The Journal of Prediction Markets (2007) **1**, 43–59

COMPARING THE EFFECTIVENESS OF ONE- AND TWO-STEP CONDITIONAL LOGIT MODELS FOR PREDICTING OUTCOMES IN A SPECULATIVE MARKET

M. Sung† and J.E.V. Johnson‡

Centre for Risk Research, School of Management, University of Southampton, Highfield, Southampton, SO17 1BJ, United Kingdom

This paper compares two approaches to predicting outcomes in a speculative market, the horserace betting market. In particular, the nature of one- and two-step conditional logit procedures involving a process for exploding the choice set are outlined, their strengths and weaknesses are compared and their relative effectiveness is evaluated by predicting winning probabilities for horse races at a UK racetrack. The models incorporate variables which are widely recognised as having predictive power and which should therefore be effectively discounted in market odds. Despite this handicap, both approaches produce probability estimates which can be used to earn positive returns, but the two-step approach yields substantially higher profits.

I. INTRODUCTION

Establishing the extent to which a financial market incorporates information provides important clues to the manner in which it operates and it is widely recognised that horserace betting markets, which share many features in common with wider financial markets, can provide a valuable window on speculative market behaviour (e.g. Snyder, 1978; Hong and Chiu, 1988; Law and Peel, 2002). Sauer (1998, p 2021), for example, observes:

> "wagering markets are especially simple financial markets, in which the pricing problem is reduced. As a result, wagering markets can provide a clear view of pricing issues which are complicated elsewhere."

Two of the distinctive features of horserace betting markets which make it possible to understand behaviour more clearly than in other financial markets is the generation of an unequivocal outcome (a winner) within a definitive period and the availability of a large number of essentially similar markets (races) for analysis. These features enable market efficiency to be tested by measuring the appropriateness of the asset's price (market odds) against race outcome. As a result, an extensive literature has developed addressing weak-, semi-strong- and strong-form efficiency in horserace betting markets. However, there have been relatively few studies exploring the extent to which horserace bettors discount, in market odds, a *combination* of

†Tel: +44(0) 23 8059 9248 Fax: +44(0) 23 8059 3844 Email: ms9@soton.ac.uk
‡Tel: +44(0) 23 8059 2546 Fax: +44(0) 23 8059 3844 Email: jej@soton.ac.uk

fundamental variables associated with horses, their jockeys and trainers and the race conditions. The majority of these studies employ a one-step modelling process which involves regressing measures of past performance (e.g. past finish position) on information derived from fundamental variables *alongside* market generated probabilities (e.g. Bolton and Chapman, 1986; Chapman, 1994; Gu, Huang and Benter, 2003). However, Benter (1994) advocated the use of a two-step procedure, which involves developing a model based solely on fundamental variables to predict winning probabilities. These probabilities are then used as inputs to a second stage model which also incorporates market generated probabilities. He argues that such a process produces more accurate predictions. Benter's highly successful betting operation in Hong Kong provides anecdotal evidence to support this view and two stage models developed by Edelman (2003) and Sung, Johnson and Bruce (2005) have produced encouraging results. However, to date no comparison of the accuracy of winning probabilities from one- and two-step modelling procedures has been undertaken. This is clearly an important omission, since a modelling technique which captures the full information content of fundamental and market-generated variables is more likely to demonstrate the true degree of market efficiency. This paper aims to fill this important gap by evaluating the effectiveness of these two modelling approaches in predicting winning probabilities at a racetrack in the UK and their ability to reveal market inefficiency.

To achieve this, the paper is structured as follows: The one- and two-step modelling procedures on which this study focuses are outlined in section II, along with their relative strengths and weaknesses. The data and explanatory variables used to develop parallel one- and two-step models are described and justified in section III, together with the procedures used to assess the relative predictive power of these two approaches. In section IV, the results of model estimation and out-of-sample testing are reported and discussed. Some implications and conclusions are developed in section V.

II. ONE- AND TWO-STEP MODELLING

(a) One- and Two-step Conditional Logit Modelling Procedures

The modelling procedure which forms the basis of the one- and two-step procedures which are compared here is the most widely used in assessing the degree of semi-strong-form efficiency in racetrack betting markets, namely, conditional logit. The aim of a conditional logit model is to predict a vector of winning probabilities $\boldsymbol{p}_{ij}^{e} = (p_{1j}^{e}, p_{2j}^{e}, \ldots, p_{n_j j}^{e})$ for race j, where p_{ij}^{e} is the estimated model probability of horse i winning race j and n_j is the number of horses in the race j. These probabilities are estimated on the basis of a vector of m variables: $\mathbf{x}_{ij} = [x_{ij}(1), \ldots, x_{ij}(m)]$, capturing information associated with each horse. Since horse races are competitive, an efficient probability estimate of horse i's chance of winning race j is more likely to be obtained if

its chance of winning is regarded as being *conditional* on the information available for the other runners in race j. To achieve this, a 'winningness index' W_{ij} for horse i in race j is defined as follows:

