weak sense. A market is weakly efficient as long as no persistent profit opportunities exist that could be exploited based on all available information. While there is no doubt that these types of individuals exist in society, and account for the efficiency of many other markets, the persistence of the favorite-longshot bias, and of its opposite, imply that either: (1) they simply don’t participate to a great enough extent in these markets, or (2) they don’t have the necessary information or at least not in a timely enough manner to utilize it before the betting window closes for the race.
There is substantial evidence in Sobel and Raines (2003) that these arbitrageurs are absent to a large extent at the two tracks they study based on the clear and persistent presence of arbitrage opportunities at these tracks. These occurrences actually increase on nights (or during later races) dominated by the serious or regular bettors. Their data suggest that these arbitrage opportunities are less frequent on higher-attendance weekend nights, not because arbitrageurs are there to arbitrage these away, but because of the presence of casual bettors who tend to bet on longshots. The serious bettors at these tracks appear to allocate their time and money toward betting in the exotic bets, and attempting to pick winners better than everyone else, with little regard for attempting to arbitrage away even visible opportunities. In fact, evidence from other studies appears to be mounting that the vast majority of bettors tend to make bets independent of the track odds.

Camerer (1998), for example, finds that track bettors did not respond to large changes in track odds. Camerer went to a local track and made large bets on single horses to change their odds. He then withdrew the bets just before the betting on the race closed. By comparing the remaining bets in these races with bets made on similar horses when he did not interfere, he was able to see how the large changes in odds influenced other bettors. He found no statistically significant difference. He interprets this result as that it did not manipulate the bettors who saw through the incorrect odds, but in fact arbitrage would suggest that bettors should have responded to this horse being overbet by betting on other entrants, who have now become under-bet. The bettors did not. His results also suggest that these average racetrack bettors do not actively search out or participate in arbitrage activities. In a similar study, Hanson et al. (2006) find experimentally that prediction market prices are unaffected by agents wishing to distort them. Wolfers and Zitzewitz (2004) survey studies in price manipulation and conclude that no effect exists on prices beyond the short term. A historical study by Strumpf and Rhode (2003) finds, as well, that efforts to manipulate prices are ultimately unsuccessful.

It is important to note that arbitrage involves both discovering profit opportunities, which often requires data and time-intensive statistical analysis, and then acting on this information. This is precisely the type of entrepreneurship modeled by Kirzner (1973). The entrepreneur is someone who spots previously unnoticed profit opportunities and acts on them. This not only takes a certain type of individual to perform, but also requires a market return from their actions.

One must keep in mind that the arbitrage opportunities available at racetracks are not only minor, but are also uncovered. Unlike triangular exchange rate arbitrage, race-track bettors must still take risk to exercise arbitrage. In addition, like any other pure return-maximizing strategy, returns will be driven down to zero economic profit, not zero accounting profit, on the time and investment needed to conduct the arbitrage itself. While the return to betting on favorites exceeds the return on longshots in most studies, the absolute level of the return on favorites must be high enough to overcome the track-take which is usually between 15% and 20%. However, using experimental methods, Hurley and McDonough (1995) find that changing the track-take has no effect on the severity of the favorite-longshot bias. In addition, the track minimum payout of $1.10 on
a $1 bet creates a lower bound on the return from a bet on an extreme favorite, perhaps at a level that is positive. Swidler and Shaw (1995) point out that the total daily monetary return on the arbitrage opportunities they identify would not be enough to cover the cost of admission, parking, the program, and a minimum wage rate per hour given the price setting ability of large individual bets. One might similarly ask why economists who have found and published numerous examples of profitable arbitrage opportunities in the stock market have not used their own money to arbitrage these inefficiencies away. The reason, of course, is the opportunity cost of economists’ time and the risk involved. With so many arbitrage opportunities abounding in the world, it would seem to be consistent to say the arbitragers have plenty of work and can concentrate on the least costly and most covered opportunities, while leaving the smaller inefficiencies abounding in other markets.

As Leeson et al. (2006) discuss, the presence of forgone profit opportunities is much more likely than the presence of persistent losses in markets. Economic losses diminish resources under the control of the decision-maker, and lead to business failure. When a potential profit opportunity goes unexploited, while there is an opportunity cost, it does not create the same strong incentive for self correction. In fact, an individual may never know he or she passed up a potentially profitable opportunity, but will certainly know if he or she earns a loss. According to this theory, then, the weaker incentive for market correction can lead to a greater presence and persistence of unexploited profit opportunities in markets.

