SMARTsig Confidential 8.11, November 2001

In this month's issue . . .

NEXT ISSUE

December 2001, issue 8.12, is scheduled for posting on Thursday, 29 Nov. 2001.

SMART UP FRONT

The intelligent choice

Football Game Ratings

Last soccer season and through these pages we had a shot at spread-betting soccer superiorities. The model used for the predictions was one compiled several seasons previously and, coupled with the fact that the staking was tinkered with part-way through the project, we ended last term with a 36 point loss.

However, in the best traditions of the well informed punter, we said we would benefit from the experience, learn from the mistakes made and set about making amends in the 2001/02 football season. So, without further ado, on with our mission to make the superiority spreads pay their way.

The first job is to re-evaluate and recompile our prediction model. But rather than simply churning out a new table using the same criteria as last time, perhaps there would be benefits from examining the base from which they are calculated? The reason behind this view is that I've been experimenting with football ratings in general.

A few issues back I tested some football predictions, using what I've found to be one of the simplest, yet most reliable methods of establishing a team's merit - i.e. league points divided by games played. A match rating is then simply calculated by subtracting the away team rating from the home team's.

Match Rating = (HTpts/HTplayed) - (ATpts/ATplayed)

One of the follow-ups to the article was from Richard Scott

in Australia, who suggested using different points to the traditional 3pts win, 1 pt draw, and re-checking. I didn't respond immediately (sorry Richard - but this late reply shows I have given it some thought)

It makes sense when you think about it too. The function of the 3 points win, 1 point draw, as the current league system dictates, is primarily to compile a table allowing league positions – and thus promotion and relegation – to be assessed.

The method was upgraded several years ago from the old 2 points win, 1 point draw – the intention at the time I seem to remember was that if a win was worth three times the draw, rather than just double the value, it would promote a fighting spirit in the teams and therefore provide better entertainment for the spectators.

Whether it had the desired effect or not is a matter for your own opinion, but no football clubs fiercely objected to the rule change if memory serves, despite the fact that in some cases finishing league positions could (and would) be affected. I suppose if you know the rules from the outset, and everyone is working to the same structure, there can be little foundation to disputing season-end league positions.

And it DOES make a difference. Take the bottom of last season's (2000/2001) Division 1 table;

	P	w	d	l	f	a	w	d	l	f	a	pts
18 Grimsby T.....	46	10	4	9	26	27	4	6	13	17	35	52
19 Stockport Co..	46	6	11	6	29	26	5	7	11	29	39	51
20 Portsmouth....	46	9	8	6	31	25	1	11	11	16	34	49
21 Crystal P.....	46	6	6	11	28	34	6	7	10	29	36	49
22 Huddersfield.r	46	7	6	10	29	26	4	9	10	19	31	48
23 QPR.....r	46	6	9	8	24	28	1	10	12	21	47	40
24 Tranmere R...r	46	8	7	8	30	33	1	4	18	16	44	38

Huddersfield, QPR and Tranmere were relegated. But what would the picture have been if the old 2 points win, 1 point draw were still in place?

	P	w	d	l	f	a	w	d	l	f	a	pts
18 Stockport Co..	46	6	11	6	29	26	5	7	11	29	39	40
19 Portsmouth....	46	9	8	6	31	25	1	11	11	16	34	39
20 Grimsby T.....	46	10	4	9	26	27	4	6	13	17	35	38
21 Huddersfield..	46	7	6	10	29	26	4	9	10	19	31	37
22 Crystal P....r	46	6	6	11	28	34	6	7	10	29	36	37
23 QPR.....r	46	6	9	8	24	28	1	10	12	21	47	33
24 Tranmere R...r	46	8	7	8	30	33	1	4	18	16	44	29

It changes every position except those of the bottom-most two, who were out of it in any case. With the old points system Palace would have been relegated and Huddersfield would have enjoyed another season of division one football, courtesy of their better goal difference. No one bemoans the fact of what would have happened if . . . because they are the well established rules to which everyone adheres.

It does however illustrate the issue in question rather well. Points allocated for wins and draws are perfectly okay for compiling a league table, but are they a suitable vehicle for ranking the ability of one team against another? On the other hand, isn't a league table intended essentially to indicate the ability of each team? A brief glimpse at the above two league table illustrations and one wonders which was effectively the better team – Grimsby or Portsmouth? Palace or Huddersfield? Their respective WDL records are identical in both tables, it's merely the method of points allocation that changes their respective rankings.

In his excellent book "The Essential Football Betting Guide", Paul Steele compares the predictive powers of various methods and his research suggests that one of the best ratings methods overall is the simple points per win, points per draw methods examined here. Paul's approach though checks various values including one which gives more points for away wins/draws than for wins and draws at home.

Having examined this evidence closely, thorough though it is, it falls into what I see as a trap which will always be evident when whole numbers are employed.

As an example imagine two teams who have each played 8

games.
TeamA has a record of won 1, drawn 0, lost 7
TeamB has w0, d3, l5

Using a w=2, d=1 method shows ***TeamB*** to be superior to ***TeamA***. Okay, ***TeamB*** have lost fewer games, but they have yet to win a game!

Adopting a w=3, d=1 approach on the other hand, shows two teams with equal ratings, but do they have equal ability? Putting aside for one moment the quality of opposition they've met, can such judgements ever be assessed by processing results using exact whole numbers?

Using 2pts win, 1 point draw, we could have 5 teams with different WDL records, yet all sharing equal points . . .

	Played	W	D	L	Pts
Team A	8	4	0	4	8
Team B	8	3	2	3	8
Team C	8	2	4	2	8
Team D	8	1	6	1	8
Team E	8	0	8	0	8

Is it really feasible that the above five teams are of equal ability as the points system suggests? Team A have won more than the others, but they've lost more too. Team E may well be undefeated, but they've also played eight games and failed to win one yet. The only commonality with the teams is that they've all won the same number of games that they've drawn.

A similar table could be drawn up using 3 points win and 1 point draw, the WDL patterns would be a little different than the example above, but it would look equally as unrealistic.

As I see it the only way to solve the problem of x number of draws equalling y number of wins is to try using fractional amounts and testing the outcome. 3 points per win with, for example 1.4 points for a draw, will greatly reduce those

occasions where a team with multiple draws can ever equal the points total of another team with more wins but fewer drawn games. This is probably as it should be, the different make-up of wins & draws may be a little better, a little worse – but exactly equal? Somehow I doubt it.

To test out my theories I processed the last 4 seasons of English league data using 1 point per win, with draw points ranging from 0.0 to 1.0 in steps of 0.1

After each 0.1 point pass I checked the 'performance' of the (HTpts/HTplayed) - (ATpts/ATplayed) game rating method by noting the percentage of home wins in the highest 5%, 10% & 20% of ratings, and the Away win percentage in the lowest 5%, 10% and 20% of the ratings.

Although 'trends' can be spotted, i.e. better to give a proportionally higher number of points for a draw that a lower number, it was difficult though, using my tests alone to give a definitive "best".

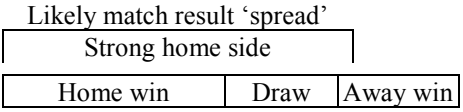
e.g.
Best home predictor for highest 5% of ratings was 1pt for win, 0.7pts for draw (this would be just the same for win=2, draw=1.4, or win=3, draw=2.1, etc)
Best for highest 10% though was with win=1 and draw=0.5, whereas best for highest 20% was tied equally with each of draw= 0.6pts, 0.7pts AND 0.9pts (all with win=1)

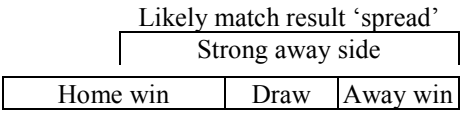
Away wins:
lowest 5%, best was draw=1.0pt (equal draw and win points!!)
lowest 10% best was draw = 0.8pts,
and 20% was also 0.8pts.

Something was evidently wrong here. I don't believe for one minute that one method can be better at detecting home wins, and another points allocation is better at forecasting draws.

As far as I'm concerned a football game is a direct battle of strengths between opposing teams, and as such the final result MUST be linear. A strong home side must give added weight to a home win with the balance tipping towards an away win as the relative merits of the visiting team outweigh those of the home side.

A drawn game is one where the strengths of neither team were enough to swing the match their way.





The illustrations also shows why the percentage of draws remain fairly constant throughout the range of game predictions. And, that strong away prediction is more prone to a totally opposite result than is a strong home prediction.

Examining a range of points awarded for games won & drawn over the past four full seasons it was clear that 3 points per win, 1 point per draw was not the optimum if we wish to use such a system for assessing the relative strengths and weaknesses of two opposing sides.

According to my test results over the past four seasons at least, the older method of 2 points win, 1 point draw was actually nearer the mark in giving a more accurate 'power' rating. (For assessment purposes my initial research actually used 1 points win, 0.5 points draw – but this equates to exactly the same as the 2pts/1pt method, as would 6pts/3pts or any other variation where the points awarded for a win are twice those awarded for the draw)

Always with one eye upon such research being equally useful to our non-computer using readers, a method was adopted whereby 3 points were awarded for a win, as is currently the case. Using such an approach would allow those rating games with pen and paper a simpler method of conversion from the published league tables. Using the 3 point per win as my fixed benchmark, I then trawled through a range of point values for the draw from 0 (zero) to 3 points (draw points = win points) in increments of 0.05, every value throughout the range was charted, each with view to its usefulness in assessing the potential of individual teams.

The graph plotted from the above research peaked at the 1.2 mark. i.e. awarding 3 points for a win and 1.2 points for a draw was the level where the points system had the better predictive powers. All the pen & inkers need do to convert to this approach is, using the current league tables take note of the number of draws recorded for each team and increase the 'official' league points by a sum of (drawn games * 0.2)

From the published league tables, a team's record may be as follows;

P	w	d	l	f	a	w	d	l	f	a	pts
12	3	2	1	8	4	2	2	2	6	6	19

Converting the above to our SMARTsig rating is a simple job of multiplying the number of drawn games (2 at home + 2 away = 4) by 0.2 . The resulting 0.8 is added to the published league table points. Therefore the *SMARTsig* points total rating for this team is 19.8

Once this new points total has been calculated for each team, the SMARTsig 'game-rating' can be easily calculated, it effectively is the away side's points-per-game deducted from the home side's points per game. (points here are 'new points value' as recalculated above).

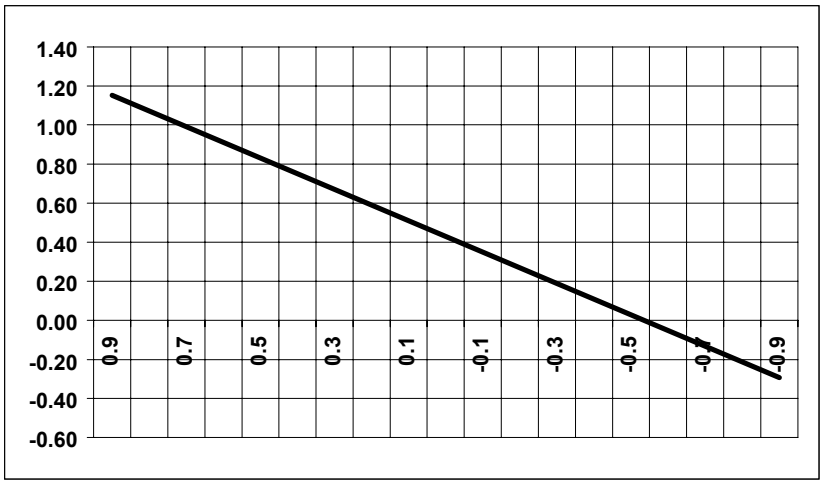
As an example, the above record points-per-game would be

19.8 divided by 12 = 1.65 The whole game-rating procedure therefore can be written as . . .

> (Home team 'new' points divided by home team games played) minus (Away team 'new' points divided by away team games played)

Once this optimum points system had been established it was now time to re-run the four seasons of data where the SMARTsig game rating is generated for each match and each games subsequent superiority to the home side recorded. (Game superiority is merely home goals scored minus away goals scored, as offered weekly in the spread-betting markets)

The subsequent 'smoothed' graph line plotting the superiority progression through the range of SMARTsig game ratings is conveniently linear, allowing users to locate the predicted game superiority for any match to be located quickly once the game rating is known.



Because a team's points-per-game can range from a maximum of 3.0 to zero, it follows that a game-rating can range from +3.0 (home team 3 points per game, away team

0[zero] per game) to –3.0 (home side 0, away side 3) However, it will be noted that the above graph plots only within the range of +0.9 to –0.9

An explanation is in order. All the research carried out only examined matches where both teams had played a minimum of 8 games each in the current season, so in reality it is extremely rare for any team to have a record of 100% wins or 100% defeats at any point, rarer still for two such teams to meet in a league game, and even then there's a 50-50 chance that each team has experienced opposite fortunes.