$$(1) \qquad W_{ij} = \sum_{k=1}^{m} \beta_k x_{ij}(k) + \varepsilon_{ij}$$

where β_k is a coefficient which measures the relative contribution of information $x_{ij}(k)$ to horse i's chance of winning and ε_{ij} is unperceived information. If W_{ij} is defined such that the horse with the highest value of the index wins race j then it can be shown that, if error terms ε_{ij} are independent and distributed according to the double exponential distribution, the probability of horse i winning race j is given by the following conditional logit function (McFadden, 1974):

$$(2) \qquad p_{ij}^{e} = \frac{\exp\left(\sum_{k=1}^{m} \beta_k x_{ij}(k)\right)}{\sum_{i=1}^{n_j} \exp\left(\sum_{k=1}^{m} \beta_k x_{ij}(k)\right)}$$

where the β_k are estimated using maximum likelihood procedures. The conditional logit model has been successfully employed for a range of discrete choice problems (McFadden, 1974) including a number of studies estimating the winning probability of racehorses (e.g. Figlewski, 1979; Bolton and Chapman, 1986; Edelman, 2003).

The one- and two-step procedures differ in terms of when and how *fundamental information* concerning the previous performances and preferences of the horse, its jockey or trainer and race conditions (represented below in equation (3) as $m - 1$ fundamental variables, $y_{ij}(k)$) are combined with *market-generated information* (usually employed in the form of normalised probabilities derived from closing market odds, p_{ij}^{s}). In a one-step procedure the two types of information are combined in a single conditional logit model of the following form:

$$(3) \qquad p_{ij}^{e} = \frac{\exp[\beta_m \ln p_{ij}^{s} + \sum_{k=1}^{m-1} \beta_k y_{ij}(k)]}{\sum_{i=1}^{n_j} \exp\left[\beta_m \ln p_{ij}^{s} + \sum_{k=1}^{m-1} \beta_k y_{ij}(k)\right]}$$

Benter (1994) was the first to develop a computer model for predicting winning probabilities of horses in two-steps. The two-step procedure involves first developing a conditional logit function of the form given in equation (2), and simply employing the $m - 1$ fundamental variables $y_{ij}(k)$ (Benter, 1994). This provides an estimate of the probability of horse i winning race j, p_{ij}^{f}, which is based solely on the fundamental variables. However, according to Benter (1994), the fundamental probability consistently diverges from the observed win percentage for each odds category. As a result, an adjustment from the fundamental probability to the unobserved true winning chance of a horse is necessary. Consequently, a second-step is required, incorporating the natural logarithm of the fundamental model probability, $\ln(p_{ij}^{f})$, as well as

the natural logarithm of the normalised closing odds probability, $\ln(p_{ij}^{s})$, based on a second set of races. Consequently, the final estimated model probability for horse i in race j is obtained as follows:

$$p_{ij}^{e} = \frac{\exp\left(\alpha \ln\left(p_{ij}^{s}\right) + \gamma \ln\left(p_{ij}^{f}\right)\right)}{\sum_{i=1}^{n_j} \exp\left(\alpha \ln\left(p_{ij}^{s}\right) + \gamma \ln\left(p_{ij}^{f}\right)\right)} \tag{4}$$

where α and γ are parameters to be estimated using maximum likelihood procedures. The second-step model is designed to capture the subtle relationship between these two explanatory variables and the outcome of a race.

(b) Strengths and Weaknesses of One- and Two-step Conditional Logit Modelling

In order to test the hypothesis that racetrack betting market is semi-strong form efficient, it is clearly necessary to develop a model which combines both fundamental and market-generated information. The merits of combining this information in a one-step model are not only its simplicity but also the opportunity it affords for examining the significance of each individual fundamental variable in an explicit manner; since probabilities derived from the final market odds also appear in the model. This facilitates understanding of how bettors utilise publicly available information. In particular, the fundamental information which has been ignored or under-weighted by the betting public can be identified. Another technical advantage of the one-step modelling process is that it permits the use of a larger training sample and this can improve model accuracy. This is particularly true for a conditional logit model as it treats one race rather than one horse as an observation during estimation. The two-step modelling procedure, on the other hand, requires that the training sample is split in two, one for each step; this is required in order to overcome the potential problem of over-fitting (Benter, 1994). The one-step model might therefore have particular merit when only a limited sample of races is available for analysis.

Despite the potential benefits of the one-step procedure indicated above, it could also be argued that combining all the fundamental variables with the odds variable may result in counterintuitive signs for the model parameters due to the variables being highly correlated. This may increase the difficulty of interpreting the results (Benter, 1994). A second important practical difficulty associated with the one-step process is that in order to use the probabilities generated by this process to bet it is necessary to have a good estimate of the final market odds. Benter (1994) suggests that these can be obtained from market odds prevailing one or two minutes prior to the start of the race. However, this would not allow sufficient time in which to run a complete model incorporating all fundamental variables and a variable derived from final odds estimates, and then place the bet. It can, therefore, be

argued that the one-step model would be more difficult to implement in real time. Step one of the two-step model procedure involves estimating the probability derived from fundamental variables; these data are available several hours before the race begins. Step two, which combines probability estimates derived from the fundamental model and from the final market odds (or those prevailing one or two minutes before the race start) could be developed in a few seconds, which would permit the model probabilities to be used to bet. This may therefore represent a more practical method for implementing a betting strategy in real time. Therefore, it could be argued that the two-step model is the one which really tests whether a market is inefficient.

