

**APPLICATION OF AN SPRT-LIKE TEST TO DETERMINE THE BEST
PROBABILITY FORECASTING MODEL FOR THE WIN, QUINELLA AND
EXACTA WAGER**

by

Lysa M. Porth

A thesis

submitted to the Faculty of Graduate Studies

in partial fulfillment of the

requirements for the degree of

Master of Business Administration

I.H. Asper School of Business

University of Manitoba

Winnipeg, Manitoba

**THE UNIVERSITY OF MANITOBA
FACULTY OF GRADUATE STUDIES

COPYRIGHT PERMISSION**

**APPLICATION OF AN SPRT-LIKE TEST TO DETERMINE THE BEST
PROBABLITY FORECASTING MODEL FOR THE WIN, QUINELLA AND
EXACTA WAGER**

BY

Lysa M. Porth

**A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of
Manitoba in partial fulfillment of the requirement of the degree
of
Master of Business Administration**

Lysa M. Porth © 2006

Permission has been granted to the Library of the University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to University Microfilms Inc. to publish an abstract of this thesis/practicum.

This reproduction or copy of this thesis has been made available by authority of the copyright owner solely for the purpose of private study and research, and may only be reproduced and copied as permitted by copyright laws or with express written authorization from the copyright owner.

Abstract

The win, quinella and exacta represent three of the most prevalent bets wagered at the racetrack. Their victory depends on the probability that the horse(s) chosen are correct. It is commonly assumed in literature that the win pool is largely efficient and its probabilities serve as good estimates for calculating subjective probabilities in other wagering pools in association with such models as Harville (1973) and Henery (1981). Ziemba and Hausch (1985) and Asch and Quandt (1987) develop this assumption to uncover wagers with theoretical positive expectation in the exacta and quinella pools. In this thesis, an SPRT-like statistical test is used to determine which probability models are most accurate for forecasting win, quinella and exacta wager outcome. Results establish that probabilities extracted from the exacta pool are superior for predicting win outcomes, while data acquired from the win pool corrected for the favorite-longshot bias using Asch and Quandt's regression method, are best for approximating quinella and exacta results. By harnessing these implied probabilities, a value betting strategy is employed to locate bets with positive expected values in each wagering pool under consideration.

Table of Contents

List of Figures	iii
List of Tables	iv
List of Tables	iv
Chapter 1: Introduction	1
Chapter 2 : Pari-Mutuel System	5
Chapter 3: Overview of Racetrack Probability Models and Theory	8
A. Win Pool Efficiency	8
B. Favorite-Longshot Bias	11
Chapter 4: Exacta Wagering	15
Chapter 5: Quinella Wagering	30
Chapter 6: Win Wagering	42
Chapter 7: Results	51
Chapter 8: Wagering Strategy Using Positive Expected Values	53
Appendix A- Exacta Data	61
Appendix B- Quinella Data	65
Appendix C- Win Data	68
References	69

List of Figures

Figure 4.1: Exacta Model Probability Comparison	28
Figure 5.1: Quinella Model Probability Comparison	40
Figure 6.1: Win Model Probability Comparison	49

List of Tables

Table 3.1: Expected Value Per Dollar Bet for Different Odds Levels	11
Table 4.1: Asch and Quandt's Regression Data	21
Table 4.2: Exacta Model Probability Data	61
Table 5.1: Quinella Model Probability Data	65
Table 6.1: Win Model Probability Data	68

Chapter 1: Introduction

Horse racing is one of the oldest sports known to mankind, originally held at the Olympic Games in Greece dating back to 700 B.C. In North America, horse racing began in 1665 in Hampton Plains, Long Island. Since this time, horse racing has evolved from an activity to amuse the leisure class into an immense industry providing public entertainment around the world. As this business grew, wager varieties expanded from simple win bets to complex exotics. As a result of wagers becoming more intricate, the available methods for calculating probability outcomes soared. Despite the vast number of models that have emerged to calculate probability estimates, the goal is inherent: establish the most accurate probability approximation and gamble only on those shown to generate a positive expected return. Correspondingly, the objective of this thesis is to determine the best probability forecasting model for each of the win, quinella and exacta wagers at the racetrack.

In situations where only one observation for each estimation method exists and the forecasting procedure can be repeated several times, an SPRT-like hypothesis test can be used to decipher the best of k estimation procedures as developed by Rosenbloom (2000) and (2003). By definition, “an estimation procedure i^* will be declared the best of estimation procedures $\{1,2,\dots,k\}$ at a significance level α (α is a parameter between 0 and 1, typically .10, .05 or .01) if the posterior probability of estimation procedure i^* being correct is above $1-\alpha$ (assuming one of the estimation procedures $\{1,2,\dots,k\}$ is correct).”

This thesis employs an SPRT-like test to explore estimation models in the context of horse racing. Specifically, for selecting a win wager in race t , three methods by which probabilities can be estimated are considered. These include straight win pool probabilities, win pool probabilities corrected for the favorite-longshot bias using Asch and Quandt's regression model and implied win probabilities from the exacta pool. Given that there are three methods to estimate the probability of horse j winning race t , the SPRT-like hypothesis test can be used to determine which of the three models under consideration produces the most accurate probabilities. Similarly, for selecting an exacta wager in race t , four methods by which probabilities can be forecasted are compared. These include win pool probabilities using the Harville method, win pool probabilities using the Henery method, win pool probabilities corrected for the favorite-longshot bias using Asch and Quandt's regression model and exacta pool probabilities. Since there are four methods for calculating the probability of an i - j exacta in race t , the SPRT-like test can be applied to determine which of the four models under consideration is superior. Finally, for selecting a quinella wager in race t , four models can be contrasted including win pool probabilities using the Harville model, win pool probabilities using the Henery model, win pool probabilities corrected for the favorite-longshot bias using Asch and Quandt's regression method and quinella pool probabilities. Again, four distinct methods exist from which quinella probabilities can be calculated, meeting the criteria of the SPRT-like test which identifies the most accurate model under consideration.

Beyond the framework of horse racing, the SPRT-like statistical test has a variety of additional applications. One example of competing multinomial parameter estimation

procedures is the sport of baseball. Specifically, three models can be compared that forecast the probability of which competing team will win the game: egrandslam.com; newsfutures.com and tradesports.com. Each one of these sources provide probability estimates at their respective websites: www.egrandslam.com, www.newsfutures.com and www.tradesports.com. Which model is most accurate at predicting winning outcomes?

A further example of how an SPRT-like test can be applied is in deciding whether prediction markets are better than 'bookies' at establishing the correct probabilities in the NFL. Two sources of probabilities come from www.thegreek.com and www.tradesports.com. Although these sources quote odds rather than probabilities, it is possible to extract probabilities. For example if the Patriots are shown at -200 and the Colts at +170, this translates into the fact that people who bet \$200 to win on the Patriots will win \$100 if successful, while those who bet \$100 to win on the Colts will win \$170 if successful. From this, it can be deduced that if there was no house edge, the price would be somewhere between -185 and +185 respectively. As such, odds of 185-100 equates to a probability of $185/285 = 0.649$ in favor of the Patriots. Which model is most accurate at predicting winning outcomes?

Real-money exchanges such as www.tradesports.com and play-money exchange such as www.newsfutures.com offer probabilities on events ranging from sports betting to election outcomes, and offer a third example of an SPRT-like application. Tradesports and Newsfutures serve as two estimation procedures that can be repeated several times to determine whether real-money or play-money prediction markets are more accurate. For

more information, see “Statistical Tests of Real-Money versus Play-Money Prediction Markets” (Rosenbloom and Notz, 2006).

Once the ‘best’ multinomial parameter estimation procedure has been established, the probabilities produced by the corresponding model can be used to calculate the expected return per dollar wagered. Accordingly, in the context of horse racing, this information can be applied to a value betting strategy to maximize returns at the racetrack.

Chapter 2 : Pari-Mutuel System

All bets at North American racetracks are placed under a pari-mutuel system, developed in the late 19th century by Pierre Oller, whereby winners divide the total amount bet, after deducting track operating expenses, racing purses and taxes, in proportion to individual wagers. Collectively, these deductions comprise the commission rate (also referred to as the track-take) which fluctuates depending on the racetrack and the wager in question. For example, in North America the track-take is roughly 17% for straight bets (i.e. win, place and show) and 22% for exotic bets (i.e. exacta, quinella, triactor, daily double etc). As such, approximately 80% of all money wagered goes back into the pockets of bettors, making this an exceptionally unique system where players bet against each other, not the racetrack and consequently 'beat the races' only when they 'beat the crowd.'

In order to beat the crowd, bettors must accurately choose the winning entry or grouping of entries. This is not an easy task however given the plethora of wager types and techniques to play each bet. Consider the following wagers and associated methods:

Straight Bets:

- ⊗ **Win:** Pick the first horse (win) to cross the finish line
- ⊗ **Place:** Pick the first or second horse (win or place) to cross the finish line
- ⊗ **Show:** Pick the first, second or third horse (win, place or show) to cross the finish line

Exotic Bets:

- ⊙ **Exacta (intra-race):** Pick the top two finishers in a given race in the exact order. A box allows you to reverse the order of finish, but doubles the cost of the bet. A wheel allows you to play a horse in either the first or second position with all others in the race. A 'partial wheel' means you can select two or more finishers to finish first or second with your single.
- ⊙ **Quinella (intra-race):** Pick the top two finishers in a given race in either order. This is similar to an exacta box but does not double the cost of the bet and payoffs are usually half of what the exacta pays out.
- ⊙ **Triacta/trifecta (intra-race):** Pick the top three finishers in a given race in the exact order. Similar to the exacta but takes it one step further. Boxes, wheels and part-wheels are permitted.
- ⊙ **Daily Double (inter-race):** Pick the winners of consecutive races on a single ticket. A wheel allows you to pick a single runner in either the first or second race and combine it with all others in the next race. A partial wheel restricts your picking to a single entrant in one race and several others in the remaining race.
- ⊙ **Pick Three, Four, Six (inter-race):** Pick the winners of three or more consecutive races. Wheeling with this type of wager allows you to choose as many finishers in each leg of the sequence, but increases the cost of the ticket exponentially.

As inferred from the definitions above and reflected in the corresponding higher track take-out, exotic wagers carry additional complexity. Exotic wagers challenge the bettor to link the outcomes of two or more entrants in a given race or alternatively to select the

winners in multiple consecutive races on a single ticket. As a result, exotics carry complex series of numerical combinations, with most bettors wagering several horses in each finish position to hedge their bet. Considering the basic theory of risk versus reward, a successful exotic wager generates an appreciably larger return on investment. Furthermore, due to the diversity offered by exotic wagers, straight wagers have diminished in popularity accounting for 33 percent of the pari-mutuel handle. The remaining 67 percent is comprised of exotic wagers: 58 percent intra-race exotics (i.e. quinella, exacta, triacta, etc) and 9 percent inter-race exotics (i.e. daily double, pick three, pick six etc).¹

¹ www.ntra.com

Chapter 3: Overview of Racetrack Probability Models and Theory

Over time, many methods have emerged to calculate probabilities at the racetrack. Some methods boast simplicity and offer ‘close’ approximations to true probabilities, while others require overwhelmingly complex calculations that require a spreadsheet and a statistician alongside at the racetrack. Whichever method is employed, the goal remains the same: to calculate the most accurate probability estimate that correctly selects the entrant or group of entries to win a race.

A. Win Pool Efficiency

One widely circulated method of calculating racing probabilities has been presented by researchers including Ziemba and Hausch (1985) and Asch and Quandt (1987). They conclude that win pool probabilities are reasonably accurate estimates for win bets, so the probability of horse “i” winning a race is equivalent to the fraction of the win pool bet on that horse. Given the unique characteristics of the pari-mutuel system, each wager variety is segregated into its own pool, in essence forming a distinct market. Therefore, to calculate the probability of an entry winning a race, the following formula can be applied:

$$P_w = \frac{1 - \textit{Take}}{(1 + \textit{Odds} + \textit{Breakage})} \quad (3.1)$$

where:

- Track take refers to the percentage deducted from mutuel pools and kept by the track (usually between 13% to 25%)

- ◉ Odds are displayed on the tote board. Odds on an entry are calculated from the percentage bet on the entry.
- ◉ Breakage refers to the downward rounding of odds for easy payout (It is imperative that breakage is not overlooked as it has a surprisingly significant impact on overall profits. For example, at Canadian tracks nickel breakage is applied. This has the effect of rounding payouts down to the nearest multiple of five cents for each \$1 wager. Similarly, at American tracks dime breakage is applied having the effect of rounding payouts down to the nearest multiple of ten cents for each \$1 wager). Consider the following example of how breakage is calculated at a Canadian track: in a race where a \$2 exacta pays \$11.20, because of breakage all payoffs between \$11.20 and \$11.29 are rounded down to \$11.20. Subsequently, to account for breakage when determining P_w , an average estimate of \$11.25 is used.

This method of calculating probabilities concurs with the theory of market efficiency which implies that “information is widely available to participants and all relevant and ascertainable information is reflected in prices” (Gabriel and Marsden 1990). Defined in terms of racetrack betting, efficiency describes the inability to pursue a betting strategy that yields a return significantly above the average loss to all bettors (i.e. track take). Researchers who subscribe to this theory are plentiful including Snyder (1978), Fabricand (1979), Hausch, Ziemba & Rubinstein (1981); Asch, Malkiel & Quandt (1982; 1984); Ziemba & Hausch (1984); Crafts (1985), Zuber, Gander & Bowers (1985) and Hausch, Lo and Ziemba (1994). If it is in fact true that the win pool is an efficient market, then the basic premise behind market efficiency would make wagering in the win pool futile.

Given this, numerous researchers and bettors have looked to other wagering pools to seek profitable strategies to apply at the racetrack. Specifically, Ziemba and Hausch (1985) and Asch and Quandt (1987) draw the parallel between an efficient win pool and the exotic pools. They maintain that if win pool probabilities are reasonably accurate estimates for win bets, then win pool probabilities alongside such models as Harville (1973) or Henery (1981) can be applied to exotic bets such as the quinella or exacta to capitalize on potential inefficiencies.

One major assumption of both the Harville and Henery models is that their formulas rely on probabilities generated solely from the win pool to establish estimates of probabilities in the exotic pools (i.e. quinella, exacta etc.). Computationally however, the Harville method is much simpler than the Henery model and subsequently is much more commonly utilized.

Since both models rest exclusively on the assumption that win pool probabilities are reasonable estimates for win bets, it is logical to examine this premise further. This assumption is based largely on data collected over thousands of races that shows entries that have fraction p of the win pool as a group, win approximately fraction p of the races. But as contested by Rosenbloom (1999), to support this assumption would be equivalent to assuming that because it snows 36 days a year in Winnipeg, the probability of snow on any given day is 10% even on a day in July. Furthermore, the well-known occurrence of the favorite-longshot bias casts additional doubtfulness on the premise underlying these two models.

B. Favorite-Longshot Bias

The favorite-longshot bias is probably the most incessant empirical regularity in racetrack literature. This phenomenon demonstrates that betting odds provide biased estimates of the probability of each horse winning. Specifically, despite the fact that it is actuarially unfair to bet at any odds, data shows that this is accentuated by long-shots. That is, long-shots (low-probability horses) are significantly over bet while favorites (high-probability horses) are under bet relative to their objective probabilities. This occurrence is confirmed in table 3.1 where it is shown that the longer the odds, the lower the average returns and the shorter the odds, the higher the average returns.