The +0.9 to –0.9 game-rating spread covers the majority of games in reality, outside of this range and data becomes a little too sparse to be reliable. In practice I intend to extrapolate the data range another 0.2 in both directions, making it +1.1 to –1.1. This then covers 97% of all matches seen over the four-year research period, the remaining 3%, represent fly-away league leaders playing bottom-of-table no-hopers, often a punter's graveyard and as such are probably best ignored.

Because of the linear nature of the graph and predicted superiorities, a formula can also be given for those wishing to automate and/or computerise the method.

Using a fixed benchmark for the SMARTsig match-rating of 0 (zero) indicating a 0.44 home team superiority, a multiplier of 0.8 can be employed to calculate any particular game-rating, as follows;

Predicted home team goal superiority = 0.44 + (SMARTsig game rating * 0.8)

e.g. game rating (+ 0.916) = 0.44 + (0.916 * 0.8)
= 0.44 + 0.7328
= 1.1728
(1.1728 home side superiority)

game rating of (– 1.087) = 0.44 + (– 1.087 * 0.8)
= 0.44 + (– 0.8696)
= – 0.4296
(0.4296 away side superiority)

Using the mathematical method it is important to remember that I would recommend ignoring all *game* ratings higher than 1.1 and all those lower than – 1.1 (minus 1.1)

Maybe good news for those who prefer to keep their workload as simple as possible and would rather not to make the suggested amendment of 1.2 points for a draw, I also generated the relevant tables and predicted superiorities for the traditional 3 points win, 1 point draw system.

Interestingly, exactly the same output graph and numeric calculations can be used, viz;

Simple *game* rating = (home team total points/home team games played) – (away team points/played)

Predicted home team goal *superiority* = 0.44 + (Simple game rating * 0.8)

High percentage performances

For those of you wishing to use these ratings as a basis for traditional home, away or draw betting, the following tables indicate the percentage success rates for the highest and lowest game ratings;

The entire test sample of 6,627 games yielded;

- 3,067 Home wins (46.28%)
- 1,835 Draws (27.69%)
- 1,725 Away wins (26.03%)

HIGH Match ratings (should indicate improved home win chances)

Simple Match Ratings (using 3 points win, 1 points draw)

Criteria	Games	Homes	Draws	Aways
Rating greater than 0.875	333 (near 5% of sample)	210 (63.1%)	68 (20.4%)	55 (16.5%)
Rating greater than 0.691	665 (near 10% of sample)	437 (65.7%)	132 (19.9%)	96 (14.4%)
Rating greater than 0.5	1,147 (near 17% of sample)	717 (62.5%)	249 (21.7%)	181 (15.8%)

SMARTsig Match Ratings (using 3 points win, 1.2 points draw)

Criteria	Games	Homes	Draws	Aways
Rating 0.85 & greater	370 (near 5.6% of sample)	237 (64.1%)	76 (20.5%)	57 (15.4%)
Rating 0.69 & greater	660 (near 10% of sample)	429 (65%)	133 (20.1%)	98 (14.9%)
Rating 0.5 & greater	1,168 (near 18% of sample)	729 (62.4%)	252 (21.6%)	187 (16%)

LOW Match ratings (should indicate improved away win chances)

Simple Match Ratings (using 3 points win, 1 points draw)

Criteria	Games	Homes	Draws	Aways
Rating – 0.91 & lower	331 (near 5% of sample)	92 (27.8%)	88 (26.6%)	151 (45.6%)
Rating – 0.709 & lower	663 (near 10% of sample)	195 (29.4%)	182 (27.5%)	286 (43.1%)
Rating – 0.55 & lower	1,072 (near 16% of sample)	324 (30.2%)	299 (27.9%)	449 (41.9%)

SMARTsig Match Ratings (using 3 points win, 1.2 points draw)

Criteria	Games	Homes	Draws	Aways
Rating 0.9 & lower	343 (near 5% of sample)	96 (28%)	90 (26.2%)	157 (45.8%)
Rating 0.7 & lower	681 (near 10% of sample)	198 (29.1%)	191 (28%)	292 (42.9%)
Rating 0.55 & lower	1,074 (near 16% of sample)	323 (30.1%)	300 (27.9%)	451 (42%)

Sharp-eyed readers will have noticed the sample sizes varying from one table to the next, making direct comparisons difficult, if not impossible.

1	0.892857	1
2	0.9	1
3	0.9	0
4	0.9	1
5	0.9	1
6	0.9	2
7	0.902381	2
8	0.902632	1
9	0.903226	1
10	0.903659	1
11	0.904762	1
12	0.905316	1
13	0.90625	1
14	0.908034	1
15	0.909091	0
16	0.909091	2
17	0.909091	2
18	0.909091	1
19	0.910256	2
20	0.911765	1
21	0.911765	1
22	0.911905	1

This effect is due to the frequent clusters of equal match ratings throughout the files. Although for example, 331 games would have given the nearest point to 5% of the whole, an accurate conclusion cannot be drawn if we do not include every example at the extremes of the sample, where the match ratings may be equal. To accommodate this effect the sample size was enlarged or reduced.

e.g. If we were attempting to take the lowest 15 rating samples with match results from the table on the left we would be including only one of the 0.909091 ratings. There are in fact 4 such games with an equal rating though. The first one is a draw (0), the next two are away wins and the fourth is a home win.

Examining a sample 15 without

recourse to the rating values would therefore give a false picture. It would be necessary therefore in this example to examine either the lowest 14 – or – the lowest 18.

Match ratings ranges with best draw performances

Locating where to look for the best draw ratings along a linear model with home wins and away wins at opposite ends of the scale is not so straightforward. A game between equally rated sides would have a game rating of around the zero mark, but of course that viewpoint ignores the advantage the home side has with familiar surroundings and a partisan crowd.

Over 46% of our entire sample were home wins, which would suggest looking at the range from game rating of zero through the early negative values – but at what point should we begin? And where do we end?

For illustration purposes only therefore, and to avoid the shouts of back-fitting, it is easier to trawl through various sample ranges of around 5%, 10% and 15% to discover in which areas we get the highest percentage of games where the two teams finished level.

The following tables then show the best ranges where draws were found, using this information may, or may not reinforce your own views on which areas would be best for matches finishing all-square.

Simple Match Ratings (using 3 points win, 1 points draw)

Criteria	Games	Draws
Rating less than – 0.0066 but greater than – 0.085	331 (near 5% of sample)	34.14%
Rating less than 0.0454 but greater than – 0.095	663 (near 10% of sample)	32.73%
Rating less than 0.095 but greater than – 0.1166	999 (near 15% of sample)	32.03%

SMARTsig Match Ratings (3 points win, 1.2 points draw)

Criteria	Games	Draws
Rating less than zero but greater than – 0.072	320 (near 5% of sample)	34.69%
Rating less than 0.0995 but greater than – 0.015	663 (near 10% of sample)	32.88%
Rating less than 0.15 but greater than – 0.0649	994 (near 15% of sample)	32.09%

These tables are for guidance only and are unable to show the full picture clearly. In fact using the 5%, 10% and 15% ranges for analysis was found to be both ineffective and often quite misleading in testing and pinpointing the best overall points-for-a-draw value to employ.

The method I finally employed is a straightforward enough procedure, but would take more space than I have available for the subject this month. I'll give a more in-depth account of the method and the results it produced next issue.

SMARTsig

TWO-YEAR-OLD REVIEW 2001
by Dr. Peter May

Available in both electronic format (Adobe Acrobat) for computer download, and printed version (A4 size)

Normal prices Acrobat £5.00, printed £6.95 + £1 p&p

SPECIAL DISCOUNTS FOR SMARTsig MEMBERS

SAVE £2.00!

Acrobat version £3.00
Printed version £4.95 + £1 p&p
Order direct from Peter - ***but don't forget to quote SMARTsig to obtain your discount***

Cheques or postal/money orders only to:
PJM Publications,
44 Blandy Avenue, Southmoor, Oxon OX13 5DB
Tel: 01865 820167 ~ email: peter@pjmr.freeserve.co.uk

Exploring the theme of reducing the number of bets whilst maintaining the good priced winners. Increasing the quality and profitabilty of your betting account.

ACTION OR REWARD?

Russell Hart

Someone said to me recently that people would rather place bets rather than win money. His opinion, based on years of selling systems, was that a system that created ten bets a day would sell better than one that created one bet every ten days.

When researching some staking options recently, I came across a situation where it made much more sense to bet on horses whose prices fell into a certain range of SPs rather than all the system selections. This article takes the argument one step further to examine the phenomenon that restricting your bets can more than double the return from the system.

Which system?

The SMARTsig email group had a discussion recently on a system that included filters like last losing distance and best finish in last six. I put the filters into RSB and came up with a different set of filters that built on the original idea.

The following table gives a step by step approach to horse selection.

Step No	Filter	Action
1.	Going	Go to step 2 if the going is heavy, very soft, soft, firm or hard *We are only interested when the going is at extremes.*

Step No	Filter	Action
2.	Position, best of last 6	Go to step 3 if the horse's best finish has been a second *We are only interested in horses that have been hidden away yet have been close to winning.*
3.	Trainers, strike rate	Go to step 4 if the trainer's strike rate is less than 20% *We are interested in the type of horse that does not interest the top trainers.*
4.	Position, 2 races ago	Go to step 5 if the horse finished 3rd to 8th *We are not interested in horses that showed too much in this race.*
5.	Odds, last race	Go to step 6 if the SP was between 6/4 and 63/1. *We are not interested in horses that had very short or very long odds*
6.	Last losing distance	Go to step 7 if the last losing distance was less than 4 lengths *We are not interested in horses that were close to the winner last time out*
7.	Weight Allotted	Go to step 8 if the horse is set to carry 8st 8lbs to 9st 0lbs *We are not interested in horses with too little or much weight*
8.	Odds forecast rank	Have a bet if horse is 2nd to 6th favourite *We want better odds by avoiding the forecast favourite*

I built the system for years 1993 to 1999. The results were as follows:

Year	Winners	Runners	Strike Rate	LSP	LSP%
1993	18	79	22.78	30.25	38.29
1994	13	64	20.31	30.13	47.08
1995	21	60	35.00	48.69	81.15
1996	10	57	17.54	0.01	0.02
1997	14	53	26.42	50.25	94.81
1998	16	65	24.62	23.50	36.15
1999	20	85	23.53	58.51	68.84
Total	112	463	24.19	241.34	52.13

Well, I know there's a bad year in there but 1996 was a nightmare for betting. 52% level stakes profit and an Archie over 9 gave me confidence in using the system for this article. The system also made a 12.24% level stakes profit in 2000.

How to judge the system?

The foremost test for a system is its return on investment. How much money did the system give me for the money I put in? I added two more criteria: did the system make a profit every season and how many bets did I have.

For all three tests I got Excel to calculate results based on betting on horses whose SP was greater than a set of odds – starting at evens and ending at 15/1.

Starting Price greater than	Return on Investment	Starting Price greater than	Return on Investment
1/1	52.13%	9/1	150.00%
2/1	54.54%	10/1	167.61%
3/1	60.27%	11/1	170.00%
4/1	69.52%	12/1	229.27%
5/1	83.77%	13/1	277.27%
6/1	102.42%	14/1	277.27%
7/1	110.13%	15/1	375.00%
8/1	130.08%		

As you increase the SP the Return on Investment increases. Thus if you only bet on horses with an SP of 10/1 or more you would end up with a return on investment of 167.61%. Good enough for most people.

What does this do the bank?

Applying the same set of odds, I used Excel to calculate the likelihood of the system losing money over a season. The following table summarises the findings with the profit being expressed in units:

Starting Price greater than	Minimum Profit	Starting Price greater than	Minimum Profit
1/1	0.10	9/1	-9.00
2/1	-1.00	10/1	-7.00
3/1	0.50	11/1	-6.00
4/1	0.50	12/1	-6.00
5/1	3.50	13/1	-4.00
6/1	9.50	14/1	-4.00
7/1	4.00	15/1	-2.00
8/1	-4.00		

From 8/1 onwards, the system loses money in one of the years. Thus if you fancied the 167% RoI from betting on 10/1 and above shots, then you would have to be prepared to lose money in one season.

How much action do I get?

I used Excel to count the number of bets for each of the starting prices. The following table summarises the findings:

Starting Price greater than	Number of bets	Starting Price greater than	Number of bets
1/1	463	9/1	88
2/1	429	10/1	71
3/1	378	11/1	50
4/1	315	12/1	41
5/1	265	13/1	22
6/1	207	14/1	22
7/1	158	15/1	8
8/1	123		

As you would expect the number of bets decreases as you take the longer odds.

Deciding what to do?

Given the three separate criteria, I believe you need to look at them in relation to each other. Combining the three tables you get the following:

Starting Price greater than	Return on Investment	Minimum Profit	Number of bets
1/1	52.13%	.1	463
2/1	54.54%	-1	429
3/1	60.27%	0.50	378
4/1	69.52%	0.50	315
5/1	83.77%	3.50	265
6/1	102.42%	9.50	207
7/1	110.13%	4.00	158
8/1	130.08%	-4.00	123
9/1	150.00%	-9.00	88
10/1	167.61%	-7.00	71
11/1	170.00%	-6.00	50
12/1	229.27%	-6.00	41
13/1	277.27%	-4.00	22
14/1	277.27%	-4.00	22
15/1	375.00%	-2.00	8

Action or Money?

