LINEAR PARAMORPHIC PROBABILITY ESTIMATES OF PARI-MUTUEL DECISION-MAKING.

OLEKSANDER. MITZAK
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
https://scholar.uwindsor.ca/etd/2618

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l’objet d’un droit d’auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.


LA THÈSE A ÉTÉ, MICROFILMÉE TELLE QUE NOUS L'AVONS REÇUE
LINEAR PARAMORPHIC PROBABILITY ESTIMATES
OF PARI-MUTUEL DECISION-MAKING

by

Olehskander Mitzak
B. A. (General Arts), York University, 1972
B. A. (Honours Psychology), York University, 1973
M. A. (Personality Psychology), University of Windsor, 1976

A Dissertation
Submitted to the Faculty of Graduate Studies
Through the Department of Psychology
in Partial Fulfillment of the
Requirements For the Degree
of Doctor of Philosophy
at the
University of Windsor

Windsor, Ontario, Canada
1981
ABSTRACT

Although numerous psychological studies of pari-mutuel decision-making have demonstrated a high level of congruence between estimated win probabilities and observed win probabilities, they have nevertheless reported an "under/over" bettor bias at the extremes of the subjective probability distribution whereby highly successful horses are underbet and highly unsuccessful horses are overbet. The purpose of this study was to explore this bettor bias through the construction of a paramorphic representation, which is produced by a linear weighting procedure that embodies the information base used by decision-makers. Dawes and Corrigan (1974) have concluded that a linear model based on the input of a judge will often produce superior criterion results even though the model is based on the judge and not the criterion. This superiority is attributed to the fact that a linear representation of a decision-maker's behavior removes unreliability from his decisions while retaining the essence of his expertise. Here, it was assumed that the bettor bias was a form of unreliability and that superior mathematical expectations could be obtained if the public's subjective win probabilities were replaced (bootstrapped) by their linear paramorphic counterparts.

Computer analyses were performed on a data base consisting of detailed historical and descriptive information published in the racing form on 1994 horses that competed in 240 harness races. An under/over bettor bias was found and the data base was judged comparable to those
reported in the literature. Using a least squares regression procedure, this author was able to construct a twelve-predictor model that accounted for 65% of the linearized variance in the public probabilities. The actual and predicted probabilities were then ranked in descending order and hypothetical unit wagers were placed on each set of ranks.

Preliminary results indicated that the linear estimates could not be used to produce expectations greater than those of the betting public. For example, unit wagers on the public favorites produced an expectation of .12, while similar wagers on paramorphic favorites produced an expectation of -.13. However, it was noted that an expectation of .33 could be achieved if a unit wager was placed on the horse that produced the largest positive error in linear fit in each race. This suggested that the public probabilities were determined by information which was not in the public domain. A "fair game model" was used to classify the races into two subsamples according to the degree that their estimated probabilities reflected the available information pool, giving 106 races deemed "fair" and 134 races deemed "unfair."

Recalculated expectations in the two subsamples revealed that the under/over bettor bias was not present in the fair subsample, while it was extremely pronounced in the unfair subsample. In the fair subsample, the paramorphic favorites and second favorites produced expectations of .30 and .40, respectively, and thereby vindicating the efficacy of paramorphic bootstrapping. In addition, the horses with the highest positive error of prediction produced a large negative expectation of -.46. In the unfair subsample, the paramorphic favorites
and second favorites produced large negative expectations of -.48 and -.13 respectively. However, the horses with the highest positive error in linear fit produced an extremely high positive expectation of .95.

It was concluded that the reported systematic disturbances between subjective and objective probability distributions can be attributed to a subset of races in which one of the competitors is overbet despite an apparent mediocrity in published records. In other words, the under/over bettor bias reported in the literature was shown to be an artifact of information availability. Thus it was argued that a model of judges will outperform the judges themselves if the information pool conforms to a fair game model. However, it was noted that if the information pool is affected by exogenous information then the deviations and not the predictions themselves would reveal the impact of nonlinear influences. It was recommended that the paramorphic bootstrapping / fair game strategy be generalized to other research areas such as clinical prediction.
ACKNOWLEDGEMENTS

I would like to thank the people whose guidance and encouragement made the completion of this dissertation possible. I begin by offering my deepest gratitude to the members of my examining committee: Drs. Igor Kusyszyn, Raymond H. Daly, and Meyer W. Starr. To Igor, I am much indebted for your generosity, in encouraging the feasibility of this project, financing the construction of the data base, and sharing freely your knowledge and experience. To Ray, I offer my deepest respect for your wisdom and kindness, and in particular for showing me that psychology can be a most graceful adventure when the mind is disciplined by the heart. To Meyer, my mentor and friend, I owe more to you than to anyone else. Your guidance, your faith, and your loyalty will always be remembered and cherished in my heart.

I am also very grateful for the detailed statistical advice and criticisms from Dr. A. Arthur Smith, Mr. Michael D. Baillargeon, and Mr. Peter F. O'Neill. I am particularly indebted to Mr. O'Neill for his companionship and insights during those long winter nights at the computer centre as we explored yet another facet of the data base.

I would also like to thank Dr. Akira Kobasigawa and Mr. Joseph Paul Younder for taking the time to read preliminary drafts of this manuscript and offering invaluable editorial comments and suggestions. My sincere thanks is also given to Mrs. Carol Hallman for her careful and prompt typing of this manuscript.

Finally, I would like to offer my deepest love to my two closest friends and confidants, Mickey and Holly.
ПРІСВЯЧУЮ
МОЇМ ДОРОГИМ РОДІЧАМ
АЛЕКСАНДРОВІ І АГАФІЇ
З ЛЮБОЮ І ГЛИБОКОЮ ПОШАНОЮ

...ОДЕСЬ
TABLE OF CONTENTS

CERTIFICATE OF EXAMINATION ........................................... i

ABSTRACT ........................................................................... ii

ACKNOWLEDGEMENTS ........................................................... v

LIST OF TABLES .................................................................. viii

LIST OF FIGURES ................................................................ ix

CHAPTER

I INTRODUCTION: ................................................................. 1

Definition of the Problem Domain ........................................ 2
The Pari-Mutuel Wagering System ........................................ 6
The Pari-Mutuel Wagerer .................................................... 11
Literature Review ............................................................. 19
The Research Problem ....................................................... 26
Linear Models ................................................................. 27
Efficient Markets Hypothesis ............................................. 30
Summary of the Problem ................................................... 36

II METHOD .......................................................................... 37

Data Base ........................................................................... 37
Preliminary Considerations ............................................... 38
Analytic Procedures ........................................................ 41

III RESULTS ......................................................................... 44

Development of a Paramorphic Representation .................... 47
The Paramorphic Representation ....................................... 50
The Bootstrapping/Fair Game Issue .................................... 56

IV DISCUSSION .................................................................. 66

Structure of Paramorphic Probabilities ............................... 66
Function of Paramorphic Probabilities ............................... 71
Future Research Directions .............................................. 76

REFERENCES ..................................................................... 79

APPENDIX

A DESCRIPTION OF DATA BASE .................................... 86

VITA AUCTORIS .................................................................. 92
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Example of Pari-Mutuel Win Betting in a Fictitious Six-Horse Race.</td>
<td>7</td>
</tr>
<tr>
<td>2. Mean Order of Finish Position as a Function of Odds Ranked within Race.</td>
<td>45</td>
</tr>
<tr>
<td>3. Expectations of Unit Wagers as a Function of Odds Ranked within Race.</td>
<td>46</td>
</tr>
<tr>
<td>4. Summary of Final Regression Model Selected by the Stepwise Selection Procedure.</td>
<td>51</td>
</tr>
<tr>
<td>5. First Order Correlations for the Twelve Selected Predictors with the Logit EPW Criterion and the FIN Criterion.</td>
<td>52</td>
</tr>
<tr>
<td>6. Intercorrelations Among Major Constructs for Total, Fair, and Unfair Samples.</td>
<td>58</td>
</tr>
<tr>
<td>7. Expectations and (Win Percents) for EPW, EPW', and DPW in Total, Fair, and Unfair Samples.</td>
<td>60</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure | Page
--- | ---
1  Frequency distributions of EPW and Logit EPW | 48
2  Frequency distributions of Logit EPW and EPW' | 57
A-1  An example of sources used to construct the data base | 87
CHAPTER I

INTRODUCTION

Since its inception in the early 1870's, the pari-mutuel system of betting has become the single most popular method of redistributing wagered capital in multi-entry competitions such as standardbred and thoroughbred racing. Apart from its dispensational function, the pari-mutuel system is capable of providing psychologists with a rich source of real-life data concerning probability estimates generated within the context of group decision-making under conditions of risk and uncertainty. Unfortunately, the relevant studies have been limited to simple frequency-count designs whereby comparisons have been made between ex ante subjective probabilities (i.e., closing dollar odds) and ex post objective probabilities (i.e., success rates at various odds-levels). Although this approach has produced consistent findings, other important issues remain outstanding. For example, what information does the public attend to when assigning probability estimates, and how is this information weighted? Second, can linear models, and in particular paramorphic models, be used to produce more accurate probability estimates? Third, to what extent can deviations from paramorphic predictions be used to test the "fair game" or "efficient markets" hypothesis? Finally, can a paramorphic efficient-markets approach be used as a data-analytic technique in other decision-making research? This study will address these and related issues in detail.
Definition of the Problem Domain

Human decision theory can be defined as the scientific explication of the proper and actual principles of choice behavior. Interest in this area has developed in two distinct, but not mutually exclusive, directions: normative and descriptive decision theory. Normative decision theory is concerned with exploring the various strategies available to ideal decision-makers in terms of optimizing objectively defined criteria. This line of inquiry has developed within two disciplines: mathematics and economics. The mathematical approach is concerned with finding solutions for extrema based on rules or strategies that are required to optimize a given set of numerical consequences. Examples of work in this area can be found in Ramsey (1931) and Savage (1954). Similarly, the economic approach emphasizes the ideal forms of choice behavior which maximize either objective monetary value (profit) or subjective monetary value (utility). Examples of work in this area can be found in Friedman and Savage (1948) and von Neumann and Morgenstern (1953). In both mathematical and economic decision theory, major emphasis is placed on the choices men should make rather than the choices they actually do make. In contrast to this normative approach, descriptive decision theory is concerned with documenting, explaining and predicting the actual choice behaviors of real people. To this extent descriptive decision theory falls within the domain of psychology.

Psychological (or behavioral) decision theory aspires to relate various personal and situational characteristics to the common and unique choice behaviors of individuals and groups. A selection of the
various research topics within this orientation could include: information processing (e.g., Schroder, Driver, & Streufert, 1967; Shepard, 1967); risk taking (e.g., Kogan & Wallach, 1964); probability learning (e.g., Edwards, 1961; Estes, 1950); riskless choice (e.g., Adams & Fagot, 1959; Yntema & Torgerson, 1967); gambling behavior (e.g., Bergler, 1957; Cohen, 1960, 1964); subjective probability (e.g., Edwards, 1962; Tversky, 1967); and game theory (e.g., Deutsch & Krauss, 1962; Rapaport, 1966). Comprehensive introductions to the many facets of behavioral decision theory are offered by Edwards and Tversky (1967) and Lee (1971).

The distinguishing feature of psychological decision theory is its emphasis on using empirical evidence. Although behavioral decision theory borrows extensively from normative decision theory, it does not attempt to place man on a procrustean bed of optimization principles. For this reason it can be viewed as an important counter-balance to the economic and mathematical orientations. Unfortunately, this spirit of empiricism has not guaranteed unambiguous or generalizable results. Many fundamental constructs (e.g., rationality, utility, variance preference, subjective probability) have not been convincingly demonstrated or adequately validated (see Lee, 1971). All too often important variables have been studied in isolation and in a piecemeal fashion -- generating much data and little insight into human decision-making.

It is felt that the major weakness in behavioral decision-making research stems from a general preoccupation by investigators with laboratory experimentation and a misguided eschewal of naturalistic
field studies. This bias against in vivo data originates from psychology's emulation of the experimental procedures used in the natural sciences which are assumed to provide rigor and control in the research situation. Although laboratory research methods are capable of producing high levels of internal validity, or reliability, they do not necessarily ensure reasonable levels of external validity, or generalizability (Campbell & Stanley, 1966). In light of the expressed aims of behavioral decision theory, one might expect that a premium would be placed on achieving generalizability through the use of naturalistic data-gathering procedures. Furthermore, a growing number of critics (e.g., Argyris, 1968; Aronsmith & Carlsmith, 1968; LaFave, 1971; Orne, 1962; Rosenthal, 1966) argue that the philosophical presuppositions of physicalism in the laboratory setting are untenable and inappropriate when applied to human subjects. Issues such as motivation, demand characteristics, reactivity, impact versus control, and individual differences exemplify the difficulties faced by social scientists who employ classical manipulative methods.

In addition to these problems, there exists among researchers a naive belief that the fixed-model analysis of variance (and covariance) is the only legitimate data-analytic procedure that should be used to explicate relationships among independent and dependent constructs. Concomitant to this view is the mistaken belief that the primary function of multiple linear regression is prediction, and that it is inappropriate for problems related to rigour and control. In reality, fixed-effects analysis of variance is a special case of linear multiple regression analysis (i.e., the general linear model) which contains features of analytic flexibility not found in the analysis of variance.
Multiple regression is also capable of providing the necessary statistical controls demanded by laboratory researchers, while at the same time producing high levels of external validity. For a discussion on the advantages of multiple linear regression in relation to the analysis of variance, see Cohen (1968), Cohen and Cohen (1975), and Draper and Smith (1966).

In light of these issues, this author believes that behavioral decision-making research can best profit by: (a) gathering data from naturalistic environments; and, (b) fully exploiting the analytic capabilities of the general linear model. Working within such a framework, this study will explore a popular form of real-life decision-making that has received scant attention from behavioral decision-making researchers: group wagering behavior at the pari-mutuel race-track. This area of interest was selected for several reasons. First, the relevant data are electronically recorded, verified by law, and publicly available. Second, unlike typical laboratory studies, the subjects are neither coerced nor recruited; rather, they freely participate of their own volition. Third, participants respond to pari-mutuel wagering with high degrees of motivation, since the incentives are meaningful and the realism of the activity is salient. And finally, the pari-mutuel data are amenable to the analytic assumptions and capabilities of the general linear model.

The central argument of this study will be that a linear regression model of pari-mutuel wagering can be used to explicate and validate principles of human decision-making. More specifically, the author believes that linear analysis can be used to identify and evaluate the relevant factors that influence race-track bettors when they assign probability estimates under the constraints of the pari-mutuel system.
Before such an argument can be proposed and supported by the relevant literature, the reader must first be familiarized with: (a) the descriptive features of the pari-mutuel wagering system; and (b) the underlying assumptions about the participants in the pari-mutuel setting. These two preliminaries will precede the literature review.