One final explanation for the lack of arbitrage is that the bettors capable of this activity simply do not have the necessary information or at least not in a timely enough manner to utilize it before bets close for the race. Because the returns on bets in economic studies are based on the final odds, these are not the odds visible during the entire betting period to potential bettors. At some tracks, for example, late money bets, accompanied by the delay in updating the odds on the tote board, and the unpredictable length of time needed to wait in line to place the bet, means that it is almost impossible to make a bet based on the final market odds.

4. IS IT RISK OR INFORMATION?

Thus we are left with two potential theories to explain the favorite-longshot bias (or its opposite). It is either the result of an equilibrium outcome driven by arbitrageurs with a preference for risk, or it is the result of a weighted average of casual and regular bettors’ strategies without offsetting arbitrage occurring. And while the risk-preference model dominated the literature for decades, modern evidence seems to clearly point in the other direction.

The most convincing evidence on this comes from Sobel and Raines (2003), who gathered not only overall odds and payout data for the win market, but also data on all other markets, as well as the day of the week, track attendance, number of racing entrants in the race, the grade (or quality level) of the race, and the race number. In addition, data were collected on the various types of bets placed, from simple to exotic.
This breadth of supplemental data allowed them to see how the favorite-longshot bias changed across different bets existing at the same time at the track, as well as how these biases changed across different nights that had significantly different proportions of serious and causal bettors.

They conclude that the bias can be explained (and its presence in certain situations, and the opposite in others) as a reflection of changing combinations of serious and casual bettors. This information-based model of serious and casual bettors has clearly different empirical predictions from the risk-preference model, most notably the linear versus non-linear nature of the betting percentages relative to their true values, and also in terms of how the bias would change based on the number of racing entrants in the race, and across different types of bets.

Figure 4 shows the relationships between objective (true) probabilities of an entrant winning the race and the subjective probability of the entrant winning, gauged by true betting percentages. In an efficient market, subjective probabilities match perfectly with the objective probabilities, which would be reflected by a 45° line from the origin.

**FIGURE 4** Information vs. risk in subjective and objective probabilities.
No bias exists in this case. The existence of less-informed bettors, who bet too evenly across racing entrants, can be illustrated as a rotation of this line around the average probability of $1/N$. This is shown by the line $I$ in the figure. The degree to which casual bettors dominate the betting market will determine the degree to which the line rotates.

The risk model does not involve a linear rotation as is true in the case of the information-based model. Instead, pre-existing desires for riskier bets dictate betting patterns. A typical risk-preference would result in a betting pattern illustrated by the curve $R$ in Figure 4 according to the relationship

$$
\rho_i = \lambda \pi_i^o
$$

where $\rho_i$ represents the subjective probability for racing entrant $i$, $\pi_i^o$ represents the objective probability for racing entrant $i$, and $\lambda$ and $\tau$ are parameters of the utility function that measure preference for risk that can be estimated using betting market outcomes. The existence of the favorite-longshot bias in this theory comes from a desire of bettors to choose the riskier bets associated with longshots.

Now, for comparison, it is worthwhile to briefly derive the mathematical properties of an information-based model for risk-neutral bettors with Bayesian updating. Assuming that prior to arriving at the track, each bettor has no knowledge about the racing entrants, thus all entrants are equally likely to win in the diffuse prior. The mean of the bettor’s prior belief about the entrants probability of win, $\rho_p$, must be $\rho_p = 1/N$.

Once at the track, the bettor collects a set, $k$, of information signals regarding the racing entrants’ true probability of win. Let $\rho_I$ 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, $\rho$, on which the bettor makes bets, is given by

$$
\rho = (1 - h) \rho_p + h \rho_I,
$$

where $h$ is a precision parameter that allocates the relative weight placed on the new information and the prior belief. 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 of the bet, $c$.