All systems are open to this type of analysis. Excel is a great help. Now you have a base for deciding your approach to betting on the system. You could bet on everything better than evens, which would give you 60 plus bets a year with a profit in every year and a 52.13% return on investment.

On the other extreme, you could have 1 or 2 bets per season with a season making a loss but a 375% RoI. Being a little more conservative, I prefer the middle ground of betting on horses with SPs of 7/1 and more. 110.13% RoI is double the Level Stakes Profit; you never make a loss and have some 26 bets a season.

The more bets you make the lower your return on investment. Each system has a rich vein – you need to find it. If you want lots of bets then have lots of systems and find the band of gold that each contains.

I'd rather have the money than the action – staying patient is the difficulty.

The perfect place to recycle your unwanted racing books and betting paraphernalia.

Do your bit to save the planet - and do someone a favour at the same time!

SWAPSHOP is a FREE service to members
So please make FULL use of this facility

WANTED:
Having much of my cherished archive of racing papers reduced to pulp by floodwater I'm now looking for old copies of the Racing Post, particularly January, February & March 2000. But any issues from before that time also considered.
Bryan Shreeve 01992 710975 (Waltham Abbey, Essex)

WANTED:
Past annuals from Racing Research (Whitley) wanted. Realistic prices please.
Philip Alexander 01709 377628 (South Yorkshire)

WANTED:
Horseracing systems from the past. Systems and racing magazines, booklets, manuals, etc. Anything considered.
Brian Blackwell, Australia
email: <briblack@onthenet.com.au>

Having a clear-out or looking for something difficult to find by other means?

Make use of SWAP-SHOP, a **FREE** to members service from SMARTsig

In an effort to perhaps curb the occasional sky-high payouts the CSF has evolved into a complex formula. Its exact details remain shrouded in secrecy, so it takes a brave man to attempt to unravel the mystery.

CALCULATING THE CSF

Tony Watson

There has been much publicity recently about the CSF and its calculation. The purpose of this article is to attempt to review the steps of the calculation of the CSF and to comment on its ingredients. It is not a complete explanation and is certainly not designed as a bookie's justification of it. I hope that wiser heads than mine can explain it better or more deeply.

Let me say up front that the difficulty in reviewing the CSF used for horseracing is that details are not published. It is however possible to get a close idea of the methodology used by BOLA from other sources, namely the CSF as used in BOLA-sponsored BAGS racing and the draw adjustment details published in 1996 in relation to the now defunct Trifecta bet.

First though a bit of theory. This is necessary as one of the big myths of the CSF formula is that it is 'so complicated that even Einstein would not understand it'. Yes, there are difficulties, but some parts are easier then others.

1. THEORY

In essence, the CSF is a double bet on the winners of two separate races, namely the race itself (the winner) and a notional race involving all the other horses apart from the winner (i.e. the second placed horse).

If the true probabilities of all the horses in a race were p1, p2, p3, etc, the probability of horse P1 beating horse P2 is

p1 * p2 / (1 - p1)
(the probability of P1 winning) (the probability of P2 winning the race for second.)

The (1 - p1) adjustment reflects the fact that in the race for second, probability of p1 has been 'taken out of the field' in excluding the winner.

Bookies think in terms of odds and in a perfect book with no overround, the odds are related to the probability through the well-known formula

Odds = (1 - p) / p,
a 0.25 probability shot is 0.75/0.25 = 3/1. Put another way,

p = 1 / (1 +Odds)

Putting all this together and having the odds of the winner as x and the odds of the second as y, the probability of the 1-2 result, using the formula above, is

$$(1/1+x) * (1/1+y / (1 - 1/(1+x)))$$

. . . which simplifies to $1/(x\ (1+y))$.

Converting this to odds using the usual formula gives the odds as

$$1 - ((1/x\ (1+y) / (1/x\ (1+y))$$

. . . which again luckily simplifies down to $x\ (1+y)$ -1.

If these are the odds for the CSF, the payout would be this plus the stake of 1, namely $x\ (1+y)$

To summarise, in a pure race with no overround and no adjustments for draw, tax etc, the CSF payout would be the odds of the first horse multiplied by one more than the odds of the second horse.

e.g. If a 12/1 shot beats a 3/1 shot, the 'pure' CSF would be 12 * (3+1) = £48 TO £1.

Another way of reaching this answer is to work out the return on the double and reduce it by the 'extraction from field' factor for the second race, as in

Return for race for winner = 12 + 1 =13
Return for race for second = 3 + 1 = 4

Return for 'double' . . . 13 * 4 = 52,
. . . to be reduced by the element of the probability covering the winning horse (1- 1/13)
52 * 12/13 = 48.

More generally, with x as the odds of the winner and y as the odds of the second horse, the CSF payout in this 'pure' race would be $(1 + x)\ (1 + y)\ (\ 1 - 1/(1+ x))$. This is the starting point for the BAGS calculation.

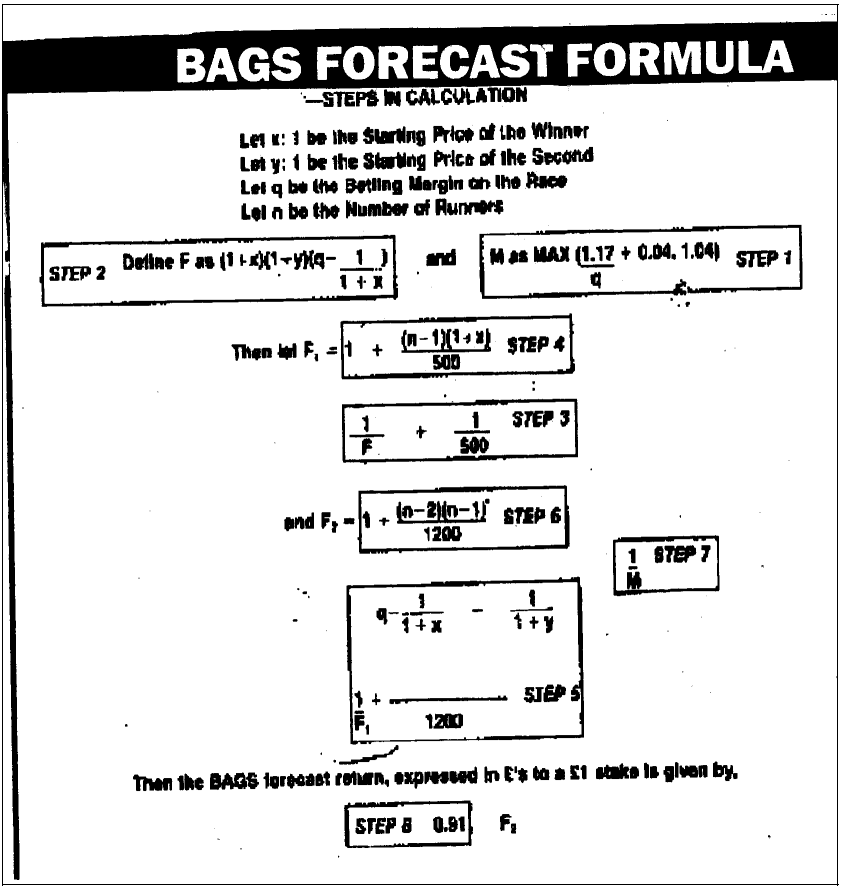
2. THE BAGS FORMULA

It helps to understand what I have referred to as the pure model in order to appreciate the steps (and insertions) which constitute the BAGS formula which, according to the Sporting Life of 11 June 1994 was as re-printed opposite.

Let me say at the start that neither the layout nor the printing do the reader any favours here and if it was not for an example which was added to the formula, it is unlikely that the reader would be any better off in understanding the process. However, with the benefit of the example, let us tiptoe gently through the CSF calculation for a 6 dog race where the SPs are

4/6, 2/1, 8/1, 12/1, 33/1 and 50/1

(a) The first thing to notice is that we have already moved away from the pure model, in that the total of the probabilities is not 1, but rather 1.17 (denoted by q in the



formula), i.e. there is an overround on this race of some 117%. However, a first stab at the CSF odds would reasonably be to take the formula at the end of section 1, namely the combined odds of the first and second dogs, adjusted by the extraction from the second horse field.

Since the total of the odds is now not 1 but 1.17, the first cast for the CSF would read, as it does in Step 2 of the BAGS model (we will return to Step 1 later):-

$$(1 + x)\ (1+y)\ (\ q - 1/(1+x))$$

So first stabs at the CSF payout, known as F in the published formula, are, say

2/1 beats 4/6 - 3 * 10/6 * (1.17 - 1/3) = £4.18
2/1 beats 33/1 - 3 * 34 * (1.17 - 1/3) = £85.38
50/1 beats 4/6 - 51 * 10/6 * (1.17 - 1/51) = £97.78
50/1 beats 33/1 - 51 * 34 * (1.17 - 1/51) = £1994.78

Note that at this stage, the payouts are linear, in the sense that the payout for 50/1 beats 4/6 is 23.39 times the payout for 2 beats 4/6, and the payout for 50/1 beats 33/1 is 23.39 times the payout of 50/1 beats 4/6. Also, the probabilities of horse A beating each of the rest of the field add up to the probability of Horse A winning outright.

Example

Table A

Forecast with (say) 2/1 winning	CSF = F	Odds of CSF (F-1)	Probability (1/ (Odds + 1)
2/1 beats 4/6	£4.18	3.18	0.239
2/1 beats 8/1	£22.59	21.59	0.044
2/1 beats 12/1	£32.93	31.03	0.031
2/1 beats 33/1	£85.38	84.38	0.011
2/1 beats 50/1	£128.01	127.01	0.008

Total probabilities = 0.333, the probability of a 2/1 shot winning. This demonstrates that at this stage, there is no currently overround built in to the notional race for second place. As we will see, this is not an acceptable solution to the people whose formula this is!

(b) What happens next is the first of the two *'Roulette'* adjustments (as the formula calls them) to the payouts.

Without the benefit of an official explanation, I can only guess that this adjustment is designed to reflect the reality that overround is not spread equally over all dogs, with those with the tightest odds having less scope for bookie profitability.

The use of the word 'roulette' suggests the Monte Carlo method for establishing likely probabilities, which in this case might involve running simulations of races with known SPs many thousands of times in order to see what the occurrence of various CSF combinations turn out to be.

However, the fact is that the adjustment has the clear effect of reducing all CSF combinations, but especially the larger ones. I suspect that it was this adjustment which the Racing Post was referring to when it said that the relative odds of the winner and second place in the Wokingham caused the CSF to be adjusted downwards, ensuring a typical blustering reply from BOLA.

The adjustment (Step 3) is that the previously calculated figure is amended to 1 / (1/F + 1/500)

Basically, this adjustment increases the CSF probability by a factor of 1/500th or 0.002. Tiny if the probability is already significant, in that 0.5 increased by 0.002 makes next to no difference, but something that is already a 500/1- forecast gets its payout halved by this tiny adjustment. The CSFs of our four selected possibilities are now

2/1 beats 4/6 - was 4.18,
now modified to 1 /(1/4.18 + 1/500) = 4.15
2/1 beats 33/1 - was 85.38,
now modified to 1 /(1/85.38 + 1/500) = 72.93
50/1 beats 4/6 - was 97.78,
now modified to 1 /(1/97.78 + 1/500) = 81.78
50/1 beats 33/1 - was 1994.78,
now modified to 1 /(1/1994.78 + 1/500) = 399.79

(c) The adjustment in (b) has reduced the probabilities of every forecast combination, including in particular, each of the 5 possibilities in Table A above.

The sum of the probabilities of the 5 separate forecasts now no longer add up to the probability of the winner.

The probability shortfall is 5
i.e. the number of forecasts * 0.002 = 0.01
. . and this represents in the case of a 2/1 winner a proportion of probability of 0.01 / 0.333 or 0.03. The CSFs involving the 2/1 winner should be increased by this proportion to ensure that the race for second place is still neither under- or over-round.

In general terms, and in a race of n runners, there will be n-1 forecasts involving a stated winner, and the probability shortfall will be $(n - 1) * 0.002 / (1/(1 + x))$,
which simplifies to $(n - 1)(1 + x) / 500$. This is the next step in the general calculation, signified by step 4 in the formula, whereby the new F created by step (b) is turned into F_1, calculated as $F_1 = F * (1 + (n - 1)(1 + x) / 500)$

This gives our new F_1 for our examples as

2/1 beats 4/6	4.15 * (1 + 5 * 3/500)	= 4.27
2/1 beats 33/1	72.93 * (1 + 5 * 3/500)	= 75.12
50/1 beats 4/6	81.78 * (1 + 5 * 51/500)	= 123.49
50/1 beats 33/1	399.79 * (1 + 5 * 51/500)	= 603.68

(d) A second and more subtle roulette factor is applied next in step 5. Whereas the first factor just arbitrarily reduced all probabilities by 1 / 500, this adjustment makes a similar reduction of all probabilities, but this time, various probabilities are reduced by different fractions. Specifically, the probability of each CSF is reduced by a factor of :

(1.17 (i.e. the race overround) less the probabilities taken out of the field by both the winner and the second) divided by 1200.