The Pari-Mutuel Wagering System

Pari-mutuel betting ("to wager among ourselves") was developed in 1872 by M. Oller, a Parisian businessman (Buck, 1962), as a method of horse-race betting whereby a bookmaker or racing association acts as a broker for the betting public and reimburses the wagered capital in an inverse proportion to the level of endorsement on the winning\(^1\) horse. The pertinent details of pari-mutuel betting can best be explained by example.

Table 1 represents the win betting patterns in a fictitious six-horse race. Individual members of the betting public first evaluate the win potential for each of the six contestants by examining past performance data published in the track program. For reasons that will be discussed in the next section, the betting public places wagers in varying amounts on the selection they prefer. The first numeric column in Table 1 shows the money bet (MB) on each i\(^{th}\) contestant, along with the total moneys wagered (\(\Sigma\text{MB}_i\)), commonly referred to as the "pool." The second numeric column in Table 1 shows the socially determined probabilities of winning (\(\text{PW}_i\)) for each potential winner, where:

\[
\text{PW}_i = \frac{\text{MB}_i}{\Sigma\text{MB}_i}
\]  

\(^1\)Although the pari-mutuel system can accommodate performance criteria other than winning (e.g., placing or showing), or performance criteria involving more than one competitor (e.g., daily double, exactor, quinella, or triactor), this paper will only consider win betting since sufficient information on these other forms of pari-mutuel betting is not typically published.
Table 1
Example of Pari-Mutuel Win Betting
in a Fictitious Six-Horse Race

<table>
<thead>
<tr>
<th>Horse</th>
<th>MB</th>
<th>PW</th>
<th>DO</th>
<th>TO</th>
<th>1/(DO+1)</th>
<th>EPW</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1,000</td>
<td>.006</td>
<td>140.10</td>
<td>99:1</td>
<td>.007</td>
<td>.005</td>
</tr>
<tr>
<td>B</td>
<td>2,000</td>
<td>.012</td>
<td>69.55</td>
<td>60:1</td>
<td>.014</td>
<td>.011</td>
</tr>
<tr>
<td>C</td>
<td>3,000</td>
<td>.018</td>
<td>46.00</td>
<td>45:1</td>
<td>.021</td>
<td>.017</td>
</tr>
<tr>
<td>D</td>
<td>15,000</td>
<td>.090</td>
<td>8.40</td>
<td>8:1</td>
<td>.106</td>
<td>.089</td>
</tr>
<tr>
<td>E</td>
<td>38,000</td>
<td>.228</td>
<td>2.70</td>
<td>5:2</td>
<td>.270</td>
<td>.227</td>
</tr>
<tr>
<td>F</td>
<td>107,000</td>
<td>.644</td>
<td>.30</td>
<td>1:5</td>
<td>.769</td>
<td>.647</td>
</tr>
<tr>
<td>Total</td>
<td>166,000</td>
<td>1.000</td>
<td>.30</td>
<td>1:5</td>
<td>1.188</td>
<td>.996</td>
</tr>
</tbody>
</table>

MB: money bet.
PW: probability of winning.
DO: dollar odds.
TO: tote odds.
1/(DO+1): inverse of dollar odds plus unity.
EPW: estimated probability of winning.
The PW represent the cumulative probabilities of the betting public reflecting how closely the public believes the six contestants are matched in ability. Thus, in our example, the public believes that if this race were repeated over a 1,000 identical trials, horse "A" would win approximately 6 races. Unfortunately, pari-mutuel betting information (published in race "charts") does not typically report MB, \( \Sigma MB \), or PW; however, "dollar odds" are provided. Dollar odds (DO) can be described as follows:

All the money bet on prospective winners of a race is pooled and a percentage (ranging from 7 to 15) known as the 'take' is deducted as revenue for the track and for the state and local government. The ('dollar') odds on any horse...are computed by subtracting the sum bet on the horse from the pool after take and dividing by the sum bet on the horse...No pennies, and in some states, no nickles are returned...this is termed "breakage" to the nickle or dime. (Griffith, 1949, p. 291)

In other words, dollar odds represent the expected profit for each prospective winner based on a dollar unit wager. Using algebraic notation, dollar odds become:

\[
DO_i = \left\{ \frac{\Sigma MB_i - [ MB_i + (T \times \Sigma MB_i) ]}{MB_i} \right\}
\]

Here, "T" refers to the track take, and the DO are rounded down to the first nickel or dime. The third numeric column in Table 1 shows the DO for each horse based on a 15\% take and breakage to the first nickel. For example, the dollar odds for horse "F" equal .319, which is then rounded to .30:

\[
\left\{ \frac{166,000 - [ 107,000 + ( .15 \times 166,000 ) ]}{107,000} = .30
\]

Hence, a successful $2.00 wager (the typical minimum wager) on horse "F" would return $2.60, representing $0.60 in profit (2 x .30) plus the original $2.00 wager.
During the actual betting period (20 minutes prior to each race) the public is provided with up-to-the minute estimates of the current dollar odds from previous betting displayed on a centrally positioned scoreboard known as the "totalizator" or the "tote." The tote flashes the emerging betting patterns in 60-90 second cycles using approximations of current dollar odds by means of "tote odds." Tote odds (TO) are simplified versions of dollar odds, and they represent ratios of expected profit per arbitrary unit wager. In our example, the fourth numeric column in Table 1 represents the TO which range from the least popular competitor (the race "longshot") to the most popular (the race "favorite"). When current dollar odds are greater than, or equal to 99.00, the corresponding tote odds display 99:1 (see horse "A"). When the current odds are greater than 49.95 and less than 99.00, the corresponding tote odds are rounded down to the nearest ten-dollar unit (see horse "B"). In a similar fashion, current dollar odds which are greater than 24.95 and less than 50.00 will produce tote odds that are rounded down to the nearest 5-dollar unit (see horse "C"). When the current dollar odds are greater than 4.95 and less than 25.00, the corresponding tote odds round the DO down to the nearest dollar integer (see horse "D"). Finally, when the current dollar odds are less than 5.00, finer discriminations of the tote odds are provided. Typically, totalizators are capable of generating 18 unique ratios for dollar odds-intervals between 4.95 and .05 (the legally set minimum payoff). These ratios are: 9:2, 4:1, 7:2, 3:1, 5:2, 2:1, 9:5, 8:5, 3:2, 7:5, 6:5, 1:1, 4:5, 3:5, 1:2, 2:5, 1:5, and 1:9. Dollar odds in this lower range are rounded down to the nearest appropriate tote odds ratio (see
horse "E"). At the conclusion of betting, successful bettors redeem their winning tickets according to the DO which are referred to as "closing dollar odds" or simply "dollar odds."

Initially, our concern was in calculating the PW, the public probabilities of winning; however, because necessary information was unavailable, our focus shifted to DO, the dollar odds. Fortunately, dollar odds can be transformed into public probabilities using the following relation which can be derived from equations [1] and [2]:

\[ PW_i = \left[ \frac{1}{(DO_i + 1)} \right] / \sum \left[ \frac{1}{(DO_i + 1)} \right] \]

Equation [3] reveals that the publically determined probabilities of winning can be calculated from the normalized inverse of, the dollar odds plus unity. This relation is demonstrated in the last two columns of Table 1. The fifth numeric column shows the reciprocal of the dollar odds plus one: \( \left[ \frac{1}{\sqrt{(DO+1)}} \right] \). The sum of these "probabilities" would have equalled 1.00 had \( T \) been set to zero in equation [2]. To correct this overestimation, each entry in the fifth column is normalized by dividing through by \( \sum \left[ \frac{1}{(DO+1)} \right] \). The results of this operation are presented in the final numeric column as the estimated public probabilities of winning (EPW). A comparison between corresponding entries for each PW and EPW shows that they are not precisely isomorphic; specifically, EPW tends to be slightly lower than PW. This minor discrepancy can be attributed to breakage and rounding error, and for purposes of analysis this underestimation can be ignored.

In summary, the pari-mutuel system of wagering is a method of collating the collective decision-making behaviors of the betting public expressed in terms of dollar odds which are inversely proportional to any contestant's assumed probability of winning. The next section will
examine the underlying assumptions about the decision-making constructs involved in producing dollar odds.

The Pari-Mutuel Wagerer

This section will outline the major personal and situational characteristics of an "ideal" pari-mutuel wagerer as a representative of the betting public (cf., Weitzman, 1965). These characteristics will be gleaned primarily from normative decision theory, and they will prepare the reader both for the literature review and the rationale for the present study.

We begin by assuming that the ideal pari-mutuel wagerer is a decision-maker who can be portrayed as an "economic man." Economic man behaves according to three principles (Edwards, 1967, pp. 14-16). The first states that he is fully informed. This principle implies that the ideal pari-mutuel wagerer is capable of using all the available and relevant information prior to making his decision. The second principle states that economic man is infinitely sensitive. This suggests that an ideal decision maker is capable of perceiving and analyzing a decision problem in terms of infinitely divisible gradients. Consequently, our pari-mutuel bettor is assumed to be capable of responding to his choice situation through the use of continuous measurement. The third, and most important, principle states that economic man is rational. This entails that a decision-maker is both hedonistic and purposeful (Parsons, 1949, pp. 135-136). He is hedonistic to the extent that he will seek decision-making consequences which will, in some sense, provide the most happiness and least pain. Thus, in the pari-mutuel

---

2 This principal subsumes the traditional economic assumption that an individual ranks all possible alternatives, chooses the highest ordered alternative available, and conforms to the transitivity assumption (Tsukahara & Brumm, 1974, p. 2).
setting our typical wagerer will select those alternatives which he believes will provide maximum gratification. Although the specific features of what constitutes maximum gratification will be discussed at a latter point, it can be hypothesized that an important dimension of personal gratification at the race-track would include some variant of maximizing long-term monetary profit. In addition to being hedonistic, rational man is assumed to behave purposefully. Faced with \( n \) discrete options, a decision-maker will consistently employ a strategy, or decision principle, that will indicate which option will be most likely to produce the highest level of personal gratification. Before we extend our analogy between rational man and the pari-mutuel bettor, it is first necessary to focus more intensely on the normative features of decision strategies in terms of (a) condition of play, and (b) admissibility.

Decision principles are employed under one of three conditions of play: certainty, risk and uncertainty (Luce & Raiffa, 1957, p. 13). A decision-maker is said to employ a strategy under the condition of certainty when the probability of achieving personal satisfaction will equal unity once one of the \( n \) options is selected. For example, an individual who chooses his favorite color would be expressing his preference under the condition of certainty.

A decision-maker is said to operate under the condition of risk when he is fully aware of the true probabilities of satisfaction for each of the \( n \) options. In other words, decision-making under risk can be defined as choice behavior occurring within the context of a known probability distribution. In decision theory such choice behavior is
referred to as a "gamble." An example of a gamble occurs when an individual uses a coin toss to select between two equally attractive alternatives. Here, the individual matches his equally preferable options with a gamble of equally likely probabilities.

An individual is said to behave under the conditions of uncertainty when his decision is based on an assumed probability distribution for the \( n \) options. In contrast to the previous condition, decision-making under uncertainty deals with an individual's beliefs rather than his knowledge, and as such represents a most common class of choice behavior. An example of decision-making under uncertainty could involve an individual who believes that his chance of winning a lottery exceeds those chances of his fellow players.

It should be noted that an individual's beliefs about an underlying probability distribution does not necessarily imply a mismatch with the objective probability distribution. In many instances such as mismatch simply reflects a lack of information on the part of the decision-maker as to the necessary dimensions required to generate an accurate estimate. For this reason, some authors prefer to substitute "ignorance" or "partial ignorance" for the term "uncertainty" (e.g., see Lee, 1971, pp. 23-25).

In addition to the condition of play, one can look at decision-making in terms of the type of admissible strategy an individual may employ in order to satisfy his hedonistic motives. Normative decision theorists have proposed numerous principles to explain the purposefulness of the decision-maker (see Edwards, 1967): e.g., the maximin strategy, the minimax strategy, the dominance principle, and the Laplace criterion. However, the most prominent and
extensive rule for decision-making (excluding conditions under uncertainty) has been the "expected value" principle, commonly known as the "mathematical expectation" (Bayes principle). The expected value (EV) for any option equals the sum of all the alternatives weighted by their respective probabilities:

$$EV_i = \Sigma(P_i V_i)$$  \[4\]

where the $P_i$ refer to the probabilities of successful consequences for each of the alternatives, and the $V_i$ refer to the values of the possible outcomes (e.g., monetary profits or losses) for each of the alternatives.

According to the EV principle, a rational decision-maker will select that option which produces the highest positive EV. Taking an example from Kogan and Wallach (1967, p. 116), an individual is offered the opportunity to participate in one of two gambles. The first offers one chance in two of winning 50¢, and one chance in two of losing 50¢. The second offers one chance in four of winning $1.20, and three chances in four of losing 30¢. According to the expected value principle, a rational person should select the second gamble, since the expected value of the first shows no profit [ $(.5 \times 50\text{¢}) + (.5 \times -50\text{¢}) = 0\text{¢}$] and its expectation is lower than the second alternative [ $(.25 \times 1.20\text{ dollars}) + .75 \times -30\text{¢} = 7.5\text{¢}$].

Although the EV principle has been the most widely advocated strategy for explicating the decision-making process, the supporting evidence has been mixed (e.g., see Coombs & Komorita, 1958; Edwards, 1967; Mosteller & Nogee, 1951). The major weakness of the EV principle stems from the fact that subjects apparently do not consistently choose
the option with the highest expectation. Discontent with the EV principle has spawned three variant models, each attempting to account for a larger portion of the decision-maker's behavior. The three models differ from equation [4] according to the meaning that they attribute to $P_i$ and/or $V_i$. In the original EV, the terms $P_i$ and $V_i$ are assumed to be objective (i.e., veridical across all individuals), while in the three variant models, one or both terms are replaced by their subjective analogue (i.e., subjective probability and subjective value, commonly referred to as utility). These models are: subjective expected value \( \text{SEV} \) (subjective probability and value); expected utility \( \text{EU} \) (objective probability and utility); and, subjectively expected utility \( \text{SEU} \) (subjective probability and utility).

