TOTE ARBITRAGE AND LOCK OPPORTUNITIES IN RACETRACK BETTING

D.C. EDELMAN AND N.R. O'BRIAN

ABSTRACT. We describe a game-theoretic approach to the investigation of arbitrage opportunities based on combinations of exotic wagers for the same event. This situation is also known as a 'lock' or 'Dutch Book'. We apply the technique to recent totalisator data from Australian thoroughbred races. It appears that such opportunities appear fairly regularly, at least according to published final dividends for various bet types. We demonstrate the method in some detail on a particular example.

KEYWORDS: Dutch Book, lock, arbitrage, zero-sum game, minimax, totalisator, pari-mutuel, gambling, bet, wager, racetrack.

1. INTRODUCTION

In gambling parlance a combination of bets or other risky investments yielding a guaranteed profit, regardless of outcome, is known as a *lock* or *Dutch Book*. For obvious reasons, such opportunities are rarely publicly available, although some special situations have been indentified, even in the context of racetrack betting. See Willis [4] and Hausch and Ziemba [1] for arguments and examples based on comparison of Win, (American) Place and Show dividends.

There is a simple and well-known criterion for the existence of a lock based on a combination of Win bets alone. Suppose that n horses contest a race, and a successful unit bet on horse i to win returns T_i units (including the original stake). In order to guarantee a return of 1 unit, the gambler must bet T_i^{-1} on runner i for all $1 \leq i \leq n$. This requires a total outlay of

$$\sum_{i=1}^{n} T_i^{-1} \tag{1}$$

units. If the outlay is less than 1, the combination is a lock.

For a typical pari-mutuel or totalisator market this can never happen. Let N_i be the total number of units wagered on horse i to win, and $N = \sum_i N_i$. Then T_i is determined by the formula

$$T_i = (1 - \alpha)\frac{N}{N_i}$$

with $0 < \alpha < 1$. This results from deducting a proportion α from the total pool N and dividing the remainder equally among the N_i winning units.

Date: October 17, 2001.

Then

$$\sum_{i=1}^{n} T_i^{-1} = (1-\alpha)^{-1}\frac{\sum_i N_i}{N} = (1-\alpha)^{-1} > 1\,.$$

Typically calculation of the dividends T_i also involves rounding down the quantities N/N_i, increasing the sum of the T_i^{-1} still further. Similar argument apply to other bet types, with the sum ranging over possible distinct winning combinations.

If there are several independent markets offering Win bets for the same race, the situation may be improved by backing each runner i in the market offering the largest value of T_i. Given the typically rapid and efficient dissemination of betting information and the obvious nature of the strategy, the market is likely to quickly discover and eliminate this type of lock.

A more interesting situation occurs for combinations of Win and 'exotic' wagers, involving combinations of runners finishing in a specified order. In Australian totalisator markets these include the Exacta (first and second in order), Quinella (first two in either order) and Trifecta (first three in order) bets. Each bet type has its own pool, from which the final dividend is calculated.

We show that construction of a lock in this type of situation has a natural formulation as a two-person zero-sum matrix game, for which the required combination of bets is the usual minimax solution of the game. The minimum expected return to the gambler is the value of the game. This formulation and its solution provides a natural generalisation of the standard Win market lock described above.

We also investigate the existence of locks in the context of Win, Quinella and Exacta markets for Australian thoroughbred races, based on final dividends published by the New South Wales Totalisator Agency Board for 2667 events in a three month period in early 2000. It appears that potential locks in the above sense certainly exist in a small proportion of races. We also discuss some problems which are likely to arise in any attempt to exploit these situations.

2. Locks for Exotic Bets

When several bet types are available the above form of a lock based on Win bets can be generalised as follows. Suppose that the different betting opportunities are indexed by α, and the possible race outcomes by β. For a race with n runners the Win, Quinella and Exacta bets give a total of

$$n + \frac{n(n-1)}{2} + n(n-1)$$

possibilities for α, while β runs over the $n(n-1)$ combinations for first and second place.

If the minimum outlay needed to guarantee a return of one unit is less than 1, a lock is possible. More formally, let $x_\alpha \geq 0$ be the allocation to opportunity α and $a_{\alpha\beta} > 0$ the return to the gambler for a unit bet of type

α when the outcome of the race is β. The problem is then to minimise the total outlay

$$\sum_\alpha x_\alpha$$

subject to the requirement that

$$\sum_\alpha a_{\alpha\beta} x_\alpha \geq 1$$

for every outcome β.