Without assuming a specific form for $h$, the relationship with its arguments is intuitively obvious

$$
\frac{\partial h}{\partial k} > 0, \quad \frac{\partial h}{\partial \sigma_I} < 0, \quad \frac{\partial h}{\partial c} < 0.
$$

This simply says that as a bettor becomes more information rich, he or she will choose to place less weight on 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 prior beliefs. 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 $\rho_I = v + \phi \pi$. The posterior mean equation from which the bettor places his or her bets is $\rho = \alpha + \beta \pi$, where $\alpha = \frac{(1 - h)}{H} + h(v + \phi)$ and $\beta = h\phi$. Strong form market efficiency would exist if $\beta = 1$ and $\alpha = 0$. The relation $0 < \beta < 1$ gives the traditional favorite-longshot bias, while $\beta > 1$ produces an opposite favorite-longshot bias.

How do these differing model predictions fit actual betting data? Figure 5 shows data on the objective and subjective probabilities for all quinella bets at the tracks studied.
by Sobel and Raines, along with fitted risk and information model regression lines. In the figure, the relationship in the data is almost perfectly linear, in rejection of the risk explanation that implies a non-linear relationship between objective and subjective probabilities. The linear relationship suggested by the information model does a particularly good job in fitting the data at the two extremes relative to the risk model.

To this point we have focused on a comparison of the subjective and objective probabilities. We now turn our attention to how this relationship appears once converted into the rates of return on bets versus objective probabilities. Let $\pi_i$ denote the true win probability of racing entrant $i$. Let $\rho_i$ denote the betting market’s subjective estimate of the entrant’s probability of win. Griffith (1949), and McGlothlin (1956), were the first to formally note that this subjective probability is simply the percent of the total betting pool that is bet on this particular racing entrant $i$. The track payout per dollar if the bet wins, $S_i$, is given by $S_i = O_i + 1$, where $O_i$ is the track odds, which are defined in terms of the underlying subjective probabilities as $O_i = (1 - \rho_i)/\rho_i$. Further substitution and simplification produces $S_i = 1/\rho_i$. The expected return per dollar bet, $R$, is given by $R = S_i \times \pi_i + 0 \times (1 - \pi_i) = S_i \times \pi_i$. The rate of return, $r$, for a $1$ bet, is thus

\[ r = \frac{R}{1} = S_i \times \pi_i. \]

\[ \text{Here we abstract from the track-take for simplicity. Inclusive of a track-take of } t, \text{ the track odds are } O = (1 - \rho')/\rho' \text{ where } \rho' = \rho(1 - t). \]
r = (R - 1)/1, which through substitution yields \( r = (\pi_i - \rho_i)/\rho_i \). Figure 6 shows the data from Figure 5 transformed according to this relationship.

In the transformed data, the information model and the risk model now both predict a non-linear relationship, however these non-linear relationships are clearly different. Again, the information model clearly fits the data better than a risk-preference model.

Additional evidence in favor of the information-based explanation is found by examining bets of different complexity that exist simultaneously at the track. More complex, exotic bets function to turn informed regular bettors into bettors with less information in a Bayesian-type information model. Information-perception models are grounded in the ability of the bettor to digest information and to make an appropriate bet based on this information. As such, more complicated bets should see the same smoothing effect when compared to easier ones. Sobel and Raines’ data show that the relative lack of information available for the more complex bets creates a greater smoothing effect among the more complex bets.

Table 3 displays the regression results for each of five different types of bets for which data were available and present at the same time at the track. The regressions are ordered so that the complexity of the bet, in terms of the number of parameters or
TABLE 3  Objective vs. subjective probability regressions: all bets

<table>
<thead>
<tr>
<th></th>
<th>Win</th>
<th>Place</th>
<th>Show</th>
<th>Quinella</th>
<th>Perfecta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.018**</td>
<td>-0.003</td>
<td>0.049**</td>
<td>0.008**</td>
<td>0.003**</td>
</tr>
<tr>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.013)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Objective probability</td>
<td>1.143**</td>
<td>1.013**</td>
<td>0.869**</td>
<td>0.770**</td>
<td>0.788**</td>
</tr>
<tr>
<td>(0.041)</td>
<td>(0.020)</td>
<td>(0.032)</td>
<td>(0.019)</td>
<td>(0.035)</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.9924</td>
<td>0.9977</td>
<td>0.9919</td>
<td>0.9849</td>
<td>0.9062</td>
</tr>
<tr>
<td>Betting combinations</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>28</td>
<td>56</td>
</tr>
<tr>
<td>Observations per betting combination</td>
<td>2558</td>
<td>2558</td>
<td>2558</td>
<td>2459</td>
<td>893</td>
</tr>
<tr>
<td>$T$-ratio for slope = 1</td>
<td>3.503**</td>
<td>0.645</td>
<td>-4.082**</td>
<td>-12.262**</td>
<td>-6.136**</td>
</tr>
<tr>
<td>Estimated rotation point</td>
<td>12.59%</td>
<td>25.10%</td>
<td>37.51%</td>
<td>3.55%</td>
<td>1.45%</td>
</tr>
<tr>
<td>Predicted rotation point (null prior)</td>
<td>12.50%</td>
<td>25.00%</td>
<td>37.50%</td>
<td>3.57%</td>
<td>1.79%</td>
</tr>
</tbody>
</table>