Subjective (or psychological) probability can be defined as "the belief an individual has that... ('an') event will occur" (Hoerl & Fallon, 1974, p. 227). Although the evidence is mixed (e.g., see Cohen, Dearnaley & Hansel, 1957), subjective probabilities are assumed to possess the same properties as objective probabilities (viz., the sum of the probabilities for a set of mutually exclusive and exhaustive events will equal unity). It should be noted that subjective probabilities refer to probabilistic statements generated by real people, while the term "personal probabilities" (Savage, 1954) is reserved for probabilistic statements generated by ideal or normative players. Although this distinction is often ignored, it is useful when one recalls that decision-making under the condition of uncertainty requires a person to reconstruct an underlying objective probability distribution. Thus, in one sense, a personal probability distribution reflects an ideal case where an individual has sufficient information to
produce an isomorphic correspondence between his beliefs and the true state of nature.

Utility can be defined as "the attractiveness of a consequence to a person as judged from foresight" (Lee, 1971, p. 9). In economic theory, it is assumed that the objective unit-value of a commodity is not in a one-to-one correspondence with an individual's subjective unit-value (i.e., a "utile"). Different individuals place different levels of attractiveness on the same object. By introducing the notion of utility, economists hoped to explain the seemingly irrational behaviors of men playing casino games and buying insurance -- both of which produce negative expectations. Consequently, economists have been particularly keen in searching for general and specific utility curves that relate money to utility. For example, Markowitz (1955) argued that utility is a function of a person's customary wealth. Thus, a dollar should have more utility for a beggar than a millionaire. He proposed a positive monotonic function that was made up of alternating sequences of concave upward and concave downward components placed symmetrically about a person's level of customary wealth. Other examples of utility analysis have been offered by Friedman and Savage (1948), Mosteller and Nogee (1951), Stevens (1959), and Vickrey (1945). Unfortunately, the usefulness of the concept of utility (especially for the EU and SEU variants) has not been convincingly demonstrated. Although using utilities in place of values do in fact produce more accurate predictions of choice behavior, their superiority has yet to be shown to exceed chance levels (see Lee, 1971, p. 105).

Returning to the pari-mutuel setting, it is not clear whether horse-race betting occurs under the condition of risk or uncertainty.
Theoretically, the PW reflects only the collective judgments of the betting public in assigning probability estimates, and they do not necessarily reflect the underlying "true" probabilities of winning (PW'). Unlike games of pure chance (e.g., gambles such as roulette), multi-entry competitions such as horse-racing do not contain unique "internal" mathematical probabilities of winning (Gruen, 1976). Consequently, horse-race bettors can generate personal estimates of PW' by processing (handicapping) information they believe to be relevant to the criterion of winning. The pari-mutuel system is simply a method of collating all personal estimates of PW' to produce PW. Thus, from the perspective of the pari-mutuel bettor, the dollar odds may be treated as reflections of either an objective or a subjective probability distribution. In the former, the bettor does not calculate a personal estimate of PW'; rather, he treats dollar odds as true or given probabilities, and consequently his wagers would constitute decisions made under the condition of risk. However, if our bettor realizes that dollar odds represent collective subjective probabilities, and he constructs personal estimates of PW' then his wagers would represent decisions made under the condition of uncertainty. This issue is further complicated by the fact that subjective estimates of PW' may be veridical to PW (i.e., personalistic probabilities). For these reasons, this author prefers to define pari-mutuel wagering as decision-making under the condition of risk and uncertainty. However, as will be seen shortly, this author feels that the ideal wagerer operates predominantly under the condition of uncertainty.

The final issue to be discussed in this section concerns the type of admissible strategy our ideal bettor is assumed to employ. Although
the assumption of a general utility curve for our decision-maker would be a most fruitful avenue of study, our commitment to dollar odds prevents such an approach. Unavailable factors such as customary wealth, amounts wagered, effects of differential reinforcements would render both utility models (EU and SEU) inappropriate (however, see Gruen, 1976; Rosett, 1965; Weitzman, 1965). In the remaining two models, EV and SEV, the former would not be useful since the values (dollar odds) are algebraic transformations of the "objective" win probabilities (PW') and their expectations would always equal a constant. This leaves us with the SEV model. However, because of the special nature of the pari-mutuel system a modified version of this expectation will be presented.

We assume that our ideal wagerer behaves as if he were a professional gambler. A professional gambler is an individual who employs his superior handicapping skills to consistently produce more accurate personal estimates of the PW' in comparison to the majority of the betting public. Before wagering he intuitively constructs ratios between his personal estimates of PW' and PW (cf., Eisenberg & Gale, 1959). When a specific ratio is less than unity it is known as an "underlay" indicating an unfavorable wager; conversely, a ratio greater than unity is known as an "overlay" representing a favorable wager. Thus, a professional gambler is an individual who seeks out and wagers on the highest overlay in any race. In terms of an SEV model, the professional gambler's strategy can be defined as:

\[
SEV_i = \left( \frac{(1-T)}{PW_i} \right) - 1
\]  

An inspection of equation [5] reveals that when PW equals PW' the expected per dollar profit/loss will equal the track take (i.e., SEV
reduces to EV). In other words, if the public's cumulative determination for each PW equals PW', then a wager on any i'th competitor will produce a long-term percentage loss equal to the track take. Consequently, any better who wishes to "beat the races" must produce substantial increments in the ratios between the PW and the PW' to produce long-term profits. Interestingly, the available behavioral literature on group decision-making at the pari-mutuel race-track has implicitly followed the example of the professional gambler's concern with overlays and underlays.

**Literature Review**

The first detailed study of subjective probabilities in a pari-mutuel setting was reported by Griffith (1949). He reasoned that not all forms of gambling are psychologically equivalent; rather, they may be placed within a hierarchy of skills ranging from the conative (e.g., roulette) to the cognitive (e.g., poker). To the extent that separate irrational factors in pari-mutuel wagering would on the average cancel out, he hypothesized that betting behaviors were largely determined by rational (cognitive) factors. Consequently, he proposed that dollar odds could be used as indices of the public's handicapping skills. The major thrust of his study was to compare these decision-making skills with the win percentages within each odds-level category. The rationale for this approach was based on the assumption that, "if the psychological ('dollar') odds equalled the a posteriori given by the reciprocal of the percentage winners, the product of the number of winners and their odds would equal the number of entries at
each odds group" (p. 292). Expressed differently, the public's probability estimates would be accurate if the observed win percentage for all horses listed at, for example 3:1 odds ($PW_i = .25$) would be one in four ($PW'_i = .25$). His data consisted of the number of entries and winners within 11 collapsed odds-levels for 1,386 thoroughbred races at three New York State race-tracks. Griffith's results were presented in the form of a graph where the distribution of winners times odds was superimposed over the distribution of total entries. The degree of overlap between these two positively skewed distributions increased once the odds had been corrected for breakage and take. This result indicated a high correspondence between the ex ante subjective probabilities and the ex post objective probabilities, and it validated his assumption that pari-mutuel wagering is regulated by skill rather than chance factors. (For a discussion on the chance versus skill dimension in decision-making, see Kogan and Wallach, 1967, pp. 124-128.) In addition to this finding, Griffith reported a marked deviation of dollar odds from the empirical probabilities at both ends of the performance distribution. Specifically, bettors overestimated the probabilities for longshots and underestimated the probabilities for favorites and near-favorites.

In a similar fashion, Fabricand (1965) analyzed over 10,000 races run at various American race-tracks between the years 1955 and 1962. He partitioned the dollar odds into 22 categories corresponding roughly to the tote odds. He then compared the number of winners within each odds category to the expected number of winners in each category. His results showed that the public probabilities fell within the estimated confidence intervals for 13 of the categories. In the remaining nine
categories, six discrepancies—showed a higher than expected number of
winners for odds below 5:2, while three discrepancies showed a lower
than expected number of winners for dollar odds greater than 9:1. In
other words, the anticipated equality between the public estimations and
the observed win percentages was maintained in the central region of the
performance distribution. As further support for this finding, Fabricand used a variation of equation [5] to calculate the expectation
for each category. This procedure produced a moderately negative
monotonic distribution for the ascending odds categories with the
expectations ranging from .034 to -.54.

Harville (1973) supported Fabricand’s findings with a sample of 335
thoroughbred races run at various Ohio and Kentucky race-tracks.
Instead of focusing solely on win probabilities, he also compared the
theoretical and empirical probabilities for horses "placing" (i.e.,
finishing either first or second) and horses "showing" (i.e., finishing
no worse than third). For all three performance criteria, the results
indicated a close correspondence between theoretical and empirical
probability distributions, with the major deviations occurring in the
upper and lower regions, such that low percentage horses were overbet
and high percentage horses were underbet.

Employing a slightly different methodology, Hoerl and Fallen (1974)
used a data base comprised of 1,825 thoroughbred races for the complete
1970 season at two major New York State race-tracks. They partitioned
their data into eight categories according to the number of entries per
race (5-12). For each category, they ranked the dollar odds within each
race, and they compared the subjective probabilities of winning and the
observed proportion of wins within each odds rank. Chi square values
for each category indicated that in no case was the null hypothesis rejected (i.e., the subjective win probabilities equalled the observed win probabilities). In addition, they reported a tendency on the part of the betting public to overpredict the win potentials for high ranked odds horses. In a further analysis, these authors calculated the mean order of finish within each category for each odds-rank level. The results showed a perfect monotonic non-decreasing function between the mean order of finish and the odds-group ranking.

McGlothlin (1955) examined the stability of pari-mutuel choice behaviors within daily sequences of races. His data consisted of 9,248 thoroughbred races (approximately 1,156 racing dates) run at six California and five New York race-tracks for the years 1947-1953. The dollar odds were grouped into nine ascending odds intervals (corrected for breakage and take), with odds greater than 25:1 collapsed into the last interval. The treatment of the data base was similar to Griffith's except that the total sample was also partitioned into eight subsamples according to the ordinal position the race assumed in the daily program. The results for the total sample showed that: odds less than 3:1 produced an expectation of .08; odds between 3:1 and 8:1 produced a zero expectation; and odds greater than 8:1 yielded an expectation of -.10. Both non-zero expectations exceeded the .05 level of significance. The results for each of the first six subsamples (races one through six) displayed a similar pattern, characterized by a higher than chance expectations for low odds horses and a lower than chance expectations for high odds horses. The seventh race in each daily program also showed increasingly negative expectations for higher odds horses; however, the data failed to show positive expectations for horses in the
lower odds range -- rather near zero expectations were observed. McGlothlin attributed the uniqueness of this subsample to the fact that the seventh race was typically the feature race, and that the public would have a relatively increased familiarity with the win potentials for the well publicized competitors in these races. Turning to the eighth and final subsample, McGlothlin's results displayed marked deviations from the earlier subsamples. Horses below 2:1 odds produced a highly significant expectation of .22; horses between 2:1 and 3:1 odds produced an equally surprising expectation of .11; while horses between 3:1 and 4:1 odds produced a negative expectation of -.10. The middle range odds showed no significant deviations from a zero expectation, while long odds horses produced large negative expectations exceeding -.21. McGlothlin interpreted the unique pattern of the results for the eighth race in terms of a shifting marginal utility on the part of the betting public. Specifically, he reasoned that the public avoids favorites and near-favorites and looks to the longshots and horses in the 7:2 odds range to recoup earlier losses.

Snyder (1978) confirmed McGlothlin's finding of a bettor bias in the last race. His study analyzed the results of a three month 1975 racing meet at an Illinois race-track comprised of 7,657 thoroughbred horses. In this sample there were nine races per day. For the first eight races, Snyder reported small positive expectations for odds classes (adjusted for take) up to 10:1, and increasingly negative expectations for higher odds classes. The rates of return ranged from .03 for odds less than 3:1, to -.15 for the highest odds class at 30:1. Turning to the last race, Snyder found significant positive expectations in the low odds classes and extremely significant negative expectations
in the high odds classes. These rates of return ranged from .25 for odds less than 3:1 up to -.60 for odds at 30:1 or greater. As a further test of the better bias in the last race, Snyder collected additional data from the predicted probability estimates of five public handicapping services for the same series of races. Public handicappers are expert prognosticators whose purpose is to predict (using tote odd values) the subjective probabilities they believe the public will assign to each competitor. Snyder repeated his analysis by comparing the experts' odds with the public's odds for the first eight races and the last. For two of the public handicapping services (Daily Racing Form, track handicapper) the root-mean-squared deviations from the public's rate of returns were not significantly different for the first eight races. The remaining handicappers (from three Chicago newspapers -- Daily News, Sun Times, and Tribune) produced slightly significant root-mean-squared deviations. The five root-mean-squared deviations for each of the prognosticators in the first eight races were: 2.7, 6.7, 7.9, 12.4 and 20.8, respectively. In general, Snyder concluded that the experts were as adept in correctly assigning probability estimates as was the betting public. The slight tendency of the three newspapers to deviate from the public's estimates was attributed to the fact that they frequently employed part-time (and therefore inexperienced) employees. Interestingly, for all five handicapping services there appeared to be a slight tendency to underestimate the chances of favorites and overestimate the changes of longshots beyond the public's own levels of under-estimation. This effect was most pronounced when Snyder calculated the average deviations between public and expert expectations.
for the last race. The root-mean-squared deviations in the last race were approximately three times greater than those in the first eight races. These values were: 17, 18, 22.9, 31.4 and 20.5, respectively. Unfortunately, Snyder did not offer a clear explanation for this effect. However, subsequent interviews with the handicappers suggested that "they did not quote the favorites at lower (and more realistic) odds or longer shots at higher (and more realistic) odds because they didn't want to influence the public's betting behaviour by 'pointing a finger' at those horses more likely to win or lose" (p. 207).

Unlike the previously reported studies which were limited to thoroughbred races, Ali (1977) analyzed over 20,000 standardbred (harness) races. In contrast to a thoroughbred horse who carries a jockey and is permitted to gallop at full gait, a standardbred horse is harnessed to a two-wheeled sulky and is driven by a teamster who insures that his horse maintains a specified pacing (or trotting) gait. As was the case in the Hoerl and Fallen study, the dollar odds were ranked from favorite to longshot. Deviations between the subjective and objective win frequencies for each odds rank were normalized by their expected standard errors. The results indicated the now familiar differences between the subjective and objective probabilities, where the favorites' win probability were significantly underestimated and the longshots' win probabilities were significantly overestimated. Ali proposed that this effect was caused in part by an inverse relation between available capital and risk-taking tendencies. In other words, a less affluent wagerer is more likely to accept a high risk gamble than his more affluent neighbour. In support of this proposition, Ali showed that an empirically derived risk coefficient varied inversely with the average
unit wager per person at three different sized New York race-tracks. The reader will have noticed that this proposition was partially supported by McGlothlin's previously reported finding that the overlap between objective and subjective probabilities was most discrepant in the final race during which the majority of patrons would be expected to have less wagering capital as compared to the first race. However, Snyder's finding that non-betting prognosticators also produce such discrepant estimates weakens Ali's contention.