Again, I can only speculate on the thinking that led to it, but the form of the adjustment appears to concentrate on the field without the first and second dogs, and this is supported by the form of the corresponding adjustment in (e) below.

The effect of the adjustment is not only that larger forecasts

are affected to a greater extent than smaller forecasts, as in the first fixed roulette adjustment, but that the that small odds forecasts are reduced by an smaller absolute factor than the larger ones. For instance, the probability increase factor in the 2/1 beats 4/6 forecast is 0.000197, whereas the similar factor for the 50/1 beats 33/1 forecast is 4.7 times bigger at 0.000934. Returning to our examples, the figures after Step 5 are :-

2/1 beats 4/6
1 / (1/4.27 + ((1.17 – 1/3 – 6/10) / 1200) = 4.26
2/1 beats 33/1
1 / (1/75.12 + ((1.17 – 1/3 – 1/34) / 1200) = 71.50
50/1 beats 4/6
1 / (1/123.49 + ((1.17 – 1/51 – 6/10) / 1200) = 116.87
50/1 beats 33/1
1 / (1/603.68 + ((1.17 – 1/51 – 1/34) / 1200) = 386.01

(e) As with (c) above, this apparent allocation of probability towards the losers has created a distortion which should be corrected. It is here that I am at my weakest mathematically, but I can see that the reduction in (d) has been across all combinations and while I cannot yet satisfy myself as to its detailed mathematical propriety (any guidance welcome!) , I can see in principle that something like the next step is justified.

The adjustment in Step 6 of the formula increases the payout by a factor of 1 + (n - 1)(n - 2)/1200, and it may (or may not) be significant that in a race of n runners, there are (n-1)(n-2) forecasts which do not involve the actual winner coming first or second and the actual second coming second. In any event, in our 6 dog race, the adjustment amounts to 1 + 5 * 4/1200 = 1.0167, the effect of which is as follows:

2/1 beats 4/6	4.26 * 1.0167	= 4.33
2/1 beats 33/1	71.50 * 1.0167	= 72.69
50/1 beats 4/6	116.87 * 1.0167	= 118.82
50/1 beats 33/1	386.01 * 1.0167	= 392.45

(f) We are nearly there(!) The only thing missing is that the race for second has throughout been calculated as a 100% book with no under/over-round. This is deemed unrealistic and CSFs are at this stage reduced by a margin element for this second race in the deemed 'double' which comprises the CSF. Confusingly, this reduction is calculated as Step 1 in the formula, and this margin, known as M, is defined as

M = 1.04 or, if the overall margin q for the race is greater than 1.17, 0.04 + 1.17/q.

Thus if the actual betting margin for the race is 20% so q = 1.20, M, the deemed margin for the race for second, would be 0.04 + 1.17/1.20 = 101.5

In our example, the betting margin on the race is 1.17, so the deemed margin M for the race for second is 1.04 (i.e. 104% overround). Step 7 of the formula effects this new margin by reducing the payout by 1/M :

2/1 beats 4/6	4.33 / 1.04	= £4.16
2/1 beats 33/1	72.69 / 1.04	= £69.89
50/1 beats 4/6	118.86 / 1.04	= £114.25
50/1 beats 33/1	392.14 / 1.04	= £377.35

The 1995 formula has a step 8, which is to reduce further the payout by tax (10% in those days). Hopefully, the current CSF formula does not do this! So that was the BAGS formula, as least in 1995. The effect of the various margin and roulette adjustments can be seen from our worked examples, namely;

Forecast	'Gut reaction' payout based in SP of winner x (SP of second + 1)	First step of real calculation before roulette adjustments	Final CSF
2/1 beats 4/6	£3.33	£4.18	£4.16
2/1 beats 33/1	£68.00	£85.35	£69.89
50/1 beats 4/6	£83.33	£97.78	£114.25
50/1 beats 33/1	£1700.00	£1994.78	£377.35

For the BAGS formula at least, this demonstrates that at the

lower end of the scale, the detailed formula gives payouts which are higher than would be intuitively expected. However, the formula does have the effect of reducing markedly the intuitive payouts when the SPs of the two horses concerned increase.

2. THE HORSERACING FORMULA

As I say, I have not been able to lay my hands on the CSF formula used for horseracing, but I should be astonished if it involved a completely different methodology. I would however expect differences in the details, especially

- the 500 figure in Step 3
- the 1200 figure in Step 5
- the deemed margin to be incorporated in the race for second in Steps 1 and 7.

Who knows, if these specific questions are asked of BOLA, perhaps they might be persuaded to respond?

3. THE EFFECT OF THE DRAW

BOLA recently said that everyone knows that the effect of the draw is now incorporated into the CSF calculation. How everybody knows this is not clear, since I cannot trace any announcement or press release stating this. What there is of course is the press release of February 1996 which spelled out how the effect of the draw is to be incorporated into the now defunct tricast (naming the first three in correct order), but there is no mention in the package of any similar changes to the CSF.

However, I would again expect that having devised the tricast methodology, it would have been adapted as appropriate for the CSF calculation, so lets see what we can learn from the tricast adjustment. The February 1996 changes introduced several new concepts, as follows:-

a) The Stall Difference (SD), defined as the difference between lowest & highest stall numbers of the tricast.

b) The Stall Difference Parity (SDP), defined as one third of the number of runners, rounded up, with a maximum of 6. Applying the SDP to the field splits the field into 3 sections, which I have referred to below as the outer and middle 'thirds', although the outer thirds can never consist of more than the 6 highest- or lowest-drawn horses, regardless of field size.
c) The Price Parity (PP), defined for a tricast as the combined betting percentage of 3 4/1 shots = 60
d) The TakeOut (TO), defined as the combined betting percentage of the tricast selections eg 3/1, 9/1 and 20/1 would have a TO of 25+10+5 = 40
e) The Runner Factor (RF), defined at 70 plus 2 per runner, with a maximum of 106

Armed with these new concepts, the process of adjusting the tricast went as follows:-

I. Calculate the tricast in the normal way
II. Apply the draw adjustments if the race has 8 or more actual runners and the distance is 9f or less
III. Compare SD with SDP. Depending on the answer, the calculation goes off in three different ways

i) If SD is below SDP with all of the tricast horses in one or other of the outer thirds of the course, calculate the TO of the three horses and deduct it from the PP, 60. Calculate the difference as a percentage of the RF. The tricast is reduced by this percentage. Example, with the first three horses in an 18 horse race starting at 4/1, 9/1 and 9/1, the TO is 40, the RF is 106 and the tricast would be reduced by a factor of 106 * (60 – 40)/100 = 21.2%
ii) If SD is equal to SDP or below SDP with at least one of the horses drawn in the middle third of the course, there is no adjustment to the calculated tricast
iii) If SD is greater than SDP, a calculation is made of all possible tricasts. For each tricast with an SD below SDP, the potential reductions are totalled. This total of potential reductions is then spread over the number of

tricasts with SD above SDP. For example, of the potential percentage reductions add up to 1780.26 and there were 616 combinations where SD exceeded SDP, the winning tricast would be increased by 1780.26 / 616 = 2.89%

The press release makes a great deal of play about how recalculating all tricasts over a three year period led to 69% being increased by the new adjustments and only 7% being decreased, but this is not relevant to our analysis. Let us instead try and interpolate how this technique might be used currently to adjust the CSF formula. Basically, I would expect the same qualifying criterion, namely that draw adjustments will only be made in races of 8 or more actual runners and with a distance of 9f or less.

I'd expect the definitions of SDP and SD to be the same, although the latter only looking at the two horses concerned in the forecast. I'd also expect the PP to be the takeout of two 4/1 shots, i.e. 40. If this is correct, then the CSF will be adjusted downwards for draw effect if both first and second are in the same outer third of the draw, and if the combined takeout of the two horses is less than 40. I would expect the percentage reduction to be represented by the TO shortfall compared to 40 / the RF, as defined for the tricast.

Ideally, this article would have finished with a demonstration of a real calculation, perhaps even of the controversial Wokingham result where a 7/1 shot beat an 8/1 shot but since they were drawn 4 and 3 out of 30, the declared CSF was only £39.13. Unfortunately, there are too many missing figures from my understanding of the BOLA Horseracing version of the CSF, but if my guesses on the draw element are correct, this payout was reduced by 106.8 * (40 – 23.6) / 100, or 17.5%, because of the draw.

SMARTsig

Tony's courageous stab at cracking the CSF formula is to be applauded. Can any other members assist in getting us any closer to a solution?

- Stef

Rather than dismissing the horserace meeting big-bet reports as 'from an unknown source', Bob looks at what can be gained.

BIG HITTERS

Bob, Edinburgh

JJ Egan, writing in the October issue disagrees with my view that the biggest bets placed in the on-course betting market were made by professional backers.. Such bets are probably bookmakers' (particularly Ladbrokes) hedging money, says JJe.

Quite right, says Stef agreeing with JJe, and 'if we checked all bets of £3,000+ from individuals, the majority of these would not be from punters whose main income is from betting.'

In the first place it seems to me that such remarks are entirely guesswork and no reasons were given why I should prefer such a version to my own. Take it or leave isn't good enough for me. On the other hand, while I admit that my opinion is speculative I at least tried to rationalise what I wrote.

Nevertheless, perhaps I should have explained my reasoning in greater detail.

A bookmaker places money in the ring for many different reasons. First, there are the theoretically correct book-making reasons.

- SP layers may wish to bring down the on-course price of a particular animal in order to reduce their off-course liabilities.

- The bookmaker has a liability on a horse larger than he wants and decides to lay this off with others.

- The bookmaker may want to take a profit. For example, if he has taken a bet at a lower price than he can obtain elsewhere then by betting this he crystallises his profit.

There are however also theoretically incorrect reasons for a bookmaker betting.

- The bookmaker might simply want to bet an animal which he himself believes will win.

The big drawback for a bookmaker betting himself in the ring is that he is reducing his own takings. It does not ultimately matter whether it is money from a 'face' or money from a 'mug'. The point is that a reduced turnover is the end product of doing so. Sure, the margin may be different on the different kinds of business but ultimately, there is a margin and it is the bookie's job is to ensure that the margin is in his favour.

Incidentally bookmakers do not normally balance their books in the way that the layman punter thinks. Obviously bookies normally have their largest liability on the favourite, less so for the second favourite and only dribs and drabs on the remainder. Few, if any, real-life punters want outsiders to big money, because if they do no such bets are ever reported! This means that individual books rarely are in balance.

Bookmakers regard turnover as king and bookmaking is essentially longer term business. Just like pro punters, bookies rely on their winners paying for their losers in the course of a year. The over-round works in their favour in the long term. It does not guarantee the ring a profit in every race, or even every meeting. This is a very naïve view of the business.

Barry Dennis for one has revealed that he doesn't waste time on working out the margin every time he bets. And a great many other bookmakers acknowledge that what they are after every time is income and they never, but never,

lay money elsewhere. To do so is tantamount to reducing their turnover.

It is different with companies such as William Hill and Ladbrokes who know what they believe the exact price of a horse ought to be. They certainly aim to achieve a balanced market but there is absolutely no reason for them to fine tune it to the extent punters believe that they do. They too depend on the long term factor for their profitability.

In any case the chains do not need to place large sums to bring the price down to the required level. In the normal way, as I think it was Phil Bull who confirmed, they only need to threaten to bet money with the market leaders to bring the prices down to the required level. Their actual bets are not normally obvious from betting reports other than by being referred to as 'office money'.

All that said, it seemed immediately obvious to me that neither JJe nor Stef took my advice to look in detail at the bets I mentioned in issue 8.09.

I am not going to go through the whole list of ten but perhaps if space permits we can look at a couple of the ten grand jobs, one winner and one loser.

£10,000 was bet on CURRENCY to win £17,500.

This was Raceform race no. 2753, a 16 runner, 6 furlong sprint handicap on a Thursday at Haydock. The horse was the most popular press choice and Postmark's nap. It was forecast by the Post to start as 9/4 favourite with the second favourite at 9/2.

A pretty clear form pick, one would think. Now what was the pattern of betting and how did the ten grand fit into the picture?

Well here are the first half dozen bets reported in the Post.

£2000-£1000	i.e. 2/1
£1000-£500	2/1
£3750-£2000	7/4
£2625-£1500	7/4
£17500-£10000	7/4
£3500-£2000	7/4

And after many more bets of smaller amounts (from market sheep with all the value gone?) the horse started at 6/4.

Please note that the ten grand was simply part of a series, albeit the largest in the series! It was the overall weight of opinion that drove the price down. Why Ladbrokes would want to bet a disproportionate ten grand on such an animal to no effect is not clear to me and I can't accept that this bet was a chain bookmaker acting in the market.