This is a straightforward linear programming problem. It is easily seen to be feasible and can be solved by standard techniques. We can also consider the dual problem, which is to find the values y_β which maximise

$$\sum_\beta y_\beta$$

subject to constraints

$$\sum_\beta a_{\alpha\beta} y_\beta \leq 1 \quad \text{and} \quad y_\beta \geq 0$$

for all α, β. We see below that the y_β, after normalisation, have a natural interpretation as a probability distribution on the outcomes β, as implied by the combined market for the different types of wager. Similarly the constraints involve expectations of the various bet types with respect to this distribution.

If x^*_α and y^*_β are solutions to the original optimisation problem and its dual, the duality theorem of linear programming [3] shows that the value of the objective function is the same in both cases, so

$$\sum_\alpha x^*_\alpha = \sum_\beta y^*_\beta = v$$

say.

If we consider only Win bets, both α and β index the runners themselves. In this case $a_{\alpha\beta} = 0$ unless $\alpha = \beta$ and $a_{\alpha\alpha}$ is the return T_α for a unit bet on α to win. The solution to the linear programming problem and its dual is easily seen to be

$$x^*_\alpha = y^*_\alpha = T_\alpha^{-1}\,,$$

in agreement with the Dutch Book calculation described above.

3. Game-Theoretic Formulation

The standard two-person zero-sum matrix game is typically formulated in terms of two players, R (for rows) and C (for columns), each with a finite range of possible plays. The parameters of the game are set by the *payoff matrix* $a_{\alpha\beta}$, where α and β are row and column indices respectively. A single play involves a simultaneous choice of row α by R and column β by C. Player C is then required to pay R an amount $a_{\alpha\beta}$. In this context the main result is the existence of a *minimax* strategy for both players. Any departure from this strategy by one player may be exploited by the other to increase expected return.

This strategy is usually formulated in probabilistic terms, requiring each player to play randomly, choosing among the possible plays with specified probabilities. In this interpretation, if R chooses α with probability p_α and C chooses β with probability q_β, then the expected transfer from C to R is

$$\sum_{\alpha,\,\beta} a_{\alpha\beta} p_\alpha q_\beta \tag{2}$$

units per play.

In a mathematically equivalent formulation, one or both players construct a 'portfolio' by spreading risk across the various plays in the same proportions p_α and q_β. If both players do this then the transfer from C to R is the same on each play, and again given by formula (2). The von Neumann minimax theorem then asserts that

$$\max_p \min_q \sum_{\alpha,\,\beta} a_{\alpha\beta} p_\alpha q_\beta = \min_q \max_p \sum_{\alpha,\,\beta} a_{\alpha\beta} p_\alpha q_\beta \tag{3}$$

where p and q range over the possible portfolios for R and C, with

$$p_\alpha,\ q_\beta \geq 0, \quad \sum_\alpha p_\alpha = \sum_\beta q_\beta = 1\,.$$

There are a variety of proofs for this result. See [2] for a recent textbook treatment. Here we follow [3] and relate it to the duality of the linear programming problems above. The gambler plays the part of the row player R, constructing a portfolio p_α based on the different bet types. The column player C is represented by the outcome of the race, which we suppose follows the distribution q_β.

As above let v be the common value of the objective functions for the solutions to the linear programming problem and its dual and set

$$p^*_\alpha = x^*_\alpha / v\,, \qquad q^*_\beta = y^*_\beta / v$$

so that $\sum_\alpha p^*_\alpha = \sum_\beta q^*_\beta = 1$. For any other portfolios q_β we then have

$$\sum_{\alpha,\,\beta} a_{\alpha\beta} p^*_\alpha q_\beta = \sum_\beta q_\beta \sum_\alpha a_{\alpha\beta} x^*_\alpha / v \geq \sum_\beta q_\beta / v = 1/v\,.$$

A similar calculation using the solution of the dual problem shows that

$$\sum_{\alpha,\,\beta} a_{\alpha\beta} p_\alpha q^*_\beta \leq 1/v$$

for all portfolios p_α. Taken together, these inequalities show that

$$\sum_{\alpha,\,\beta} a_{\alpha\beta} p_\alpha q^*_\beta \leq \sum_{\alpha,\,\beta} a_{\alpha\beta} p^*_\alpha q^*_\beta \leq \sum_{\alpha,\,\beta} a_{\alpha\beta} p^*_\alpha q_\beta \tag{4}$$

for all portfolios p_α and q_β, with the middle term equal to $1/v$.