The Research Problem

The preceding review of the relevant literature suggests three major conclusions. First, the high levels of congruence between dollar odds and observed win frequencies attest to the public's decision-making skills in processing pertinent information, and it substantiates our assumption of rationality for the pari-mutuel bettor. Second, the "under/over bettor bias" at both ends of the subjective probability distribution appears to be a highly consistent and ubiquitous phenomenon in decision-making under conditions of uncertainty. Exogenous evidence for such an effect can be drawn from various tasks involving averaged group estimations of known objective probability distributions (e.g., see Attnavee, 1953; Cohen, 1960; Cohen, Dearnaley & Hansel, 1956; Howard, 1963; Preston & Baratta, 1948; Mosteller & Nogee, 1951). This phenomenon appears to be an analogue of the "central tendency effect" found in psychophysical judgments (see Helson, 1964, pp. 94-102). And third, the literature review suggests that public estimations of objective probability distributions may be influenced by situational factors such as effects of prior gains and losses, available capital, and quality of information.
Linear Models

Although the literature review demonstrates the applicability of pari-mutuel data to the study of human decision-making, several shortcomings related to linear models are evident.

One shortcoming in the literature is that in no case did a researcher investigate which of a multitude of variables the public attends to prior to making its decisions, and how this information is processed. In a typical harness race, for example, the public is provided with more than 150 unique sources of information for each competitor. It would seem unrealistic to assume that all or most of these variables are evaluated within the short period of time between races.

Psychological evidence provided by Miller (1956) and Shepard, Hovland, and Jenkins (1961) suggests that man is severely limited in the number of conceptual units he can process. Further evidence by Desoto (1961) and Osgood, Suci and Tannenbaum (1957) suggests that an individual will collapse multi-attribute dimensions into smaller manageable dimensions. In a review on man's ability to make optimum selections among multi-attribute alternatives, Shepard (1967) concluded that man's "evolutionary background has poorly prepared us for the increasing load of logical and combinatorial manipulation that must intervene between the interpretation of a given situation and the initiation of final action" (p. 261). Shepard contends that although man is capable of evincing high degrees of perceptual analysis to raw sensory inputs, he is nonetheless limited in his ability to apply combinatorial rules to the components of a choice situation. As Dawes
(1977) puts it, "people (expert and nonexpert) are much poorer at processing information than at perceiving and coding it" (p. 335). Given these limitations, the identification of the variables used in group probability estimations at the pari-mutuel race-track, and the development of their combinatorial weights would be a most promising avenue of research. A fruitful beginning to such an approach would involve the use of linear models.

Dawes and Corrigan (1974) reviewed the role of multi-attribute decomposition (linear analysis) in behavioral decision-making and concluded that linear models based on multiple codable inputs of actuarial information consistently outperform human judgments. In their review, they distinguished four functions for linear models in decision-making: normative, competitive, paramorphic, and paramorphic bootstrapping. A normative linear model is a numerical procedure used to aid the decision-maker in reaching an optimal choice (e.g., see Einhorn, 1972, Sawyer, 1966; Shepard, 1967). A competitive linear model (in the "clinical versus statistical" sense) is a weighting procedure of a criterion which is used in contrast to a human judge (e.g., see Meehl, 1954). Both of these models are constructed and validated in terms of an objectively defined criterion. In contrast, the remaining two models are constructed and validated in terms of the decision-maker himself. Paramorphic linear representation is a term coined by Hoffman (1960) to represent any weighting procedure which is used to describe the policies of a decision-maker (see Slovic & Lichtenstien, 1971 for a review). Paramorphic bootstrapping, or simply bootstrapping (Bowman, 1963, Yntema
& Torgerson, 1961) is defined as a procedure in which a paramorphic linear model is used to replace the decision-maker. Dawes and Corrigan (1974) describe the bootstrapping effect in the following manner:

When there are actual criterion values against which the predictions of both the judge and the linear model of the judge can be compared, the paramorphic linear model often does a better job than does the judge himself. That is, the correlation between output of the model and criterion is often higher than the correlation between the decision-maker's judgment and criterion, even though the model is based on the behavior of the decision-maker. (p. 101.)

Evidence for the bootstrapping effect is provided by Bowman (1963), Dawes (1971), Goldberg (1970), Wiggins and Kohen (1971), and Wiggins, Gregory, and Diller (reported in Dawes & Corrigan, 1974). Bootstrapping appears to be most successful when: (a) predictor variables are conditionally monotone with the criterion; (b) there is unreliability in either, or both, the independent and dependent variables; and (c) deviations from optional weighting do not produce serious estimation errors. The success of bootstrapping can be attributed to its ability to embody (but not necessarily mirror) the distinguishing features of the decision-maker without the human error component. Although Dawes and his associates (e.g., Dawes & Corrigan, 1974; Howard & Dawes, 1976) propose alternative explanations for the bootstrapping phenomenon (related to the robustness of the general linear model) its heuristic value for the present study will be significant. In other words, paramorphic bootstrapping as an error-free representation, would be a most appropriate technique to evaluate the decision-making capabilities of pari-mutuel wagerers.
Specifically, this author suggests the use of linear multiple regression analysis to study the dollar odds as a function of the historical variables published in the track program. The partial regression sums of squares will proportionally indicate which variables the public attends to when assigning probability estimates. The resulting first-order linear model will be used as a paramorphic representation of the public's decision-making process at the race-track. The bootstrapping effect will be tested by inspecting the correlational structure of both the paramorphic representation and dollar odds with the finish criterion. In light of the literature review, it is expected that the correlation between the paramorphic representation and the finish criterion will be significantly higher than the correlation of dollar odds with the finish criterion. It is also expected that the paramorphic representation can be used to identify true overlays and thereby produce substantial positive expectations. However, this author believes that the effectiveness of such a paramorphic model must be tempered by the presence or absence of information efficiency in the pari-mutuel market. This issue will be clarified in the following section.

Efficient Markets Hypothesis

The second major shortcoming evident in the literature review concerns the implicit assumption held by researchers that pari-mutuel wagerers operate within an efficient information market (i.e., all participants are privy to the same information pool). Economists employ the "efficient capital markets" hypothesis to classify capital (stock) markets according to the degree their prices fully reflect the available
information (for a major economic review, see Fama, 1970). They identify three points on the continuum of information/price efficiency: strong form, semi-strong form, and weak form. A market is said to be strongly efficient (or equitable) if the expected value of a commodity is calculable from all the available information, and this value responds instantaneously to new information. In other words, a system is considered strongly efficient if all prior information can be used to generate an objective probability distribution for some future event (cf., decision-making under risk). At this level of efficiency no individual can aspire to produce substantial long-term profits beyond the expected norms. A market is said to be semi-strongly efficient if a subset of pertinent information is only available to a small favoured group, and the effect of new information on prices requires short periods of time to adjust. At this level, prices will not fully reflect all the information and a few individuals will be capable of producing expectations beyond the norm (cf., decision-making under uncertainty). A weakly efficient market is similar to the semi-strong form with the exception that the former requires considerable more time for the prices to fully adjust to the import of new information. In this market "insiders" are capable of producing abnormally high gains in comparison to those who use actuarial data.

Although the methodological techniques used by economists to test the efficient markets hypothesis are beyond the scope of this presentation, most can be characterized by a "fair game model" (Fama, 1970, pp. 384-385). According to this model, unexpected returns (UR) for any period, equal the difference between the actual returns (AR) for the same period and the expected returns (ER) for the same period given
the total pool of available information (AI) at the beginning of the period (from Basu, 1977, p. 8). Expressed algebraically:

$$ UR = [ AR - ( ER | AI ) ] $$

[6]

Using the relationship in [6], a market would be classified as strongly efficient if the expectation of the unexpected returns, given all the available information, equalled zero:

$$ \text{EXPECTATION} ( UR | AI ) = 0 $$

[7]

Conversely, a market would be classified as weakly efficient if only historical data (but not "insider" information) produced a constant expectation.

Race-track wisdom has it that the pari-mutuel wagering market operates under conditions of weak efficiency. Since race horses do not normally compete more than once a week, the public must base its decisions predominantly on actuarial data which is at best several days old. However, horsemen (e.g., trainers, jockeys, owners, track veterinarians) would be privy to the most recent information concerning the conditioning and health for each of the competitors. Assuming a "normal" love for money, it would not be unrealistic to propose that these individuals participate in the pari-mutuel betting market (see Kusyshyn, 1979, pp. 64-65). This participation would, in turn, influence the character of the pari-mutuel betting market towards weak efficiency. Consequently, the reported correspondence between public and observed probability distributions may be even higher than reported, if one allows for the impact of "insider" information. However, empirical support for weak efficiency is both limited and inconclusive.
Dowie (1976) collected data on approximately 29,000 thoroughbred races run at numerous British race-tracks during 1973. Unlike their American counterparts, British horse bettors eschew on-course pari-mutuel wagering in favor of legalized bookmaking services. British bookmakers offer bettors "starting prices" (SP) at specific, but changing, odds. Returns on SP for successful bettors are unaffected by subsequent SP changes, although bookmakers protect themselves from preceding SP deviations by adjusting subsequent estimates accordingly. Dowie's data base included those SP "at which a 'sizable' bet could have been made on the course just before the off" (p. 140). In addition, his data base included a corresponding set of forecast prices (FP) which are estimates produced by daily racing papers based on the early activities of the major and minor SP markets (cf., Snyder, 1978). The primary focus of Dowie's study was to test the efficiency of the betting market. He compared the objective win frequencies within each odds category for both SP and FP and reported correlations within each market price at .91 and .88, respectively. (Both markets demonstrated the under/over estimation bias discussed earlier.) For unexplained reasons, Dowie considered the SP to incorporate superior or inside information, while the FP were considered to reflect historical or actuarial information. Consequently, he interpreted the non-significant differences between these two correlations as evidence for a strongly efficient market in horse-race wagering.

Fabricand (1975) pointed out that the burden of proof rests with Dowie to show that the information base of the SP was in fact substantially different from that of FP. In his own studies, (reported in Fabricand, 1979, pp. 69-150), he offers evidence in support of weak
efficiency in the pari-mutuel market. His approach was based on three considerations. First, the favorite in any race is potentially the best available wager since the public is most likely to underestimate its true win potential. Second, the high win percentage for favorites (approximately 35%) would buffer investors from financial ruin by reducing the possibility of long series of losses to a minimum. And third, the betting public is susceptible to Fabricand's "Principle of Maximum Confusion." According to this principle "the public is most likely to err in determining the winning probabilities of the favorite in those races wherein the historical record of the favorite is surprisingly similar or inferior to that of one or more other horses in the race...('Consequently') we assume that many bettors, unaware of hidden factors, opt for the bigger payoff available on longer odds horses" (p. 72). Fabricand's test of market efficiency consisted of empirically recalculating the expectations for favorites after competing horses with similar or superior historical records to the favorite were assigned zero probabilities of winning. This procedure was performed on two separate samples of 593 and 649 thoroughbred races run in the period 1974-1976 at various American race-tracks. The recalculated expectations for the favorites in both samples were then partitioned into 10 intervals and each was compared to its original expectation in light of the observed win frequencies. The results showed that for each sample, significant profits ranging from 11% to 95% could be realized by betting on favorites with recalculated expectations beyond .42. Fabricand concluded that since these favorites displayed no obvious superiority over other competitors their popularity in the betting pool could only be attributed to exogenous information not available to the general public.
Although Fabricand's results lend support to weak information efficiency in horse racing, this author believes that the validity of his findings hinge on the methods he used to define "similarity in historical records." Fabricand employed a subjectively derived point method (i.e., a normative linear model) which he assumed would identify those factors that appear alike to the wagering public (i.e., a subjective paramorphic representation). In light of our discussion on the limitations of man's ability to accurately process large sets of information, it would seem that a statistically derived model of the belief systems of the wagering public would produce superior results.

It is this author's contention that the studies by Dowie and Fabricand were limited by their inability to operationalize and covary the impact of historical information on the behaviors of the betting public. To this extent, the use of a statistically derived paramorphic representation would appear to be a more rigorous avenue to test the efficiency of the pari-mutuel market. This study will use the previously proposed paramorphic representation as an index of the publically available information pool in the pari-mutuel market. Consequently, the presence of major deviations from the derived paramorphic probabilities in the positive direction (i.e., unexplained underlays) will be used to identify those races where exogenous information was present. It was hypothesized that in those races which do not conform to the fair game model, the largest positive errors of prediction will produce significantly positive expectations.
Summary of the Problem

The major thrust of the present study will be to exploit the full capabilities of the general linear model to extend our knowledge of behavioral decision-making at the race-track. In light of the literature review, this study has three possible areas of contribution. First, it will be used to further demonstrate the close correspondence between dollar odds and observed probability frequencies within the context of an "under/over" bettor bias. Second, it will use correlational analysis to identify the relevant historical variables that contribute to the decision-making policies of pari-mutuel wagerers in producing dollar odds. Finally, it will employ linear regression procedures on these salient predictors to produce a paramorphic representation of the public's probability estimates. This representation will serve a twofold purpose. By examining the residuals within each race, the paramorphic representation will be used to classify the efficiency of the information pool within each racing event as either strong (fair) or weak (unfair). In the former, it is hypothesized that bootstrapping the paramorphic-representation will produce expectations significantly greater than those imposed by the pari-mutuel system (i.e., expectations greater than negative take). In the latter case, it is hypothesized that when the paramorphic probabilities are treated as a covariate, that the errors of prediction, and not the predictions themselves, will produce substantial positive expectations.
CHAPTER II
METHOD

Data Base

The analyses were performed on a data base commissioned by International Gaming Incorporated and constructed by this author. The data base consisted of detailed information on 1,994 pacing horses\(^3\) that competed in 240 harness races at various competitive levels during the 1977 racing season on the Ontario Jockey Club circuit (Mohawk Raceway, Campbellville; and Greenwood Raceway, Toronto). The commission at these two major Canadian race-tracks for this period was 17.1% and breakage was taken to the first nickel. The raw data were taken from: (a) historical records published in the track program; and (b) official results published in race charts. The former consisted of 154 codable variables for each horse that were publicly available prior to the running of a race. The bulk of these included: performance details for each horse's last six races; his win-place-show record and gross earnings in each of the last two years; eligibility conditions; and various situational factors such as driver ability, post position and track condition. The data also included two codable variables that became available after the running of a race: finish position and closing dollar odds. Appendix A provides an example of the raw data and descriptions for each variable. The raw data were transcribed onto coding sheets, keypunched onto cards, verified, and transferred as card images onto magnetic tape.