A Ladbrokes bet to reduce its own liability on the horse would be a different matter. Frankly, it would also be incredible. No bookmaking company would manage its business in such a way that it would have to do this. (The company would be reducing its turnover in effect.) Also since the horse got beat, as a Ladbrokes shareholder I'd like to think that someone got sacked for it!

Turning next to a ten grand winner.

Race 2842 at Sandown was a two-year-old maiden event for 4 runners. BRAGADINO won at 4/5.

Again, this was the most tipped horse, Racing Post's *Spotlight* pick, and at 9/4, the forecast favourite. Since the 3/1 third favourite, CORTON, 'the likely danger' according to Spotlight was withdrawn you would expect the betting to reflect this which it did.

The betting reported went as follows.

£11,000-£10,000	11/10
£5,500-£5,000	11/10

then 10 more smaller bets at 11/10 then another whopper;

Even £10,000 evens

. . . followed by an avalanche of smaller bets, including "office money".

So the on-course betting did reflect expectations. I am afraid that it is a feat of the imagination to visualise these big bets as bookmakers manipulating the market. There is simply no need for it.

I do of course concede that it is entirely possible that a chain bookmaker could have laid these bets. But this would only be in the case of the chain wanting to divest itself of liabilities in respect of money from faces. In such a case this is effectively money from the faces themselves, albeit at second hand. This seems an unlikely event to me, however, on the grounds that the chains have a marked dislike of big bets from such punters.

If you care to examine the betting patterns as I have done you will see big volumes as you would expect at the big festivals, public holidays and at the weekend. Personally, I have no difficulty in recognising many of the big bets on the hyped horses in the big races at such times as basically mug punter money. It's how the bookies make their living after all.

The real crunch for my critics are the big bets in the small races. These need explaining.

Two bets in excess of £10,000 were reported recently in one small egg & spoon race at Musselburgh, my local track. And I mentioned previously that the second biggest bet this year, £50,000 was in a four runner event in an evening race up at Hamilton.

Ladbrokes hedging money? With all due respect, I think this is nuts.

Many Scots punters can guess who it was that was doing the betting. Surely this is the same with local punters all over the country. If so, I rest my case.

Today it's the 5th October. So perhaps a glance at the Racing Post reporting yesterday's racing at Newmarket might help.

The first race was a nursery handicap for 19 runners and I would have been astonished to see a pro putting his money down. Correct. There weren't any big bets in the race and the 13/2 favourite could only finish third.

The 2.20 was won by Aidan O'Brien's BECKETT, the 3/1 favourite. One punter had two grand on the winner to win seven. Good luck to him but this would have been a huge price for a pro. O'Brien and the Spotlight nap so why wasn't there more money for it? Ask yourself. Go on. Go on. Ah, go on. (I thought it was too tough to call.)

The third race, JOHANNESBURG won at 30/100. The biggest bet was £18,500 to win £5,550. (Guess who?) This is an exceptionally large stake but it was never in danger. Let's not beat about the bush on this, the winning punters got the value. The losers didn't! If you appreciate this, then you are half way there. Again, the chains didn't need to intervene in this market to drive down the price.

The 3.25, a listed race, was won by a 33/1 shot, INDIAN PRINCE. At first sight, this belongs in the category of lucky outsider. It was friendless in the tipster box and Spotlight said it had plenty to find.

No way could any rational punter pick it but astonishingly there were a number of decent bets on this with one punter betting £10,000 to £150 each way. I expect the systemites and the tipster ads will be claiming it but the real lesson from this race, though, is that the winner is exactly the kind of horse that is a must for the notebook!

In the fourth I did bet TISSIFER to win the 27 runner claiming race at 2/1. It did. There were a number of bets around two grand for this winner. The case for betting in claimers was made recently by Alan Rogers in issue 8.08 and as far as I am concerned Alan's research was rock solid.

TISSIFER would have been a confident bet for me at evens. At 2/1 it was cracking value.

MILLENNIUM DRAGON won the 4.30, a 21 runner maiden, but with so many unexposed horses in it I couldn't see any justification for betting in the race. The punters made the un-raced O'Brien/Kinane CENTURY CITY favourite at 13/8. You won't get a clearer example of hype making the favourite! Still, one big hitter (and I can't believe that it was 'guess who?') had four grand on to win eleven on the winner!

Two lessons from this race:

- The hyped stable doesn't always saddle the winner.

- The big money isn't always on the favourite!

The last race, the Getting-Out Stakes, was a 21 runner five-furlong handicap and no place for a risk fearing punter like me. Spotlight fans would have been delighted with his comeback pick of the 11/1 winner, KINGSCROSS.

And yes, at Nottingham on the same day I was on ASHNAYA the 2/1 winner of the 2-miler by 7 lengths. It really couldn't be missed although Spotlight napped ORANGERIE and made it his forecast favourite too. ORANGERIE was beaten 13 lengths.

The punter who does study the big bets as suggested in issue 8.09 will achieve an entirely new level of understanding.

There are three immediate advantages.

- Such a punter always bets confidently. This is because he knows (statistically and no guessing about it) that in any race he is more likely to win than to lose!

- It follows then that he will also always know that, win or lose, he was on the statistically 'right' horse! How many can claim that?

- Finally, such a punter gains an extra edge in recognising value in a really big way.

The punter with all three advantages can't help but make money. But it is something that everyone has to do for themselves and I know from long experience that only the dedicated few will actually manage to do it!

SMARTsig

Our 2001/02 football season attempt to make a profit from soccer superiority spread betting gets underway.

SOCCER SUPERIORITY SPREADS

SMARTsig

This is the first appearance of this feature this season and will be updated every month throughout the current campaign.

Very few games to assess as yet, in keeping with our policy of not getting involved in any games where either of the two teams have played fewer than 8 league games during the current season.

continued over the page . . .

Superiority spreads in the main are quoted for all Premiership games. Televised Nationwide league matches are also quoted but these are in the minority. Our match ratings are calculated from each team's record in **league** games. The only superiority quotes we're interested in therefore are those for league games - **superiorities quoted for European and domestic Cup competitions are NOT played.**

A trade is indicated in cases where the SMARTsig match rating (see formula given in *Up Front*, page 10 of this issue) allows us a minimum of a 0.4 goal advantage over or under the spread quote.

Example;
Aston Villa v Charlton, superiority quote of 0.8-1.1 goals
In order for a trade to qualify in the above game our match rating would need to be 0.4 (or more) goals above the high quote or below the low quote. If the SMARTsig match rating was 0.4 or less, we could SELL, a rating of 1.5 or above, then we'd BUY. All match ratings between 0.5 and 1.4 are too close to the quote and would be NO TRADE.

Under these rules, only one game has qualified to date, and it has given us a winning start!

		Pl	Pts		Pl	Pts	Match Rating
20 Oct 2001	Man Utd	8	17.4	Bolton	9	12.6	0.775

(17.4/8) - (12.6/9) = 2.175 - 1.4 = 0.775

Applying the conversion formula, 0.44 + (rating * 0.8) gives a projected home superiority of 1.06

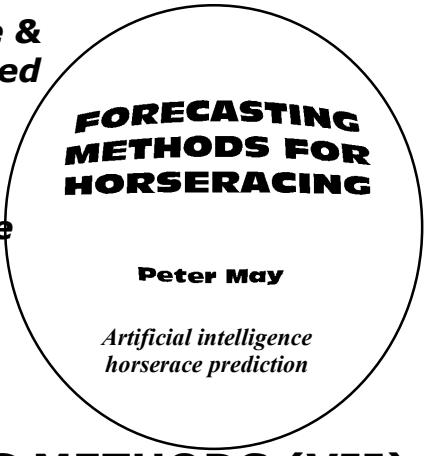
SpreadEx quoted 2.3-2.6 in Utd's favour giving us a 1.24 edge to SELL the 2.3

The game result Man Utd 1 Bolton 2 gave us a final make-up of -1, and a +3.3 profit.

. . . updated again next month . . .

Artificial Intelligence & Neural Networks applied to horseracing.

More from Chapter Three examining race analysis methods.



AI FORECASTING METHODS (VII)

Peter May

Chapter 3 continued:
The Effect of the Draw

The size and shape of Britain's racecourses varies enormously, from the small almost circular Chester to the figure of eight course at Windsor. One attribute that all courses share though is the use of starting stalls. These were introduced in Great Britain in the 1960's after continued pressure from the founder of the Timeform organisation, Phil Bull.

The aim of starting stalls is to produce an even break for all the runners, which is especially important in sprint races. The allocation of stall numbers to the runners at the overnight stage determines their starting stall positions, and to a certain extent the part of the track over which they will run: near side, far side or centre.

Obviously, the ultimate position of the horse relative to the other runners and the running rail is determined by the jockey, however the stall number (or *draw*) plays an important part in the positioning of the runners. And this can have a marked effect on the eventual race result.

There are several ways in which the draw can affect the outcome of a race.

Firstly, moisture is not necessarily retained uniformly across the width of the course, consequently the horses running on the softer, or slower, side will be at a considerable disadvantage. Naturally, non-uniform moisture retention is especially apparent on cambered tracks with the lower part of the course retaining water for a longer period of time. Such variability in the ground conditions becomes more significant in sprint races where horses tend not to group together. Over longer distances the runners tend to bunch together soon after the start thus minimising the effect of the softer ground.

Secondly, horses often run better when positioned close to a running rail. The rail prevents the horse from wandering across the track (at least in one direction) keeping its path to the line straighter. And thirdly, on round courses, horses drawn on the inside will benefit from taking the shortest route from starting stall to finishing line. All of the these factors can influence the outcome of a race.

The effect of the draw is published in most racing publications (daily papers, formbooks etc.), however, whilst these assessments indicate whether high or low numbers are preferred, they do not attempt to quantify the effect.

Consequently, it is difficult to incorporate this information into forecasting models. For example, if for a particular course high numbers are considered to hold the advantage, would a horse starting from stall 15 in a race with 20 runners be classified as well drawn?

Furthermore, how much is the horse in stall 20 favoured over the horse in stall 10? To answer these questions, and provide information which can be included into forecasting procedures, it is necessary to know by how much one stall is favoured over another in terms of distance (i.e. 2 lengths better).

In order to quantify the draw effect it is necessary to analyse a sample of races for every race distance at each track. The population from which this sample is selected for analysis in this section constitutes all non-juvenile handicap races run during the four seasons from 1993-1997.

Handicap races are chosen in preference to non-handicaps since the ability range of the runners is normalised by the allocation of different weights by the BHB handicappers, therefore only a small external effect will make a noticeable change to the result of the race. In non-handicaps the ability range is likely to be much wider, and even with the best draw of the race and a ten lengths start a large proportion of the runners would still fail to win.

In other words, the draw will not, in most cases, be a significant determining factor in non-handicaps, the crucial element being the ability of the horses. A similar argument can be applied to nurseries, although not to quite the same extent. Unexposed two-year-old horses are exceptionally difficult to handicap and therefore the differing abilities will not be accounted for as well as they are for older horses.

Whilst an analysis of the success rates of horses running from each stall will indicate whether high or low numbers are favoured on a particular track, this will not provide an explicit numerical value of the advantage afforded to each runner. In order to determine the effect in terms of distance advantage it is necessary to consider the distance between the horses at the finish by stall number.

However, simply analysing the average distance each horse is beaten by the race winner on a stall-by-stall basis would yield biased results. Since the allocation of stall numbers always starts from one, in small fields the winner must be drawn low; and in large fields for the horses drawn high to win they must overcome more opponents.

It is possible to remove this bias, though, by considering the distance each horse is beaten by the runner drawn in a

specific stall, say stall one. On this basis the average distance beaten for all horses in stall one will be zero.

For the other stalls, a positive average distance beaten indicates that these stalls are disadvantaged by the draw compared to stall one. A negative value indicates a draw advantage compared to stall one.

In figure 3.14 this analysis is presented in graphical form for the 5 furlong course at Chester.

From figure 3.14 it can be seen that, on average, the horses drawn high are at a disadvantage compared to the stall one runner. For instance, the horses in stall 16 beaten, on average, by about 6 lengths by the horses running from stall one for the races analysed.

This is not surprising since Chester is a tight, circular left-handed course, and consequently the horses drawn wide have further to race than those drawn on the inside. Fitting a trend line to these data indicates that the disadvantage to horses drawn in stalls other than stall one is equivalent to 0.45 lengths per stall. In other words a horse drawn in stall 5 has a disadvantage of approximately 2¼ lengths (i.e. 0.45x5) to overcome which equates to 6¾lbs using the distance to weight conversion given in Chapter 2.