Table 3.1: Expected Value Per Dollar Bet for Different Odds Levels in 35,285 Races Run During 1947-1975; Source Snyder (1978)

Rates of Return on Bets to Win by Grouped Odds, Take Added Back

Study	0.75	1.25	2.5	5.0	7.5	10.0	15.0	33.0
Fabricant	11.1 ^a	9.0 ^a	4.6 ^a	-1.4	-3.3	-3.7	-8.1	-39.5 ^a
Griffith	8.0	4.9	3.1	-3.1	-34.6 ^a	-34.1 ^a	-10.5	-65.5 ^a
McGlothlin	8.0 ^b	8.0 ^a	8.0 ^a	-0.8	-4.6	-7.0 ^b	-9.7	-11.0
Seligman	14.0	4.0	-1.0	1.0	-2.0	-4.0	-7.8	-24.2
Snyder	5.5	5.5	4.0	-1.2	3.4	2.9	2.4	-15.8
Weitzman	9.0 ^a	3.2	6.8 ^a	-1.3	-4.2	-5.1	-8.2 ^b	-18.0 ^a
Combined	9.1 ^a	6.4 ^a	6.1 ^a	-1.2	-5.2 ^a	-5.2 ^a	-10.2 ^a	-23.7 ^a

^aSignificantly different from zero at 1% level or better

^bSignificantly different from zero at 5% level or better

While betting on favorites is an almost certain guarantee that you will cash more tickets given the entries solid track records, there is a considerably reduced return on investment

since many other handicappers will have backed the same winners. On the other hand however, betting on long-shots implies that you will cash fewer tickets given the entries inherently remote chance of winning based on past track records. Again we have the same risk versus return scenario where the return on investment will be significantly improved because only a small fraction of the betting public will share equally in the total amount of money bet at the race. Similar to stock markets, wagering strategies are a function of a persons desire for a small, steady return on investment or alternatively the willingness to withstand a fair amount of losing before the more occasional 'jackpot' is obtained. There are a number of explanations that attempt to uncover the uneven distribution of wagers on long-shots including risk love, the thrill of winning big and the influence of casual bettors.

- I. **Risk Love:** In their paper, Asch and Quandt (1990) describe risk love as the process of an individual giving up expected return to acquire more risk, because this risk is desirable. This corresponds to the basic premise of neoclassical theory which depends on individuals operating rationally and each seeking to maximize their individual utility or profit. In essence, patrons gain extra utility from betting on long-shots which provides rationalization for them to accept lower average returns. Specifically, bettors have a risk-loving utility function.
- II. **Winning Big:** A second explanation for a gamblers preference for long-shots lies in the prospect of winning 'big' and the associated ego that is gained. In their paper, Hausch, Ziemba and Rubenstein (1980) position that luck and entertainment are largely absent in betting favorites and so long-shots offer the thrill of successfully detecting a moderate or long odds winner. Since this achievement

confirms one's ability to outperform other bettors, added excitement is obtained making the low probability wager worthwhile.

- III. **Casual Bettors:** A further validation is that casual, less informed bettors wager too evenly across the board. This lack of knowledge, or lack of attention directed towards probabilities distorts the pool since wagers are distributed more evenly across all racing entries. The result is that long-shots receive more bets and favorites receive too few, relative to their true chances of winning.

Regardless of the explanation that seems most suitable, the inefficiencies that arise present potential opportunities in leveraging returns at the racetrack. The next chapters outline three experiments that focus on determining possible inefficiencies at the racetrack including:

1. Which is the more accurate estimate for the exacta probability?
 - a. Exacta probabilities from the exacta pool
 - b. Exacta probabilities from the win pool applying the Henery method
 - c. Exacta probabilities from the win pool applying the Harville method
 - d. Exacta probabilities from the win pool applying Asch and Quandt's regression method to correct for the favorite-longshot bias
2. Which is the more accurate estimate for the quinella probability?
 - a. Quinella probabilities from the quinella pool
 - b. Quinella probabilities from the win pool applying the Henery method
 - c. Quinella probabilities from the win pool applying the Harville method

- d. Quinella probabilities from the win pool applying Asch and Quandt's regression method to correct for the favorite-longshot bias
3. Which is the more accurate estimate for the win probability?
- a. Win probabilities from the exacta pool
 - b. Win probabilities from the win pool
 - c. Win probabilities from the win pool applying Asch and Quandt's regression method to correct for the favorite-longshot bias

Chapter 4: Exacta Wagering

Belonging to the collection of exotic wagers, the exacta corresponds to the first two horses to cross the finish line in the exacta order (win and place), with wagers being placed in either a straight, boxed or wheeled format. In recent years, exotic wagers have become more common at the racetrack, soaring in popularity over the historically favorite straight pools (win, place and show). Specifically, the exacta is estimated to account for approximately a 30% share of total wagers placed at the racetrack².

One attempt to explain why the exacta wager has become more universal, is due to the ability to incorporate multiple combinations in the form of boxes and wheels. In essence, this provides alternative wagering options on races with no clear favorites or on races with a dominant choice while at the same time offering the possibility of major rewards with minor investments. Furthermore, the increased interest in the exacta (and other exotic wagers), has been further explained by a number of researchers who look to the notion of 'smart money'. In the racing context, smart money can be defined as money bet or invested by experienced gamblers or investors who have inside information. In an effort to protect their wagers, bettors will avoid straight plays where odds are continuously displayed on the tote board and instead place their wagers in the exotic pools where odds are displayed less frequently. In effect, this limits the freeloading attempts of bettors who are less informed and bet simply on the odds generated from other racetrack patrons.

² www.churchhilldowns.com

In order to capitalize on any inefficiencies, an accurate estimate of the win and place combination yielding the highest probability is required. The objective thus becomes to determine the method that calculates the most accurate probabilities for win and place combinations. This supposition is explored, and in doing so is the first to empirically test win pool probabilities using Henery, Harville and Asch and Quandt's corrected favorite-longshot bias model against the probabilities produced from the exacta pool. Specifically, four methods are compared from which exacta probabilities can be calculated, namely Harville; Henery; corrected favorite-longshot bias and the exacta pool to determine which model is statistically best.

The Harville Model

The Harville model is the simplest and probably the most commonly used model by academics and bettors alike. This method calculates the probability for exacta combinations in terms of win probabilities only. This model assumes that if horse 'i' wins the race, the conditional probability of horse 'j' coming in second is given by

$$P_{j|i} = \frac{P_j}{(1 - P_i)} \quad (4.1)$$

where P_i and P_j are the probabilities of horse 'i' and 'j' winning the race. The probability of a horse winning the race is estimated by using the fraction of the win pool bet on that horse. Combining probabilities for the win and place horse, the probability of an i-j exacta becomes

$$P_{ij} = \frac{P_i \times P_j}{(1 - P_i)} \quad (4.2)$$

To illustrate the above method, consider the following data from the 9th race at Belmont Park on Sunday September 26th, 2004:

Wager Type	Winning Number	Paid	Pool	Win Take
Exacta	2-3	\$26.00	\$270,499.00	14.00%

Horse No.	Odds	Breakage	P_i
2	0.30	0.025	0.649
3	30.75	0.125	0.027

From the above data, P_{ij} can be calculated:

$$P_{23} = \frac{0.649 \times 0.027}{(1 - 0.649)} = 0.0498$$

The Harville formula is a natural and obvious way of estimating the probability of an exacta. However, there are problems with this model including the previously discussed favorite-longshot bias where longshots tend to be over bet while favorites tend to be under bet.

In addition a further shortcoming of the Harville formula is it is consistent with running times being distributed as independent exponential random variables. While the independence assumption may be reasonable, empirical distributions are far from exponential.

The Henery Model

Henery (1981) makes the more logical assumption that running times are independent normal with unit variance (i.e. the time for horse 'i' is normally distributed with mean θ_i and standard deviation 1). Unfortunately, with this assumption there is no closed form solution to the probability of the i-j exacta, P_{ij} . However, P_{ij} can be approximated. In doing so, θ_i must be estimated by solving the equation

$$P_i = \Phi \left(\frac{Z_0 + \theta_i \mu_{1,n}}{((n-1)\phi(Z_0))} \right) \quad (4.3)$$

for θ_i where Φ is the cumulative distribution function for the standard normal, ϕ is the density function of the standard normal, n is the number of betting entries in the race, $\mu_{i,n}$ is the expected value of the i-th standard normal order statistic in a sample size of n and $z_0 = \Phi^{-1}(1/n)$. To determine $\mu_{i,n}$, we apply values from tables obtained from Teichrow (1956) that approximate $\mu_{i,n}$ for n up to 20. Finally, after all θ_i are estimated, P_{ij} is approximated with the following formula:

$$P_{ij} = \Phi \left[a + \gamma \left\{ \frac{\theta_i \mu_{1,n} + \theta_j \mu_{2,n} + (\theta_i + \theta_j)(\mu_{1,n} + \mu_{2,n})}{(n-2)} \right\} \right] \quad (4.4)$$

Where

$$a = \Phi^{-1}\left(\frac{1}{n(n-1)}\right)$$

And

$$\gamma = \frac{1}{(n(n-1)\phi(a))}$$

Additionally, to satisfy the fact that the sum of the probabilities should add up to one, the P_{ij} 's are normalized.

Consider the following numerical example from the 9th race at Belmont Park on Sunday September 26, 2004:

Finish	Odds	Breakage	Probability
1	0.30	0.025	0.626
2	30.75	0.125	0.026
3	7.10	0.050	0.102
4	5.60	0.050	0.125
5	13.70	0.050	0.056
6	19.60	0.050	0.040

where

n	6
Z_0	-0.96742
$\mu_{1,n}$	-1.2672
$\mu_{2,n}$	-0.64175
$\varphi(Z_0)$	0.249851
a	-1.83391
γ	0.449042

Given the 1-2 finish, the matrix below is consulted to obtain the corresponding probability estimate:

	1	2	3	4	5	6
1		0.088	0.151	0.164	0.118	0.104
2	0.024		0.003	0.003	0.002	0.001
3	0.071	0.005		0.014	0.008	0.006
4	0.084	0.006	0.015		0.010	0.008
5	0.044	0.002	0.006	0.007		0.003
6	0.034	0.002	0.004	0.005	0.003	

$$P_{1,2} = 0.088$$

Using data from Hong Kong, Lo and Bacon-Shone (1994) found that the Henery model is more accurate than the Harville model. However, likely due to the fact that the Henery model is so complex, it has rarely been implemented.

The Win Pool Model Corrected for the Favorite-Longshot Bias

In their paper, "Efficiency and Profitability in Exotic Bets," Asch and Quandt (1987) propose a corrected win pool model for calculating exacta probabilities. To obtain the regression equation, Asch and Quandt calculated the average relationship between the objective and subjective probabilities using data from 705 races containing 6729 horses aggregated into 20 classes, as displayed in Table 4.1 below.

Table 4.1: Asch and Quandt's Regression Data

Objective (p_i) and the mean subjective (s_i) winning probability estimates for Meadowlands data

p_i	s_i
0.003	0.0068
0.003	0.011
0.0059	0.151
0.0119	0.0201
0.0178	0.0258
0.0445	0.0323
0.0297	0.0398
0.0386	0.0489
0.0415	0.0581
0.0804	0.0683
0.0804	0.0787
0.0923	0.0908
0.131	0.1036
0.1161	0.1185
0.125	0.1368
0.1637	0.1587
0.1548	0.1841
0.2024	0.222
0.3006	0.2747
0.4554	0.4031

To use this method, a number of sequences are required beginning with determining the fraction of the win pool bet on each horse $s_i = W_i / W$. Next, the subjective probabilities

are 'corrected' by applying a regression equation, determined by regressing objective probabilities gathered from the exacta pool on subjective probabilities determined from win pool estimates. To obtain the true probability estimate for the win and place finishers, the implied probability for each horse s_i , is substituted into the regression equation below,

$$P_i = -0.0100 + 1.0959 s_i \quad (4.5)$$

where a P_i value between 0 and 1 is achieved when S_i is in the range of 0.009125 and 0.921617.

Next, the Harville formula is applied to the objective probabilities to determine the exacta probability estimate for the first and second place horses, P_i and P_j :

$$P_{ij} = \frac{P_i \times P_j}{(1 - P_i)} \quad (4.6)$$

Consider an illustration of Asch and Quandt's corrected exacta probability estimate, using data from the 9th race at Belmont Park on Sunday September 26th, 2004:

Wager Type	Winning Number	Paid	Pool	Win Take
Exacta	2-3	\$26.00	\$270,499.00	14.00%

Step one: determine the subjective probabilities for horse 'i' and 'j':

Horse No.	Odds	Breakage	S_j
2	0.30	0.025	0.6491
3	30.75	0.125	0.0270

where

$$S_2 = \frac{(1-.14)}{(1+0.30+0.025)} = 0.6491$$

$$S_3 = \frac{(1-.14)}{(1+30.75+0.125)} = 0.0270$$

Step two: substitute the subjective probabilities s_i into Asch and Quandt's regression equation to obtain corrected values for horse 'i' and 'j'.

$$P_2 = -0.0100 + 1.0959(0.649) = 0.7013$$

$$P_3 = -0.0100 + 1.0959(0.027) = 0.0196$$

Step three: apply Harville's formula to the objective probabilities to obtain exact probability estimates corrected for the favorite-longshot bias.

$$P_{23} = \frac{(.70124) \times (.01959)}{(1-.70124)} = 0.046$$

The Exacta Pool Model

A fourth possible model is to assume that the probability of the i-j exacta is given by the fraction of the exacta pool bet on the i-j exacta. In effect, this assumes the exacta pool is an efficient market. Surprisingly this very simple model has been deemed inaccurate. While the ability to wheel and box exotics may explain the increased popularity of these wagers, Ziemba and Hausch claim these same factors contribute to inaccurate probabilities. For instance, many bettors make a "wheel" bet where the wager includes every combination, with a favorite in the first position of the exacta bet. It can be argued that with wheel bets, combinations with a favorite in the first position and a long-shot in the second position are over-bet. Alternatively, another popular bet is the "box" bet. If a bettor believes the best three horses in order are A,B and C, then the bettor makes equal bets on all exacta combinations involving A,B and C. The result is that even though the bettor believes the AB exacta is more likely than the CB exacta, the bettor makes equal wagers on these combinations and therefore distorts the exacta pool. While these arguments are logical there has never been an empirical test of this argument.

To calculate the probability of an exacta using probabilities generated from the exacta pool, we take the fraction of the exacta pool bet on the i-j exacta factoring in track takeout and breakage. To illustrate this calculation consider the following example:

Wager Type	Winning Number	Paid	Breakage	Pool	Exacta Take
Exacta	2-3	\$26.00	\$0.10	\$270,499.00	17.50%

$$P_{23} = \frac{(1 - .175)}{\left(\frac{(26 + 0.10 - 2)}{2} \right) - 1} = 0.0632$$

An SPRT-Like Test

The sequential probability ratio test (SPRT) due to Wald (1947), is a test of two hypotheses for a simple hypothesis H_1 against an alternative simple hypothesis H_2 . After each sampling stage, the likelihood ratio L_1/L_2 (where suffixes 1 and 2 correspond to hypotheses 1 and 2 respectively and L is the likelihood function of all sample members drawn to date), is calculated. This sampling procedure is repeated while $B < L_1/L_2 < A$, and terminates when $L_1/L_2 \leq B$ (where H_2 hypothesis is accepted) or when $L_1/L_2 \geq A$ (where hypothesis H_1 is accepted) where the two positive constants A and B are determined by the prescribed requirements of Type I and Type II errors.

Wald proved that under certain regularity conditions, the SPRT will terminate in a finite number of steps with probability one, provided that the data is independent and identically distributed. For an experiment where the data is not identically distributed however, and there are two or more parameter estimation methods, Rosenbloom's (2000) and (2003) SPRT-like statistical test can be applied.

The SPRT-like method provides a Bayesian interpretation of the stopping rule and guarantees termination in deciding the best of k probability forecasting models. If it is assumed before data collection that each hypothesis has an equal chance of being correct,

then R_j can be viewed as the posterior probability of Hypothesis j being correct. Given this interpretation, the expression for R_j is simply Bayes' formula

$$R_{j,t} = \frac{B_j L_{j,t}}{\sum_{i=1}^K B_i L_{i,t}} \quad (4.7)$$

which can be rewritten for computational purposes as

$$R_{j,t} = \frac{1}{\sum_{i=1}^K (B_i / B_j)(L_{i,t} / L_{j,t})} \quad (4.8)$$

where $L_{j,t}$ = Likelihood under Hypothesis H_j of obtaining data x_1, x_2, \dots, x_t , and B_j = Prior probability that probability forecasting model j is correct (typically a uniform prior is assumed).