\(^3\) Since horses compete on a regular basis, a large number of the 1,994 sample points (approximately 40%) represented repeat appearances of the same horses at different time lags.
Preliminary Considerations

Before the sequence of analyses is presented, the reader must first be familiarized with several unique features and problems associated with the scaling of pari-mutuel data.

Throughout the preceding text, the author has used the terms dollar odds (DO) and estimated probability of winning (EPW) interchangeably. Although each can be derived from the other, using equations [2] and [3], this should not suggest that one is a linear transform of the other, since the proportion of variance that each of these variables produces is not equal across all regressors. As a result, the counted fraction MB/MB (i.e., EPW) was used as the dependent measure throughout the analyses, even though it may be referred to as dollar odds in the text.

The use of EPW, however, posed an additional problem. Proportions, or probabilities such as EPW, are not highly amenable to the linear assumptions of least squares analysis. This is especially true if they are used as dependent measures. Samples of proportions do not produce constant variance; they often violate the assumption that the correlation between the numerator and denominator is linear; and their unit of measurement is rarely constant over the full range of scale values (see Cohen & Cohen, 1975, pp. 254-459). Distributions of proportions are thus more likely to be non-normal and thereby violate the assumptions of fixed sampling theory in linear analysis. The reader will recall that Griffith (1949) reported that the resulting distribution of the public probabilities was positively skewed. In order to linearize the EPW distribution, the author employed a logit
transformation. This nonlinear transformation equals half the natural logarithm (base e) of the ratio, of a proportion to its compliment:

\[ \text{Logit EPW} = \{ \frac{1}{2} \log_e \left( \frac{\text{EPW}}{1 - \text{EPW}} \right) \} \]

Since the \( \frac{1}{2} \) was superfluous for least squares estimation, it was ignored. Although the construction of the paramorphic representation was based on the logit of EPW, the predicted values were rescaled back to their original scale values (i.e., proportions) by taking the antilog of the inverse of the predicted Logit EPW plus unity:

\[ \text{EPW}_i' = \text{antilog}_e \left[ \frac{1}{\text{Logit EPW}_i + 1} \right] \]

The predicted \( \text{EPW}_i' \) were subsequently proportionalized within each race which allowed the author to directly compare and evaluate any \( \text{EPW}_i \) with its paramorphic representation (i.e., \( \text{EPW}_i' \)) on the same scale.

The independent (i.e., historical) variables also posed methodological problems in scaling. Because of the ipsative nature of dollar odds (i.e., \( \sum \text{EPW}_i = 1 \)), the race was treated as the primary unit, or block, in the analyses. With the exception of certain qualitative variables such as post position or gender of horse, the majority of the independent variables were adjusted or normalized to enable meaningful comparisons within each race block. Normalization was required in the present study since the date base sampled various quantitative and qualitative levels of competition. This approach was based on the assumption that a rational wagerer would be more concerned with the relative merits of the competitors within a race, than with their absolute status along some underlying dimension of equine speed potential.
Because of the exploratory nature of this study, the identification of an appropriate technique to normalize the attributes in the 1,994 sample points (horses) with reference to the 240 blocks (races), was eclectically based. One method of normalization was to rank each attribute within each block. Economists have long proposed that ordinal representation is a minimum requirement of rational man (see footnote 2). A second normalization method was to standardize each attribute by converting it to a z-score based on the mean attribute within each block. Third, each variable was rescaled to a distance score representing its deviation from the maximum (and minimum) score in each block. Another method converted each attribute to a proportion representing the ratio of the variable score to the sum of the attribute scores within that race. A logit transformation was also employed on these proportion scores.

In addition to the preceding methods, the author used dummy variable coding to "nominalize" variables whose interval scale properties were suspect. Although nominalization is typically used to produce quantitative distinctions with an n level qualitative variable, it may also be used to rescale the n points along an "interval" continuum (see Cohen, 1968; Cohen & Cohen, 1975, pp. 171-211, and pp. 241-242). Nominalization is achieved through dummy variable representation which partitions observations into n-1 orthogonal groups corresponding to the first n-1 interval points along a numbered scale. A criterion is then regressed onto these n-1 independent variables and the raw score partial regression coefficients from this analysis can then be used to replace the original values for the n-1 interval points. The final interval serves as a reference point and its coefficient value
is set at zero. Subsequently, a Duncan Multiple Range test can be used to indicate if the \( n-1 \) intervals are significantly different, and can then be used to collapse across nonsignificant levels.

In the present study, the author used nominalization to evaluate the scale properties of rank scores. For example, if a ranked attribute within a nine-horse race produces a significant correlation with Logit EPW, the resulting nine rank values for the total sample points may not be efficiently represented as falling along nine equally spaced intervals. By nominalizing the first eight ranks and then inspecting the intervals given by the Duncan Multiple Range values, the results could suggest that the public responds differentially to the first three ranks, while the remaining six ranks are not distinguishable. Thus, the relation of the ranked attribute to dollar odds would be more efficiently represented by four rather than nine interval values. Nominalization was also employed on qualitative variables such as ordinal race positions, track conditions, post positions, and the like.

**Analytic Procedures**

The Statistical Analysis System (Helwig & Council, 1979) was used to analyze the data in a four-stage sequence. First, the representativeness of the sample odds was determined by calculating expectations for the ranked odds and determining if the under/over bettor bias was present. Second, each of the 154 historical variables and their transforms were correlated with Logit EPW. The correlational structure was examined to see which normalization techniques were most appropriate to characterize the variance of dollar odds. Third, the identified predictors of Logit EPW were used in a first-order stepwise
linear regression procedure. This least squares technique balance the needs for parsimony and comprehensiveness in the model building process by selecting a manageable set of independent variables that produces a satisfactory multiple correlation with the criterion (see Draper & Smith, 1966). Because of the large number of degrees of freedom in the data base, a .01 "entry" criterion and a .0001 "stay" criterion were specified. In the former, only variables producing partial F statistics beyond the .01 level were selected from the predictor pool; while in the latter, only variables producing partial F statistics beyond the .0001 level were retained in the model. The procedure terminates when the significance level of the highest remaining partial correlation in the predictor pool exceeds the entry criterion. On completion of the stepwise screening, the sums of the weighted predictors were rescaled by the previously mentioned antilog transformation, and the resulting values were proportionalized within each block to preserve the ipsative nature of the win probabilities. These values were deemed as paramorphic representations of the betting public's judgments.

In the fourth stage, difference scores between predicted and actual probabilities were calculated and ranked in descending order within each race. Each race was then classified as either "fair" or "unfair" according to the magnitude of the largest positive deviation. Fair races were defined as those in which the largest positive discrepancy did not exceed five percentage points, thus reflecting races where the evidence for exogenous information was at a minimum. Following the rationale presented in the previous chapter, the paramorphic bootstrapping effect was tested on this subsample. This was accomplished in a twofold fashion. First, correlations of dollar odds
and their paramorphic representations with the finish criterion (FIN) were calculated. As was previously mentioned, the bootstrapping effect would be evident if the correlation between the paramorphic probabilities and finish position was significantly greater than the latter's correlation with dollar odds. As a second test for the bootstrapping effect, the paramorphic probabilities and dollar odds were ranked in descending order within each race and hypothetical unit wagers were placed on both sets of ranks. Evidence for the bootstrapping effect would consist of a decreasing monotonic function between the rank of the paramorphic expectation and the empirical expectation, with the lower paramorphic ranks producing expectations greater than negative take.

Turning to the subsample classified "unfair," the presence of information inefficiency was tested by ranking (in descending order) the paramorphic deviations within each race and placing hypothetical unit wagers on each rank. Monotonically decreasing profits according to rank were considered as validating evidence for the presence of weak information efficiency in this subsample.

In summary, bivariate and multiple correlation analysis were used to: (1) identify the major determinants of dollar odds; (2) construct and evaluate a paramorphic linear model of the wagering public's decision-making processes; (3) test for the presence of the bootstrapping effect, and (4) examine the degree of information efficiency in the pari-mutuel setting.
CHAPTER III

RESULTS

Prior to the development of the paramorphic linear model, the representativeness of the sample odds were examined. The 1,994 sample points of dollar odds were ranked within each of the 240 races and mean orders of finish were calculated for each rank. Table 2 shows a strong positive monotonic function between the judged popularity of a competitor and its subsequent performance. This relationship was supported by a significant first order correlation of .42 ($p < .0001$) between FIN and Logit EPW. A Duncan Multiple Range test revealed that the Logit EPW means at each FIN level were all significantly different from each other ($p < .0001$).

The results presented in Table 3 confirm the presence of an under/over bettor bias in the sample. Unit wagers on short odds horses yielded a high percentage of successful bets that produced small but positive rates of return (corrected for the 17.1% track take). Horses with increasingly higher odds produced lower success rates and moderately severe negative expectations. The apparently high positive return rate for horses ranked seventh in odds can be attributed to six winners in this category whose odds were abnormally high, ranging from 30:1 to 60:1. (As a consequence, the author's interpretations of subsequent expectation patterns will be tempered by this uncharacteristic deviation.) Nevertheless, the overall pattern of expectations reconfirms the findings of Griffith (1949), Hoefl and Fallin (1974), and others, that despite the public's adroit handicapping skills there exists a tendency to underestimate high probability events.
### Table 2

Mean Order of Finish Position as a Function of Odds Ranked within Race

<table>
<thead>
<tr>
<th>Entries Per Race</th>
<th>Races</th>
<th>Odds-Rank 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>6</td>
<td>2.6</td>
<td>2.7</td>
<td>2.8</td>
<td>4.0</td>
<td>3.8</td>
<td>5.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>38</td>
<td>2.3</td>
<td>3.5</td>
<td>3.6</td>
<td>4.1</td>
<td>4.2</td>
<td>4.8</td>
<td>5.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>72</td>
<td>3.2</td>
<td>3.4</td>
<td>4.1</td>
<td>4.6</td>
<td>5.1</td>
<td>4.4</td>
<td>5.2</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>124</td>
<td>3.1</td>
<td>3.6</td>
<td>4.4</td>
<td>4.7</td>
<td>5.1</td>
<td>5.7</td>
<td>5.0</td>
<td>6.5</td>
<td>6.6</td>
</tr>
</tbody>
</table>
### Table 3
Expectations of Unit Wagers as a Function of Odds Ranked within Race

<table>
<thead>
<tr>
<th>Odds-Rank</th>
<th>Races</th>
<th>Winners (%)</th>
<th>Winnings</th>
<th>Expectation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>240</td>
<td>89 (37)</td>
<td>226.65</td>
<td>.12</td>
</tr>
<tr>
<td>2</td>
<td>240</td>
<td>55 (23)</td>
<td>219.45</td>
<td>.09</td>
</tr>
<tr>
<td>3</td>
<td>240</td>
<td>25 (10)</td>
<td>128.10</td>
<td>-.30</td>
</tr>
<tr>
<td>4</td>
<td>240</td>
<td>25 (10)</td>
<td>190.25</td>
<td>-.04</td>
</tr>
<tr>
<td>5</td>
<td>240</td>
<td>14 (6)</td>
<td>147.15</td>
<td>-.22</td>
</tr>
<tr>
<td>6</td>
<td>240</td>
<td>11 (5)</td>
<td>169.45</td>
<td>-.12</td>
</tr>
<tr>
<td>7</td>
<td>234</td>
<td>19 (8)</td>
<td>387.40</td>
<td>.80</td>
</tr>
<tr>
<td>8</td>
<td>196</td>
<td>1 (.5)</td>
<td>29.50</td>
<td>-.68</td>
</tr>
<tr>
<td>9</td>
<td>124</td>
<td>1 (.5)</td>
<td>89.90</td>
<td>-.11</td>
</tr>
</tbody>
</table>

**Note:** Expectations (but not winnings) were corrected for the 17.1% track take.
and overestimate low probability events. In light of these results, the data base was judged to be comparable with the pari-mutuel data bases reported in the literature review.

Earlier it was pointed out that distributions of proportions are typically non-normal and their use as dependent measures seriously violate the assumptions of fixed sampling theory in linear analysis. In anticipation of a strong positively skewed EPW distribution, the author proposed a logit transformation to achieve normalization. Consequently, the EPW and Logit EPW distributions were examined prior to the construction of the paramorphic representation. The reader can readily see in Figure 1 that the EPW distribution was in fact positively skewed and that the logit transformation produced a more than satisfactory linearization. The legitimacy of using this nonlinear transformation was strengthened by the fact that for virtually every one of the 154 independent variables (and their transformations) the correlations were higher with Logit EPW than with EPW.

Development of a Paramorphic Representation

An inspection of the first order criterion correlations revealed several important patterns which influenced the construction of the paramorphic representation. First, in the majority of cases, the normalization of predictors within race blocks produced consistently superior criterion correlations. In particular, ranked or proportionalized attributes accounted for larger portions of variance in Logit EPW than their unnormalized counterparts. Consequently, the majority of variables used in the model building process were first transformed into either ranks or proportion scores.
Figure 1. Frequency distributions of EPW and Logit EPW.
Second, variables representing repeated points in a time sequence showed a time-dependent lag structure such that more recent information correlated higher with the criterion than previous information. Specifically, information from a horse's last two races was weighted more heavily by the public than information provided from the remaining four races. For example, the criterion correlations with the last six finish positions were: - .34, - .22, - .11, - .09, - .05, - .04, respectively. A similar time dependent pattern was observed when the predictor pool was correlated with the FIN criterion. As a result of these patterns, variables beyond the second time lag were not included in the initial predictor pool.

Finally, the correlation pattern indicated the presence of "classical" suppression in those variables related to a horse's speed performance. Classical suppression is evident when one or more exogenous variables significantly correlate with a predictor variable while correlating near zero with the criterion. These exogenous variables can be considered covariates of the predictor which attenuate the relationship between the independent and dependent variables. Consequently, the shared variance between the predictor and the covariate(s) must be partialled. Two variables -- track condition and track speed -- were identified as classical suppressors of a horse's race times. Track conditions reflect the degree to which the racing surface has been impeded by weather conditions, while track speeds (published in the racing form) provide comparative speed ratings for all major race-tracks. Both suppressors accounted for a combined 25% of the variance in final times, yet they accounted for less than 1% of the criterion variance. First order criterion correlations with the speed
predictors (prior to ranking) showed significant increase from -.33 to -.44 in the last race \((p < .0001)\), and from -.29 to -.38 in the second-last race \((p < .001)\) after partialling both suppressors. Thus, race speeds were adjusted for track condition and track speed, and then ranked prior to being entered into the least-squares solution.