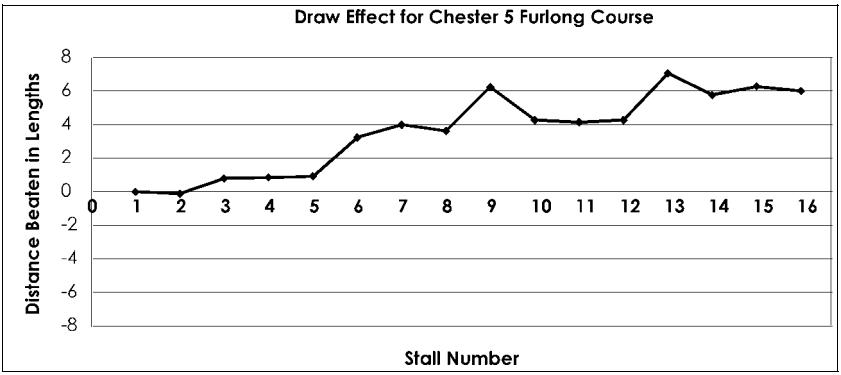


Figure 3.14: Analysis of the draw effect

On some tracks the positioning of the starting stalls and the state of the going can also effect the draw. However, this is easily accounted for in the analysis by isolating particular track conditions and considering each set separately. For instance analysing races run on soft or heavy ground separately from fast ground events.

The average winning distance in 5 furlong handicap races (i. e. between the winner and the second) is a little over 1 length, so a draw effect similar to the one at Chester is clearly significant. Once quantified this factor can either be included in the model as a separate variable or used to adjust the ability ratings of the runners.

This factor is of such significance though, that it can be used without considering any other data to form a successful prediction model, and this is discussed further in Chapter 5.

Fitness

Not all horses are 100% fit to race when they reach the starting stalls. In the trade press trainers will often be quoted as saying their horse will "need the run" before it returns to the course after a break, and jockeys regularly report that particular horses will "improve for the run" after an introductory race.

Many people find this hard to believe; why would a trainer enter an unfit horse?

Apparently, it is extremely difficult to get a horse to complete race fitness on the gallops and hence the first race back after a break forms part of the training programme. However, this brings into question whether the horse is trying 100% to win. After all a horse which is considered to need the run is clearly not running to its best ability since it is short of peak fitness. Furthermore, the jockey is unlikely to give a partially fit horse the same ride as a fully fit animal and will tend to handle the horse more tenderly.

To support this theory that a horse may lack race fitness after a break from racing, two analyses of race performance against time between races is given in figures 3.15 and 3.16 for juvenile and non-juvenile runners.

Days Since Last Run	Number of Winners	Number of Runners	% Strike Rate
1-7	272	2096	13.0%
8-15	1042	8897	11.7%
16-21	624	5331	11.7%
22-28	412	3546	11.6%
29-56	479	4793	10.0%
57-84	98	1131	8.7%
85-200	51	756	6.7%
201+	0	6	0.0%
All	**2978**	**26556**	**11.2%**

Figure 3.15: Analysis by days since last run (juveniles)

Days Since Last Run	Number of Winners	Number of Runners	% Strike Rate
1-7	1861	15732	11.8%
8-15	3822	38607	9.9%
16-21	2048	20525	10.0%
22-28	1267	13230	9.6%
29-56	1499	16382	9.2%
57-84	377	4527	8.3%
85-200	682	8755	7.8%
201+	444	6878	6.5%
All	**12000**	**124636**	**9.6**

Figure 3.16: Analysis by days since last run (non-juveniles)

Clearly, it is preferable for the horse to have run recently, and consequently this time variable should be included in any prediction model. However, there are other factors which can be considered when evaluating the fitness credential of a runner.

Firstly, some trainers seem more able at getting a horse back to peak race fitness after a period of inactivity.

For example, John Gosden has an overall success rate of about 20%, however, with horses returning from a break of 100 days or more this strike rate improves to almost 24%. Clearly, John Gosden is capable of producing horses fully fit after a break from racing.

In contrast David Elsworth's strike rate of approximately 11% drops to just over 4% for horses unraced for at least 100 days. In appendix A2 a list of trainers is given with their success rates for all horses and those unraced for more than 100 days.

It is possible to combine these two factors to produce an effective guide to the likely fitness of horses. The trainer's data can be categorised into four groups: unknown, poor, moderate, and good, which refer to the trainer's ability to produce a horse race fit after a break.

Defining these categories is somewhat arbitrary and the chosen method in this text is to use the ratio of success rates given in appendix A2.

Specifically:

$$\text{Ratio} = \frac{\text{Success rate for horses running after a break}}{\text{Success rate for all horses}}$$

Using this ratio the trainers can be classified as follows:

Trainer Classification	Rule
Good	Ratio >= 1.0
Moderate	0.5 <= Ratio < 1.0
Poor	Ratio < 0.5
Unknown	Runs of horses returning from a break < 10

Figure 3.17: Trainer classifications

Additionally, the time variable can be divided into two groups by introducing a time threshold. Horses which have raced more recently than the threshold number of races are considered to be race fit, the fitness of the remainder is considered questionable. The threshold for the following examples has been set at 57 days. Using these categories the following rules can be generated:

if *number of days since last run < 57*
then *fitness is OK*

if *number of days since last run < 57*
and *trainer's classification is good*
then *fitness is OK*

if *number of days since last run < 57*
and *trainer's classification is moderate*
then *fitness is doubtful*

if *number of days since last run < 57*
and *trainer's classification is poor or unknown*
then *fitness is unlikely*

There are now three categories of fitness: *OK*, *doubtful* and *unlikely*. An analysis of these conclusions by success rate for all race types is given in figure 3.18. The horses for whom fitness is not in doubt seem to have a higher success rate than the others, with the unfit horses performing quite poorly.

Fitness Classification	Number of Winners	Number of Runners	% Strike Rate
Fitness OK	13812	133179	10.4%
Fitness doubtful	916	9869	9.3%
Fitness unlikely	1166	18013	6.5%

Figure 3.18: Analysis of fitness classification

Naturally, there are special cases. Some horses are easy to get fit and always seem to run well after a break. Information which can help isolate this type of horse is contained in the race commentaries. In each race commentary a reference is included regarding the paddock appearance of the horse.

Fit horses are given the comment *looked well*, whist those in need of the race would be reported as *backward*. An analysis of previous paddock comments for a specific horse would indicate whether the horse is likely to be fit or not following a long absence for the track.

In general, though, the above rules appear to classify between the three categories of fitness quite well and can be easily incorporated into forecasting models.

Improvement and Previous Success Rate

Unexpected results of horseraces can usually be explained by one or both of the following reasons:

a) horses performing unaccountably poorly, and
b) horses running well in advance of their public form.

There are many reasons for horses running poorly, but broadly these can be categorised into two groups: unsuitable race conditions and physical disability.

For instance, the horse may be unsuited by the race distance or the going, or the horse may be suffering from some illness which cannot be detected until after the race. The reverse can be said for horses running better than before. This improvement can be a result of the horse encountering ideal racing conditions for the first time, or as a consequence of physical improvement.

Younger horses, two-year-olds and three-year-olds, are improving throughout the season, due mainly to their continued physical development.

However, determining which of these will improve more than usual is not an easy task. Fortunately, help is available via the published ratings.

Both *Raceform* and *Timeform* supplement their ratings with indicators to alert the subscriber to the probable improvement likely from some animals.

Raceform use a "+" symbol, whereas Timeform use a "p" to indicate that greater than normal improvement is likely, and a "P" to flag horses which they believe are definitely capable of better.

The horses most likely to show improved form are those which are difficult to rate. For example, unexposed animals which have only raced a few times, or those which are unbeaten. Unbeaten animals are difficult to rate because an assumption of the superiority over the other runners needs to be made since a horse of this type has not necessarily been fully extended.

Horses with the potential to improve on previous ratings need to be considered very carefully and would normally be selected over horses with similar race credentials but without the same degree of potential. In a study of non-juvenile pattern races it was found that the success rate for improving horses was 1.8 times greater than for all other runners.

The previous success rate of an animal (i.e. the number of times the horse has won compared to the number of times it has raced) is linked to improvement. Specifically, a previous success rate of 100% indicates that the horse is as yet unbeaten which is one of the criteria given for potential improvement.

However, it also shows whether a horse is capable of winning a reasonable proportion of races.

In figures 3.19 and 3.20 the success rate for horses with

differing historical strike rates is presented.

From these tables it can be seen that the likelihood of success in a race can be predicted by the previous success rate, and therefore, this component can prove useful in the development of forecasting systems.

Previous Success Rate	Number of Winners	Number of Runners	% Strike Rate
0..10%	1988	19793	10.0%
11..20%	208	2014	10.3%
21..30%	136	1158	11.7%
31..40%	183	1319	13.9%
41..50%	235	1251	18.8%
51..60%	10	59	16.9%
61..70%	47	219	21.5%
71..80%	11	41	26.8%
81..90%	0	0	-
90..100%	160	702	22.8%
All	**2978**	**26556**	**11.2%**

Figure 3.19: Analysis of previous success rate for juveniles

Previous Success Rate	Number of Winners	Number of Runners	% Strike Rate
0..10%	2988	38433	7.8%
11..20%	1988	19034	10.4%
21..30%	728	5716	12.7%
31..40%	340	2393	14.2%
41..50%	174	1127	15.4%
51..60%	19	111	17.1%
61..70%	30	166	18.1%
71..80%	10	52	19.2%
81..90%	2	8	25.0%
90..100%	62	255	24.3%
All	**6341**	**67295**	**9.4%**

Figure 3.20: Analysis of previous success rate for non-juveniles

The VDW methodology centred around the ability or class of a horse. Is this ability alone enough to give us all the answers we're looking for?

DOES ABILITY WORK?

Hedgehog

For some time now I have been a believer in the VDW approach. Anyone that can claim 29 winners from 32 selections must know something! In support of this view, VDW follower John Bingham also claims a very high strike rate. If all this is true there must be something to these VDW ideas!

In the VDW methodology, to my understanding, *ability* is a key component. The Ability rating is total winning prize money divided by number of wins divided by 100 with the decimal part ignored. For example if a horse has won 2 races one for £12,550.00 and the other for £13,530.00, it has an Ability rating of 130.

My expectations were that the better the Ability rating the greater the chance of success but I realised I had never tested this assumption. This article is the result of that testing.

I tested Ability on a sample of 213 races in the period 1st May 2000 to 16th September 2000. This sample involved Flat races with a penalty value of £10,000 or above and NH races with a penalty value of £7,500 or above.

The sample was restricted to the first 5 named in the betting forecast in the Racing Post and encompassed 1,042 horses.

Of this sample the first 5 named in the betting forecast found 160 winners which represents 75.1% of races with an ROI for betting all in the first 5 of -0.2175.

This is very similar to the results from a previous article. I considered Ability in 2 ways, by absolute value and by relative value.

For *absolute* value I tabulated the results for ranges of Ability value. For example I looked at all horses in the sample with an Ability rating of 1000 or better.

For *relative* value I tabulated the results in ranges governed by the percentage of the highest Ability in the first 5 named in the betting forecast. For instance if the highest Ability in the first five named was 100 and the next highest was 80 then I would tabulate the result for the 80 in a range including 80/100 = 80%.

Let me explain the column headings of the following tables:

Range - Range considered whether absolute or relative.
W - Winners for range.
R - Runners for range.
SR% - Percentage strike rate of range (given as W/R).
Av. Odds - Average winning odds.
ROI - Return on investment after 9% tax where 0 is break even.

The Absolute Table

Range	W	R	SR%	Av. Odds	ROI
1000+	3	10	30.00	1.861	-0.213
1000 - 500	9	15	60.00	2.892	+1.142
500 - 250	9	31	29.03	4.511	+0.468
250 -100	42	294	14.29	3.934	-0.353
100 - 50	48	349	13.75	4.805	-0.268
50 - 25	42	315	13.33	5.267	-0.233
25 - 1	1	11	9.09	2.500	-0.708
0	6	17	35.29	6.167	+1.321
All	160	1042	15.36	4.555	-0.2175

After creating this table I was pleased by the results. In general the better the Ability rating, the better the strike rate. This clearly shows that Ability is a good measure of success.

In fact blindly backing any horse with an Ability rating of 250 or more would result in a strike rate of 37.5% and an ROI of +0.527. That's almost 53p profit for every pound staked!

The two things I found curious were the success of zero Ability horses and the lack of it for very high Ability horses. The zero Ability horses by the nature of the sample had to be in the first 5 named in the betting forecast, so they must have shown good form. Yet because they are maidens they go off at longer odds then they should.

The highest Ability range horses still have a good strike rate but make a loss. This is because the betting public recognises them as good things and send them off at very cramped odds.

With such a powerful indication that Ability was a good measure I constructed the relative table with high hopes of a similar result.

The Relative Table

Range	W	R	SR%	Av. Odds	ROI
100%	44	216	20.37	4.433	+0.015
99.95% - 80.00%	19	129	14.73	4.588	-0.245
79.95% - 60.00%	26	172	15.26	4.224	-0.276
59.95% - 40.00%	33	219	15.07	4.145	-0.289
39.95% - 20.00%	25	205	12.20	5.573	-0.265
19.95% - 0.05%	7	84	8.33	3.375	-0.666
0.00%	6	17	35.29	6.167	+1.321
All	160	1042	15.36	4.555	-0.2175

When I looked at this table for the first time I felt that Ability as a relative measure was not as clear an indicator as in the absolute table. Again in general the higher the relative Ability, the better the strike rate. But the profitability has been reduced to an ROI of +0.015 on the highest relative Ability horses. By the nature of the exercise there would always be a 100% horse for each race but the result doesn't quicken the heart.