In this experiment, the Henery model, Harville model, Asch and Quandt's regression model and Exacta pool model represent four multinomial parameter estimation procedures when only one observation per race is possible but the four estimation procedures can be repeated many times for different races.

The underlying multi-hypothesis test is as follows:

H_1 : Henery model probabilities are correct

Versus

H_2 : Harville model probabilities are correct

Versus

H₃: Asch and Quandt's regression model probabilities are correct

Versus

H₄: The exacta pool probabilities are correct

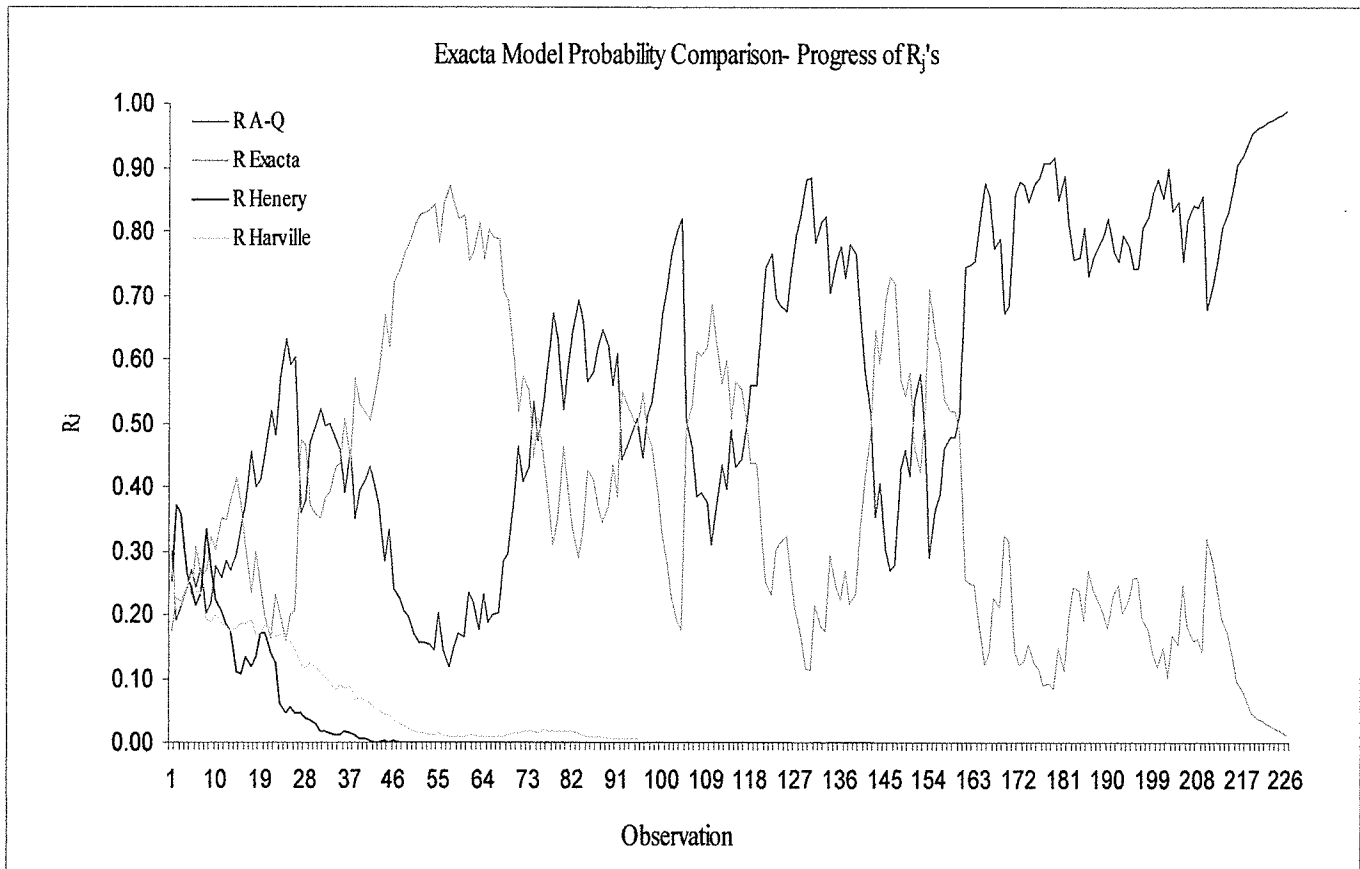
The resulting SPRT-like test is as follows:

- The significance level α was set equal to $\alpha = 0.01$ and the maximum sample size M was set equal to 1,000. If the maximum sample size is reached the test is inconclusive.
- Data was collected from Belmont Park beginning with the first race on Sunday September 26th, 2004 and continued thereafter until the experiment concluded.
- After each race, the exacta probability was calculated for each of the four methods under consideration, namely Harville, Henery, Asch and Quandt's regression method and Exacta Pool. The likelihood ratios and revised probabilities were updated.
- Once a revised probability is above $1-\alpha$, or 0.99, the associated hypothesis was accepted.

The data from the experiment is contained in Appendix A, Table 4.1: Exacta Model Probability Comparison.

A graphical representation of the revised probabilities for each of the four models under consideration is depicted below in Figure 4.1.

Figure 4.1: Exacta Model Probability Comparison



Results

Exacta probabilities produced by the four forecasting methods (Henery, Harville, regression and Exacta Pool) were often quite different. After 226 races, the SPRT-like test found significant differences between the four methods: $H_3 =$ Asch and Quandt's regression model probabilities are correct, was accepted. We can therefore conclude that

the win pool corrected for the favorite-longshot bias through Asch and Quandt's regression equation produced exacta probabilities that are more accurate than those produced by Henery or Harville win pool methods, or those calculated from the exacta pool itself.

In addition, from the data contained in Appendix A: Table 4.1, we can see that if this was a test comparing only the exacta pool and the Henery and Harville win pool methods, the experiment would have terminated after 53 races (significance level $\alpha = 0.01$). Under this scenario, the exacta pool method is superior to the Henery and Harville models.

Chapter 5: Quinella Wagering

Falling under the umbrella of exotic wagering, the quinella is akin to the exacta box, considering the first two horses to cross the finish line without regard to order. Despite the comparability of the two, each wager occupies its own pool as defined by the pari-mutuel system, and thus generates a unique payout. Unlike the exacta box, the quinella does not double the cost of the wager and payoffs are usually half of what the exacta pays out. Like the exacta, the quinella's affiliation to the exotic market, makes this a very popular bet.

As demonstrated previously, Ziemba and Hausch (1985) and Asch and Quandt (1987) state probabilities generated from the win pool (alongside methods such as Henery and Harville) are more accurate estimates than the probabilities generated from the exotic pool in question. As such, they propose that win pool probabilities can be used to leverage returns in the exotic market. In the first experiment on exacta wagering, it was shown that implied win pool probabilities using Asch and Quandt's regression equation corrected for the favorite-longshot bias produced results that were superior to other win pool probabilities using Henery, Harville, or the Exacta pool itself. Given this result, and the many similarities between the exacta and quinella wager, quinella probabilities can be compared to determine the most accurate method. In testing this premise, four methods from which quinella probabilities can be extracted are compared, including Harville; Henery; Asch and Quandt's regression method and the Quinella Pool.

The Harville Model

The Harville model has been identified as one of the simplest and most commonly used models, calculating probabilities in exotic pools in terms of win pool probabilities only. Given that the quinella wager has the same effect as boxing an exacta, the same formula that was used to calculate the exacta probability under the Harville model can be used by summing the probabilities of an AB and BA exacta. Specifically, this model assumes that if horse 'i' wins the race, the conditional probability of horse 'j' coming in second is given by

$$P_{j|i} = \frac{P_j}{(1 - P_i)} \quad (5.1)$$

where P_i and P_j are the probabilities of horse 'i' and 'j' winning the race. Since the quinella wager pays out for the top two finishers in either order, the reverse must also be accounted for. If horse 'j' wins the race, the conditional probability of horse 'i' coming in second is given by

$$P_{i|j} = \frac{P_i}{(1 - P_j)} \quad (5.2)$$

Similar to the calculation for the exacta probability, the probability of a horse winning the race is estimated by using the fraction of the win pool bet on that horse. Combining this with the probability for the place horse, and factoring in the reverse order, the probability of an i-j quinella becomes

$$P_{ij} = \frac{P_i \times P_j}{(1 - P_i)} + \frac{P_j \times P_i}{(1 - P_j)} \quad (5.3)$$

To illustrate the above formula, consider the following data from the 1st race at Delta Downs on May 8th, 2005:

Wager Type	Winning Number	Paid	Pool	Win Take
Quinella	4-6	\$14.00	\$1,089	17.00%

Horse No.	Odds	Breakage	P _j
4	2.10	0.025	0.2656
7	4.80	0.050	0.1419

From the above data, P_{ij} can be calculated:

$$P_{ij} = \frac{(.2656) \times (.1419)}{(1 - .2656)} + \frac{(.1419) \times (.2656)}{(1 - .1419)} = 0.0952$$

The same shortcomings for the Harville model hold true under the quinella as for the exacta including the favorite-longshot bias and inconsistent running times being distributed as independent exponential random variables.

The Henery Model

A second method to estimate quinella probabilities is the Henery model. As demonstrated previously, this model is quite complex requiring significant calculations to

extract the sought after probability. The P_{ij} can be determined by approximating θ_i by solving the equation

$$P_i = \Phi \left(\frac{(Z_0 + \theta_i \mu_{1,n})}{((n-1)\varphi(Z_0))} \right) \quad (5.4)$$

for θ_i where Φ is the cumulative distribution function for the standard normal, φ is the density function of the standard normal, n is the number of betting entries in the race, $\mu_{i,n}$ is the expected value of the i -th standard normal order statistic in a sample size of n and $z_0 = \Phi^{-1}(1/n)$. To determine $\mu_{i,n}$, we apply values from tables obtained from Teichrow (1956) that approximate $\mu_{i,n}$ for n up to 20. Finally, after all θ_i are estimated, P_{ij} is approximated with the following formula:

$$P_{ij} = \Phi \left[a + \gamma \left\{ \frac{\theta_i \mu_{1,n} + \theta_j \mu_{2,n} + (\theta_i + \theta_j)(\mu_{1,n} + \mu_{2,n})}{(n-2)} \right\} \right] + \Phi \left[a + \gamma \left\{ \frac{\theta_j \mu_{1,n} + \theta_i \mu_{2,n} + (\theta_j + \theta_i)(\mu_{1,n} + \mu_{2,n})}{(n-2)} \right\} \right] \quad (5.5)$$

Where

$$a = \Phi^{-1} \left(\frac{1}{n(n-1)} \right)$$

And

$$\gamma = \frac{1}{(n(n-1)\varphi(a))}$$

Additionally, to satisfy the fact that the sum of the probabilities should add up to one, the P_{ij} 's are normalized.

Consider the following numerical example from the 9th race at Belmont Park on Sunday September 26, 2004:

Finish	Odds	Breakage	Probability
1	2.10	0.025	0.2656
2	4.80	0.050	0.1419
3	5.20	0.050	0.1328
4	2.50	0.025	0.2355
5	5.70	0.050	0.1230
6	20.70	0.050	0.0382
7	10.60	0.050	0.0712

where

n	7
Z₀	-1.06757
U_{1,n}	-1.35217
U_{2,n}	-0.75737
$\phi(Z_0)$	0.225645
a	-1.98075
γ	0.424415

Given the 1-2 finish, the matrix below is consulted to obtain the corresponding probability estimate:

	1	2	3	4	5	6
1		0.049	0.047	0.069	0.044	0.022
2	0.038		0.022	0.035	0.021	0.010
3	0.036	0.022		0.032	0.019	0.009
4	0.066	0.042	0.040		0.038	0.019
5	0.033	0.020	0.019	0.030		0.008
6	0.011	0.006	0.006	0.010	0.005	

$$P_{ij} = 0.049 + 0.038 = 0.087$$

The Win Pool Model Corrected for the Favorite-Longshot Bias

Asch and Quandt's corrected win pool model serves as a third method for calculating quinella probabilities. The first step to extracting quinella probabilities using this method is to determine the fraction of the win pool bet on each horse $s_i = W_i / W$, for the win and place finishers. To obtain the true probability estimate for the win and place finishers, the implied probability for each horse s_i , is substituted into the regression equation below:

$$P_i = -0.0100 + 1.0959 s_i \quad (5.6)$$

Next, the Harville formula is applied to the objective probabilities taking into account that a successful quinella wager allows for the top two finishers in **either** order.

$$P_{ij} = \frac{P_i \times P_j}{(1 - P_i)} + \frac{P_j \times P_i}{(1 - P_j)} \quad (5.7)$$

To consider an illustration of Asch and Quandt's corrected quinella probability estimate, consider the following data from the 4th race at Bay Meadows on January 7th, 2006:

Wager Type	Winning Number	Paid	Pool	Win Take
Quinella	2-6	\$32.20	\$7,949	15.40%

Step one: determine the subjective probabilities for horse 'i' and 'j':

Horse No.	Odds	Breakage	S _i
2	050	0.025	0.555
6	23.00	0.050	0.035

where

$$S_2 = \frac{(1 - .154)}{(1 + .50 + .025)} = .555$$

$$S_6 = \frac{(1 - .154)}{(1 + 23.00 + .055)} = .035$$

Step two: substitute the subjective probabilities s_i into Asch and Quandt's regression equation to obtain corrected values for horse 'i' and 'j'.

$$P_2 = -0.0100 + (1.0959 \times 0.555) = 0.598$$

$$P_6 = -0.0100 + (1.0959 \times 0.035) = 0.028$$

Step three: apply Harville's formula to the objective probabilities to obtain quinella probability estimates corrected for the favorite-longshot bias.

$$P_{26} = \frac{(.598) \times (.028)}{(1 - .598) + (.028 \times .598)} = 0.060$$

$$(1 - .028)$$

The Quinella Pool Model

The fourth model this experiment considers, calculates the probability of a quinella based on the fraction of the quinella bet on the i-j quinella. As in the exacta pool method, the quinella pool model corresponds to the efficient market hypothesis. To date, this very simple model has not been widely applied. The most logical reason for the resistance of this models lies in the fact that the quinella wager gives equal weight for both horses A and B in the wager despite the fact that the bettor may believe A has a greater chance of winning. In theory, this results in distortion of the pool. While this argument intuitively has merit, this argument has yet to be tested for the quinella.

To calculate the probability of a quinella using probabilities generated from the quinella pool, we take the fraction of the quinella pool bet on the i-j quinella factoring in track takeout and breakage. To illustrate this calculation consider the following example:

Wager Type	Winning Number	Paid	Breakage	Pool	Quinella Take
Quinella	4-6	\$14.00	\$0.10	\$1,089	20.5%

$$P_{23} = \frac{(1 - 0.205)}{\frac{(14.00 + .10 - 2)}{2} - 1} = 0.1128$$

An SPRT-Like Test

The Harville, Henery, regression and Quinella pool models represent four multinomial parameter estimation procedures when only one observation per race is possible but the four estimation procedures can be repeated many times for different races.