In addition to the preceding considerations, the author was guided by the criterion correlations in constructing several rate scores and combinatorial scores that were included in the initial variable pool. Although descriptions of these variables are too numerous to mention, most were tempered by the author's experience and readings in the art of handicapping harness races. In sum, the pre-selected predictor pool consisted of 68 variables which were entered in a stepwise regression procedure.

The Paramorphic Representation

The summary results of the final stage in the stepwise regression procedure are presented in Table 4. The paramorphic model accounted for 65% of the variance in Logit EPW and was significant at the .0001 level. Furthermore, each of the twelve selected predictors achieved significance at the .0001 level. Subsequent first order correlations for each of these predictors with the FIN criterion showed that in no case did the coefficient signs reverse from those produced by the EPW criterion. This pattern is presented in Table 5 and it provides external support for the validity of the final predictor set by demonstrating that the components of dollar odds were congruent with subsequent empirical performance. Descriptions for each of these twelve selected predictors are as follows:
Table 4
Summary of Final Regression Model Selected by the Stepwise Selection Procedure

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1522</td>
<td>12</td>
<td>126.83</td>
<td>309*</td>
<td>.651</td>
</tr>
<tr>
<td>Speed1</td>
<td>235</td>
<td>1</td>
<td>1</td>
<td>571*</td>
<td></td>
</tr>
<tr>
<td>Speed2</td>
<td>56</td>
<td>1</td>
<td>1</td>
<td>137*</td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>91</td>
<td>1</td>
<td>1</td>
<td>220*</td>
<td></td>
</tr>
<tr>
<td>Odds1</td>
<td>89</td>
<td>1</td>
<td>1</td>
<td>215*</td>
<td></td>
</tr>
<tr>
<td>Odds2</td>
<td>24</td>
<td>1</td>
<td>1</td>
<td>59*</td>
<td></td>
</tr>
<tr>
<td>Driver</td>
<td>84</td>
<td>1</td>
<td>1</td>
<td>203*</td>
<td></td>
</tr>
<tr>
<td>Driver1</td>
<td>14</td>
<td>1</td>
<td>1</td>
<td>34*</td>
<td></td>
</tr>
<tr>
<td>Earnings77</td>
<td>43</td>
<td>1</td>
<td>1</td>
<td>104*</td>
<td></td>
</tr>
<tr>
<td>Earnings76</td>
<td>18</td>
<td>1</td>
<td>1</td>
<td>44*</td>
<td></td>
</tr>
<tr>
<td>Kick1</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>31*</td>
<td></td>
</tr>
<tr>
<td>Class1</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>21*</td>
<td></td>
</tr>
<tr>
<td>Class2</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>16*</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>814</td>
<td></td>
<td>1981</td>
<td></td>
<td>.41</td>
</tr>
<tr>
<td>Total</td>
<td>2336</td>
<td></td>
<td>1993</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .0001.
Table 5
First Order Correlations for the Twelve Selected Predictors
with the Logit EPW Criterion and the FIN Criterion

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Logit EPW</th>
<th>FIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed1</td>
<td>-.50***</td>
<td>.22***</td>
</tr>
<tr>
<td>Speed2</td>
<td>-.46***</td>
<td>.19***</td>
</tr>
<tr>
<td>Post</td>
<td>.22***</td>
<td>-.21***</td>
</tr>
<tr>
<td>Odds1</td>
<td>.48***</td>
<td>-.20***</td>
</tr>
<tr>
<td>Odds2</td>
<td>.33***</td>
<td>-.13**</td>
</tr>
<tr>
<td>Driver</td>
<td>.35***</td>
<td>-.16***</td>
</tr>
<tr>
<td>Driver1</td>
<td>.28***</td>
<td>-.11**</td>
</tr>
<tr>
<td>Earnings77</td>
<td>.47***</td>
<td>-.23***</td>
</tr>
<tr>
<td>Earnings76</td>
<td>.18***</td>
<td>-.10**</td>
</tr>
<tr>
<td>Kick1</td>
<td>-.35***</td>
<td>.18***</td>
</tr>
<tr>
<td>Class1</td>
<td>-.22***</td>
<td>.11**</td>
</tr>
<tr>
<td>Class2</td>
<td>-.18***</td>
<td>.08**</td>
</tr>
</tbody>
</table>

** p < .001
*** p < .0001
1. **Speed1.** This variable refers to the rank of a horse's mile speed in his last race after adjusting for track condition and track speed. A Duncan Multiple Range test showed that criterion means at each rank level were all significantly different from each other ($p < .0001$). This suggests that all rank levels are salient to the betting public, and that the scale properties of these ranks do not deviate from linearity.

2. **Speed2.** This predictor refers to the adjusted rank of a horse's speed in his second last race. As was the case with Speed1, the Duncan Multiple Range test indicated that all rank levels were significantly different from each other.

3. **Post.** This variable is related to the post position a horse is assigned in a race. As is described in Appendix A, a random draw determines how far from the inside rail a competitor starts a race. Preliminary analysis showed that treating this predictor as a qualitative variable (i.e., through nominalization) produced marginal increments in the criterion correlation in comparison to using scale values representing ordinal displacement from the rail position. After much experimentation, it was noted that a superior criterion correlation could be obtained by substituting for each positional value the proportion of winners that each post position produces. These rate scores were provided by International Gaming Inc., and represented ten-year success rates for each post position on the Ontario Jockey Club circuit. The following are the proportion values used in this study according their respective ascending post position: .145, .135, .125, .125, .155, .110, .075, .055, and .075.
A. Odds1. This predictor is related to a horse's dollar odds in his last race. Preliminary manipulation of this variable strongly suggested a two-stage transformation in order to maximize criterion correlation. First, dollar odds were converted to proportions by calculating the ratio of one minus track take, to the dollar odds plus unity (from equations 2 and 3). In the second step, a logit transformation was applied to these proportions.

5. Odds2. This variable is related to a horse's dollar odds in his second last race. An identical transformation sequence to that of Odds1 was employed for Odds2.

6. Driver. This rate score was constructed to measure the competence level of the driver. It consisted of the ratio of the driver's previous year's win record to his total number of races in that year. Preliminary analysis showed that a logit transformation did not improve the criterion fit.

7. Driver1. This predictor measures the competence level of the driver in a horse's last race. In harness racing it is not uncommon for the trainer to change drivers from one race to another. Often a trainer will request the services of a highly competent driver when he believes his horse has an above average chance of winning a particular race. Consequently, it is assumed that the betting public is cognizant of significant changes in driver ability from one race to another. Support for this assumption comes from the fact that the signs of the parameter estimates for Driver and Driver1 were in opposite directions. This suggests that the public uses some form of difference score to evaluate driver changes when assigning probability estimates (see Cohen & Cohen, 1975, p. 67).
8. **Earnings77.** This predictor is related to a horse's gross earnings in the current year. Preliminary manipulations suggested two adjustments in order to maximize the criterion correlation. Gross earnings were first divided by the number of pari-mutuel starts, and this rate score was then proportionalized within each race. This procedure enabled the predictor variance to be realigned across and within the range of qualitative and quantitative differences among race blocks. A logit transformation on this predictor did not significantly improve the criterion correlation.

9. **Earnings76.** This predictor is related to a horse's gross earnings in the previous year and it was modified in the same fashion as Earnings77.

10. **Kick1.** In race-track parlance the term "kick" refers to the vigor a horse displays in the final portion of a race. This concept was numerically represented by the number of lengths a horse gains or loses from the top of the stretch to the race finish (see Appendix A). This combinatorial variable, representing a horse's vigor in his last race, provided a better criterion fit than by using the horse's most recent finish position.

11. **Class1.** This variable represents the number of discrete purse levels a horse is moving from its last race to the current race. Generally, horses that perform well in their last race move up the competitive ladder; on the other hand, horses that perform poorly move down the competitive ladder. In claiming races, the competitive levels are determined by the trainer or owner; while in conditioned races, the track's racing secretary specifies the conditions for class movement.
12. **Class2.** This variable is identical to Class1, except it refers to the number of purse levels a horse has moved from its next-to-last race to the current race.

Following the construction of the paramorphic representation, the weighted sums of the predictors were rescaled into proportions using the antilog transformation presented on page 39. Figure 2 shows the effect of this procedure. The reader can readily see that the normally distributed predicted scores (Logit EPW) were effectively rescaled into paramorphic proportions (EPW') such that their distribution properties matched the positively skewed shape of EPW as shown on page 49.

**The Bootstrapping/Fair Game Issue**

Subsequent to the proportionalization of the predicted scores, each race block was classified as either "fair" or "unfair" according to the maximum positive deviation (DPW) between the public probabilities (EPW) and their corresponding paramorphic probabilities (EPW'). Earlier it was suggested that a DPW score greater than .05 would reflect the presence of exogenous information in those races where the assumption of strong information efficiency in the fair game model is violated. Preliminary analysis showed that maximum results could be obtained if the classification criterion for DPW was set at .07. Accordingly, this criterion classified 106 races as "fair" and 134 races as "unfair."

Table 6 shows the intercorrelations of the major constructs for the total sample, and the two subsamples. The reader will recall that the efficacy of paramorphic bootstrapping would be supported if the correlation between EPW' and FIN exceeded the correlation between EPW and FIN. As Table 6 shows, this pattern was not observed in the total sample space which suggests that the linear model was not sufficiently
Figure 2. Frequency distributions of Logit EPW and EW'.
Table 6
Intercorrelations Among Major Constructs for
Total, Fair, and Unfair Samples

<table>
<thead>
<tr>
<th></th>
<th>Total sample</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EPW'</td>
<td>DPW</td>
<td>FIN</td>
</tr>
<tr>
<td></td>
<td>-.81*</td>
<td>-.54*</td>
<td>-.40*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-.05*</td>
<td>-.34*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.19*</td>
</tr>
<tr>
<td></td>
<td>EPW</td>
<td>EPW'</td>
<td>DPW</td>
</tr>
<tr>
<td></td>
<td>Fair sample</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EPW'</td>
<td>DPW</td>
<td>FIN</td>
</tr>
<tr>
<td></td>
<td>-.87*</td>
<td>-.55*</td>
<td>-.38*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-.06*</td>
<td>-.41*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.08*</td>
</tr>
<tr>
<td></td>
<td>EPW</td>
<td>EPW'</td>
<td>DPW</td>
</tr>
<tr>
<td></td>
<td>Unfair sample</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EPW'</td>
<td>DPW</td>
<td>FIN</td>
</tr>
<tr>
<td></td>
<td>-.75*</td>
<td>-.55*</td>
<td>-.41*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-.14*</td>
<td>-.27*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.27*</td>
</tr>
<tr>
<td></td>
<td>EPW</td>
<td>EPW'</td>
<td>DPW</td>
</tr>
</tbody>
</table>

*p < .01
**p < .001
***p < .0001
accurate to capture the expertise of the betting public. However, an inspection of the correlation structure within the fair subsample reveals that the $EPW'$ correlation with FIN did exceed the correlation between EPW and FIN although the difference was not significant. Conversely, the $EPW'$ correlation with FIN in the unfair subsample was significantly lower ($p < .001$) than the correlation between EPW and FIN. The reader should also note that the errors of prediction (DPW) in the unfair sample produced a correlation with FIN equal to that of the linear model. Although these correlations offer preliminary support for the partitioning of the sample space, it could be argued that they are simply consequences of the classification criterion. However, the reader is reminded that the races were partitioned according to the discrepancies produced by horses and not races. Thus, if the linear fit between an actual and predicted probability in a nine-horse race was quite close for eight of the entries and one horse's discrepancy exceeded the .07 tolerance of error, that race would still be classified as unfair. This would explain why the FIN correlations with the public probabilities were not significantly different between the two subsamples. For this reason, it was difficult to use correlational patterns to evaluate the role of paramorphic representation in the bootstrapping/fair game distinction.

Table 7 provides a sharper focus on the bootstrapping/fair game issue by showing the expectations and win percents at each rank level for EPW, $EPW'$, and DPW, in the total sample, and the two subsamples. In the total sample, a comparison between corresponding EPW and $EPW'$ entries suggests that the latter does not provide higher expectations or rates of return than the former. Specifically, the win percents for the
Table 7
Expectations and (Win Percents) for EPW, EPW', and DPW
in Total, Fair, and Unfair Samples

<table>
<thead>
<tr>
<th>Rank</th>
<th>Total sample</th>
<th>Fair sample</th>
<th>Unfair sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EPW</td>
<td>EPW'</td>
<td>DPW</td>
</tr>
<tr>
<td>1</td>
<td>0.12 (37)</td>
<td>-0.13 (25)</td>
<td>0.33 (37)</td>
</tr>
<tr>
<td>2</td>
<td>0.09 (23)</td>
<td>0.12 (22)</td>
<td>0.40 (11)</td>
</tr>
<tr>
<td>3</td>
<td>-0.30 (10)</td>
<td>-0.04 (15)</td>
<td>0.04 (12)</td>
</tr>
<tr>
<td>4</td>
<td>-0.04 (10)</td>
<td>0.01 (10)</td>
<td>-0.46 (5)</td>
</tr>
<tr>
<td>5</td>
<td>-0.22 (6)</td>
<td>-0.01 (10)</td>
<td>-0.31 (3)</td>
</tr>
<tr>
<td>6</td>
<td>-0.12 (5)</td>
<td>0.29 (8)</td>
<td>0.56 (11)</td>
</tr>
<tr>
<td>7</td>
<td>0.80 (8)</td>
<td>-0.01 (5)</td>
<td>-0.34 (7)</td>
</tr>
<tr>
<td>8</td>
<td>-0.68 (5)</td>
<td>-0.10 (2)</td>
<td>0.06 (9)</td>
</tr>
<tr>
<td>9</td>
<td>-0.11 (5)</td>
<td>-0.67 (1)</td>
<td>0.49 (12)</td>
</tr>
</tbody>
</table>