The two-step modelling procedure clearly offers some important advantages over a one-step procedure but it can only be used to examine the *collective* significance of all the fundamental information (by observing the significance level of the coefficient of the fundamental probability term in the second-stage model); the marginal significance of each of the individual fundamental variables cannot be discerned.

As discussed, both the one- and two-step modelling procedures have their own strengths and weaknesses. However, there has been no investigation which compares the predictive power of these modelling approaches. The next section outlines the data and procedures employed here to undertake such an investigation.

III. DATA AND PROCEDURES

(a) Data

One of the distinguishing features of a good model for making probability predictions is that it is able to extract the full value from underlying information. In particular, to make an accurate assessment of horserace betting market efficiency, it is vital that the model fully utilises the fundamental and market-generated information. In this regard, in order to set a difficult test for the one- and two-step modelling procedures evaluated here, it was decided to limit the set of fundamental variables included in the models to those included in Bolton and Chapman's (1986) seminal paper on multi-covariate modelling in a horseracing context. This paper inspired several other researchers and practitioners to develop more sophisticated models (the most notable and successful example being Benter (1994)). It is clear, therefore, that the fundamental variable set which Bolton and Chapman (1986) employed has been in the public domain for a considerable period and according to the efficient market hypothesis it is likely that the betting public now discounts much of this information in market odds. Consequently, a modelling procedure which is able to use information extracted from *this* fundamental variable set to make predictions which yield abnormal returns might be regarded as of particular merit. A full list of the variable set used in the current study is given in Table 1.

Bolton and Chapman's (1986) study predicted winning probabilities at US racetracks whose topography, configuration and surface are highly standardised, whereas the current study employs data from the UK where racetracks are far less uniform. Consequently, to minimise this discrepancy, data is drawn from races run at one racetrack, *Wolverhampton*, whose configuration, topology and surface are similar to that of the US tracks. The dataset contained details of 16431 horses which ran in 1675 flat races during the period January 1995 to August 2000. The number of horses in each race varies from 2 to 13, with a mode of 12. The final market odds for horses in the sample range from 0.17/1 to 100/1 with a mean value of 13.64/1.

The dataset is split into two parts. The first part, involving 1110 races (10856 horses) run between January 1995 and December 1998, is used to develop the conditional logit models. The second part, involving 565 races (5575 horses) run between December 1998 and August 2000, is preserved for out-of-sample testing. The one-step model is estimated using all the observations in the training dataset, but the two-step model requires the training dataset to be further divided: the fundamental model being estimated using races run between January 1995 and December 1996 (555 races, 5524 horses) and the model combining fundamental model probabilities and market-generated probabilities being estimated using races run between December 1996 and December 1998 (555 races, 5332 horses).

The UK betting market consists of parallel bookmaker and pari-mutuel markets with odds being generated in both markets. However, the bookmaker

TABLE 1
DEFINITIONS OF THE INDEPENDENT VARIABLES EMPLOYED IN THE ONE- AND TWO-STEP MODELS

Independent variable	Variable definitions
Market-generated variable	
$\ln(p_{ij}^{s})$	The natural logarithm of the normalised final odds probability
Fundamental variables	
pre_s_ra	Speed rating for the previous race in which the horse ran
avgsr4	The average of a horse's speed rating in its last 4 races; zero when there is no past run
draw	Post-position in current race
eps	Total prize money earnings (finishing first, second or third) to date/Number of races entered
newdis	1 indicates a horse that ran three or four of its last four races at a distance of 80% less than current distance, and 0 otherwise
weight	Weight carried by the horse in current race
win_run	The percentage of the races won by the horse in its career
jnowin	The number of wins by the jockey in career to date of race
jwinper	The winning percentage of the jockey in career to date of race
jst1miss	1 indicates when the other jockey variables are missing; 0 otherwise

market is by far the larger and it has been suggested that informed bettors are more likely to bet in this market (e.g. Bruce and Johnson, 2005). Consequently, in this study, final market odds in the bookmaker market are used for model development.

(b) Procedures

The sample involves a choice set of observations, whereby 'nature' chooses a winner of each race but the sample size (555 races) is relatively small. In traditional conditional logit modelling only information concerning which horse wins the race is employed but Chapman and Staelin (1982), describe an 'explosion process' which can be used to exploit extra information from the original ranked choice sets (i.e. from horses in a race finished 2nd, 3rd etc.) without adding too much random noise. This method involves considering the finishing position of each horse in a given race as a set of mutually independent choices. Consequently, it is assumed that the horse which finished second would have won the race if the horse finishing first had not participated in the race. For example, an explosion from depth one (the original race) to three can produce two 'extra races' by sequentially eliminating the 'winner' from the pared down races (i.e. a race where the original winner is eliminated and a race where the original winner and second are eliminated). This is clearly a valuable process as it increases the number of independent choice sets, which results in more precise parameter estimates.