The underlying multi-hypothesis test is as follows:

H₁: Henery model probabilities are correct

Versus

H₂: Harville model probabilities are correct

Versus

H₃: Asch and Quandt's regression model probabilities are correct

Versus

H₄: The quinella pool probabilities are correct

The resulting SPRT-like test is as follows:

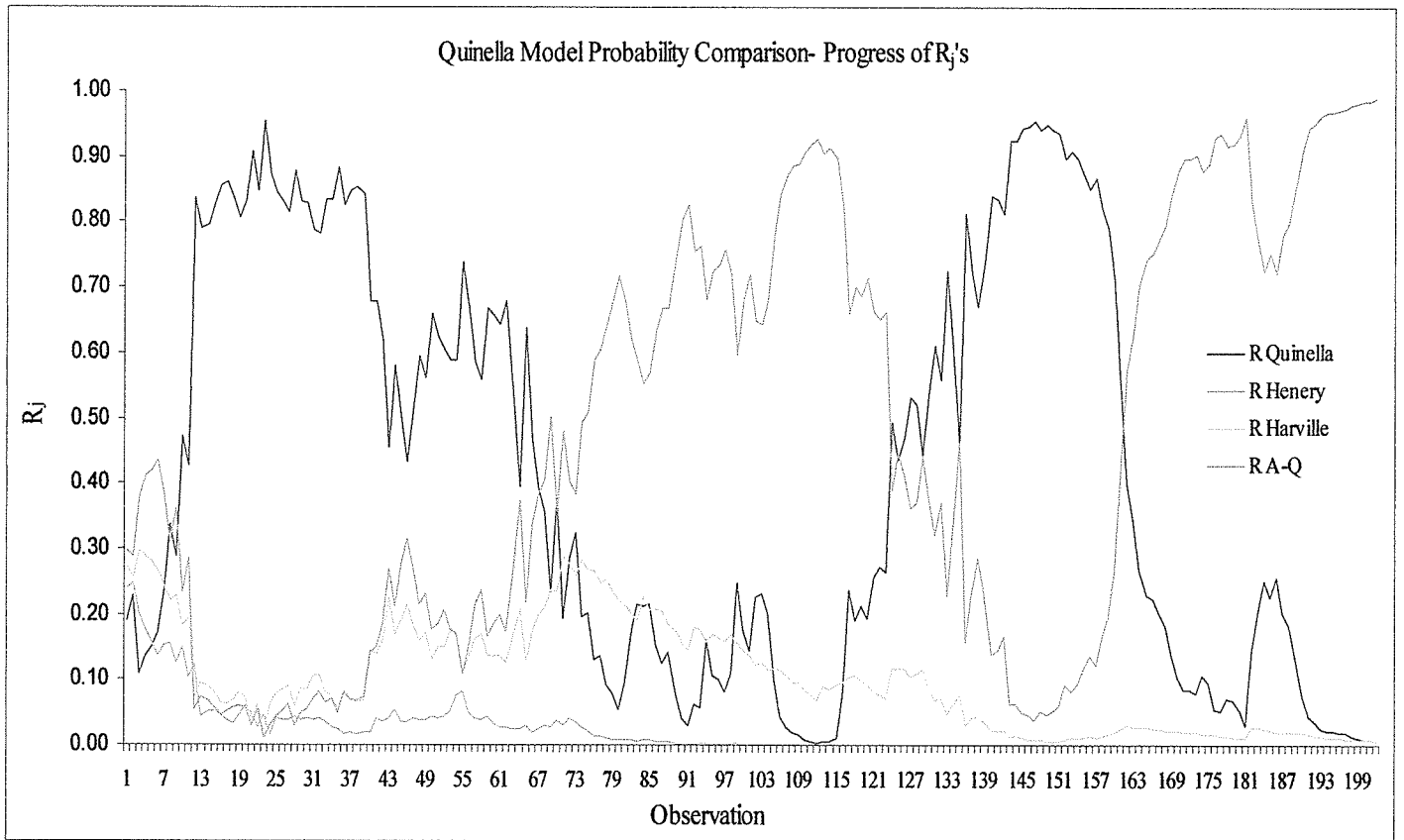
- The significance level α was set equal to $\alpha = 0.01$ and the maximum sample size M was set equal to 1,000. If the maximum sample size is reached the test is inconclusive.

- Data was collected from Delaware Park, Bay Meadows and Evangeline Downs beginning in October 2004 and continuing thereafter until the experiment concluded.
- After each race, the quinella probability was calculated for each of the four methods under consideration, namely Henery, Harville, Asch and Quandt's regression and Quinella Pool methods. The likelihood ratios and revised probabilities were updated.
- Once a revised probability is above $1-\alpha$, or 0.99, the associated hypothesis was accepted.

The data from the experiment is contained in Appendix B, Table 5.1: Quinella Model Probability Comparison.

A graphical representation of the four models under consideration is presented in Figure 5.1.

Figure 5.1: Quinella Model Probability Comparison



Results

Quinella probabilities produced by each of the four models result in probabilities that are significantly dissimilar. After 202 races, the SPRT-like test found significant differences among the four methods: $H_3 = \text{Asch}$ and Quandt's regression probability model is correct, was accepted. Based on this dataset, it can be concluded that win pool probabilities generated from Asch and Quandt's regression method corrected for the favorite-longshot bias are more accurate than those produced by the win pool calculated by the Henery and Harville models and those calculated from the quinella pool itself.

Looking at the dataset in more detail, it can be concluded that if this was a test looking at the quinella pool model and Henery and Harville win pool models in isolation, the Quinella pool model would be statistically superior after 94 races (significance level $\alpha=.01$). Similarly, we can partition the dataset and look only at the Henery and Harville models. In this scenario, the Harville model would prove superior to the Henery model after 89 data points (significance level $\alpha=.01$). Both of these findings concur with those obtained in the previous chapter on exacta wagering.

Chapter 6: Win Wagering

The win wager is the most widely known and correspondingly the simplest bet at the racetrack. Belonging to a group of wagers termed straight wagers (win, place or show), a winning ticket picks the first horse to cross the finish line, or win the race. Typically, the win pool draws first-time and casual bettors who seek the thrill of winning but necessitate a simple wagering format. This pool is also attractive to more serious gamblers, particularly in races with a clear favorite since these entries can be wagered 'straight-up' without having to be coupled with other entrants as dictated by the exotic market.

Numerous studies have been conducted to determine the accuracy of the win pool. Many researchers agree that the win pool provides a reasonably good estimate for win bets: probabilities determined by the betting crowd are very close approximations to true racing outcomes. In fact, several researchers rely on this assumption so heavily, that many of their models rest solely on the supposition that the win pool generates accurate probabilities and can therefore be used to imply probabilities in other wagering pools. In the previous experiments on the exacta and quinella exotic wagers it was shown that the most accurate method for determining corresponding probabilities was through the application of the win pool using Asch and Quandt's regression method correcting for the favorite-longshot bias. These findings demonstrate the ability to leverage returns in an exotic pool by using implied probabilities generated from a straight pool. When looking to capitalize on the win wager, it is appropriate to evaluate various methods for computing win probabilities. As a result, this study contrasts probabilities produced

directly from the win pool against win pool probabilities corrected for the favorite-longshot bias and the exacta pool method to determine which model produces the most accurate probability estimate for the win bet.

Win Pool Model

The win pool model is a straightforward method that assumes the probability of horse ‘i’ winning the race can be estimated by the fraction of the win pool bet on horse ‘i’, factoring in track takeout and breakage.

$$P_i = \frac{(1 - \textit{Take})}{(1 + \textit{Odds} + \textit{Breakage})} \quad (6.1)$$

This corresponds to a number of research studies that support this assumption concluding that the betting public is very accurate in setting the odds in the win pool: “The odds set by the crowd, despite the vast differences among the betting styles, are actually very accurate; on average the chance that a given horse will win a race is very close to what the crowd as a consensus thinks it is” (Ziembra and Hausch (1985)).

To illustrate the calculation, consider the following data from Woodbine Harness on April 5th, 2005:

Finish	Winning Number	Paid	Payout	Break	Track Take
Win	1	\$9.70	\$9.70	\$0.00	16%

Finish	Odds
Win	$(\$9.70 / 2) - 1 = 3.85$

$$P_w = \frac{(1 - .16)}{(1 + 3.85)} = .1728$$

Exacta Pool Model

The exacta pool model can be used to extract the implied win pool probability of horse 'j' by taking the total amount of money bet on exacta combinations with horse 'j' in the first position divided by the total exacta pool. Given that the exacta pool is much larger than the win pool, it is reasonable to assume that the implied win pool probabilities are potentially more accurate from the exacta pool than the win pool itself. To arrive at the win pool probability estimate for horse 'j', exacta payouts for each racing combination must be obtained from the totalizer. This can be difficult to acquire since these combinations are displayed infrequently throughout post-time. With diligence however, a matrix can be used to calculate the odds for each racing combination using the following formula:

$$Odds = \frac{Exacta\ Price}{2} - 1 \quad (6.2)$$

Once the odds have been established, the probability for each racing combination can then be calculated based on the following equation:

$$P_i = \frac{(1 - Take)}{(1 + Odds + Breakage)} \quad (6.3)$$

It follows that by normalizing the data and then summing the probabilities for each racing entry, a corresponding win probability can be extracted. To demonstrate this model, consider the following example from Woodbine Harness on April 5th, 2005:

NOTE: \$2 Exacta final payouts were obtained directly from the totalizer at post-time courtesy of Assiniboine Downs in Winnipeg, Manitoba. The following chart, illustrates the final exacta payout for each racing combination.

Step one: collect the exacta payouts for each racing combination as demonstrated below.

Prices	1	2	3	4	5	6	7	8
1		\$ 33.90	\$ 43.00	\$ 146.90	\$ 125.80	\$ 96.60	\$ 43.30	\$ 136.10
2	\$ 29.80		\$ 35.30	\$ 125.60	\$ 91.10	\$ 51.80	\$ 33.30	\$ 106.50
3	\$ 39.90	\$ 43.10		\$ 179.30	\$ 110.00	\$ 58.70	\$ 46.60	\$ 165.50
4	\$ 216.30	\$ 192.80	\$ 267.50		\$ 304.70	\$ 192.80	\$ 262.30	\$ 621.00
5	\$ 201.80	\$ 191.30	\$ 227.60	\$ 380.80		\$ 243.40	\$ 204.50	\$ 338.00
6	\$ 74.90	\$ 77.20	\$ 79.70	\$ 193.90	\$ 147.40		\$ 74.50	\$ 200.30
7	\$ 38.80	\$ 33.60	\$ 40.20	\$ 147.10	\$ 114.30	\$ 53.10		\$ 145.50
8	\$ 234.50	\$ 177.60	\$ 349.40	\$ 1,072.90	\$ 349.30	\$ 318.70	\$ 335.80	

Step two: From the exacta prices above, the odds for each combination can be extracted:

Odds	1	2	3	4	5	6	7	8
1		15.95	20.5	72.45	61.9	47.3	20.65	67.05
2	13.9		16.65	61.8	44.55	24.9	15.65	52.25
3	18.95	20.55		88.65	54	28.35	22.3	81.75
4	107.15	95.4	132.75		151.35	95.4	130.15	309.5
5	99.9	94.65	112.8	189.4		120.7	101.25	168
6	36.45	37.6	38.85	95.95	72.7		36.25	99.15
7	18.4	15.8	19.1	72.55	56.15	25.55		71.75
8	116.25	87.8	173.7	535.45	173.65	158.35	166.9	

Step three: Calculate corresponding probabilities for each entry by factoring in track takeout, breakage and odds:

Probabilities	1	2	3	4	5	6	7	8
1		0.0432	0.0341	0.0100	0.0117	0.0152	0.0338	0.0108
2	0.0492		0.0415	0.0117	0.0161	0.0283	0.0440	0.0138
3	0.0367	0.0340		0.0082	0.0133	0.0250	0.0314	0.0089
4	0.0068	0.0076	0.0055		0.0048	0.0076	0.0056	0.0024
5	0.0073	0.0077	0.0064	0.0038		0.0060	0.0072	0.0043
6	0.0196	0.0190	0.0184	0.0076	0.0099		0.0197	0.0073
7	0.0378	0.0436	0.0365	0.0100	0.0128	0.0276		0.0101
8	0.0063	0.0083	0.0042	0.0014	0.0042	0.0046	0.0044	

Step four: To ensure all probabilities add to one, the data is normalized and the combination probabilities for each racing entry is summed:

Entry	Probability	Normalized Probability
1	0.1587	0.1732
2	0.2045	0.2231
3	0.1575	0.1718
4	0.0402	0.0439
5	0.0427	0.0466
6	0.1014	0.1107
7	0.1783	0.1945
8	0.0332	0.0362

Step five: Finally, the win probability for the horse in question, entry 1 can be determined:

$$P_1 = 0.1732$$

The Win Pool Model Corrected for the Favorite-Longshot Bias

Asch and Quandt's corrected favorite-longshot bias method is a third method that can be contrasted for estimating win probabilities. By substituting implied probabilities s_i into the regression equation, the probability estimate from the win pool is obtained:

$$P_i = -0.0100 + (1.0959 \times S_i) \quad (6.4)$$

As an example, consider the following data from Woodbine Harness on April 5th, 2005:

Step one: determine the subjective probability for horse 'i':

Finish	Winning Number	Paid	Payout	Break	Track Take
Win	1	\$9.70	\$9.70	\$0.00	16%

Finish	Odds
Win	$(\$9.70 / 2) - 1 = 3.85$

$$S_i = \frac{(1 - .16)}{(1 + 3.85)} = .1728$$

Step two: substitute the subjective probability s_i into Asch and Quandt's regression equation to obtain the corrected value for horse 'i':

$$P_1 = -0.0100 + (1.0959 \times .1728) = 0.1794$$

An SPRT-Like Test

The win pool, exacta pool and Asch and Quandt's regression models represent three multinomial parameter estimation procedures when only one observation per race is possible but the three estimation procedures can be repeated many times for different races.

The underlying multi-hypothesis test is as follows:

H_1 : win pool probabilities are correct

Versus

H_2 : exacta pool probabilities are correct

Versus

H_3 : Asch and Quandt's regression equation probabilities are correct

The resulting SPRT-like test is as follows:

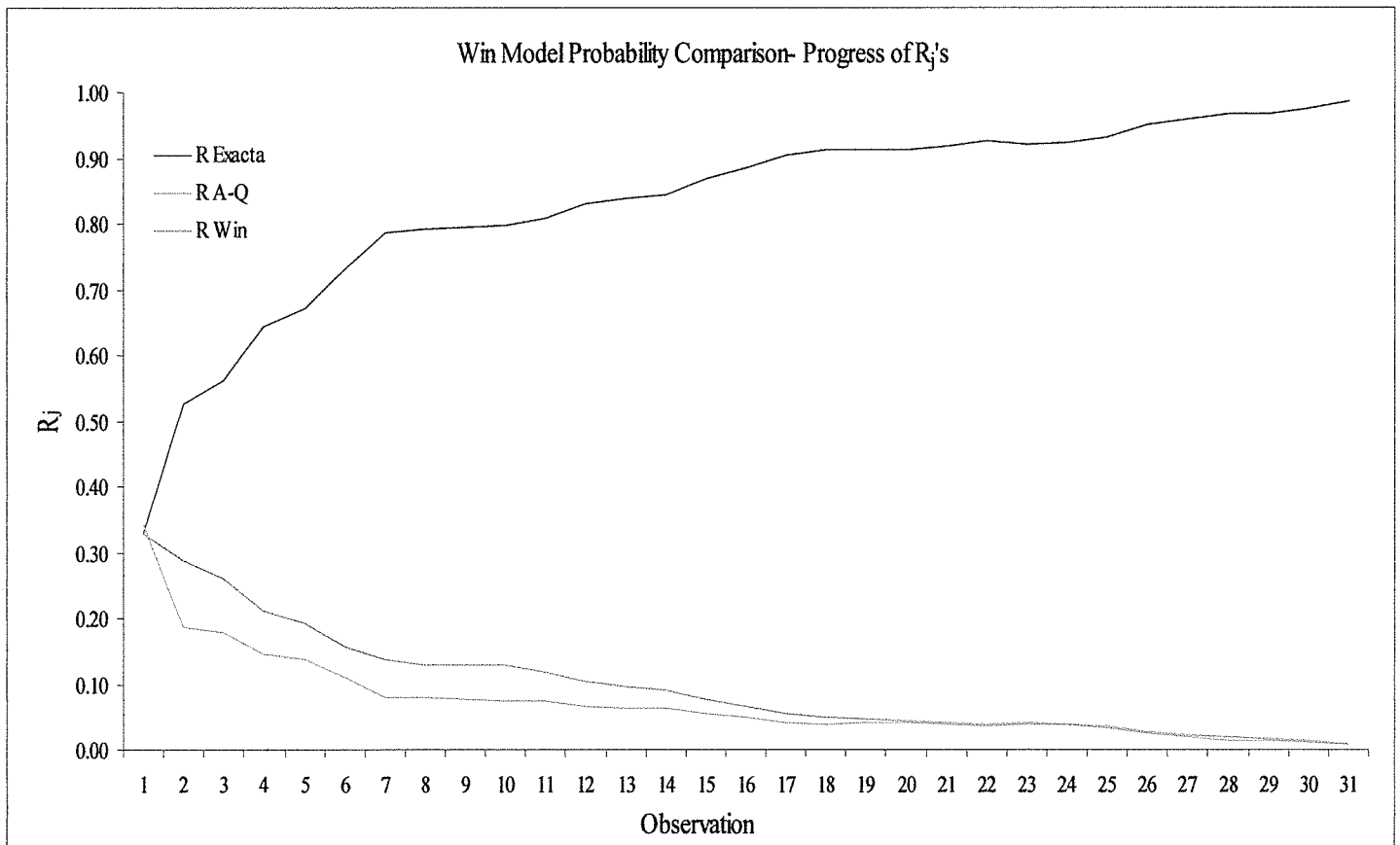
- The significance level α was set equal to $\alpha = 0.01$ and the maximum sample size M was set equal to 1,000. If the maximum sample size is reached the test is inconclusive.
- Data was collected from a variety of racetracks including Woodbine Harness, Western Fair, Turf Paradise, Tampa Bay, Kawartha Downs, and Belmont Park beginning May 9th 2005 and continued each race thereafter until the experiment concluded. This data was provided courtesy of Assiniboine Downs.