NOTE: Taken from Sobel and Raines (2003). ** indicates statistical significance at the 1% level, * at the 5% level. The test of market efficiency is not just that the slope equals one, but also that the intercept term equals zero. The $F$-statistics for this joint test strongly rejected this hypothesis for all markets except the place market. The predicted rotation point is the average probability across all betting combinations (i.e., one of eight entrants win, two of eight place, three of eight show, one of 28 quinellas win, and one of 56 perfectas win). The rotation point estimated by the model is equal to the constant divided by one minus the slope coefficient. This is the probability at which any bet above has the opposite bias as any bet below in the information model.

Data points needed to estimate the objective probability, increases as we move rightward among the columns in the table. The more complex bets show lower regression coefficients—in other words, the slope of the line becomes flatter. This is the exact result predicted by an information-based model—more flattening due to the smoothing effect of casual bettors. A further strike to the risk-preference theories is the fact that the least risky market (the show bet) and the most risky one (the perfecta) demonstrate the usual favorite-longshot bias, while the middle-risk bets show either little bias or an opposite favorite-longshot bias. These biases, however, line up in order when they are ranked by complexity.

In another test, Sobel and Raines examine how the bias changes as the number of racing entrants changes. Within a pure risk model, the over (or under) betting of a particular entrant is entirely dependent on the entrant’s true probability of win. This is not true in a model with causal bettors causing the bias by betting too evenly across racing entrants. Because casual bettors bet too evenly, the direction of the bias will depend on whether the racing entrant has an above (or below) average probability given the number of entrants in the race. They indeed find this to be the case, that a racing entrant with, say, a 1/7 probability of win would be overbet with six entrants, but under-bet with eight
Chapter 8 • Unifying the Favorite-Longshot Bias

entrants. Returning to the bottom row in Table 4, these rotation points (the points at which it switches from entrants being overbet to under-bet) are presented, along with the value predicted by the information model (1/N according to a null prior). They coincide almost perfectly with the predictions of the information model. The risk model, on the other hand, has no prediction as to the rotation point, which would depend on the severity of the preference for risk. However, the risk model would unambiguously imply that the rotation point, whatever it is, should be the same across all bets existing at the track, given that it is a reflection of risk-preference related to the probability structure of the bets. This can clearly be rejected in the data.

5. CAN THE MODEL EXPLAIN THE BIASES IN OTHER MARKETS?

No firm understanding of why sports betting markets contain the opposite bias from most racetracks is known. According to our theory, however, one possible explanation is that Vegas sports betting markets tend to be dominated by a higher proportion of serious or regular bettors (and fewer casual bettors) than are racetracks. In football and basketball, the bets studied have been the point spread and over/under (total game score) bets. These bets do differ somewhat from racetrack bets made relative to market odds, however. Bets on baseball, like in racetrack wagering, are made relative to odds, and because the opposite favorite-longshot bias is found in baseball as well, a difference in the types of bets cannot be responsible for causing an opposite bias in sports wagering than the one present in racetrack wagering. In addition, the regular favorite-longshot bias in racetrack wagering has been found by Gabriel and Marsden (1990, 1991) in the UK where bets are made against a bookie-based fixed odds system, rather than the pure market odds structure as in the US.

A systematic difference in the betting behavior of serious and casual bettors might also explain why some racetrack studies find opposite favorite-longshot biases in contrast to most other studies on racetrack betting. The higher betting per person at the Hong Kong track suggests that it is dominated more heavily by serious bettors than the average US racetrack used in other studies, and similarly, one might argue that the smaller Class II racetrack draws a relatively higher proportion of serious bettors than the tracks used in other studies.