However, there is a limit to the depth to which races can be exploded since the latter finishing positions may not truly reflect the competitiveness among the remaining horses. This arises since it may become obvious to a jockey that his/her mount will not finish in the first three (where prize money is awarded); the jockey then has little incentive to ensure that the horse achieves its best possible finish position. In fact, there may be positive incentives not to do this, as it helps to conceal the horse's true ability, which increases the value of the horse's connections' (owners, trainers etc) private information (which they can exploit in the betting ring in subsequent races). As a result, the maximum depth of explosion to which a race is exploded is restricted to three in this study (Bolton and Chapman, 1986).

For certain sets of races it may not even be appropriate to explode to level three (since this process may introduce too much random noise) and a statistical measure which can be used to determine the appropriate depth of explosion is suggested by Watson and Westin (1975). This method involves iteratively testing the hypothesis that the maximum likelihood estimates for each individual subgroup of races are equal; the explosion process is continued until the hypothesis is rejected. This procedure involves determining the log-likelihood (*LL*) values for models estimated on the following separate subgroups of races: (i) all runners included: (E = 1); (ii) runners which finished first excluded: (E = 2) – (E = 1); (iii) runners which finished first and second excluded: (E = 3) – (E = 2); (iv) races falling into categories (i) and (ii) pooled: (E = 2); (v) races falling into

categories (i), (ii) and (iii) pooled: (E = 3). The statistic used to test the hypothesis that the maximum likelihood estimates for each individual subgroup of races are equal compares, for example, the *LL* of explosion depth two (E = 2) with that from the subgroups [(E = 1) and (E = 2) – (E = 1)] combined. This statistic is, therefore, defined as $-2\{LL(E=2) - [LL(E=1) + LL((E=2) - (E=1))]\}$ (Chapman and Staelin, 1982), and follows the chi-square distribution with the degrees of freedom equal to the number of parameters in the conditional logit model (Wald, 1943). A particular subset of races is only exploded to a depth where the hypothesis that the maximum log-likelihood estimates for all the subgroups of races are equal is not rejected.

The aim of the paper is to compare the predictive ability of one- and two-step conditional logit models. The approach is, therefore, to use the exploded data sets to build both types of model. For the one-step model, the appropriate depth of explosion is determined by calculating the test statistic discussed above for the whole test sample of 1110 races (January 1995- December 1998). For the two-step model, the subset of 555 races (January 1995– December 1996) used to estimate the fundamental model and the subset of 555 races (December 1996– December 1998) used to estimate the model incorporating fundamental and market-generated probabilities are both tested separately to determine the appropriate depth of explosion.

(c) Model Evaluation

The predictive ability of the two models is compared by developing a betting strategy based on model predictions. In particular, a Kelly wagering strategy (Kelly, 1956) is employed, based on probabilities derived from the one and two-step models. The Kelly strategy ensures that total wealth grows optimally at an exponential rate in the long run with zero possibility of ruin. A Kelly betting strategy involves betting a proportion of total wealth on a given runner. As wealth levels increase later in the sequence of bets, very large absolute value bets can be recommended. To ensure that the success of one of the modelling procedures is not biased by one or two large wins later in the sequence of bets, a Kelly strategy without re-investment is employed; the wealth level is therefore returned to unity after each bet, whatever the outcome of the previous bet. The accuracy of the winning probabilities estimated by the one- and two-step modelling procedures are assessed by comparing the rates of return obtained by using these probabilities in a Kelly strategy to bet on the out-of-sample races.

IV. RESULTS AND DISCUSSION

(a) Model Estimates and Model Fit: Two-step Model

1. Step One: Fundamental Model

The coefficients for the ten fundamental variables in the conditional logit model were estimated for depths of explosion one, two and three, respectively,

TABLE 2

COEFFICIENTS AND TEST STATISTICS OF CONDITIONAL LOGIT MODELS INCORPORATING FUNDAMENTAL VARIABLES FOR EXPLOSION DEPTHS OF 1, 2, AND 3 (STEP-ONE OF A TWO-STEP PROCEDURE)