Note: Expectations were corrected for the 17.1% track take.
first two rank levels of EPW' (25% and 23%) indicate an apparent inability of the paramorphic probabilities to outperform the success rates of the public probabilities (37% and 23%). Furthermore, the puzzling negative expectation of -.13 for the paramorphic favorite seems to suggest that bootstrapping a linear model of the public's betting behavior is at best a questionable procedure. The key to understanding these negative results lies in the pattern of expectations and win percents for the DPW. If one assumes, for the moment, that the paramorphic probabilities do in fact reflect ideal public estimations, then the high positive expectations and success rates at both ends of the DPW scale would suggest that both underlays and overlays play significant roles in determining the character of pari-mutuel expectations. Thus, the high expectation of .33 and high success rate of 37% for the largest DPW in any race suggest that a rational wagerer should seek out and wager on those horses that are significantly overbet (i.e., underlays). Conversely, it may be argued that a rational wagerer should follow the example of the professional gambler, and wager on those horses that appear to be severe overlays as indicated by the high return rates and moderate win percents for entries ranked eighth and ninth on the DPW scale. On the surface, the mapping of this set of options appears to be a non-nullifiable statement, akin to stating that the sum of mutually exclusive and exhaustive probabilities will equal unity. Such a statement provides a complete description of the sample space, yet it provides no explanatory or predictive power. However, by using the .07 classification rule (which is nullifiable) to redefine the sample space into two distinct subsamples, it is possible to specify under what conditions a rational wagerer should attend to overlays or underlays.
A thorough inspection of the recalculated expectations and success rates in the two subsamples presented in Table 7 would suggest the preliminary conclusion that overlays produce substantial profits in those races that conform to the fair game model, while underlays produce equally attractive profits in those races where exogenous information is present. In the subsample deemed "fair," the reader will notice that although the success rates for the public's endorsements are still positively skewed, the expectations are all near zero or negative. 4

This pattern suggests that the public's estimations of the win potential for each contestant are so keenly accurate that no substantial increments in profit can be realized. Not only does this pattern lend support to the validity of the .07 classification rule, but it also supports the conclusion that, for these races, the public generates probability estimates under the condition of risk (i.e., the majority of the PW' estimates are veridical to PW). However, the most striking feature in the fair subsample is the pattern of expectations and success rates for the paramorphic probabilities (EPW'). It is evident that superior rates of return can be obtained if one utilizes a linear weighting of the public's decision-making policies. This conclusion is supported by the fact that over two-thirds of the winners in this subsample were paramorphic favorites or second favorites, and that the mean expectation for both rank levels exceeded .38. Conversely, the pattern for the remaining categories consisted of low win percents, and expectations that were near zero or negative. Clearly, the EPW' pattern

---

4 The reader is reminded that all high positive expectations in the sixth and seventh rank positions have been deemed to be statistical artifacts of the data base.
in the fair subsample supports one of the major hypotheses in this study -- that a statistically derived model of decision-makers will outperform the decision-makers' judgments. Complimentary support for this conclusion can be drawn from the monotonically increasing expectations produced by the DPW index. The relatively high expectations and success rates at the eighth and ninth rank levels of DPW indicate that many of the paramorphic favorites that produced large positive expectations on the EPW' scale were also major overlays.

In contrast to the fair subsample, Table 7 shows a dramatic shift in the expectation patterns in the unfair subsample. Unlike the near constant (zero) expectations for the EPW ranks in the fair subsample, the public expectations in the unfair subsample strongly resembled the under/over bettor bias which has been documented in the literature review. Of particular interest is the fact that the EPW favorites and near favorites in the unfair subsample produced higher expectations than their counterparts in the fair subsample. The reasons for this increase can be developed by examining the characteristics of the EPW' and DPW expectations. If we once again assume that the paramorphic probabilities accurately embody the ideal betting behaviors of the general public, then the extremely low success rates and negative expectations for the low ranked EPW' suggest that the information base of the paramorphic representation in this subsample was not sufficient to capture the betting policies of a subset of pari-mutuel wagerers. Consequently, it may be assumed that members of this subset employ information which is not only exogenous to the public domain, but it is also superior in its ability to identify potential winners. Support for
this position is clearly shown in Table 7 where nearly 55% of the winners in the unfair subsample were dominant underlays which produced a remarkably high expectation of .95. Unfortunately, the negative expectation of -.30 for the second largest deviation fails to give additional support for the efficacy of betting underlays. However, due to the high percentage of winners in the first rank level of DPW, less than half of the races in this subsample (p=61) were available to be correctly classified in the remaining eight ranks. Perhaps with a larger sample size a smoother pattern of monotonically decreasing expectations for the paramorphic deviations could be obtained. Nevertheless, one can reasonably conclude that a bettor faces severe negative expectations if he fails to evaluate the information character of a race. This position clearly supports the second major hypothesis of this study that the identification of violations of the fair game model can be used to produce positive expectations.

Returning to the under/over bettor bias in this subsample, it can now be argued that the high return rates for favorites and near favorites are products of informed betting. In other words many of these favorites and near favorites were not sufficiently endorsed by the general public because the historical records for these horses did not display any obvious superiority over other competitors. This conclusion is based on the fact that over 70% of the favorites in this subsample were also extreme DPW underlays. Thus, by virtue of being overbet by insiders these "paramorphically inferior" horses would tend to approach favorite or near favorite status. Given the ipsative nature of dollar odds, it can also be argued that because of this overbetting on non-obvious horses that "paramorphically superior" horses would be under-represented in the betting pool, and thereby approach longshot
status. By following this line of reasoning, the following non-obvious conclusion can be offered: The under/over bettor bias reported in the literature is an artifact of informed betting. For the unconvinced reader, the following scenario is offered. Imagine three behavioral decision-making researchers who are each independently asked to evaluate the presence of an under/over bettor bias in one of the three EPW expectation patterns presented in Table 7. The researcher presented with the total sample would tentatively confirm the presence of a bettor bias, although he would point out that the deviations at both ends of the expectation scale are modest. The researcher presented with the fair subsample would have to conclude that no bettor bias is present. Rather he would suggest that horse-race betting occurs under a condition of risk rather than uncertainty. Finally, the research who is shown the EPW expectations in the unfair sample would have to conclude that a strong bettor bias exists whereby favorites are underbet and longshots are overbet. Yet, from the privileged perspective of the paramorphic model, what this third researcher labels an underlay is in fact an overlay, and vise versa. From this example, it should be clear that the reported systematic disturbances between subjective and objective probability distributions (cf., the first researcher) can be solely attributed to a subset of races in which one of the competitors is overbet despite an apparent mediocrity in published records. The implications of this finding will be discussed in further detail in the next chapter.
CHAPTER IV
DISCUSSION

This study provides several substantive and methodological contributions to our understanding of group decision-making in the pari-mutuel setting. The major findings will be discussed in terms of the structural and functional aspects of the paramorphic representation of dollar odds. In the former, the emphasis will be upon the types and forms of information a person uses to generate probability estimates, while in the latter, the focus will be upon the advantages and implications of using a numerical weighting of the public's betting behaviors.

Structure of Paramorphic Probabilities

A preliminary aim of this study was to identify the codable inputs that the betting public uses to assign probability estimates to race horses. An inspection of the form and content of the twelve identified predictors of dollar odds, strongly suggests that the pari-mutuel bettor behaves knowledgeably and rationally. For example, the reported congruence between the coefficient signs of the predictors with both EPW and FIN demonstrates that the bettor is fully aware of the monotonic directionality between past and future events. Furthermore, the time-dependent correlation patterns between the total predictor pool and the two major dependent criteria indicate that the typical wagerer justifiably assigns more weight to recent information than to outdated data. This recency pattern can be traced to the fact that performances of race horses tend to follow a four to eight week sinusoidal pattern.
characterized by high levels of physical form for several weeks, followed by a steady tapering off in physical conditioning in subsequent weeks (see Kusyshyn, 1979, p. 32).

Not only is the public's belief system highly veridical to the true state of nature, it also appears to be the product of an active process of comparative evaluation. Virtually every predictor selected by the stepwise regression procedure was, in varying degrees, a modified or transformed version of the raw input values provided in the racing form. The high frequency of ranked or proportionalized predictors indicate that the rational wagerer rescales the available information so as to enable meaningful comparisons within race blocks. This should not imply that the pari-mutuel bettor actually calculates proportions, ranks, or logit transformations; rather, these metrics capture the flavor of his decision-making policies. Similarly, the reader is cautioned from assuming that the elements of the final predictor pool are somehow unique characterizations of the measures used by our ideal wagerer. Instead, they reflect those dimensions of historical information which are most salient to the decision-maker. The predictor pool can roughly be classified into three dimensions: recent performances, situational constraints, and earning power.

In the first dimension, three predictors -- Speed1, Speed2, and Kick1 -- embody the decision-maker's concern for demonstrable evidence that a horse has recently exhibited a competitive performance. Within the context of horse racing, it should not be surprising that our bettor is highly cognizant of speed differences among competitors. What is surprising is the extent to which raw speeds can be modified in order to maximize the criterion correlation. The results suggest that the bettor does not accept speed information at face value; rather, he attempts to
covary the impact of external influences such as track speeds and weather conditions before making a comparative judgment. Clearly, Speed1 was the most powerful predictor as it accounted for nearly 35% of the explained variance in dollar odds. The fact that criterion means at each rank level for Speed1 (and Speed2) were all significantly different from each other, also suggests that our bettor is capable of responding to rescaled indicators in terms of highly sensitive gradients. The reader will recall that sensitivity of measurement was one of the major characteristics of an "economic man" as suggested by Edwards (1967), and assumed by this author. 

In the second dimension, seven predictors -- Post, Driver, Driver1, Odds1, Odds2, Class1, and Class2 -- reflect the bettor's concern for contextual information which can be used to modify the significance of past performance indicators such as speed and vigor. The first three predictors in this dimension represent factors which may alter the actual speed performances of a competitor. Our typical bettor apparently recognizes that a situational constraint such as driver ability or post position may enhance or impede a horse's expected racing capabilities. In contrast, the remaining four predictors in this dimension represent situational factors related to a horse's competitors. Both class levels and prior odds enable the bettor to evaluate the quality of competition a horse has recently encountered. Often an apparently superior performance can be attributed to an inferior field of competitors, and vice versa. For example, an average performance at very high odds in a higher class may be considered a stronger indicator of good racing form, when compared to a superior performance at very low odds in a lower competitive class.
In the final dimension, two related predictors -- Earnings77 and Earnings76 -- suggest that the bettor is cognizant of long term comparative speed differences in the quality of competitors each horse has encountered. Because of breeding factors, each horse possesses a specific, though unknown, speed potential which is assumed to covary with earning potential. Consequently, the adjusted earnings in the last two years reflect the bettor's use of molar indicators of a horse's inherent speed capabilities and also the speed potentials of its past competitors.

In sum, the structural characteristics of the identified predictors of dollar odds suggest that the typical pari-mutuel wagerer assigns probability estimates to race horses in a purposeful and rational manner. He appears to be especially concerned with the comparative speed potentials and performances among the competitors from both short term and long term perspectives. The form of the predictors also suggests that these comparative differences are produced through an active process of evaluation in which the bettor attempts to covary the impact of situational constraints that may obfuscate a horse's true speed potential.

Before turning to the functional aspects of combining the identified predictors, it should be pointed out that a very powerful predictor of dollar odds -- the "morning line" estimates -- was deliberately excluded from the predictor pool. As is mentioned in Appendix A, race-track management provides the public with mutuel ratings or morning lines (ML) for each contestant. These values are
produced by the track handicapper(s) and they represent advance estimates of the closing dollar odds for each entry based on the information provided in the track program (cf., forecast prices, reported in Dowie, 1976). Although these estimates are expressed by a tote odds notation, they are not subjected to the ipsative constraints of dollar odds; i.e., \( \Sigma \frac{1}{(ML+1)} \neq 1 \). Generally, a horse that is predicted to be the race favorite is given a 5:2 morning line, while the second morning line favorite is given a 3:1 rating, and so forth. Hence, it was assumed that the betting public perceives ML values as ordinal estimates. This assumption was supported by a significant increase in the criterion correlation from .50 to .71 once these estimates were ranked within race blocks \( (p < .0001) \). The reader is reminded that the purpose of ML estimates is to predict the relative endorsements of the entries and not their win potentials. Thus, it can be argued that the high criterion correlation with ML reflects a self-fulfilling prophecy, whereby the public’s betting preferences are influenced by estimates of their own behaviors. Nevertheless, there are compelling methodological and theoretical grounds for excluding ML estimates from consideration.

Methodologically, it can be argued that the inclusion of ML estimates in the model building process would produce an artificially high level of multicollinearity with the remaining predictors of dollar odds. Since the track handicapper bases his estimates on the same information pool as the least squares procedure, each historical predictor would be forced to share a common portion of criterion variance with the ML predictor. This incestuous relationship between the initial predictor set and the ML variable would, in turn, force the
stepwise regression procedure to partition the criterion variance in each predictor according to its independence of the ML variance. In other words, the relative criterion contribution of each predictor would be calibrated with reference to the track handicapper's behavior and not the betting public's. Consequently, the stepwise regression procedure would tend to identify and weight those predictors that the track handicapper eschews. The problem of using ML estimates is similar to the issue of calculating correlations between elements and totals which include the elements (see Cohen & Cohen, 1972, pp. 66-67).

There is also a theoretical reason for excluding the ML variable from the initial predictor pool. The reader is reminded that a paramorphic model is by definition based on actuarial inputs (p. 28); and since the ML variable is not an actuarial variable but, rather, a subjectively derived index, it must be deemed exogenous to the information pool. In a sense, an ML can be characterized as a subjective analogue to the paramorphic-linear representation constructed in the present study. In other words, the track handicapper is concerned with calculating an index which attempts to describe the policies of the decision-makers. The fact that the objectively constructed representation in this study accounted for 15% more variance that its subjective counterpart, provides preliminary support for the conclusion offered by Dawes and Corrigan (1974) that linear models can provide superior results when compared to human judgments.

**Function of Paramorphic Probabilities**

The primary aim of this study was to demonstrate that a linear model of a decision-maker's judgments can be used to provide superior predictions if, (a) the model is sufficiently accurate in capturing the
decision-maker's policies, and (b) the decision-maker's judgments are partially unreliable. The results clearly support the efficacy of using this approach in the pari-mutuel setting if, (c) the assumed error variance (i.e., unreliability) of the betting public's decisions is purged of exogenous influences which are not random. This section will address each of these three issues in detail.

The initial task faced by this author was to construct a representative linear model of the win probabilities produced by the betting public. This entailed the estimation of the parameters of a linear regression model in which the original dependent variable was a counted fraction or proportion. Often, a researcher encounters proportions which will sum to one across all cells in the raw data. When this proportion is a dependent measure in a predictive model, a researcher would prefer that the predicted scores also sum to unity. Within this context, Cook (1971) provides the only available report on the conditions and techniques which are necessary to ensure that the predicted scores will sum to one. Unfortunately, Cook does not specify whether his recommended methods will ensure that the predicted scores will have all the mathematical properties of the original proportions; viz, that the probability for any predicted score in the sample space is no greater than one or no less than zero. In addition, Cook's recommendations were restricted to data bases where the sum of the proportions summed to one across all cells in the raw data. No mention was made of data bases where the proportions sum to one across cells within blocks. Finally, Cook fails to remind his readers that samples of proportions are typically non-normal and that they are more likely to violate the assumption of fixed sampling in linear analysis.
Faced with these shortcomings, this author expanded the logit transformation so that the predicted EPW scores were linear, possessed all the mathematical properties of proportions, and summed to one within each blocking factor. By using an anti-logit transformation on the linearized Logit EPW scores, this author was able to avoid negative estimates and thereby set a lower bound of zero on the predicted scores. By normalizing these scores within race blocks, he was also able to set an upper bound of one. (This technique could also have been used on a data base without a blocking factor, since the normalization of the anti-logit scores can be performed on the total sample which can be considered a blocking factor unto itself.) Aside from the substantive contributions of this study, this author believes that this study has provided a powerful methodological tool for researchers who are required to predict proportion scores.