- After each race, the win probability was calculated for each of the three methods under consideration, namely win pool, exacta pool and Asch and Quandt's regression methods. The likelihood ratios and revised probabilities were updated.
- Once a revised probability is above $1-\alpha$, or 0.99, the associated hypothesis was accepted.

The data from the experiment is contained in Appendix C, Table 6.1: Win Model Probability Comparison.

A graphical representation of the three probability forecasting models is presented below in Figure 6.1.

Figure 6.1: Win Model Probability Comparison



Results

Win probabilities produced by each of the three models result in probabilities that are significantly different. After 31 races, the SPRT-like test found significant differences among the three methods: H_2 = the exacta pool method is correct, was accepted. Based on this dataset, it can be concluded that the extraction of win probabilities from the exacta pool provides a more accurate estimation for win probabilities than those produced by the win pool itself.

Chapter 7: Results

A variety of wagers exist to play at the racetrack. In this thesis, three wagers in particular are examined to determine the most accurate method for predicting probabilities for application in a value betting strategy. For these wagers, it was found that probabilities produced from outside pools are more accurate than those produced in the wagering pool in question. Specifically:

- When wagering on the exacta, it was found that probabilities applied from the win pool corrected for the favorite-longshot bias using Asch and Quandt's regression equation are superior to win pool probabilities using the Henery method, win pool probabilities using the Harville method as well as exacta pool probabilities themselves. Since payoffs in the exacta pool are determined by the crowd's probability estimates for the exacta, utilizing Asch and Quandt's regression method can exploit inefficiencies since a necessary condition for a winning betting strategy is to have better probabilities than the crowd.
- When playing the quinella, probabilities from the win pool corrected for the favorite-longshot bias using Asch and Quandt's regression equation are more accurate than those produced from the win pool using the Henery method, win pool using the Harville method and the quinella pool. Given that payouts on the quinella are calculated from the crowd's probability estimate for the quinella, application of Asch and Quandt's regression method gives way to exploiting inefficiencies.
- For the win wager, implied probabilities from the exacta pool are more accurate than straight win probabilities and win pool probabilities corrected for the favorite-

longshot bias using Asch and Quandt's regression equation. Again, since win bet payouts are determined from the probabilities in the win pool, the use of implied probabilities from the exacta pool allows a bettor to exploit inefficiencies by achieving probability estimates that are superior to the crowds.

Although paradoxical, a model using data from the exacta pool proves superior for estimating win probabilities, and a model using data from the win pool is shown to be better for estimating exacta and quinella probabilities. Intuitively however, these results can be substantiated and thus are quite realistic. Specifically, it is logical that the exacta pool generates better probabilities than the win pool since the popularity of the exacta wager has led to the increase of its' pool size relative to that of the win pool. For example, there is usually more money bet implicitly on horse 'j' to win in the exacta pool than explicitly in the win pool. Furthermore, it can be rationalized that the Ash and Quandt correction model generates more accurate exacta (quinella) probabilities estimates than the exacta (quinella) pool: although more money may be bet on horse 'j' in the exacta pool, the effects of wheels, boxes, and betting 'favorite numbers' may distort the probability estimate in the exacta pool of the j-k exacta permutation- this would not be true of the Asch and Quandt correction method.

Chapter 8: Wagering Strategy Using Positive Expected Values

The Efficient Market Hypothesis (EMH) states that at any point in time, prices reflect all available information: no amount of data mining can predict future prices. In the context of horse racing, no betting system can consistently earn super normal returns. Taking into consideration that track takeout accounts for a significant portion of the winning payout (i.e. 17% for straight bets and 22% for exotic bets), many studies surrounding market efficiency at the racetrack have concluded an inability to earn above average returns in the long run. For example, in 1978 Snyder established that albeit the win pool contains known inefficiencies like the favorite-longshot bias, once the track take is added back in, the bias is not large enough to beat the market consistently. In light of this, bettors continue to search for the optimum betting strategy in an attempt to yield a positive return, at least in the short term.

One betting strategy is to wager only those propositions providing positive expectation or expected value (often referred to as 'value betting'). In other words, value betting consists of betting only those opportunities that provide a positive average. The equation for expected value is provided below:

$$\text{Expected Value} = \text{Probability of winning} \times \text{Payoff} - \text{Cost} \quad (7.1)$$

To demonstrate, suppose you were to bet \$2 to win on a horse you considered had a 30% chance of winning. Also suppose that the odds on this horse were listed as 2.50-to-1. Would a bet on this horse be considered a value bet?

$$\text{Expected Value} = 0.30 \times (\$7/2) - \$1 = \$1.05 - \$1.00 = \$0.05$$

Since five cents is obviously positive, such a bet would be considered a value bet or in more familiar terms, an overlay (a horse “going off” at a higher price than he appears to warrant based on his past performances: also known as an under bet). In this example, for each \$1 wagered, you would expect to earn a net positive value of five cents. Since this is a positive expected value, a value betting strategy would recommend wagering on this horse.

Intuitively, the best wagering strategy is one that utilizes the most precise probabilities. Moreover, if probabilities from one pool prove more accurate in determining probabilities in another pool, then not only does this imply inefficiencies exist, but this also presents a possibility for a winning strategy. Specifically, given the pari-mutuel systems uniqueness where players bet against each other and not the track, the ability to apply probabilities from an outside pool that are more accurate and significantly different than the wagering pool in question, bestows an edge over other racetrack patrons who are likely using probabilities from the pool in which they are betting. For example, it is demonstrated in chapter six for the win wager that implied probabilities from the exacta pool are significantly different and more accurate than straight probabilities from the win pool and those corrected for the favorite-longshot bias. Applying these inefficiencies and then

calculating expected values for each betting combination, wagering on only those ones with a net positive return, gives way to a betting strategy that is likely superior to what most others are using.

Applying results from the previous chapters, consider how probabilities from the exacta can be used to find bets with positive expected values in the win pool utilizing data from Race 2, Monmouth Park on June 29th, 2005:

Step One: For each racing combination, record exacta payouts from the tote board as follows:

Prices	1	2	3	4	5	6	7	8	9
1		\$319.00	\$184.90	\$861.40	\$201.80	\$331.20	\$958.10	\$962.90	\$470.50
2	\$253.90		\$97.40	\$244.20	\$95.10	\$155.20	\$1,504.20	\$544.60	\$151.00
3	\$67.90	\$48.90		\$33.90	\$16.20	\$26.50	\$207.40	\$74.50	\$25.50
4	\$951.80	\$359.90	\$73.30		\$157.00	\$280.70	\$1,313.60	\$1,293.10	\$316.50
5	\$89.50	\$54.20	\$16.40	\$73.20		\$36.90	\$290.60	\$109.10	\$45.90
6	\$236.60	\$140.60	\$58.90	\$344.90	\$59.30		\$1,273.30	\$275.00	\$90.90
7	\$954.90	\$1,879.00	\$758.90	\$2,751.80	\$756.00	\$1,532.00		\$1,532.00	\$836.60
8	\$507.90	\$475.90	\$162.00	\$1,293.10	\$258.20	\$285.40	\$1,273.30		\$343.50
9	\$289.80	\$127.80	\$42.00	\$253.60	\$69.40	\$83.70	\$538.60	\$291.60	

Step two: Once exacta payouts for each combination have been established, corresponding odds can be calculated as demonstrated below:

Probabilities	1	2	3	4	5	6	7	8	9
1		0.0063	0.0108	0.0023	0.0099	0.0060	0.0021	0.0021	0.0043
2	0.0079		0.0205	0.0082	0.0210	0.0129	0.0013	0.0037	0.0132
3	0.0295	0.0409		0.0590	0.1235	0.0755	0.0096	0.0268	0.0784
4	0.0021	0.0056	0.0273		0.0127	0.0071	0.0015	0.0015	0.0063
5	0.0223	0.0369	0.1220	0.0273		0.0542	0.0069	0.0183	0.0436
6	0.0085	0.0142	0.0340	0.0058	0.0337		0.0016	0.0073	0.0220
7	0.0021	0.0011	0.0026	0.0007	0.0026	0.0013		0.0013	0.0024
8	0.0039	0.0042	0.0123	0.0015	0.0077	0.0070	0.0016		0.0058
9	0.0069	0.0156	0.0476	0.0079	0.0288	0.0239	0.0037	0.0069	

Step three: To determine individual probabilities, sum each row and normalize the data:

	Sum of Probabilities	Normalized Probability
1	0.0438	0.034
2	0.0888	0.068
3	0.4432	0.341
4	0.0642	0.049
5	0.3315	0.255
6	0.1270	0.098
7	0.0142	0.011
8	0.0442	0.034
9	0.1413	0.109
Total	1.2981	

Step four: Once win probabilities are extracted from the exacta pool, they can be used in conjunction with odds and corresponding payouts from the win pool to calculate expected values for each entry (as shown below):

$$Payout = (Odds + 1) \times 2 \quad (7.2)$$

Step five: Once the payout is determined for each racing entry, the resultant expected value per \$1 wager can be determined:

$$EV = \frac{\text{Probability}_{\text{exactapool}} \times \text{Payout}_{\text{winpool}}}{2} - 1_{\text{cost}} \quad (7.3)$$

	Odds	Payout	Expected Value
1	15.8	\$ 33.60	-0.43
2	16.3	\$ 34.60	0.18
3	1.2	\$ 4.40	-0.25
4	17.3	\$ 36.60	-0.10
5	2.3	\$ 6.60	-0.16
6	9.2	\$ 20.40	0.00
7	68.2	\$ 138.40	-0.24
8	30.1	\$ 62.20	0.06
9	5.8	\$ 13.60	-0.26

In this race, horses 2 and 8 exhibit positive expected values that correspond to net returns of \$0.37 and \$0.12 respectively for each \$1 wagered. As such, a value betting strategy would suggest that either of these horses be wagered on. Results from this race show horse 2 running first, with a corresponding payout of \$34.60.

When considering the exacta, the most accurate betting strategy utilizes probabilities from the win pool corrected for the longshot bias through Asch and Quandt's regression equation (as shown in chapter four). To exhibit this process, consider the following example from Race 5, Monmouth Park on June 29th, 2005:

Step one: Calculate win probabilities using Asch and Quandt's regression equation correcting for the favorite-longshot bias.

Win Pool	Odds	Breakage	Probability
1	3.9	0.025	0.1746
2	3.7	0.025	0.1820
3	8.7	0.025	0.0884
4	1.6	0.025	0.3276
5	11.4	0.025	0.0692
6	5.5	0.025	0.1318
7	10.4	0.025	0.0753

Probabilities Corrected for Favorite-Longshot Bias from win pool	1	2	3	4	5	6	7
1		0.042	0.019	0.077	0.015	0.030	0.016
2	0.042		0.020	0.082	0.015	0.031	0.017
3	0.017	0.018		0.033	0.006	0.013	0.007
4	0.097	0.102	0.047		0.035	0.072	0.039
5	0.013	0.013	0.006	0.025		0.009	0.005
6	0.028	0.029	0.013	0.054	0.010		0.011
7	0.014	0.015	0.007	0.027	0.005	0.011	

Step two: Obtain exacta payouts for each racing combination from the tote board, and record the data as follows:

Exacta Payouts from Exacta Pool	1	2	3	4	5	6	7
1		\$32.80	\$52.30	\$36.30	\$88.40	\$61.40	\$65.60
2	\$32.00		\$108.70	\$38.40	\$137.00	\$52.80	\$68.60
3	\$92.70	\$134.20		\$77.80	\$156.70	\$115.90	\$147.20
4	\$25.10	\$22.80	\$46.70		\$59.00	\$28.10	\$53.90
5	\$136.60	\$222.50	\$155.90	\$103.10		\$194.00	\$179.10
6	\$57.00	\$56.70	\$98.90	\$39.80	\$124.80	\$124.80	\$80.70
7	\$71.30	\$78.80	\$110.20	\$91.70	\$150.40	\$100.30	

Step three: Calculate expected values for each \$1 wagered on the exacta using probabilities from the win pool corrected for the favorite-longshot bias and payouts from the exacta pool.

Expected Values / \$1 wagered	1	2	3	4	5	6	7
1		-1.31	-1.50	-0.60	-1.36	-1.09	-1.47
2	-0.32	-1.00	0.10	0.57	0.05	-0.17	-0.42
3	-0.20	0.21	-1.00	0.29	-0.51	-0.26	-0.49
4	0.22	0.16	0.09	-1.00	0.04	0.01	0.05
5	-0.13	0.49	-0.52	0.27	-1.00	-0.08	-0.54
6	-0.20	-0.17	-0.33	0.08	-0.36	-1.00	-0.55
7	-0.49	-0.42	-0.63	0.25	-0.61	-0.47	-1.00

The above chart on expected values demonstrates that there are 15 combinations with positive expected values. In this race, the winning exacta is 4-6, which does have a positive expected value, but only of \$0.01 per \$1 wager. The corresponding payout for the 4-6 exacta is \$28.10.

To calculate quinella probabilities, win probabilities corrected for the favorite-longshot bias can be applied in a similar fashion as the exacta as demonstrated above, only quinella payouts are used in place of exacta payouts.

In summary, probabilities from one pool can be used to determine bets with positive expected values in another pool to be used in a value betting strategy. Since it has been shown that probabilities from the wagering pool in question are not necessarily the most

accurate estimate of true probabilities, the above process highlights strategies that can be employed when wagering the win, exacta or quinella wager.