Variable[1]	Explosion Strategy						*t*-ratio	*p*-val.[2]
	E = 1		E = 2		E = 3			
	Coef.	Std. error	Coef.	Std. error	Coef.	Std. error		
pre_s_ra**	0.1968	0.0684	0.2003	0.0490	0.2005	0.0405	4.96	0.000
avgsr4**	0.3288	0.0858	0.3751	0.0612	0.3423	0.0505	6.78	0.000
draw**	0.2481	0.0488	0.2462	0.0347	0.2117	0.0282	7.51	0.000
eps	0.1244	0.0907	0.1143	0.0654	0.0881	0.0550	1.60	0.109
newdis**	−0.2191	0.0617	−0.1648	0.0435	−0.1968	0.0358	−5.50	0.000
weight**	0.1631	0.0641	0.1644	0.0456	0.1549	0.0371	4.17	0.000
win_run	0.0237	0.0722	0.0018	0.0523	0.0117	0.0444	0.26	0.792
jnowin**	0.2433	0.1278	0.3116	0.0916	0.2454	0.0775	3.17	0.002
jwinper**	0.0776	0.0308	0.0631	0.0226	0.0602	0.0190	3.17	0.002
jst1miss	−0.0496	0.0577	−0.0799	0.0443	−0.0175	0.0272	−0.64	0.521
Summary Statistics								
No. races	555		1,109		1,659			
$L(\theta = 0)$	−1,254		−2,444		−3,558			
$L(\theta = \widehat{\theta})$	−1,134		−2,210		−3,263			
Pseudo$\widehat{R}^2$	0.0958		0.0957		0.0829			

[1]**Significant at the 5% level;
[2]*The test statistics are taken from the data for an explosion depth of three.

based on the first subset of 555 races. These results are presented in Table 2. To evaluate which rank ordered explosion is appropriate, sequential tests of the hypothesis that the maximum likelihood estimates for individual subgroups of races are equal are undertaken. These results are reported in Table 3. The chi-square test statistics for explosion depth two (4.97) and three (17.71) are both less than the 5% critical value (18.31), suggesting that the hypotheses that the exploded rank ordered samples follow the same distribution as the population cannot be rejected at the 5% level. It is, therefore, valid to explode the choice set to a depth of 3 for the purpose of model estimation. The value of increasing the number of observations using the exploding procedure is demonstrated by an increase in model precision; the estimated standard errors of the model coefficients decrease on average as the depth of explosion increases by 28 percent (from depth explosion one to two), and 19% (from depth two to three).

A *LL* ratio test comparing the fundamental variable model estimated for explosion depth three (shown in Table 2) with one where no explanatory predictor is incorporated demonstrates that the ten fundamental variables, collectively, have a significant amount of explanatory power (*LL* ratio $= 590$, $\chi^2_{10}(0.05) = 18.31$). Of the ten variables, seven are significant at the 5% level. Two of these are associated with the situation in the current race (i.e. post-position and the weight carried by the horse), one with the horse's preferences (i.e. whether the horse is running at a new distance), two with the historical performances of the horse (variables involving past speed ratings), and two with jockey-related variables. The model clearly demonstrates that these variables have an impact on which horse wins a given race. In addition, the model appears sensible, since all of the significant variables have coefficients with the anticipated signs.

2. Step Two: Combining Fundamental and Market-Generated Information

A second-step conditional logit model, including the natural logarithm of (i) the estimated fundamental probability from the first-step model and (ii) the normalised probability implied by the closing bookmaker market odds is developed, based on the second subset of 555 races. The second-step model is

TABLE 3
LOG-LIKELIHOOD VALUES AND TEST STATISTICS FOR DETERMINING THE OPTIMAL EXPLOSION DEPTH FOR FIRST-STEP MODEL ESTIMATES

Subgroup of races	No. of Races	*LL* Value	*LL* ratio test statistic	χ^2_{10} (.05) critical value
(E = 1)	555	−1,134		
(E = 2) − (E = 1)	554	−1,073		
(E = 3) − (E = 2)	550	−1,044		
(E = 2)	1,109	−2,210	4.97	18.31
(E = 3)	1,659	−3,263	17.71	18.31

TABLE 4

COEFFICIENTS AND TEST STATISTICS OF CONDITIONAL LOGIT MODELS INCORPORATING FUNDAMENTAL AND MARKET-GENERATED VARIABLES FOR EXPLOSION DEPTHS OF 1, 2, AND 3 (STEP-TWO OF A TWO-STEP PROCEDURE)

Variable	Explosion Strategy						*t*-ratio	*p*-val.[1]
	E = 1		E = 2		E = 3			
	Coef.	Std. error	Coef.	Std. error	Coef.	Std. error		
$\ln(p_{ij}^{s})$**	0.7977	0.0632	0.7798	0.0454	0.6951	0.0372	17.19	0.0000
$\ln(p_{ij}^{f})$**	0.1658	0.0625	0.1386	0.0446	0.1742	0.0366	3.11	0.0020
Summary Statistics								
Number of Races	555		1,110		1,663			
$L(\theta = 0)$	−1,234		−2,400		−3,488			
$L(\theta = \widehat{\theta})$	−1,060		−2,090		−3,090			
Adj $\widehat{R}^2$	0.1408		0.1293		0.1141			

[1]The test statistics are taken from the data for an explosion depth of two.

estimated by exploding the 555-race sub-sample to depths of one, two and three (see Table 4 for detailed results). The resulting log-likelihood values and the number of races for each depth explosion are displayed in Table 5. The value of the Watson and Westin (1975) sequential pooling test statistic for E = 3 is 11.60, which is larger than the chi-square critical value (with 2 degrees of freedom: 5.99) at the 5% level of statistical significance. The hypothesis that the exploded rank ordered sample at explosion depth 3 follows the same distribution as the population is therefore rejected. Consequently, it is only appropriate to explode the 555-race sub-sample to a depth of two (1110 races).