Returning to the substantive aspects of the present study, the logit/anti-logit transformation served a dual purpose in developing and evaluating the paramorphic representation of the betting public's behaviors. First, by normalizing the dependent measure, the author was able to account for an exceptionally large proportion of linear variance in the criterion. In light of the fact that most predictive studies of human behavior report multiple correlations in the .20 to .40 range (e.g., see Epstein, 1979) the multiple correlation of .81 for Logit EPW indicates that the weighted predictor set was extremely accurate in embodying the public's betting behaviors. Second, by applying the anti-logit normalization sequence to the predicted Logit EPW scores, the author was able to directly compare the paramorphic probabilities and the original EPW scores through the construction of the DFW "residuals."
These residuals, in turn, enabled the author to explicate the necessary conditions under which a paramorphic representation could be used successfully in the pari-mutuel setting.

The fundamental advantage of a paramorphic representation can best be summarized by the following quotation from Dawes (1971):

A mathematical model ... is an abstraction of the process it models; hence, if the decision maker's behavior involves following valid principles but following them poorly, these valid principles will be abstracted by the model -- as long as the deviations from these principles are not systematically related to the variables the decision maker is considering (p. 182).

In other words, a linear representation of a decision-maker's behavior, by its very nature, removes random error or unreliability from a person's judgments, while retaining the essence of his expertise. Thus, by definition, an accurate model of the decision-maker's behavior should provide superior decisions to those of the decision-maker by virtue of its mechanical reliability.

In the present study, the author assumed that the under/over bettor bias was the primary source of unreliability in pari-mutuel decision-making. Consequently, he hypothesized that a model of the betting public's behaviors would not be contaminated by the various determinants of this unreliability, which in turn would enable the model to outperform the betting public. However, by examining the expectations for the errors of prediction (DPW) within each race block, the author was able to detect a systematic pattern in those races where the linear fit between an expected and observed win probability exceeded a .07 tolerance threshold. This pattern consisted of a strong under/over bettor bias contingent on the presence of weak linear prediction.
It was demonstrated that extremely successful rates of return could be achieved in these races by focusing on the errors of predictions and not the predictions themselves. These results lead the author to conclude that the bettor bias discussed in the literature is a product of information availability. Conversely, the author was able to demonstrate that the errors of prediction in the remaining subset of races were in fact true measures of unreliability. When these fair races were analyzed separately, the efficacy of paramorphic bootstrapping was vindicated by the higher expectations for paramorphic favorites than public favorites.

The reader should readily appreciate that the key role of the paramorphic model was to evaluate the extent to which the available information pool was reflected in final dollar odds. The results suggest that when there is moderate agreement between actuarial data and subjective estimates, a mechanical specification equation will outperform human judges. However, when the subjective estimates seriously deviate from published information, a mechanical weighting procedure is rendered useless. Under this condition, the weighting model must be used as a covariate in order to expose hidden information. External support for this conclusion can be partially drawn from McGlothlin's (1955) examination of the under/over bettor bias. The reader will recall that McGlothlin calculated expectations for various odds-levels across eight subsamples corresponding to the ordinal position a race assumed in a daily sequence of eight races. With one exception, he was able to demonstrate the presence of the under/over bettor bias in each subsample. However, McGlothlin was unable to find a
clear under/over better bias in the seventh race subsample. Although longshots in this subsample did produce small negative expectations, the entries below 5:1 produced near zero expectations. McGlothlin attributed this negative finding to the fact that the seventh race was typically the feature race, and that horses in this type of race "usually have impressive records and are highly publicized in the local newspapers. Other horses are seldom mentioned outside of charts showing entries and results" (p. 610). Apparently McGlothlin did not realize the full implications of this statement. Clearly, he is suggesting that betting patterns are linked to the quality of information concerning the conditioning and health for each of the competitors; while, in ordinary races this type of information would remain exogenous to the information pool available to the betting public. It is precisely this problem of information availability that a paramorphic model solves by summarizing and comparing the available information pool to the dollar odds.

Future Research Directions

This study has raised serious doubts about the nature of the under/over better bias reported in the pari-mutuel literature. Virtually every study has attributed this phenomenon to the public's inability to properly estimate the extremities of an unknown objective probability distribution. In contrast, this study has shown that the betting public's estimations are remarkably veridical across all levels of the observed performance distribution, once the pari-mutuel market is purged of exogenous influences that the public is not privy to. Thus, the author strongly recommends that future studies in this area control
for the impact of information inefficiency. One method of achieving this control would be to develop a database where the extent of information inefficiency is very low. This could be done by replicating this study on a set of races similar to McGlothlin's feature race subsample, where the racing capabilities of the competitors are very much in the public domain. The author strongly suspects that the expectation patterns of the paramorphic deviations in this type of database will be overshadowed by those of the paramorphic probabilities.

In addition to replication, this author recommends that the strategy employed in this study be generalized to other research areas concerned with the prediction of outcome variables. One area that could benefit from the findings provided in this study would be clinical prediction. Critical reviews by Brody (1972, pp. 196-226), Meehl (1954), and Sawyer (1966), among others, have provided substantial evidence to conclude that statistical prediction of clinical data is superior to clinical synthesis. Specifically, linear additive combinations for outcome variables such as response to therapy, recidivism, and success in college, provide equal or superior predictions when compared to clinical predictions. As was discussed earlier, this superiority can be attributed to the absence of unreliability in the mechanical weighting procedure which the clinician is assumed to possess. However, this argument is based on the assumption that the errors of prediction for all clinicians are in fact linear (i.e., random). If, on the other hand, it can be demonstrated that the discrepancies from a linear model for some of the clinicians are predictive, then it can be argued that these clinicians are either
employing nonlinear combinatorial rules, or they are privy to exogenous information. In either case, a valid comparison between statistical and clinical prediction would have to include clinical decision-makers who are not mere shadows of the statistical method. The author believes that these clinicians can be identified through the use of a paramorphic representation whereby each clinician's judgment would be compared to a model of his peers, and the deviations from this model would be compared to the criterion. The author suspects that a certain subset of clinicians will produce deviations from a paramorphic model that will outperform the output of the model. This belief is based on the present study which successfully employed a paramorphic linear model to evaluate (and covary) the impact of historical information in relation to future events.
REFERENCES


La Fave, L. *Psychological methodology: Should it differ from that of natural science.* Canadian Psychologist, 1971, 12, 513-525.


Miller, G. A. *The magical number seven, plus or minus two.* Psychological Review, 1956, 63, 81-97.


Orne, M. T. *On the social psychology of the psychology experiment: With particular reference to demand characteristics and their implications.* American Psychologist, 1962, 17, 776-783.


Slovic, P., & Lichtenstein, S.  *Comparison of Bayesian and regression approaches to the study of information processing in judgment.*  *Organizational Behavior and Human Performance,* 1971, 6, 649-744.


APPENDIX A
DESCRIPTION OF DATA BASE

The purpose of this appendix is to familiarize the reader with the various types of information that were used to construct the database. The database included 106 claiming races grouped into six discrete levels ranging from $10,000 to $25,000 in claiming price, and 134 conditioned races grouped into six discrete levels ranging from non-winners of $2,000 in their last six starts to non-winners of $6,500 in their last six starts. With the exception of stake races, and invitational and preferred races which represent the highest levels of competition, the majority of races are either conditioned or claiming races. In the former, a horse must meet all the specified race conditions to be eligible for a particular race. These conditions are written by the track's racing secretary and they usually pertain to a horse's recent performance expressed in terms of certain amounts of money earned within a specified time interval. All competitors in a claiming race may be claimed (purchased) by any registered horseman who submits a purchase bid prior to the running of the race. The purpose of both conditioned and claiming races is to maintain high levels of competition by restricting the qualitative range of eligible horses. To this end, various allowances are also provided for certain horses according to age and gender.

The upper portion of Figure A-1 displays a reproduction of a racing form for the second race conducted at Mohawk Raceway, Campbellville, Ontario, on June 26, 1977. This race (which was not included in the
Figure A-1. An example of sources used to construct the data base: racing form; abridged actuarial data for one horse; and race result chart.
## MUTUEL RATINGS: 5 - 4 - 3

### 1 MILE PACE

**Purse $5,000**

Horses of 20,000 who are 3 years old or over will be given.

**Picks & Notes:** Those who are not winners of $5,000 or over shall be given.

### RACE 2

**Pace - Conditioned**

<table>
<thead>
<tr>
<th>Horse</th>
<th>PP</th>
<th>1/2</th>
<th>1</th>
<th>1 1/8</th>
<th>Str. Fin. Time</th>
<th>Driver</th>
<th>Odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Battles Finale</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1 1/8</td>
<td>5 5 6 6</td>
<td>Duford</td>
</tr>
<tr>
<td>Highland Jet</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4 4</td>
<td>3</td>
<td>2 2 2 2</td>
<td>WaplesR</td>
</tr>
<tr>
<td>Chinkapin</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2 2 2 2</td>
<td>Wamerj</td>
</tr>
<tr>
<td>Terry Parker</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5 1/2</td>
<td>4 2</td>
<td>WalkerL</td>
</tr>
<tr>
<td>Merrywood Sara</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3 1/2</td>
<td>4 2</td>
<td>LarkinD</td>
</tr>
</tbody>
</table>

**Time:** 29 1:00 1:30 1:59
present data base) was a conditioned race for non-winners of $6,500 in their last six starts. If fillies and mares were non-winners of $7,500 in their last six starts they too would be eligible for this race. The purse for this race was $5,000. Purse values are typically meted among the first five finishers in the following proportions: .5, .25, .12, .08, and .05, respectively.

Each horse is assigned a post position (and corresponding program number) according to a random draw. Post positions are important in harness racing since high post positions are furthest from the inside rail and require horses to travel several dozen meters further than low post positioned horses. High post positions also make it more difficult for a driver to manoeuvre his horse into a favorable racing position in the early portion of the race. On the Ontario Jockey Club circuit, horses with post positions greater than No. 8 are positioned in a second tier; e.g., a horse in post No. 9 starts behind the post No. 1 horse.

Beneath the post position numbers are listed the "morning lines" which represent the track handicapper's advance estimates of the closing dollar odds for each of the competitors. These values are expressed by tote odds notation and they are not constrained by the ipsative requirements of actual dollar odds. The "Mutuel Ratings," listed next to the race number indicate the morning line "favorites;" i.e., the three horses assigned the lowest morning line values.

The central portion of Figure A-1 shows an abridged version of the racing form for the No. 3 horse (BATTLES FINALE). This section shows only those codable variables used in the data base. Battles Finale, an eight-year-old gelding was driven by Jerry Duford. Measures of the driver ability were taken from the United States Trotting Association's
annual publication of the U.A.T.A. Roster of Harness Racing Drivers for the year 1976. This publication provides actuarial measures for all harness drivers competing in North America. The measures used in the data base included: number of starts, number of wins, seconds, thirds, and gross purse winnings. Jerry Duford's 1976 racing record consisted of 700 starts, 140 wins, 89 seconds, 82 thirds, and purse winnings of $373,706.

In a similar fashion, the upper right-hand area of the central portion of Figure A-1 shows performance measures for Battles Finale's racing record in the last two years. In the current year, he has competed in 21 races, winning six, coming third three times for cumulative purse earnings of $22,373. The previous year's record is given in the next line, and under Battles Finale's name is listed his lifetime earnings of $131,286 which does not include the current year's earnings.

The remaining portion of Battle Finale's record consisted of his last six pari-mutuel starts or "performance lines." From right to left, his most recent start was on June 19 at Mohawk Raceway, a 5/8 mile oval track. The race distance was 1 mile (1,609 meters) on a fast racing surface for a purse of $5,000. The leading horse reached the quarter-mile in 30-4/5 seconds, the half mile in 1:02-1/5, the three quarters in 1:32-4/5, and the mile in 2:02. Battles Finale started from post No. 4 and was 5th at the quarter, 6th at the half, and 6th at the three-quarter. The track program uses the letter "p" to indicate if a horse raced on the outside from the rail for each of the first three quarters. Battles Finale was 5th at the top of the stretch and 4 1/2 lengths behind the leader. He finished 4th and was 4 3/4 lengths behind the winner in a time of 2:02-1/5. In that race his closing dollar odds
was 8.20:1 and he was driven by Jerry Duford. Where applicable, the racing form will use a distinguishing symbol (e.g., <->) along side the past odds to indicate if a horse was the race favourite. For example, Battles Finale was a race favourite in his fourth last performance line.

The lower portion of Figure A-1 is a reproduction of the results chart for our sample race. The order of finish for each horse is given in the sixth numerical column. Battles Finale, the third favourite, won by 5-1/2 lengths in 1:59-2/5 at dollar odds of 2.10 to 1. Based on a two-dollar unit bet, he paid $6.20 to win, $3.40 to place, and $3.20 to show. The second favourite, Highland Jet, finished second by 5-1/2 lengths with broken equipment, and paid $3.20 to place, and $2.70 to show. The race longshot, Chinkapin, finished third by 7 lengths and paid $2.90 to show. The fractional speed quarters for the leading horse are presented at the bottom of the race chart.
VITA AUCTORIS

Oleksander Mitzak was born on July 13, 1949 in Edinburgh, Midlothian, Scotland. He received his primary and secondary education at Rosedale Public School and Saint Michael's College School, Toronto, Ontario, Canada. In September, 1968 he enrolled at York University, Downsview, Ontario. In May, 1972 he graduated with the Bachelor of Arts degree (General Arts), and in the following year he received the Bachelor of Arts degree (Honours Psychology). In September, 1973 he was accepted into the Master of Arts programme in Social-Personality Psychology at the University of Windsor, Ontario, and in May, 1976 he graduated with the Master of Arts degree (Personality). During the years 1976-1979 he was enrolled as a Doctoral psychology student at the University of Windsor, and upon completion of his comprehensive specialty examinations (Personality and Measurement) in January, 1980 he entered his Doctoral candidate year.