If a difference between the behavior of serious and casual bettors is accountable for the flipping of a favorite-longshot bias to its opposite, it might also suggest that the stock market findings are due to a relatively high proportion of serious to casual investors. An uninformed casual investor will frequently invest in mutual funds, or somehow otherwise allow an informed middleman to manage his or her portfolio. Possible explanations for the stock market bias proposed by previous authors have

10Finally, another degree to which data can be broken down is the time of day that the bet is placed. After controlling for the quality level of the race, later evening races become more heavily dominated by serious bettors (as casual bettors are more likely to leave the track early). As expected, the proportion of bets on the favorite relative to the longshot grows throughout the evening as casual bettors leave the track.
been that investors overreact to information, that the bias is due to risk-preferences (an explanation that includes the small/large firm risk differential explanation), and the presence of uninformed noise traders in the market.

While stock market research has found an opposite favorite-longshot bias in stocks and earnings forecasts, some other studies have found a regular favorite-longshot bias in options pricing. For example, the longshot or shorter maturity options have been found to be overpriced by Rubinstein (1985, 1987), and options “out-of-the-money” have found to be overpriced by Shastri and Wethyavivorn (1987), among others. Could the much more complex calculations necessary for options pricing create the different bias? Thaler (1992) has suggested that differences between biases found across different types of racetrack bets might be due to differing complexities of the bets. In addition, Elton et al. (1982) show a similar example regarding complexity affecting behavior in the stock market with respect to futures transactions. The idea is that, to some degree, as bets become more complex, all bettors become less informed relative to the level necessary to obtain accurate probability estimates.

The general theory presented in this chapter suggests a starting point for a unified model of these anomalies. This model is a three-group heterogeneous bettor model. The first group—the casual bettors—generally tends to overbet longshots (and under-bet favorites) with their strategy of using selected, and perhaps highly unreliable, signals to select their bets (due to their significantly limited information set). The second group—the serious bettors—generally tends to overbet favorites (and under-bet long-shots) with their strategy of attempting to use their knowledge and human capital to pick the winners. The third, and final, group—the arbitrageurs—are individuals who invest resources to both discover (usually through intensive data analysis) and exploit profit opportunities, but only when these opportunities generate a higher than normal (risk adjusted) rate of return from their resource investment.

In addition to allowing for heterogeneous bettors, our theory also incorporates heterogeneous bets, differing in their level of complexity. Employing Bayesian logic, a change in the complexity of a bet has a mathematically equivalent effect to changing the information level of the bettor. On more complex bets, even serious bettors become relatively more uninformed, causing them—just like the casual bettors—to bet too evenly across possible outcomes. Thus markets with simple bets and/or dominated by serious bettors will tend toward an opposite favorite-longshot bias (like sports betting), while markets with complex bets and/or dominated by casual bettors will tend toward a regular favorite-longshot bias (like exotic bets). The presence of arbitrageurs can eliminate these divergences from efficiency when they are identified and exploited, but at a cost that makes this type of activity unprofitable in most narrow and limited betting markets.

6. CONCLUSION

The ideas presented in this chapter lay a framework for a unified theory of deviations from market efficiency. If a unified model can explain these anomalies, it may allow great insights into human behavior and further the refinement of our economic
models. It appears that information, complexity, and human analytical capital are jointly responsible. Highly informed and educated bettors faced with simple situations appear to systematically overvalue favorites and undervalue longshots—they attempt to pick winners better than everyone else. As complexity increases, or the number of less informed (or less able) bettors increases, this bias moves toward the other direction, eventually overtaking the underlying opposite bias.

Opportunities to further test our model abound in other areas where these inefficiencies have been found to exist.\(^{11}\) In addition, the fundamental model presented here has broadsweeping implications for the efficiency of product markets where consumers are faced with uncertainty with regard to things such as product quality or price differences across firms, and for public sector economics in models of rent seeking, voter information, and electoral competition.

References


\(^{11}\)Try asking someone the name of the longest book they have ever read. Then ask them to try and remember (without looking) how many pages it was. Our model predicts that they will be very likely to over-estimate the true number of pages.
Russell S. Sobel and Matt E. Ryan


Chapter 8 • Unifying the Favorite-Longshot Bias