Thinking about these tables, I remembered that VDW used Ability in conjunction with the 3 form horses. Maybe this would help with profitability of the relative table.

The 3 form horses are defined as those horses with the 3 lowest form rating. The form rating is determined by adding the last 3 form figures together with any result above 9 taken as 10. So a horse with form figures 321 would have a form rating of 3+2+1=6.

I retabulated the two tables ignoring data for horses that were not form horses.

The Absolute Table for Form Horses

Range	W	R	SR%	Av. Odds	ROI
1000+	3	9	33.33	1.861	-0.125
1000 - 500	8	13	61.54	2.691	+1.084
500 - 250	4	20	20.00	2.375	-0.381
250 -100	34	198	17.17	3.843	-0.237
100 - 50	34	247	13.77	4.420	-0.316
50 - 25	34	229	14.85	4.977	-0.186
25 - 1	1	9	11.11	2.500	-0.643
0	4	8	50.00	4.000	+1.294
All	122	733	16.64	4.142	-0.215

The general trend seems to be increased strike rate but reduced average odds.

The 500-250 range performed very badly yet backing blindly any horse with an ability rating of 250 or more had a strike rate of 35.71% and an ROI of +0.127. Only 13p profit per £1 invested but still better than a building society.

The zero Ability horses have a strike rate of 50.00% and an ROI of +1.294. Not bad.

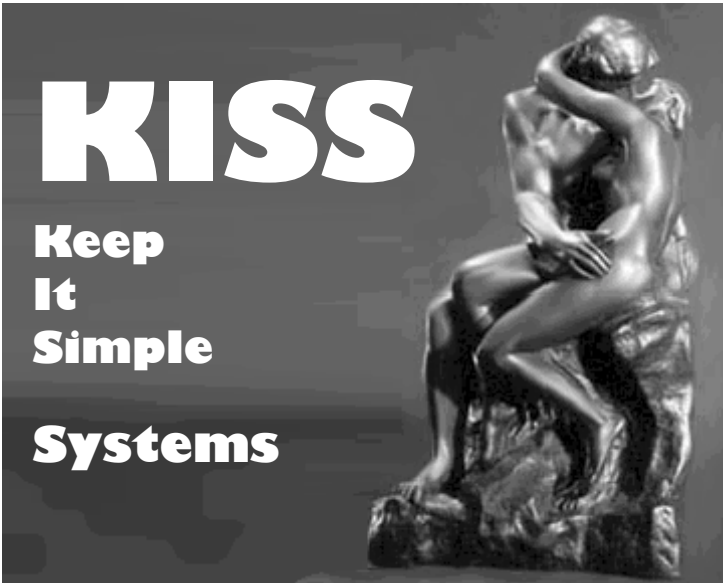
The Relative Table for Form Horses

Range	W	R	SR%	Av. Odds	ROI
100%	29	142	20.42	3.714	-0.117
99.95% - 80.00%	12	93	12.90	4.153	-0.390
79.95% - 60.00%	20	118	16.95	4.050	-0.215
59.95% - 40.00%	30	169	17.75	3.910	-0.201
39.95% - 20.00%	21	149	14.09	5.349	-0.179
19.95% - 0.05%	6	54	11.11	3.521	-0.539
0.00%	4	8	50.00	4.000	+1.294
All	122	733	16.64	4.142	-0.215

Again the general trend seems to be increased strike rate but reduced average odds. The tiny positive ROI for 100% has now become negative.

Conclusion

VDW used the 3 form horses and Ability to narrow the field of interest. I think that I have shown that Ability is a good measure of a horses chance of success. Also filtering for the form horses improves the strike rate.



Reviving the five-day-method in last month's issue also revived interest from the member who gave us the idea in the first place. Good to hear from old friends again.

FIVE-DAY FORTUNES

Jim Streek

I was pleased to find that so many have benefited from the 'five day method' which was discussed by so many and featured last issue. Not surprising though as it has performed so well for so long.

Credit must also go to Philip Alexander who helped further the system ideas, and of course to the Daily Mail's Formcast. Without hesitation I say that I would find betting on horses difficult without those ratings.

For the benefit of new members the rules that I put forward were as follows;

1. 1st., 2nd., or 3rd. last time out.

2. That run was within the last five days.

3. Top rated by Formcast (Daily Mail).

4. 1st. or 2nd in betting forecast. (joint 2nd included)

5. All codes of racing including all-weather courses.

6. Handicaps only.

I note that some are using non-handicaps, the extensive research I carried out at the time showed that selections in these races just about broke even, so were discarded for my purposes.

Several backers have tried variations using the Daily Mirror, Racing Post, etc, some use runners racing within seven days and so on, all report good results.

Unfortunately due to poor health my records were interrupted for some time, but I did restart my observations from 16th July this year.

The tax-paid profit up to October 3rd is 23.50 points. I would echo the remarks of Doug Luscombe who said "The beauty of this system is the long priced winners." The reason is that many selections are in big field handicaps, even so, it is unusual to have losing runs above five.

One alternative variation that I would strongly recommend considering is selections that comply with all the rules except number four, that is not dismissing those that are outside first or second in the betting forecast.

I have not kept an accurate record until now, but I have noticed many good winners. Today (October 5th) there were two such examples. One of these lost, but the other was Swynford Welcome running in Newmarket's 5:00 race. 4th

GIVE US A KISS!

Send SMARTsig YOUR simple system for publication!

in the betting forecast of this 28-runner handicap, with a 6lb penalty, and won at 14/1. Whatever way you use the method I am confident that it will continue to deliver the goods.

If I have overlooked anything or there are any questions I will be pleased to answer any letters sent to me via Stef

Raceform on Saturday

Whilst I'm writing I'll take the opportunity to congratulate the *Raceform on Saturdav* paper, they have certainly got their act together. Note the tips on page two under

- 'PICK' (Days best bet)
- 'Sure Things' and
- 'Key selections'

80% winners recently, I will keep a record and report later,

Until next time the best of luck and don't forget, Keep It Simple!

SMARTsig

SMARTsig Members-Only Email Discussion Group

Be there as our very own 'docu-soap', the day-to-day lives of betting folk, unfolds.

What did Nick Mordin say about Lasix? Who will Patrick upset next? Will Gino ever find out who the Ram Jam Band are? And obscure replies from that Frenchman who used to play for Manchester United . . . or at least someone who sounds very much like him!

Episodes EVERY day, the saga continues . . . Join up at www.smartsig.com

More reduced guarantee block building from John Norris who is forever digging for the perfect plan to help small stakes stretch ever further.

BUILD YOUR OWN GUARANTEE BLOCK part III

John Norris

So far I've tried to show how the unconditional 12X guarantee block yielding an unequivocal "one error minimum" guarantee can be developed into a "conditional" guarantee block. Such blocks may be based on either a two-way forecast or simply the initial assumption that no more than one particular type of result will occur.

One example might be the following block, which can be used against either four games where you think three or more will result away /home, or looked at another way, against four games in which you think that no more than two draws will occur. You will have three possibly four correct should either assumption be realised.

X	X	1	1	1	2	2	2
1	2	X	1	2	X	1	2
1	2	1	2	X	2	X	1
1	2	2	X	1	1	2	X

I'd like now to turn your attention to another type of "conditional" guarantee block, away from one built on the results *you are trying to limit* towards one built on the results which *must occur in order to trigger the guarantee.*

In so doing I hope to reduce further the number of columns necessary to achieve a viable coupon entry using another

more efficient but rather long winded way to allow for forecasting error.

That is to base a required reduction on the selective use of basic lines for elimination. This article shows a small example of such a reduction.

You'll have to take my word for it unless you read my original articles a couple of years ago on this kind of block design, but we allowed for up to one error in our initial forecast when we designed our own "unconditional" block. Such a table covering four homes for up to one error is shown here as Table A:

TABLE A

1	2	X	1	1	1	1	1	1
1	1	1	2	X	1	1	1	1
1	1	1	1	1	2	X	1	1
1	1	1	1	1	1	1	2	X

In other words, we rashly assumed that three if not all four "original column" game forecasts were going to be correct.

We were really asking for a forecasting accuracy of 75% to 100%, which is why our resulting block took so many lines. So if we reduce our initial accuracy expectations slightly, we should end up with an entry block with fewer lines.

We have seen how to use a two-way forecast to do this, with an expectation of 66%, but what happens if we expect just two of our initial forecasts to be correct? This is only 50% forecasting accuracy, and is something I'm used to (!)

For continuity and ease of calculation, we'll stay with four 3-way games. With a single forecast for each of the four, there are thirty-three chances of having up to two errors.

Assuming an initial forecast of all four home wins, Table B shows the 33 lines necessary for a correct result should our forecast contain none, one or two errors:

TABLE B

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*	*32*	*33*
1	2	X	1	1	1	1	1	1	2	2	2	2	2	2	1	1	1	1	1	1	X	X	X	X	X	X	1	1	1	1	1	1
1	1	1	2	X	1	1	1	1	2	X	1	1	1	1	2	2	2	2	1	1	2	X	1	1	1	1	X	X	X	X	1	1
1	1	1	1	1	2	X	1	1	1	1	2	X	1	1	2	X	1	1	2	2	1	1	2	X	1	1	2	X	1	1	X	X
1	1	1	1	1	1	1	2	X	1	1	1	1	2	X	1	1	2	X	2	X	1	1	1	1	2	X	1	1	2	X	2	X
0	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*1*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*	*2*

The columns are numbered along the top, and the errors allowed for are shown along the bottom.

Note that the first 9 columns are the “up to one error” Table A shown earlier, with the remaining 24 columns each containing exactly two errors.

In fact the entire table can be obtained by several methods, including writing out in full each individual column by inspection from the full perm.

A much less laborious and more interesting method, however, would be to simply align each of the " up to one error" lines against each of the full perm columns and eliminate those columns matching them in at least one position. I'll return to this idea later, but for now will try to show a simple "basic line" elimination.

Lets look at our 33 column Table B in terms of "basic lines" then:

1	1	1	1	1	1
1	1	1	1	1	1
1	1	1	2	2	X
1	2	X	2	X	X
1	4	4	6	12	6

The bottom row shows the number of variations of each type of result.

For instance, the first three columns show Table A as the number of arrangements (9) that a line of up to one error covers. Whilst the last three show those additional arrangements (24) covered by a line of two errors as in Table B.

I hope you can see that whilst there is only one variation of the original all home win forecast, for example, there are twelve variations of the 2 homes with 1 away and 1 draw. All thirty-three lines therefore are represented in a "shorthand" form.

We are trying to select, in the minimum number of lines, four results giving three correct - should one of the variations covered by the above basic lines occur.

In order to make it clearer I'll put it yet another way. If our original forecast contains 2 correct results anywhere in the 4, our reduction table is to go one better and produce a line with 3 correct results.

An Enhanced Guarantee Table, perhaps? Anyway…

Lets deal with the largest block first, the 12 variants of the basic line 112X.

The twelve lines are shown here as Table C:

TABLE C

1	1	1	1	1	1	2	X	2	X	2	X
1	1	2	X	2	X	1	1	1	1	X	2
2	X	1	1	X	2	1	1	X	2	1	1
X	2	X	2	1	1	X	2	1	1	1	1
1	2	3	4	5	6	7	8	9	10	11	12

We need to cover Table C in columns allowing for up to one error, and these columns themselves have to cover both one draw and one away win.

Looked at another way, each line above has two home wins, so why not start by using a line containing three homes? We can't try 1111 as a "one error" column for obvious reasons, so lets try the next column, 1112 as a "one error" column.

This, when set against the 12 columns above gives us three columns agreeing in three positions at columns 2, 4, and 8: 11X2, 1X12, and X112.

Eliminate those columns, and we now have:

1	1	1	1	2	2	X	2	X
1	2	2	X	1	1	1	X	2
2	1	X	2	1	X	2	1	1
X	X	1	1	X	1	1	1	1
1	3	5	6	7	9	10	11	12

Now lets try 111X and do the same thing, which gives columns 1, 3, and 7, or 112X, 121X, and 211X.

Eliminate them to give:

1	1	2	X	2	X
2	X	1	1	X	2
X	2	X	2	1	1
1	1	1	1	1	1
5	6	9	10	11	12

So we have "covered" six of the 12 lines with two lines. Now lets "invert" the 1112 and 111X.

This gives us 2111 and X111 so that our "errors" are in different positions.

1	1	X	X
2	X	1	2
X	2	2	1
1	1	1	1
5	6	10	12

Column 2111 yields columns 9 and 11, or 21X1 and 2X11, so eliminate them as illustrated on the right.