As in step one, the standard errors of the parameter estimates of the conditional logit model improve as more observations are added (i.e. at explosion depth 2: see Table 4). In addition, a *LL* ratio test comparing the likelihood of the data under the alternative hypothesis that all the parameters in the model with depth explosion two are not equal to zero, against the likelihood of the data under the null hypothesis that all the parameters in the model are equal to zero, indicates that the combination of the two variables offers significant predictive power (*LL* ratio $= 620$, $\chi^2_2(0.05) = 5.99$). In addition, the coefficients of both the log of the normalised closing bookmaker market odds probability and the log of the fundamental model probability are both significant at the 5% level (t-ratio = 17.19 and 3.11 respectively). This suggests that *both* these variables provide valuable information for estimating winning probabilities. This is confirmed by a *LL* ratio test comparing the log-likelihood of the combined model (*LL* = − 2,090) with the log-likelihood of a model simply incorporating the log of the probability derived from the final market odds (*LL* = − 2,095); (LL ratio = 10, which is significant at the 1 per cent level ($\chi^2_1(0.01) = 6.64$)).

(b) Model Estimates and Model Fit: Ono-step Model

The one-step model is estimated by exploding the 1110 races run between January 1995 and December 1998 to depths of one, two and three. The log-likelihood values of each explosion depth are summarised in Table 6. The test statistic for depth explosion two is less than the chi-square critical value but the test statistic for depth explosion three is greater than this critical value,

TABLE 5
LOG-LIKELIHOOD VALUES AND TEST STATISTICS FOR DETERMINING THE OPTIMAL EXPLOSION DEPTH FOR SECOND-STEP MODEL ESTIMATES

Choice Group	No. of Races	*LL* Value	*LL* ratio	χ^2_2 (.05) critical value
(E = 1)	555	− 1,060		
(E = 2) − (E = 1)	555	− 1,029		
(E = 3) − (E = 2)	553	− 994		
(E = 2)	1,110	− 2,090	1.20	5.99
(E = 3)	1,663	− 3,090	11.60	5.99

indicating that it is only appropriate to explode races to a depth of two for model estimation; expanding the in-sample size from 1110 to 2219 races (20,601 horses).

Table 7 reports the results of estimating a one-step conditional logit model (referred to as model A) incorporating the 10 fundamental variables together with log of the normalised probability derived from final bookmaker market odds. This model is estimated using data exploded to a depth of two.

A comparison of model A with the model including only fundamental variables, which is developed at step-one of the two-step procedure (results presented in Table 2), reveals several interesting issues relevant to betting market efficiency. Coefficients of two variables (i.e. speed rating for the previous race in which the horse ran and average career earnings) change to counter-intuitive signs in model A (suggesting that the public over-emphasise this information in their assessment of winning probabilities). In addition, four variables which were significant in the fundamental variable model (developed at step one of the two-step procedure) are not significant at the 5% level when combined with log of the normalised odds probability (i.e. weight carried by the horse in current race, speed rating for the previous race in which the horse ran, the number of wins and the winning percentage of the jockey throughout career). This suggests that, whilst these factors influence the winning probabilities, they are fully discounted in odds. Three variables, which were significant in the model including only fundamental variables (which is developed at step-one of the two-step procedure), remain significant in model A. These include post-position, average speed rating of the horse's last four runs and whether the horse is running at a significantly longer distance than in its last four races. This finding implies that post-position is not fully incorporated into the closing market prices and is consistent with existing studies which explore the role of post-position (e.g. Quirin, 1979; Canfield, Fauman and Ziemba, 1987; Betton, 1994). The other two variables which are significant in both models involve information derived from several past performances of a horse; these might, therefore, be regarded as relatively opaque variables, and the fact that the betting public does not appear to fully

TABLE 6
LOG-LIKELIHOOD VALUES FOR DETERMINING THE OPTIMAL EXPLOSION DEPTH FOR MODEL ESTIMATES, WHICH INCLUDES THE TEN FUNDAMENTAL VARIABLES AND THE LOG OF THE NORMALISED ODDS PROBABILITY

Choice Group	No. of Races	*LL* Value	*LL* ratio	χ^2_{11} (.05) critical value
(E = 1)	1,110	−2,106		
(E = 2) − (E = 1)	1,109	−2,038		
(E = 3) − (E = 2)	1,103	−1,978		
(E = 2)	2,219	−4,149	9.67	19.68
(E = 3)	3,322	−6,141	29.98	19.68