Appendix A- Exacta Data

Table 4.2: Exacta Model Probability Data

P12-reg	P12-H	P12-E	P12-He	Lh/Lc	Le/Lh	Lh/Lhe	Lh/Lr	Revised Prob Reg	Revised Prob Exacta	Revised Prob Henery	Revised Prob Harville
0.05	0.05	0.03	0.04	1.59	0.63	1.09	0.93	0.30	0.17	0.25	0.28
0.02	0.02	0.04	0.04	0.92	1.08	0.57	1.09	0.19	0.23	0.37	0.21
0.03	0.03	0.03	0.03	0.97	1.03	0.60	1.02	0.21	0.22	0.36	0.21
0.02	0.02	0.02	0.01	0.99	1.01	0.93	1.01	0.24	0.25	0.26	0.25
0.03	0.03	0.03	0.03	1.11	0.90	1.07	0.96	0.27	0.23	0.24	0.26
0.01	0.01	0.02	0.01	0.76	1.31	1.10	0.96	0.24	0.31	0.21	0.23
0.06	0.06	0.05	0.06	0.90	1.11	1.01	0.86	0.27	0.26	0.23	0.24
0.05	0.05	0.06	0.09	0.71	1.40	0.57	0.94	0.20	0.27	0.33	0.19
0.04	0.03	0.04	0.03	0.59	1.70	0.69	0.87	0.22	0.32	0.27	0.19
0.20	0.17	0.15	0.13	0.67	1.49	0.90	0.73	0.27	0.30	0.22	0.20
0.02	0.02	0.02	0.02	0.53	1.89	0.89	0.71	0.26	0.35	0.21	0.18
0.06	0.05	0.06	0.05	0.52	1.91	0.98	0.64	0.28	0.35	0.19	0.18
0.01	0.01	0.02	0.01	0.46	2.18	1.00	0.65	0.27	0.38	0.17	0.17
0.03	0.03	0.03	0.02	0.44	2.30	1.64	0.61	0.29	0.41	0.11	0.18
0.07	0.06	0.05	0.06	0.50	2.00	1.74	0.55	0.34	0.37	0.11	0.19
0.08	0.07	0.06	0.09	0.60	1.66	1.39	0.49	0.38	0.31	0.13	0.18
0.14	0.12	0.09	0.11	0.82	1.22	1.60	0.42	0.45	0.23	0.12	0.19
0.04	0.04	0.05	0.05	0.56	1.78	1.25	0.42	0.40	0.30	0.13	0.17
0.02	0.02	0.02	0.02	0.69	1.44	1.01	0.42	0.41	0.25	0.17	0.17
0.03	0.03	0.02	0.03	0.92	1.08	1.07	0.40	0.45	0.20	0.17	0.18
0.17	0.14	0.12	0.12	1.10	0.91	1.27	0.34	0.52	0.16	0.14	0.18
0.01	0.01	0.02	0.01	0.71	1.41	1.31	0.34	0.48	0.23	0.12	0.16
0.13	0.12	0.10	0.06	0.83	1.20	2.79	0.30	0.57	0.20	0.06	0.17
0.12	0.11	0.09	0.08	1.01	0.99	3.61	0.26	0.63	0.16	0.05	0.16
0.01	0.01	0.02	0.02	0.78	1.28	2.87	0.26	0.59	0.20	0.05	0.16
0.05	0.05	0.05	0.04	0.70	1.42	3.11	0.24	0.60	0.21	0.05	0.14
0.00	0.00	0.00	0.00	0.26	3.86	2.72	0.34	0.36	0.47	0.05	0.12
0.07	0.06	0.06	0.06	0.25	4.01	2.98	0.31	0.38	0.47	0.04	0.12
0.14	0.12	0.09	0.10	0.34	2.98	3.69	0.26	0.47	0.37	0.03	0.12
0.10	0.09	0.09	0.08	0.33	3.05	4.03	0.23	0.50	0.36	0.03	0.12
0.08	0.07	0.07	0.05	0.31	3.21	5.84	0.21	0.52	0.35	0.02	0.11
0.02	0.02	0.02	0.02	0.26	3.79	5.39	0.20	0.50	0.38	0.02	0.10
0.03	0.03	0.03	0.02	0.25	4.07	6.80	0.19	0.50	0.39	0.01	0.10
0.08	0.07	0.09	0.07	0.19	5.24	6.72	0.17	0.48	0.43	0.01	0.08
0.01	0.01	0.01	0.01	0.20	4.90	7.00	0.19	0.46	0.44	0.01	0.09
0.01	0.01	0.02	0.02	0.16	6.13	5.02	0.21	0.39	0.51	0.02	0.08
0.08	0.07	0.06	0.06	0.20	5.10	6.44	0.19	0.46	0.44	0.01	0.09
0.02	0.02	0.04	0.03	0.11	8.72	5.79	0.19	0.35	0.57	0.01	0.07
0.03	0.03	0.03	0.01	0.13	7.67	14.10	0.18	0.39	0.53	0.00	0.07
0.03	0.03	0.03	0.03	0.13	7.58	13.25	0.17	0.41	0.52	0.01	0.07
0.11	0.10	0.10	0.05	0.12	8.09	26.84	0.14	0.43	0.50	0.00	0.06
0.03	0.03	0.03	0.02	0.10	9.92	43.62	0.14	0.40	0.54	0.00	0.05
0.02	0.02	0.02	0.02	0.08	11.80	48.24	0.13	0.37	0.58	0.00	0.05
0.01	0.01	0.02	0.03	0.06	15.58	23.73	0.15	0.28	0.67	0.00	0.04
0.12	0.11	0.10	0.08	0.07	14.04	30.62	0.13	0.33	0.62	0.00	0.04
0.01	0.01	0.02	0.02	0.05	20.39	18.83	0.15	0.24	0.72	0.00	0.04
0.06	0.06	0.07	0.05	0.04	24.96	23.17	0.13	0.23	0.74	0.00	0.03
0.05	0.05	0.06	0.05	0.03	31.13	23.05	0.12	0.21	0.77	0.00	0.02
0.08	0.07	0.08	0.06	0.03	37.69	25.21	0.11	0.20	0.78	0.00	0.02
0.03	0.03	0.04	0.04	0.02	46.10	17.14	0.10	0.17	0.81	0.00	0.02
0.04	0.04	0.04	0.04	0.02	53.93	17.80	0.10	0.16	0.83	0.00	0.02
0.07	0.06	0.07	0.05	0.02	61.73	20.93	0.09	0.16	0.83	0.00	0.01
0.12	0.10	0.12	0.05	0.01	72.84	40.31	0.07	0.15	0.84	0.00	0.01
0.02	0.02	0.02	0.01	0.01	77.77	43.52	0.07	0.15	0.84	0.00	0.01
0.02	0.02	0.02	0.03	0.02	52.58	36.17	0.07	0.20	0.78	0.00	0.01
0.01	0.01	0.01	0.01	0.01	74.18	29.02	0.08	0.14	0.84	0.00	0.01
0.08	0.07	0.11	0.06	0.01	104.32	35.56	0.07	0.12	0.87	0.00	0.01
0.05	0.05	0.04	0.04	0.01	89.50	39.07	0.06	0.15	0.84	0.00	0.01
0.10	0.08	0.08	0.08	0.01	85.98	42.39	0.06	0.17	0.82	0.00	0.01
0.02	0.02	0.02	0.02	0.01	93.24	42.01	0.05	0.16	0.83	0.00	0.01
0.05	0.04	0.03	0.04	0.02	64.53	47.70	0.05	0.23	0.75	0.00	0.01
0.02	0.02	0.02	0.02	0.01	71.74	48.76	0.05	0.22	0.77	0.00	0.01

P12-reg	P12-H	P12-E	P12-He	Lh/Le	Le/Lh	Lh/Lhe	Lh/Lr	Revised Prob Reg	Revised Prob Exacta	Revised Prob Henery	Revised Prob Harville
0.03	0.03	0.04	0.03	0.01	99.80	44.93	0.05	0.18	0.81	0.00	0.01
0.19	0.16	0.14	0.07	0.01	82.50	105.89	0.04	0.23	0.76	0.00	0.01
0.00	0.01	0.01	0.01	0.01	97.12	109.43	0.04	0.19	0.80	0.00	0.01
0.05	0.05	0.05	0.05	0.01	95.77	112.77	0.04	0.20	0.79	0.00	0.01
0.06	0.05	0.06	0.04	0.01	103.42	131.84	0.04	0.20	0.79	0.00	0.01
0.17	0.14	0.11	0.13	0.01	79.31	139.40	0.03	0.28	0.71	0.00	0.01
0.00	0.00	0.00	0.00	0.02	63.07	127.88	0.04	0.30	0.69	0.00	0.01
0.03	0.03	0.02	0.03	0.02	42.38	133.96	0.04	0.38	0.60	0.00	0.01
0.08	0.07	0.05	0.06	0.03	33.52	160.07	0.03	0.46	0.52	0.00	0.02
0.01	0.02	0.02	0.03	0.03	33.67	90.59	0.04	0.41	0.57	0.00	0.02
0.06	0.05	0.05	0.05	0.03	33.72	92.41	0.04	0.43	0.55	0.00	0.02
0.09	0.08	0.06	0.06	0.04	25.16	120.26	0.03	0.53	0.45	0.00	0.02
0.02	0.02	0.03	0.01	0.03	32.44	208.59	0.03	0.47	0.51	0.00	0.02
0.01	0.02	0.01	0.01	0.04	22.77	320.04	0.04	0.54	0.44	0.00	0.02
0.13	0.11	0.10	0.10	0.05	20.24	371.15	0.03	0.60	0.38	0.00	0.02
0.16	0.13	0.12	0.10	0.06	17.58	495.05	0.03	0.67	0.31	0.00	0.02
0.01	0.01	0.01	0.01	0.05	19.33	501.30	0.03	0.63	0.35	0.00	0.02
0.01	0.01	0.01	0.01	0.04	28.26	366.37	0.03	0.52	0.46	0.00	0.02
0.12	0.10	0.09	0.09	0.04	24.58	446.88	0.03	0.59	0.39	0.00	0.02
0.05	0.05	0.04	0.05	0.05	20.87	480.37	0.02	0.65	0.34	0.00	0.02
0.07	0.07	0.06	0.05	0.05	18.76	589.61	0.02	0.69	0.29	0.00	0.02
0.14	0.12	0.17	0.12	0.04	26.35	608.14	0.02	0.66	0.33	0.00	0.01
0.06	0.05	0.09	0.05	0.02	43.81	590.05	0.02	0.57	0.42	0.00	0.01
0.08	0.07	0.08	0.10	0.02	46.41	444.45	0.02	0.58	0.41	0.00	0.01
0.12	0.11	0.10	0.05	0.02	44.69	881.50	0.01	0.62	0.37	0.00	0.01
0.06	0.06	0.06	0.06	0.02	44.21	768.34	0.01	0.65	0.35	0.00	0.01
0.04	0.04	0.04	0.04	0.02	52.78	772.02	0.01	0.62	0.37	0.00	0.01
0.09	0.08	0.11	0.08	0.01	77.41	787.68	0.01	0.56	0.43	0.00	0.01
0.04	0.03	0.03	0.03	0.01	67.90	801.78	0.01	0.61	0.38	0.00	0.01
0.00	0.00	0.01	0.01	0.01	97.54	541.92	0.01	0.44	0.55	0.00	0.01
0.02	0.02	0.02	0.02	0.01	90.57	749.99	0.01	0.47	0.53	0.00	0.01
0.11	0.09	0.10	0.10	0.01	96.32	725.77	0.01	0.49	0.51	0.00	0.01
0.12	0.11	0.11	0.10	0.01	99.72	776.34	0.01	0.51	0.49	0.00	0.00
0.01	0.01	0.02	0.01	0.01	126.63	1863.73	0.01	0.45	0.55	0.00	0.00
0.11	0.09	0.08	0.05	0.01	111.81	3501.21	0.01	0.51	0.48	0.00	0.00
0.05	0.05	0.05	0.04	0.01	113.32	4130.47	0.01	0.53	0.46	0.00	0.00
0.09	0.08	0.07	0.07	0.01	95.82	4519.58	0.01	0.60	0.39	0.00	0.00
0.13	0.11	0.09	0.09	0.01	81.68	5814.93	0.01	0.67	0.32	0.00	0.00
0.08	0.07	0.07	0.06	0.01	75.63	7371.54	0.01	0.71	0.29	0.00	0.00
0.03	0.03	0.02	0.03	0.02	54.64	6939.96	0.01	0.76	0.23	0.00	0.00
0.15	0.13	0.12	0.10	0.02	50.43	9298.34	0.00	0.80	0.19	0.00	0.00
0.30	0.25	0.27	0.18	0.02	54.88	12878.91	0.00	0.82	0.18	0.00	0.00
0.00	0.00	0.00	0.00	0.01	118.94	10238.64	0.01	0.50	0.49	0.00	0.00
0.01	0.01	0.02	0.01	0.01	140.42	12976.73	0.01	0.47	0.53	0.00	0.00
0.01	0.01	0.01	0.01	0.01	185.67	11389.78	0.01	0.38	0.61	0.00	0.00
0.11	0.10	0.11	0.13	0.00	210.21	8912.09	0.01	0.39	0.61	0.00	0.00
0.01	0.01	0.01	0.02	0.01	182.45	7912.25	0.01	0.38	0.62	0.00	0.00
0.02	0.02	0.02	0.02	0.00	250.32	7245.95	0.01	0.31	0.69	0.00	0.00
0.09	0.08	0.07	0.06	0.00	219.38	8766.11	0.01	0.37	0.63	0.00	0.00
0.03	0.03	0.03	0.02	0.01	178.16	12600.10	0.01	0.43	0.56	0.00	0.00
0.03	0.02	0.03	0.02	0.00	216.88	12879.89	0.01	0.40	0.60	0.00	0.00
0.06	0.05	0.04	0.05	0.01	162.07	12630.64	0.01	0.49	0.51	0.00	0.00
0.09	0.08	0.11	0.09	0.00	231.12	10434.30	0.01	0.43	0.57	0.00	0.00
0.01	0.01	0.01	0.01	0.00	220.44	10763.83	0.01	0.44	0.55	0.00	0.00
0.21	0.18	0.18	0.16	0.00	218.87	11882.31	0.00	0.49	0.51	0.00	0.00
0.21	0.18	0.16	0.15	0.01	194.54	13997.12	0.00	0.56	0.44	0.00	0.00
0.05	0.05	0.05	0.05	0.01	198.28	11820.19	0.00	0.56	0.44	0.00	0.00
0.07	0.07	0.05	0.05	0.01	142.45	14241.41	0.00	0.66	0.34	0.00	0.00
0.05	0.05	0.03	0.05	0.01	98.39	14435.68	0.00	0.74	0.25	0.00	0.00
0.04	0.04	0.04	0.04	0.01	96.08	14384.01	0.00	0.76	0.23	0.00	0.00
0.05	0.05	0.08	0.06	0.01	145.88	12005.95	0.00	0.70	0.30	0.00	0.00
0.08	0.07	0.08	0.10	0.01	169.08	8176.47	0.00	0.69	0.31	0.00	0.00
0.04	0.04	0.04	0.02	0.01	193.06	13846.82	0.00	0.67	0.32	0.00	0.00
0.12	0.10	0.08	0.09	0.01	161.79	15549.29	0.00	0.74	0.26	0.00	0.00
0.23	0.19	0.17	0.15	0.01	145.58	19114.42	0.00	0.79	0.21	0.00	0.00
0.14	0.12	0.12	0.09	0.01	140.60	26866.10	0.00	0.82	0.18	0.00	0.00
0.13	0.12	0.08	0.09	0.01	101.05	33328.69	0.00	0.88	0.12	0.00	0.00
0.04	0.04	0.04	0.03	0.01	106.48	37207.41	0.00	0.89	0.11	0.00	0.00
0.00	0.00	0.00	0.00	0.01	133.38	35234.83	0.00	0.78	0.21	0.00	0.00