1	1
2	X
X	2
1	1
5	6

Column X111 yields columns 10 and 12, or X121 and X211, so out they go too:

So only two columns of Table C remain to be eliminated.

If you reflect for a moment, you will see that these two columns will between them not only cover the remaining two variations in the basic line chosen, but also the remaining columns in the whole of Table B.

Together with the four columns already used we therefore have a guarantee table in six lines which boasts that should two homes occur in the four 12X games then the table will have a column showing three correct results.

We'll call this one Table X:

TABLE X

1	1	1	1	2	X
1	1	2	X	1	1
1	1	X	2	1	1
2	X	1	1	1	1

Similar guarantee blocks may be obtained by following the above steps.

In order to produce another four game, 3-way block for instance, just take any two "one error" variations of your initial forecast.

Then add a third column consisting of any arrangement of the basic line "two different errors only", and ensure that the "errors" occupy two positions not already occupied in the first two columns.

Then simply "reverse" or "invert" the symbols to give six lines in total.

For example, the block below was compiled using initial columns 1121, 11X1, and 2X11.

1	1	1	1	2	X
1	1	1	1	X	2
1	1	2	X	1	1
2	X	1	1	1	1

The blocks become quite useful when you realise that you may alter the 12X to selections for *any* events of ***A***,***B*** & ***C***.

A would represent your first choice or "best bet", ***B*** would be your second choice, finishing with ***C*** the result you feel to be the least likely of the three.

Table X, for instance, becomes a key:

A	A	A	A	B	C
A	A	B	C	A	A
A	A	C	B	A	A
B	C	A	A	A	A

And the guarantee becomes:
"one error maximum providing there are at least two correct A forecasts in the results".

Just ensure you maintain the actual sequence of ciphers "across" the block or the guarantee will be lost. Of course, more accurate forecasting will improve the basic guarantee.

At the start of this little series, we covered eight games in two nine-line blocks for an 81 line entry which gave us a guarantee of "6 correct no matter what the result". No forecasting necessary, really.

Then we used our forecasting skill with two eight-line blocks of our own design using either our own two-way forecast, or the limitation of a particular result we thought unlikely.

This gave us a reduction to only 64 lines for a two-block entry.

We have now managed to reduce even further, as two blocks we have just designed will now cover eight games in only 36 columns for quite a respectable guarantee of six together if two of our original forecasts occur in each block.

To put it another way, we get a 75% chance from 50% forecasting. Hopefully, I'll expand these tables in another issue.

Making a profit from betting is all about backing enough winners at the right odds. Sounds so simple to say it, but what are the mathematical implications?

BEST ODDS AT WHICH TO BET

Steve Tilley

This article gives a method of doing better in tipping competitions and a system for mug punters, such as relatives and work mates, to use on the racecourse.

When we bet we generally set ourselves profit targets, for example;
"I want to make a 20% return this month",
"I want to double my money this afternoon at the bookies",
"If this new system doesn't make a profit this month I'm ditching it".

Let's take a closer look at how the odds at which we bet affects the chance of us reaching our target.

We'll start with a very simple example, a tipping competition. In this particular competition we have to bet over 24 weeks, one horse per week. Having looked at it over previous years if I can get a 50% profit at the end of the season I will be in with a very good chance of winning it.

As we're looking at odds let's see the fewest winners I will need at various odds. We will assume for the moment all my bets are at the same odds. To get a 50% profit I need to have a total return of 36 units, 24 times 1.5. Each winner returns the winning odds plus one. So a 10/1 winner returns 11 units a 6/4 winner 2.5 units. So if I bet at 5/1 I will need. . .

$$\frac{36}{5+1} = 6 \text{ winners}$$

Now if I bet at 10/1 I will need

$$\frac{36}{10+1} = 3.27 \text{ winners}$$

Now clearly you can't have 0.27 of a winner so I will need at least 4 winners to reach my target. In fact I will exceed it as I will have 44 units or more returned. But 3 winners will only return 33 units and so I will miss my target.

In general we can write a formula to calculate the minimum number of winners we need to reach a target

$$\text{round\_up}\left[\frac{\text{runner} \times (1 + \text{target})}{\text{odds} + 1}\right] = \text{winners}$$

The target is the target profit or loss as a decimal. Breakeven is zero. Round up tells us to round the result up to the nearest whole number. This rounding up is very important as it causes some interesting results.

So if I bet at 12/1 I will need

$$\frac{24 \times (1 + 0.5)}{12 + 1} = 2.769$$

Which rounded up means at least 3 winners.

We now make another assumption (don't worry we'll come back to all the assumptions later it's just they help make things easier at this stage).

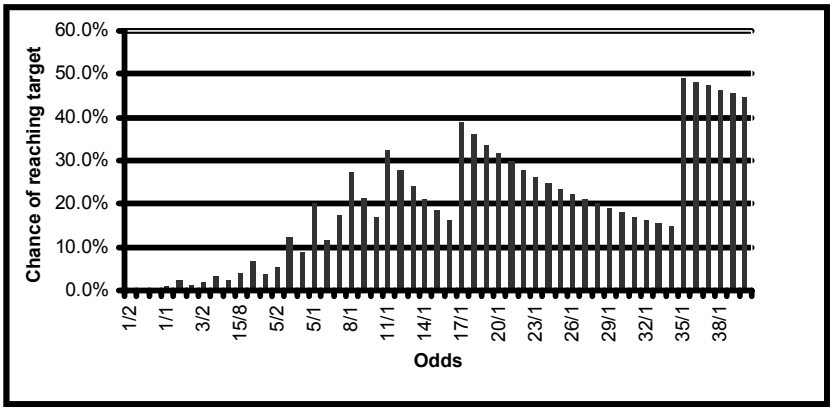
Let us assume that all our bets have the same chance of winning and that chance is given by their odds. In other words they are all fair bets.

If we bet at 2/1 we will win 33.33% of the time, if we bet at 9/1 we win 10% of the time etc.

There is a method of calculating the chance of at least W winners out of N runners with a chance of winning P. This uses the binomial distribution and can easily be done using Microsoft Excel. Contact me for a spread sheet if you are interested.

So if we were backing at 5/1 we need at least 6 winners. They win 16.67% of the time. The chance of us getting 6 winners or more out of 24 can be shown to be 20%. We can plot a graph of the chance of us reaching our target across a range of odds.

Graph 1



There are several interesting points about this graph.

If we bet at short odds the chance of us reaching our target is very small. As we move up to the longer odds we have a greater chance of success. Some longer odds are very much better than others, for example if you bet at 16/1 you chance of success is about 16% whereas at 20/1 it is about 32%.

By moving from 16/1 shots to 20/1 shots you have doubled your chances of success.

The position of the best odds is given by

$$\frac{\text{runners} \times (1 + \text{target})}{\text{winners}} - 1 = \text{odds}$$

Here winners is any whole number from 1 to the number of bets.

So if we were to set winners as 1 or 2 then the best odds are

$$\frac{24 \times (1 + 0.5)}{1} - 1 = 35 \qquad \frac{24 \times (1 + 0.5)}{2} - 1 = 17$$

This is because these are the shortest odds that let us get away with that number of winners.

Similarly the worst odds we can pick are always just less than the best odds. Because we pick shorter odds than the peak we need one more winner and the odds are the longest we can have needing that many winners.

For example the chance of getting two winners at 17/1 is much greater than getting 3 at 16/1.

Now we cannot bet at 17/1 usually but as you can see from the graph 20/1 or even 25/1 gives us a much better chance of reaching our target than 16/1.

As to the shorter priced animals less than 5/1 gives us very little chance of reaching our target.

As a quick aside - and this will be covered in another article - we can calculate what happens when a good punter enters the competition.

Let us take a good punter who bets between 4/5 on and 2/1. He bets at an advantage of 40%. This means his even money shots win 70% of the time.

It can be shown his chance of making a 50% profit or more is about 33%.

Now consider a punter betting at 40/1. His chance of reaching the same target can be shown to be about 44%.

The effect described above is extremely powerful as it can negate a punter betting with a 40% advantage. (This is why good tipsters find tipping competitions difficult)

Bet the same amount in each race, get one winner during the day and you've doubled your money

We can also use this to give a system for mug punters at the races. Mug punters like to back in every race and as a rule of thumb will consider it a good day if they've doubled their money as this will pay for drinks entry, food and still leave some over.

So, in a six race card tell them to back any horse but bet as close as they can to 11/1 but NOT less than 11/1. In a seven race card tell them as close as they can to 14/1 (it should be 13/1 but that's never offered) Tell them to bet the same amount in each race, which means if they get one winner during the day they've doubled their money.

So in six race cards they will be backing generally at 12/1 or 14/1 and in 7 race cards 14/1 or 16/1. Using RSB a random bet at these odds has a loss of about 30%. The chance of our mug doubling his money using this "system" is about 30% (Which I don't think is too bad). Can you work out how we got the odds for the system?

My next article will show why short priced horses are a more risky investment than longer priced ones.

GAMBLING BOOKS from SMARTsig
www.smartsig.com

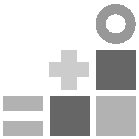
the SMARTsig Internet web site now has a direct link with Gambler's Book Store. Simply -click- on the SMARTsig bookstore link on our web site home page for the best selection of betting books around.

Look out for the frequent special offers!

RacingSystemBuilder

SMART members qualify for 10% discount RSB software packages.

Racedata Modelling Ltd.,
Upper Buckenhill
Farmhouse, Fownhope,
Herefordshire HR1 4PU
Tel: 01432 860 864



HOOF HORSERACE OPTIMUM ODDS FORECASTER

NH value ratings on the Internet
". . simply in a class of their own . ."

email: tony@hoof.demon.co.uk

Tel: 01873 811427

Soccer results on disk, weekly results & odds by email

For further information regarding the weekend odds service, soccer results files and almost every other up-to-the-minute soccer statistic you can think of – all at realistic, value for money prices, contact Bill Hunter at;

Mabels-Tables, PO Box 14555, Dunfermline KY11 4WA
Tel: 01383 721 729 email: bill@mabels-tables.co.uk
Internet: www.mabels-tables.com

SMARTsig members qualify for a further 10% discount . . . even on top of existing discounts.

SMARTsig results-on-disk

The only limit is your imagination !

9 seasons of **NH** (1992/3 – 2000/01) & **9** seasons of **Flat** (1992 –2000) PC comma separated text files on CD only £31 post paid per season.

SAVE! purchase any 5 seasons or more at the one time and qualify for 20% discount (5 seasons @ £124 post paid - 10 seasons @ £248 etc.)

Only available through SMARTsig, PO Box 44, Hayle. TR27 6YH
Credit/debit cards 'phone/fax 01736 754400 or email stef@smartsig.com

All advertisements are published in good faith, and do not imply any recommendation. SMARTsig members should satisfy themselves as to the suitability of a product/service before proceeding.

SMARTsig member benefits are for subscribers only. Payment by cheque, money order or credit/debit card. Card payments can be made by telephone or fax (01736 754400)

Tick (√) box		*UK & N. Ireland*		*Irish Republic & other European*		*Rest of the World*
3 months		£13.00		£14.05		£15.58
6 months		£25.00		£27.10		£30.16
12 months		£45.00		£49.20		£55.32

Please enrol me as a subscriber, every issue hot off the press, by first class mail.

Name		**Telephone**
Address		
		(Please PRINT)
Post Code		

Card number	***Expiry***	***Issue no (Switch)***	***Signature***

EASY PAYMENTS! Subs can be paid by bank standing order at only £3.75 monthly (UK customers). Call for details.

BACK ISSUES: *£3.50 each whilst stocks last. (Single issues add 50p postage, order of 2 or more, post free) We are no longer able to reprint out-of stock issues. However, work is underway to publish back issues in electronic book format in volumes.*

SMARTsig Confidential

Published by: **SMARTsig, PO Box 44, HAYLE. TR27 6YH**
`phone and fax 01736 754400
email stef@smartsig.com Internet: www.smartsig.com

The opinions expressed herein are not necessarily those of the Editor or publisher, but are often taken directly from members contributions. ***SMARTsig*** does not accept any liabilities for inaccuracies within the content of the magazine, nor for any consequences thereof. We will always endeavour to print replies and/or corrections by or on behalf of parties who may feel they may have been misrepresented in any way. ***SMARTsig*** encourages and welcomes contributions from its members but is unable to accept any responsibility for loss or any damage of any material, solicited or otherwise. Everything published in done so in good faith, and is the copyright of either ***SMARTsig***, the contributor, or both - subject to it not having been published elsewhere beforehand. This Journal, nor any of its contents must not, in whole or part, be copied, duplicated, loaned or distributed without the written permission of the copyright holder(s).
We recommend you exercise caution with any contacts established through our group and never speculate with money you cannot afford to lose.

© 2001 SMARTsig

ActiveGambler.Com

Where Winners Find Winners

If there were a **legal** way to put <u>your</u> betting business on **high-performance** *steroids*, would you do it?

Naturally, you would.

Visit our website daily for an **industrial-strength** approach to pre-race analysis that can substantially **beef up** your betting results!

www.ActiveGambler.com