TABLE 7
COEFFICIENTS AND TEST STATISTICS OF CONDITIONAL LOGIT MODELS INCORPORATING FUNDAMENTAL AND MARKET-GENERATED VARIABLES FOR EXPLOSION DEPTHS OF 1, 2, AND 3 (IN A ONE-STEP PROCEDURE)

Variables	Model A		
	Coefficients	Std. error	t-ratio
$\ln(p_{ij}^{s})$	**0.8091	0.0337	23.99
pre_s_ra	−0.0123	0.0362	−0.34
avgsr4	**0.1413	0.0449	3.15
draw	**0.1367	0.0251	5.44
eps	−0.0506	0.0464	−1.09
newdis	** −0.0960	0.0332	−2.90
weight	*0.0554	0.0330	1.68
win_run	0.0117	0.0394	0.30
jnowin	*0.0620	0.0369	1.68
jwinper	0.0350	0.0225	1.56
jst1miss	* −0.0708	0.0376	−1.88
Summary statistics			
No. of races	2219 (20601 runners)		
$L(\theta = 0)$	−4,844		
$L(\theta = \widehat{\theta})$	−4,149		
Pseudo$\widetilde{R}^2$	0.1436		

**Significant at the 5% level.
*Significant at the 10% level.

discount this information in odds is in line with laboratory based research which indicates that individuals take less account of data which is less readily discernable when forming their judgements (Tversky and Kahneman, 1974). Finally, one of the most striking findings associated with model A and the model incorporating fundamental probabilities and market-generated probabilities (at step-two of the two-step procedure) is that the log of the normalised odds probability appears to be highly significant, implying it has a dominant influence in predicting the race winner.

(c) Comparison of models' predictive ability based on a Kelly wagering strategy

The accuracy of winning probabilities predicted by models developed by the one and two-step procedures (both of which incorporate market-generated and fundamental variables) are tested by a simulated betting exercise. Both models are used to predict probabilities for the 565 out-of-sample races run between December 1998 and August 2000, and these are used as inputs to implement Kelly wagering strategies without re-investment of profits. The rates of return obtained for the models developed by the one- and two-step procedures over this period were 0.96 percent and 17.53 percent, respectively.

Clearly both models suggest that the market is semi-strong form inefficient, but the two-step model appears to capture significantly more information relevant for winning probability prediction than the one-step model.

V. CONCLUSION

The paper set out to compare the accuracy of probability estimates based on one- and two-step conditional logit analysis. In particular, the paper assessed the ability of these models to make accurate assessments of the winning probability of horses running in flat races in the UK. The models were set a difficult task since the only independent variables which were employed were those which have been widely publicised as having an influence on winning probability (i.e. variables employed in Bolton and Chapman, 1986).

The results suggest a number of important conclusions. Most importantly, in relation to the central objective of this paper, the analysis conducted here suggests that the two-step model captures more information contained in the independent variables; as significantly larger profits were obtained in the out-of-sample period using a betting strategy based on the predicted probabilities from this model. These results imply that tests of market efficiency which employ the one-step model may over-estimate the degree to which market odds discount fundamental information. One of the reasons for this may be that the two-step model reduces the impact of multicollinearity. Odds clearly play a dominant role in predicting the outcome of a race and under these conditions the correlations between the odds and other fundamental variables are likely to be high. A model which incorporates odds and fundamental variables in one step is, therefore, likely to produce more unstable predictions due to muticollinearity. In addition, the coefficients of the fundamental variables do not aid understanding of the relationship between winning probability and the fundamental variables since these coefficients will be affected by the degree to which bettors account for these variables in odds. A two-step modelling process separates the odds-related variable from the fundamental variables and allows these fundamental variables to compete for importance in one model. This may reduce the problems resulting from multicollinearity. An additional benefit of the two-step model, as discussed earlier, is that it can be used *in practice* to capitalise on any market inefficiency identified, since step one (the development of a fundamental variable model) can be undertaken well before the betting period starts. This enables bettors to complete step two in the last two minutes of the betting period, allowing for market odds close to the start of the race. The one-step modelling procedure would take far too long to complete, even with modern computers, to enable predictions of winning probabilities to be generated in the last one or two minutes before the race starts. Consequently, it can be argued that the two-step modelling procedure is the only one of these approaches which allows for practical application of the model in real time;

and is therefore the only one which can be applied to test for true market efficiency.

A further interesting finding reported here is that the variables which are significant in the one-step model are, to some degree, different from the variables which are significant in the fundamental model of the two-step model. For example, jockey-related variables are significant in the fundamental model of the two-step procedure but are no longer significant when they are included alongside the odds variable in the one-step model. This implies that the odds variable incorporates information in relation to the past performances of jockeys. On the other hand, if the results of the different modelling procedures had not been compared, the importance of each variable with and without the odds variable included would not be revealed. In other words, the empirical comparison between the modelling methods aids our understanding the important factors affecting the results of horse races and the extent to which this information is used efficiently.