In the terminology of game theory this shows that p^*_α, q^*_β is a 'saddle point' for the game. It is well-known that this condition characterises the minimax solution. For example, the first inequality gives

$$\max_p \sum_{\alpha,\,\beta} a_{\alpha\beta} p_\alpha q^*_\beta = 1/v\,.$$

Then certainly

$$\min_q \max_p \sum_{\alpha,\,\beta} a_{\alpha\beta} p_\alpha q_\beta \leq 1/v\,.$$

Similarly the second inequality gives

$$1/v \leq \max_p \min_q \sum_{\alpha,\,\beta} a_{\alpha\beta} p_\alpha q_\beta\,.$$

Together these inequalities show that

$$\min_q \max_p \leq \max_p \min_q\,.$$

The reverse inequality follows formally, without any need for the saddle point condition. This then implies the minimax equality (3), with both sides equal to $1/v$. See [2] for further properties of the minimax solution.

If q_β is the probability of outcome β of the race, then $\sum_{\alpha,\,\beta} a_{\alpha\beta} p_\alpha q_\beta$ is the gambler's expected return when using portfolio p_α. According to inequalities (4), choice of portfolio p^*_α ensures that this is at least equal to the value $1/v$ of the game.

Similarly the q^*_β specify the worst-case distribution for the gambler, in the sense that in this case no portfolio p_α can give an expectation greater than $1/v$.

4. Applications

We analyse published final dividends for 2667 Australian thoroughbred races in the period February to April 2000. Off-course pari-mutuel betting in New South Wales is operated exclusively by the NSW Totalisator Agency Board (TAB) and the sample represents essentially all thoroughbred races covered by the TAB in this period. Approximate dividends and pool sizes for Win, Place, Quinella and Exacta bets are available through the internet and are updated frequently until the start of each race. Final dividends for each bet type are published after the start. These were logged automatically to produce the data for analysis.

For the exotic bet types only the 60 most favoured combinations are made available. We treat missing data conservatively, assuming that the missing dividends were all equal to the maximum of the published values. although in general they will be larger. The total size of the pool for each bet type is also published.

For any single bet type we can apply (1) to determine the minimum outlay to return 1 unit. We find average values of 1.18, 1.19 and 1.24 respectively for the Win, Quinella and Exacta pools. These figures result from deduction of a fixed proportion of the pool and rounding-down of individual dividends, and give an indication of the difficulty of profitable arbitrage.

The results of the minimax strategy on the sample races are summarised in Table 1. This shows the number of samples N, the average number of starters n and the average pool sizes for different ranges of the value v. In particular this shows 31 examples in the range $v < 1$, suggesting the possibility of a lock. Note also that lower values of v tend to be strongly associated with smaller pools and smaller field sizes. All pool sizes are given in Australian dollars.

v	N	n	Average Pool Size		
			Win	Quinella	Exacta
0.75-0.80	1	14.0	9131	1325	374
0.90-0.95	5	7.4	14580	2076	718
0.95-1.00	25	8.2	25201	3652	1586
1.00-1.05	185	9.0	37495	5610	2274
1.05-1.10	918	9.6	60012	9281	3424
1.10-1.15	1298	10.4	104457	16321	5199
1.15-1.20	235	12.9	258798	40000	9986

TABLE 1. Pool and field size for different value ranges.

We further illustrate the lock concept with a specific example. The final dividends for the seven starters in Race 4 at Bunbury on 4th April 2000 are shown in Tables 2 to 4. For each runner or combination the top line gives the return for a successful unit bet, and the figure below in brackets gives the corresponding contribution in the solution to the linear programming problem, when this is non-zero. In this case all three bet types contribute to the solution, requiring an outlay of 0.9641 units to secure a return of at least one unit. All outcomes for first and second place yield one unit, except the $(3, 7)$ combination which pays one unit for each of the Win and Quinella bets. The pools in this example were relatively large for this value of v, at \$31846 for the Win, \$3970 for the Quinella and \$2965 for the Exacta.