P12-reg	P12-H	P12-E	P12-He	Lh/Le	Le/Lh	Lh/Lhe	Lh/Lr	Revised Prob Reg	Revised Prob Exacta	Revised Prob Henery	Revised Prob Harville
0.08	0.07	0.07	0.06	0.01	124.11	42522.35	0.00	0.81	0.18	0.00	0.00
0.09	0.08	0.08	0.04	0.01	131.64	86979.54	0.00	0.82	0.17	0.00	0.00
0.00	0.00	0.00	0.00	0.01	187.77	78299.57	0.00	0.71	0.29	0.00	0.00
0.03	0.03	0.02	0.02	0.01	155.29	83180.88	0.00	0.75	0.24	0.00	0.00
0.01	0.01	0.01	0.01	0.01	137.88	79005.53	0.00	0.78	0.22	0.00	0.00
0.01	0.01	0.01	0.01	0.01	155.71	51292.05	0.00	0.73	0.27	0.00	0.00
0.02	0.02	0.01	0.02	0.01	121.02	52949.77	0.00	0.78	0.22	0.00	0.00
0.01	0.01	0.01	0.01	0.01	131.23	46865.36	0.00	0.77	0.23	0.00	0.00
0.00	0.01	0.01	0.01	0.01	180.78	30021.95	0.00	0.67	0.33	0.00	0.00
0.02	0.02	0.02	0.02	0.00	256.12	30771.42	0.00	0.59	0.41	0.00	0.00
0.02	0.02	0.03	0.04	0.00	297.58	17751.38	0.00	0.51	0.49	0.00	0.00
0.01	0.01	0.01	0.01	0.00	502.99	11571.46	0.00	0.35	0.65	0.00	0.00
0.24	0.20	0.19	0.11	0.00	479.22	21307.67	0.00	0.41	0.59	0.00	0.00
0.01	0.01	0.02	0.02	0.00	694.29	13329.89	0.00	0.30	0.70	0.00	0.00
0.02	0.02	0.02	0.02	0.00	840.81	11985.81	0.00	0.27	0.73	0.00	0.00
0.04	0.04	0.04	0.04	0.00	874.54	13480.17	0.00	0.28	0.72	0.00	0.00
0.14	0.12	0.07	0.10	0.00	516.29	16700.39	0.00	0.43	0.57	0.00	0.00
0.04	0.04	0.04	0.04	0.00	501.17	17330.27	0.00	0.46	0.54	0.00	0.00
0.01	0.01	0.01	0.08	0.00	549.14	2855.24	0.00	0.42	0.58	0.00	0.00
0.05	0.04	0.03	0.03	0.00	375.24	4236.29	0.00	0.54	0.46	0.00	0.00
0.04	0.04	0.04	0.04	0.00	348.62	4199.11	0.00	0.58	0.42	0.00	0.00
0.01	0.01	0.01	0.01	0.00	498.59	4045.49	0.00	0.48	0.52	0.00	0.00
0.01	0.01	0.02	0.01	0.00	1072.20	3617.75	0.00	0.29	0.71	0.00	0.00
0.19	0.17	0.14	0.15	0.00	894.97	4088.15	0.00	0.36	0.63	0.00	0.00
0.03	0.03	0.03	0.03	0.00	853.33	3895.70	0.00	0.39	0.61	0.00	0.00
0.03	0.03	0.02	0.03	0.00	659.31	3916.47	0.00	0.46	0.54	0.00	0.00
0.10	0.09	0.09	0.07	0.00	697.06	4575.28	0.00	0.48	0.52	0.00	0.00
0.02	0.02	0.02	0.02	0.00	701.52	4592.17	0.00	0.48	0.52	0.00	0.00
0.35	0.26	0.30	0.18	0.00	804.41	6928.28	0.00	0.51	0.48	0.00	0.00
0.33	0.28	0.12	0.09	0.00	352.12	21910.56	0.00	0.74	0.25	0.00	0.00
0.05	0.04	0.05	0.04	0.00	379.08	22636.88	0.00	0.75	0.25	0.00	0.00
0.02	0.02	0.02	0.03	0.00	369.73	15219.11	0.00	0.75	0.25	0.00	0.00
0.06	0.05	0.04	0.05	0.00	292.44	16916.53	0.00	0.81	0.19	0.00	0.00
0.10	0.08	0.06	0.06	0.01	197.57	22535.89	0.00	0.88	0.12	0.00	0.00
0.02	0.02	0.03	0.03	0.00	246.06	19380.67	0.00	0.86	0.14	0.00	0.00
0.01	0.01	0.01	0.02	0.00	348.31	11190.55	0.00	0.77	0.23	0.00	0.00
0.22	0.18	0.20	0.16	0.00	396.46	11989.47	0.00	0.79	0.21	0.00	0.00
0.00	0.00	0.01	0.01	0.00	610.36	9057.27	0.00	0.67	0.33	0.00	0.00
0.02	0.02	0.02	0.04	0.00	563.87	6009.63	0.00	0.68	0.32	0.00	0.00
0.05	0.05	0.02	0.08	0.00	208.19	3507.56	0.00	0.86	0.14	0.00	0.00
0.15	0.13	0.13	0.10	0.00	205.12	4793.17	0.00	0.88	0.12	0.00	0.00
0.01	0.01	0.01	0.01	0.01	192.00	3487.47	0.00	0.87	0.13	0.00	0.00
0.01	0.01	0.01	0.01	0.00	226.11	3613.46	0.00	0.85	0.15	0.00	0.00
0.03	0.03	0.03	0.03	0.01	189.69	4077.49	0.00	0.88	0.12	0.00	0.00
0.05	0.05	0.05	0.04	0.01	191.30	4987.09	0.00	0.88	0.12	0.00	0.00
0.04	0.04	0.03	0.04	0.01	157.81	5205.25	0.00	0.91	0.09	0.00	0.00
0.02	0.02	0.02	0.01	0.01	161.61	5664.97	0.00	0.91	0.09	0.00	0.00
0.01	0.01	0.01	0.01	0.01	139.12	6339.46	0.00	0.92	0.08	0.00	0.00
0.00	0.00	0.01	0.00	0.00	235.67	7155.22	0.00	0.85	0.15	0.00	0.00
0.06	0.05	0.04	0.04	0.01	185.86	9386.29	0.00	0.89	0.11	0.00	0.00
0.00	0.00	0.00	0.00	0.00	238.63	8709.36	0.00	0.81	0.19	0.00	0.00
0.00	0.00	0.01	0.01	0.00	281.43	8212.17	0.00	0.76	0.24	0.00	0.00
0.03	0.03	0.03	0.03	0.00	285.87	8571.66	0.00	0.76	0.24	0.00	0.00
0.03	0.03	0.02	0.01	0.00	225.75	17275.05	0.00	0.81	0.19	0.00	0.00
0.01	0.01	0.02	0.01	0.00	350.41	18600.41	0.00	0.73	0.27	0.00	0.00
0.04	0.04	0.03	0.04	0.00	316.74	16317.92	0.00	0.76	0.24	0.00	0.00
0.05	0.05	0.05	0.04	0.00	308.10	21181.36	0.00	0.78	0.22	0.00	0.00
0.04	0.04	0.04	0.05	0.00	280.95	18660.22	0.00	0.80	0.20	0.00	0.00
0.10	0.09	0.09	0.04	0.00	273.52	41249.59	0.00	0.82	0.18	0.00	0.00
0.02	0.02	0.02	0.02	0.00	385.28	40399.64	0.00	0.77	0.23	0.00	0.00
0.01	0.01	0.01	0.02	0.00	397.51	30465.48	0.00	0.75	0.25	0.00	0.00
0.03	0.03	0.03	0.03	0.00	327.21	30461.23	0.00	0.80	0.20	0.00	0.00
0.00	0.01	0.01	0.01	0.00	332.03	20899.54	0.00	0.78	0.22	0.00	0.00
0.02	0.02	0.02	0.02	0.00	408.68	17712.08	0.00	0.74	0.26	0.00	0.00
0.09	0.08	0.09	0.07	0.00	465.66	20170.92	0.00	0.74	0.26	0.00	0.00
0.13	0.11	0.09	0.06	0.00	374.40	40682.26	0.00	0.81	0.19	0.00	0.00
0.02	0.02	0.02	0.02	0.00	343.35	37330.48	0.00	0.82	0.18	0.00	0.00
0.06	0.05	0.04	0.06	0.00	280.41	31593.42	0.00	0.86	0.14	0.00	0.00
0.09	0.08	0.08	0.07	0.00	266.39	34962.34	0.00	0.88	0.12	0.00	0.00

P12-reg	P12-H	P12-E	P12-He	Lh/Le	Le/Lh	Lh/Lhe	Lh/Lr	Revised Prob Reg	Revised Prob Exacta	Revised Prob Henery	Revised Prob Harville
0.00	0.00	0.01	0.00	0.00	306.96	42248.15	0.00	0.85	0.15	0.00	0.00
0.10	0.09	0.06	0.07	0.00	223.03	50058.16	0.00	0.90	0.10	0.00	0.00
0.01	0.01	0.01	0.01	0.00	372.46	43186.08	0.00	0.83	0.17	0.00	0.00
0.16	0.14	0.15	0.12	0.00	397.54	48826.91	0.00	0.85	0.15	0.00	0.00
0.01	0.01	0.01	0.01	0.00	640.74	33351.56	0.00	0.75	0.25	0.00	0.00
0.08	0.07	0.05	0.06	0.00	488.46	36145.46	0.00	0.82	0.18	0.00	0.00
0.02	0.02	0.01	0.02	0.00	423.18	37658.52	0.00	0.84	0.16	0.00	0.00
0.08	0.07	0.08	0.07	0.00	485.36	38678.23	0.00	0.84	0.16	0.00	0.00
0.04	0.04	0.03	0.03	0.00	446.69	52095.10	0.00	0.86	0.14	0.00	0.00
0.00	0.00	0.01	0.00	0.00	1106.89	77213.42	0.00	0.68	0.32	0.00	0.00
0.09	0.08	0.07	0.07	0.00	1049.41	84118.97	0.00	0.72	0.28	0.00	0.00
0.04	0.03	0.03	0.05	0.00	888.25	58647.89	0.00	0.76	0.24	0.00	0.00
0.06	0.05	0.04	0.05	0.00	707.44	62148.81	0.00	0.81	0.19	0.00	0.00
0.29	0.24	0.25	0.11	0.00	751.15	132836.58	0.00	0.83	0.17	0.00	0.00
0.10	0.09	0.07	0.07	0.00	646.37	155747.22	0.00	0.86	0.14	0.00	0.00
0.07	0.06	0.05	0.06	0.00	480.02	173943.69	0.00	0.91	0.09	0.00	0.00
0.11	0.10	0.09	0.08	0.00	467.79	198083.59	0.00	0.92	0.08	0.00	0.00
0.10	0.09	0.08	0.08	0.00	408.04	235077.79	0.00	0.94	0.06	0.00	0.00
0.17	0.15	0.12	0.11	0.00	340.32	321960.19	0.00	0.95	0.05	0.00	0.00
0.04	0.03	0.03	0.04	0.00	275.93	318637.99	0.00	0.96	0.04	0.00	0.00
0.04	0.04	0.04	0.04	0.00	272.82	333904.91	0.00	0.96	0.04	0.00	0.00
0.06	0.05	0.05	0.06	0.00	241.42	320319.59	0.00	0.97	0.03	0.00	0.00
0.20	0.17	0.18	0.13	0.00	267.29	401027.95	0.00	0.97	0.03	0.00	0.00
0.16	0.14	0.12	0.12	0.00	233.88	475845.98	0.00	0.98	0.02	0.00	0.00
0.11	0.10	0.10	0.09	0.00	231.15	494240.68	0.00	0.98	0.02	0.00	0.00
0.06	0.06	0.04	0.05	0.01	169.78	534894.98	0.00	0.99	0.01	0.00	0.00

In this table, P_{12-j} reflects the probability estimate under hypothesis H_j while L_i/L_j is the ratio of the likelihoods under hypotheses i and j respectively.

Appendix B- Quinella Data

Table 5.1: Quinella Model Probability Data

P1,2 h	P 1,2 q	P1,2 he	P1,2 r	Lh/Lq	Lh/Lhe	Lh/Lr	Lq/Lh	Revised Prob Quinella	Revised Prob Henery	Revised Prob Harville	Revised Prob Regression
0.07	0.05	0.06	0.08	1.44	1.13	0.92	0.70	0.19	0.24	0.27	0.30
0.05	0.06	0.05	0.05	1.12	1.14	0.89	0.89	0.23	0.25	0.26	0.29
0.14	0.06	0.11	0.15	2.70	1.39	0.79	0.37	0.11	0.20	0.30	0.38
0.14	0.18	0.11	0.16	2.09	1.75	0.70	0.48	0.14	0.18	0.29	0.41
0.05	0.05	0.04	0.05	1.86	1.87	0.67	0.54	0.15	0.16	0.28	0.42
0.07	0.09	0.06	0.08	1.56	2.14	0.62	0.64	0.17	0.13	0.27	0.44
0.03	0.04	0.03	0.02	1.04	1.76	0.64	0.96	0.23	0.15	0.24	0.38
0.02	0.02	0.02	0.01	0.65	1.72	0.71	1.53	0.34	0.16	0.22	0.31
0.17	0.14	0.16	0.19	0.79	1.88	0.63	1.27	0.29	0.13	0.23	0.36
0.01	0.01	0.01	0.01	0.39	1.67	0.78	2.58	0.47	0.15	0.18	0.24
0.20	0.17	0.16	0.23	0.45	2.14	0.68	2.21	0.43	0.10	0.19	0.29
0.00	0.01	0.00	0.00	0.09	2.01	1.36	11.51	0.84	0.12	0.07	0.05
0.05	0.04	0.05	0.05	0.12	2.20	1.28	8.34	0.79	0.04	0.09	0.07
0.03	0.03	0.03	0.03	0.11	2.15	1.31	8.74	0.80	0.05	0.09	0.07
0.02	0.03	0.02	0.02	0.09	2.05	1.38	10.60	0.83	0.05	0.08	0.06
0.03	0.04	0.03	0.03	0.08	2.03	1.36	13.23	0.86	0.05	0.06	0.05
0.09	0.10	0.08	0.10	0.07	2.47	1.24	13.79	0.86	0.04	0.06	0.05
0.08	0.07	0.11	0.08	0.08	1.85	1.19	12.26	0.84	0.03	0.07	0.06
0.01	0.01	0.01	0.01	0.10	1.63	1.35	9.84	0.81	0.05	0.08	0.06
0.10	0.11	0.09	0.10	0.09	1.79	1.22	11.65	0.83	0.06	0.07	0.06
0.02	0.05	0.02	0.02	0.04	1.70	1.23	23.69	0.91	0.05	0.04	0.03
0.06	0.03	0.06	0.06	0.07	1.70	1.15	13.75	0.85	0.03	0.06	0.05
0.00	0.01	0.00	0.00	0.02	1.65	2.25	42.73	0.95	0.05	0.02	0.01
0.09	0.03	0.08	0.09	0.07	1.72	2.06	13.98	0.87	0.02	0.06	0.03
0.24	0.18	0.17	0.28	0.09	2.40	1.78	10.68	0.84	0.04	0.08	0.04
0.10	0.09	0.09	0.11	0.10	2.79	1.62	9.69	0.83	0.04	0.09	0.05
0.12	0.11	0.13	0.14	0.11	2.77	1.47	8.95	0.81	0.04	0.09	0.06
0.01	0.02	0.02	0.01	0.07	1.87	2.11	14.46	0.88	0.04	0.06	0.03
0.21	0.14	0.16	0.24	0.11	2.53	1.82	9.52	0.83	0.04	0.09	0.05
0.19	0.20	0.18	0.22	0.10	2.67	1.59	9.74	0.83	0.04	0.09	0.05
0.08	0.06	0.08	0.09	0.13	2.81	1.50	7.51	0.79	0.04	0.10	0.07
0.16	0.16	0.13	0.18	0.14	3.64	1.33	7.29	0.78	0.04	0.11	0.08
0.12	0.17	0.12	0.13	0.09	3.49	1.20	10.56	0.83	0.04	0.08	0.07
0.13	0.13	0.12	0.14	0.09	3.89	1.09	10.89	0.83	0.03	0.08	0.07
0.03	0.04	0.03	0.03	0.06	3.99	1.08	16.30	0.88	0.02	0.05	0.05
0.13	0.08	0.11	0.14	0.09	4.62	0.97	10.54	0.82	0.02	0.08	0.08
0.03	0.03	0.03	0.03	0.08	4.87	0.97	12.27	0.85	0.02	0.07	0.07
0.05	0.05	0.05	0.05	0.08	4.34	0.97	12.94	0.85	0.02	0.07	0.07
0.04	0.04	0.04	0.04	0.08	4.60	0.94	12.23	0.84	0.02	0.07	0.07
0.03	0.01	0.04	0.03	0.21	3.90	1.00	4.73	0.68	0.02	0.14	0.14
0.06	0.07	0.06	0.07	0.21	4.37	0.93	4.85	0.68	0.04	0.14	0.15
0.08	0.06	0.08	0.09	0.26	4.51	0.87	3.84	0.62	0.04	0.16	0.18
0.06	0.03	0.06	0.06	0.50	4.69	0.84	2.01	0.45	0.04	0.23	0.27
0.06	0.11	0.07	0.07	0.29	4.47	0.79	3.46	0.58	0.05	0.17	0.21
0.22	0.16	0.16	0.25	0.38	6.22	0.69	2.64	0.50	0.04	0.19	0.28
0.03	0.03	0.04	0.03	0.50	5.90	0.68	2.02	0.43	0.04	0.21	0.32
0.03	0.04	0.03	0.03	0.35	5.75	0.68	2.83	0.52	0.04	0.18	0.27
0.02	0.03	0.03	0.02	0.27	4.94	0.73	3.77	0.59	0.04	0.16	0.22
0.02	0.02	0.02	0.02	0.31	4.79	0.75	3.26	0.56	0.04	0.17	0.23
0.06	0.09	0.07	0.06	0.20	3.98	0.75	5.02	0.66	0.04	0.13	0.18
0.02	0.01	0.02	0.02	0.24	3.55	0.81	4.20	0.62	0.04	0.15	0.18
0.13	0.12	0.11	0.14	0.25	4.06	0.72	4.05	0.61	0.04	0.15	0.21
0.02	0.02	0.03	0.01	0.29	2.82	0.98	3.40	0.59	0.05	0.17	0.18
0.03	0.03	0.04	0.03	0.29	2.33	1.01	3.45	0.59	0.08	0.17	0.17
0.03	0.05	0.02	0.03	0.15	2.45	1.01	6.74	0.74	0.08	0.11	0.11
0.17	0.13	0.13	0.20	0.20	3.21	0.89	4.91	0.67	0.05	0.14	0.15
0.31	0.23	0.21	0.36	0.28	4.80	0.75	3.60	0.59	0.04	0.16	0.22
0.06	0.06	0.06	0.07	0.30	4.81	0.71	3.29	0.56	0.04	0.17	0.24
0.01	0.01	0.01	0.01	0.20	4.49	0.82	4.91	0.67	0.04	0.14	0.17
0.13	0.13	0.11	0.15	0.21	5.70	0.73	4.87	0.66	0.03	0.13	0.18
0.05	0.05	0.05	0.06	0.21	6.34	0.69	4.70	0.64	0.03	0.14	0.20
0.03	0.04	0.04	0.03	0.18	5.57	0.71	5.46	0.68	0.03	0.12	0.17
0.12	0.08	0.10	0.14	0.29	6.74	0.64	3.40	0.56	0.02	0.16	0.26
0.24	0.14	0.21	0.28	0.52	7.98	0.55	1.92	0.40	0.02	0.21	0.37
0.02	0.04	0.02	0.02	0.20	7.33	0.58	5.05	0.64	0.03	0.13	0.22
0.11	0.06	0.10	0.12	0.37	8.36	0.52	2.67	0.47	0.02	0.18	0.34
0.07	0.05	0.08	0.07	0.51	7.42	0.52	1.97	0.39	0.02	0.20	0.38
0.03	0.03	0.03	0.03	0.59	7.51	0.51	1.70	0.36	0.03	0.21	0.41
0.09	0.06	0.08	0.10	1.00	8.70	0.47	1.00	0.24	0.03	0.24	0.50