Finally, the study offers important conclusions concerning the degree of semi-strong form efficiency in the UK betting market. The results demonstrate that certain types of information are not accounted for in market odds in the UK. The one- and two-step models both identified a number of significant explanatory variables derived from publicly available information. For example, the position of a horse in the starting stalls (post-position) appears to be significant at the 5% level. This is a surprising finding for two reasons: (i) post-position is normally made public the day before the race and this should provide sufficient time for the public to take this information into account, (ii) Bolton and Chapman (1986) reported post-position to have non-trivial effects on winning probabilities. In addition, it is interesting to note that the average speed rating for a horse in its last four races plays a significant role in forecasting the outcomes of unseen races. To take this variable into account the betting public need to transform the underlying data. The fact that they do not appear to account for this variable in their betting decisions suggests that data which is not readily available (i.e. requires some prior analysis) may not be acted on by the betting public. The study also confirms the strongly positive relationship between odds and the likelihood of a horse winning a race, confirming that the odds variable contains a considerable amount of information associated with a horse's relative competitiveness in a race. The return of 17.53% over the holdout sample period for a betting strategy based on probabilities predicted by the two-step model suggests that the UK bookmaker-based betting market is not semi-strong form efficient. This is surprising since the variables employed have been in the public domain since Bolton and Chapman published their article in 1986. This finding runs counter to the efficient market hypothesis which would predict that markets react to the publication of information and discount it in market prices.

In summary, the results reported here further our understanding of modelling of outcome probabilities in a speculative market and confirm that levels of efficiency identified in these markets are highly dependent on the modelling technique employed.

REFERENCES

W Benter 'Computer based horse race handicapping and wagering systems: A report' in D B Hausch, V S Y Lo and W T Ziemba (eds) *Efficiency of Racetrack Betting Markets* (London, Academic Press, 1994) pp 183–198.

S Betton 'Post-position bias: An econometric analysis of the 1987 season at Exhibition Park' in D B Hausch, V S Y Lo and W T Ziemba (eds) *Efficiency of Racetrack Betting Markets* (London, Academic Press, 1994) pp 511–526.

R N Bolton and R G Chapman 'Searching for positive returns at the track: A multinomial logit model for handicapping horse races' *Management Science* (1986) 32 1040–1060.

A C Bruce and J E V Johnson 'Market ecology and decision behaviour in state-contingent claims markets' *Journal of Economic Behavior and Organization* (2005) 56 199–217.

B R Canfield, B C Fauman, W T Ziemba 'Efficient market adjustment of odds prices to reflect track biases' *Management Science* (1987) 33 1428–1439.

R G Chapman 'Still searching for positive returns at the track: Empirical results from 2,000 Hong Kong races' in D B Hausch V S Y Lo and W T Ziemba (eds) *Efficiency of Racetrack Betting Markets* (London, Academic Press, 1994) pp 173–181.

R G Chapman and R Staelin 'Exploiting rank ordered choice set data within the stochastic utility model' *Journal of Marketing Research* (1982) 19 288–301.

D C Edelman 'Adapting support vector machine methods for horserace odds prediction' (2003) *The 12th International Conference on Gambling and Risk Taking*, Vancouver BC, June.

S Figlewski 'Subjective information and market efficiency in a betting market' *Journal of Political Economy* (1979) 87 75–89.

M G Gu, C Huang and W Benter 'Multinomial probit models for competitive horse racing', 2003, *Working paper of the Chinese University of Hong Kong*.

Y Hong and C Chiu 'Sex, locus of control and illusion of control in Hong Kong as correlates of gambling involvement' *The Journal of Social Psychology* (1988) 128 667–673.

J L Kelly 'A new interpretation of information rate' *The Bell System Technical Journal* (1956) 35 917–926.

D Law and D A Peel 'Insider trading, herding behaviour and market plungers in the British horse-race betting market' *Economica* (2002) 69 327–338.

D McFadden 'Conditional logit analysis of qualitative choice behavior' in P Zarembka (ed) *Frontiers in Econometrics* (New York, Academic Press, 1974) pp 105–142.

W L Quirin *Winning at the Races: Computer Discoveries in Thoroughbred Handicapping* (New York, William Morrow, 1979).

R D Sauer 'The economics of wagering markets' *Journal of Economic Literature* (1998) XXXVI 2021–2064.

W W Snyder 'Horse racing: Testing the efficient markets model' *Journal of Finance* (1978) 33(4) 1109–1118.

M Sung, J E V Johnson and A C Bruce 'Searching for semi-strong form efficiency in British racetrack betting markets' in L Vaughan Williams (ed) *Information Efficiency in Financial and Betting Markets* (Cambridge University Press, 2005) pp 179–192.

A Tversky and D Kahneman 'Judgment under uncertainty: Heuristics and biases' *Science* (1974) 185(4157) (Sep. 27) 1124–1131.

A Wald 'Tests of statistical hypotheses concerning several parameters when the number of observations is large' *Transactions of the American Mathematical Society* (1943) 54 426–482.

P L Watson and R B Westin 'Transferability of disaggregated mode choice models' *Regional Science and Urban Economics* (1975) 5 227–249.