	1	2	3	4	5	6	7
	2.60	3.00	49.40	13.50	9.60	39.70	4.00
	–	–	(0.0202)	–	–	–	–

TABLE 2. Win Dividends and Allocations

	1	2	3	4	5	6	7
1	*	5.50 —	81.00 (0.0123)	61.00 (0.0164)	29.00 (0.0345)	85.00 (0.0118)	7.80 —
2	4.90 —	*	82.00 (0.0122)	73.00 (0.0137)	22.00 —	60.00 (0.0167)	11.00 —
3	147.00 —	93.00 —	*	374.00 —	180.00 —	231.00 —	211.00 —
4	99.00 (0.0101)	79.00 (0.0127)	303.00 (0.0033)	*	221.00 (0.0045)	442.00 (0.0023)	162.00 (0.0062)
5	63.00 (0.0159)	37.00 —	270.00 (0.0037)	162.00 (0.0062)	*	324.00 (0.0031)	53.00 (0.0189)
6	187.00 (0.0053)	118.00 (0.0085)	694.00 (0.0014)	810.00 (0.0012)	810.00 (0.0012)	*	347.00 —
7	6.70 —	12.00 —	65.00 —	45.00 (0.0222)	56.00 (0.0179)	67.00 —	*

Table 3. Exacta Dividends and Allocations. Row = 1st, Column = 2nd.

	1	2	3	4	5	6	7
1	*	4.00 (0.2500)	34.00 —	12.00 —	9.40 —	45.00 —	5.50 (0.1818)
2	*	*	72.00 —	27.00 —	14.00 (0.0714)	30.00 —	6.60 (0.1515)
3	*	*	*	127.00 —	64.00 —	193.00 —	72.00 (0.0139)
4	*	*	*	*	44.00 —	46.00 —	33.00 —
5	*	*	*	*	*	64.00 —	14.00 —
6	*	*	*	*	*	*	76.00 (0.0132)
7	*	*	*	*	*	*	*

Table 4. Quinella Dividends and Allocations

5. Practical Considerations

There are a number of obstacles to successful exploitation of the type of example shown above.

(1) Clearly final dividends are not available at the time the calculation needs to be done. Observation of pool size variation shows that a considerable proportion of bets are placed in the last few seconds before the start. Latency in processing and communication systems will further delay the calculation of final odds.
(2) In the pari-mutuel context any bet will of itself cause a decrease in the corresponding dividend, with the effect increasing for large bets and small pools. Table 1 shows that the lock situation is typically associated with a small pool size. The possibility of large profits will then be limited by this effect.
(3) In the example above the smallest allocation required for the lock is 0.0012 units. With a unit of \$1 the \$0.50 minimum betting unit for the TAB is more than 400 times larger than this figure. This suggests that a total bet size of several hundred dollars is needed to avoid large discretisation errors in the implementation of the strategy. This will increase the chance of coming into conflict with point 2 above.

In summary, we find that the lock condition appears fairly regularly on final dividends, although the matters raised above indicate that the strategy will not be truly risk-free. Nevertheless Table 1 suggests that there is considerable value present, which may be of use in increasing the possibility of positive returns by other methods.

The gambler's least-favourable distribution q^*_β appears as a by-product of the lock calculation. In the present case this gives a distribution on all combinations of first and second place, and in some sense integrates the information available from the individual pools. It may be interesting to compare the predictive value of this distribution with the probabilities implied by individual pools.

References

[1] Donald B. Hausch and William T. Ziemba. Locks at the racetrack. *Interfaces*, 20:41–48, 1991.
[2] JianHua Wang. *The Theory of Games.* Oxford Mathematical Monographs. Clarendon Press, Oxford, 1988.
[3] Roger Webster. *Convexity.* Oxford University Press, Oxford ; New York, 1994.
[4] K.E. Willis. Optimum no-risk strategy for pari-mutuel betting. *Management Science*, 10(3):574–577, 1964.

Department of Accounting and Finance, University of Wollongong, Wollongong, NSW 2522, Australia

School of Mathematics and Statistics, University of Sydney, Sydney, NSW 2006, Australia