P1,2 h	P 1,2 q	P1,2 he	P1,2 r	Lh/Lq	Lh/Lhe	Lh/Lr	Lq/Lh	Revised Prob Quinella	Revised Prob Henery	Revised Prob Harville	Revised Prob Regression
0.11	0.14	0.06	0.12	0.01	3776.43	0.18	138.06	0.95	0.00	0.01	0.04
0.35	0.30	0.28	0.42	0.01	4795.24	0.15	118.93	0.94	0.00	0.01	0.05
0.05	0.06	0.05	0.05	0.01	4340.57	0.15	144.96	0.95	0.00	0.01	0.04
0.48	0.49	0.34	0.58	0.01	6060.06	0.12	149.12	0.94	0.00	0.01	0.05
0.10	0.09	0.09	0.11	0.01	6741.39	0.11	142.28	0.93	0.00	0.01	0.06
0.25	0.18	0.15	0.29	0.01	11008.05	0.09	101.79	0.90	0.00	0.01	0.09
0.01	0.01	0.02	0.01	0.01	8623.55	0.12	93.27	0.91	0.00	0.01	0.08
0.02	0.02	0.02	0.02	0.01	8513.22	0.12	81.37	0.90	0.00	0.01	0.09
0.14	0.12	0.10	0.16	0.01	12682.26	0.11	71.15	0.87	0.00	0.01	0.12
0.18	0.17	0.18	0.21	0.01	12833.36	0.09	67.05	0.85	0.00	0.01	0.14
0.06	0.07	0.06	0.06	0.01	11659.73	0.09	80.86	0.87	0.00	0.01	0.12
0.35	0.29	0.28	0.42	0.01	14782.02	0.07	66.72	0.82	0.00	0.01	0.17
0.06	0.05	0.10	0.06	0.02	9751.46	0.08	51.53	0.79	0.00	0.02	0.20
0.14	0.10	0.13	0.16	0.03	10685.04	0.07	38.04	0.71	0.00	0.02	0.27
0.23	0.13	0.18	0.27	0.05	13564.44	0.06	21.72	0.55	0.00	0.03	0.42
0.38	0.24	0.26	0.45	0.07	19655.65	0.05	13.84	0.40	0.00	0.03	0.57
0.29	0.26	0.24	0.33	0.08	23675.97	0.04	12.79	0.35	0.00	0.03	0.63
0.10	0.07	0.09	0.11	0.10	25492.75	0.04	9.73	0.27	0.00	0.03	0.70
0.08	0.07	0.09	0.08	0.12	22336.53	0.04	8.38	0.23	0.00	0.03	0.74
0.09	0.10	0.09	0.10	0.11	22540.89	0.03	8.86	0.22	0.00	0.03	0.75
0.13	0.12	0.13	0.14	0.11	21736.34	0.03	8.74	0.20	0.00	0.02	0.77
0.08	0.08	0.08	0.09	0.12	22642.32	0.03	8.36	0.18	0.00	0.02	0.79
0.13	0.10	0.07	0.14	0.15	41955.63	0.02	6.56	0.14	0.00	0.02	0.84
0.05	0.04	0.05	0.05	0.20	43525.00	0.02	5.00	0.10	0.00	0.02	0.88
0.08	0.07	0.07	0.09	0.23	48538.25	0.02	4.43	0.08	0.00	0.02	0.90
0.04	0.04	0.05	0.04	0.23	38088.96	0.02	4.29	0.08	0.00	0.02	0.90
0.08	0.08	0.07	0.09	0.23	44910.65	0.02	4.39	0.08	0.00	0.02	0.90
0.07	0.10	0.08	0.07	0.16	38132.41	0.02	6.33	0.11	0.00	0.02	0.88
0.09	0.08	0.06	0.09	0.17	63396.16	0.02	5.77	0.09	0.00	0.02	0.89
0.14	0.09	0.11	0.16	0.28	81236.25	0.02	3.63	0.06	0.00	0.02	0.93
0.10	0.11	0.09	0.11	0.27	93013.83	0.01	3.71	0.05	0.00	0.01	0.93
0.05	0.08	0.05	0.05	0.18	98564.59	0.01	5.54	0.07	0.00	0.01	0.91
0.08	0.08	0.06	0.08	0.18	118505.84	0.01	5.70	0.07	0.00	0.01	0.92
0.03	0.02	0.03	0.03	0.22	122905.69	0.01	4.50	0.05	0.00	0.01	0.93
0.13	0.08	0.10	0.14	0.37	158891.04	0.01	2.73	0.03	0.00	0.01	0.96
0.00	0.00	0.00	0.00	0.19	121068.29	0.03	5.19	0.14	0.00	0.03	0.83
0.02	0.03	0.02	0.02	0.14	107593.80	0.04	7.28	0.20	0.00	0.03	0.77
0.06	0.09	0.05	0.07	0.10	124579.23	0.03	10.46	0.25	0.00	0.02	0.72
0.17	0.17	0.18	0.20	0.10	120112.33	0.03	10.30	0.23	0.00	0.02	0.75
0.05	0.07	0.06	0.06	0.08	104730.04	0.03	12.90	0.26	0.00	0.02	0.72
0.34	0.29	0.28	0.40	0.09	127122.98	0.02	11.11	0.20	0.00	0.02	0.78
0.06	0.05	0.10	0.06	0.11	78699.21	0.03	9.01	0.18	0.00	0.02	0.80
0.14	0.10	0.13	0.15	0.14	82525.31	0.02	6.92	0.13	0.00	0.02	0.85
0.22	0.13	0.18	0.25	0.24	100113.05	0.02	4.12	0.07	0.00	0.02	0.91
0.36	0.24	0.26	0.43	0.36	137751.72	0.02	2.75	0.04	0.00	0.02	0.94
0.27	0.26	0.24	0.32	0.38	158253.63	0.01	2.66	0.04	0.00	0.01	0.95
0.09	0.07	0.09	0.10	0.48	163329.26	0.01	2.10	0.03	0.00	0.01	0.96
0.08	0.07	0.09	0.08	0.53	136998.81	0.01	1.88	0.02	0.00	0.01	0.97
0.09	0.10	0.09	0.10	0.48	132455.44	0.01	2.07	0.02	0.00	0.01	0.97
0.12	0.12	0.13	0.13	0.47	122088.61	0.01	2.13	0.02	0.00	0.01	0.97
0.08	0.08	0.08	0.08	0.47	121933.56	0.01	2.11	0.02	0.00	0.01	0.97
0.12	0.10	0.07	0.14	0.58	216255.13	0.01	1.73	0.01	0.00	0.01	0.98
0.05	0.04	0.05	0.05	0.73	215382.02	0.01	1.36	0.01	0.00	0.01	0.98
0.08	0.07	0.07	0.09	0.80	230361.44	0.01	1.26	0.01	0.00	0.01	0.98
0.04	0.04	0.05	0.04	0.79	173330.32	0.01	1.26	0.01	0.00	0.01	0.98
0.13	0.09	0.11	0.14	1.19	207873.91	0.01	0.84	0.01	0.00	0.01	0.99

In this table, P_{12-j} reflects the probability estimate under hypothesis H_j while L_i/L_j is the ratio of the likelihoods under hypotheses i and j respectively.

Appendix C- Win Data

Table 6.1: Win Model Probability Data

$P_{1,2w}$	$P_{1,2e}$	$P_{1,2r}$	Lw/Le	Lw/Lr	Le/Lw	Revised Prob Exacta	Revised Prob Regression	Revised Prob Win
0.17	0.17	0.18	1.00	0.96	1.00	0.33	0.34	0.33
0.02	0.04	0.01	0.54	1.54	1.84	0.53	0.19	0.29
0.20	0.23	0.21	0.47	1.47	2.14	0.56	0.18	0.26
0.11	0.15	0.11	0.33	1.46	3.05	0.64	0.14	0.21
0.21	0.24	0.22	0.28	1.40	3.52	0.67	0.14	0.19
0.09	0.12	0.09	0.22	1.42	4.65	0.73	0.11	0.16
0.04	0.05	0.03	0.17	1.73	5.79	0.79	0.08	0.14
0.22	0.23	0.23	0.16	1.65	6.13	0.79	0.08	0.13
0.09	0.09	0.08	0.16	1.68	6.14	0.79	0.08	0.13
0.08	0.09	0.08	0.16	1.72	6.25	0.80	0.07	0.13
0.36	0.39	0.38	0.15	1.61	6.83	0.81	0.07	0.12
0.11	0.12	0.11	0.13	1.61	7.92	0.83	0.07	0.10
0.34	0.37	0.36	0.12	1.51	8.69	0.84	0.06	0.10
0.30	0.32	0.32	0.11	1.42	9.27	0.84	0.06	0.09
0.17	0.21	0.17	0.09	1.37	11.46	0.87	0.06	0.08
0.16	0.18	0.16	0.07	1.33	13.43	0.88	0.05	0.07
0.13	0.15	0.13	0.06	1.31	16.49	0.90	0.04	0.05
0.30	0.34	0.32	0.05	1.23	18.84	0.91	0.04	0.05
0.27	0.28	0.29	0.05	1.16	19.51	0.91	0.04	0.05
0.49	0.51	0.52	0.05	1.08	20.25	0.91	0.04	0.05
0.09	0.10	0.09	0.05	1.09	21.69	0.92	0.04	0.04
0.21	0.24	0.22	0.04	1.04	24.26	0.93	0.04	0.04
0.11	0.10	0.11	0.04	1.04	22.92	0.92	0.04	0.04
0.11	0.11	0.11	0.04	1.03	23.80	0.92	0.04	0.04
0.11	0.12	0.11	0.04	1.03	26.64	0.93	0.03	0.03
0.05	0.07	0.05	0.03	1.14	35.40	0.95	0.02	0.03
0.05	0.06	0.05	0.02	1.28	41.77	0.96	0.02	0.02
0.10	0.13	0.10	0.02	1.28	51.75	0.97	0.01	0.02
0.44	0.46	0.47	0.02	1.19	54.37	0.97	0.01	0.02
0.25	0.34	0.27	0.01	1.13	73.49	0.97	0.01	0.01
0.12	0.20	0.12	0.01	1.11	127.29	0.99	0.01	0.01

In this table, P_{12-j} reflects the probability estimate under hypothesis H_j while L_i/L_j is the ratio of the likelihoods under hypotheses i and j respectively.

References

- Asch, P. and Quandt, R. E., "Efficiency and Profitability in Exotic Bets," *Economica, New Series*, Vol. 54, No. 215 (Aug., 1987), 289-298.
- Dolbear, F. T., "Is Racetrack Betting on Exactas Efficient?" *Economica, New Series*, Vol. 60, No. 237 (Feb., 1993), 105-111.
- Harville, D.A., "Assigning Probabilities to the Outcomes of Multi-Entry Competitions," *Journal of the American Statistical Association*, Vol. 68, No. 342 (Jun., 1973), 312-316.
- Hausch, D.B., Ziemba, W.T., and Rubinstein M., "Efficiency in the Market for Racetrack Betting", *Management Science* Vol. 27, No. 12 (1981), 1435-1452.
- Hausch, D.B. and Ziemba, W.T., "Betting at the Racetrack," (1985).
- Hausch, D.B., Lo, V. and W.T. Ziemba, *The Efficiency of Racetrack Betting Markets*, Academic Press (1994).
- Henery, R.J., "Permutation Probabilities as Models for Horse Races", *Journal of the Royal Statistical Society Series B (Methodological)*, Vol. 43, No.1 (1981), 86-91
- Lo, V.S.Y. and Bacon-Shone, J., "A Comparison Between Two Models for Predicting Ordering Probabilities in Multiple-Entry Competitions," *The Statistician*, Vo. 43, No.2 (1994), 317-327.
- Rosenbloom, E.S., "Selecting the Best of k Multinomial Parameter Estimation Procedures using SPRT," *Sequential Analysis*, 19, No. 4 (2000), 177-192.
- Rosenbloom and Notz, "Statistical Tests of Real-Money versus Play-Money Prediction Markets," *Electronic Markets*, Vol. 16, No.1, 2006.

Rosenbloom, E.S., "A Better Probability Model for the Racetrack using Beyer Speed Numbers," *OMEGA* Vol. 31, No. 5 (2003), 339-348.

Snyder, W. "Horse racing: testing the efficient markets model", *Journal of Finance* 133 (1978) 1109-1118.

Teichroew D., "Tables of Expected Values of Order Statistics and Products of Order Statistics for Samples of Size Twenty and Less from the Normal Distribution," *The Annals of Mathematical Statistics*, Vol. 27, No.2 (Jun., 1956), 410-426.