Adv. Appl. Prob. **38,** 899–915 (2006)
Printed in Northern Ireland

THE MINIMAX BOOKIE: THE TWO-HORSE CASE

DANIEL BARRY * AND
CAROLINE LYNCH,** *University of Limerick*

Abstract

A bookmaker takes bets on a two-horse race, attempting to minimise expected loss over all possible outcomes of the race. Profits are controlled by manipulation of customers' betting behaviour, which is assumed to be determined uniquely by the price quoted for each horse. We consider different strategies for choosing these prices as bets accumulate, and examine the penalty incurred by the use of a strategy other than the optimal one, in the general case where the distribution of the customers' betting probabilities is unknown.

Keywords: Dynamic programming; bookmaker; horse race; Markov decision process

2000 Mathematics Subject Classification: Primary 90B40
Secondary 90C39; 90C40; 90J27

1. Introduction

A gambler enters a bookie's (bookmaker's) shop seeking to place a wager on a horse race for which there are only two possible outcomes, which we will label A (horse A wins) and B (horse B wins). Let p denote the gambler's probability that outcome A will occur. The bookie quotes odds of O_1 against outcome A and of O_2 against outcome B. This means that a winning wager of one unit on outcome A produces a return of $O_1 + 1$ while a winning wager of one unit on outcome B produces a return of $O_2 + 1$. Hence, a wager on outcome A will be attractive to the gambler if

$$p(O_1 + 1) \geq 1$$

or, equivalently,

$$p \geq \frac{1}{O_1 + 1} = \theta_1.$$

Similarly, a wager on outcome B will be attractive to the gambler if

$$1 - p \geq \frac{1}{O_2 + 1} = \theta_2$$

or, equivalently,

$$p \leq 1 - \theta_2.$$

The quantities θ_1 and θ_2 are referred to as the bookie's quoted probabilities for outcomes A and B, respectively. Hence, the strategy for an individual gambler is simple: he places a wager

Received 19 August 2005; revision received 27 July 2006.
* Postal address: Vice President Academic & Registrar, Plassey House, University of Limerick, Limerick, Ireland.
** Current address: Cloncool, O'Callaghan's Mills, County Clare, Ireland. Email address: clynch@canada.com

https://doi.org/10.1017/S0001867800001385 Published online by Cambridge University Press

on any outcome for which his probability exceeds that quoted by the bookie. In this paper we focus on strategies that the bookie might follow.

We idealise the bookie's shop by assuming that the bookie sells two types of tickets: one which guarantees a return of one unit should outcome A occur, and costs θ_1; and one which guarantees a return of one unit should outcome B occur, and costs θ_2. We also assume that the bookie knows, before accepting wagers, that N customers will consider a wager on the horse race and that their probabilities, $p_1, p_2, \dots, p_N$, of outcome A occurring behave like a random sample from a probability distribution whose cumulative distribution function, F, is known to the bookie. Finally, we assume that customers can buy at most one of each type of ticket and that the bookie is free to alter the quoted probabilities after each customer leaves. Let θ_1^{L} be the largest value of θ for which $F(\theta) = 0$ and let θ_1^{R} be the smallest value of θ for which $F(\theta) = 1$. We restrict the bookie to use quoted probabilities in the range $[\theta_1^{\mathrm{L}}, \theta_1^{\mathrm{R}}]$ for outcome A and in the range $[\theta_2^{\mathrm{L}}, \theta_2^{\mathrm{R}}]$ for outcome B, where $\theta_2^{\mathrm{L}} = 1 - \theta_1^{\mathrm{R}}$ and $\theta_2^{\mathrm{R}} = 1 - \theta_1^{\mathrm{L}}$.

At any particular time the bookie can calculate the *state of the book* for outcome A, which is defined as the profit accruing to the bookie if he stops accepting bets at that time and outcome A eventually occurs. Let A_n and B_n respectively denote the states of the book for outcomes A and B when n of the N customers remain. Clearly, the bookie would like A_0 and B_0 to be large. In Section 2 we consider a strategy which seeks to set quoted probabilities so as to maximise $\mathrm{E}[\min\{A_0, B_0\} \mid A_N = B_N = 0]$.

2. Optimal strategy (dynamic programming)

For $i, j \in \{0, 1\}$, let $a_{ij}(\theta_1, \theta_2)$ denote the probability that a customer, faced with quoted probabilities θ_1 and θ_2 on outcomes A and B, respectively, chooses to purchase i bets on A and j bets on B. For $0 \le n \le N$, define $R_n(a, b)$ to be the maximum value achievable for $\mathrm{E}[\min\{A_0, B_0\} \mid A_n = a,\ B_n = b]$.

Lemma 1. *We have*

$$R_n(a, b) = \frac{a + b}{2} + P_n(d)$$

where $d = a - b$, $P_0(d) = -|d|/2$, *and, for* $n \ge 1$,

$$\begin{aligned} P_n(d) = P_{n-1}(d) + \max_{\theta_1, \theta_2} \{ & a_{10}(\theta_1, \theta_2)[\theta_1 - \tfrac{1}{2} + P_{n-1}(d - 1) - P_{n-1}(d)] \\ & + a_{01}(\theta_1, \theta_2)[\theta_2 - \tfrac{1}{2} + P_{n-1}(d + 1) - P_{n-1}(d)] \\ & + a_{11}(\theta_1, \theta_2)(\theta_1 + \theta_2 - 1)\}. \end{aligned}$$

We denote by $\hat{\theta}_1^n(d)$ *and* $\hat{\theta}_2^n(d)$ *the respective values of* θ_1 *and* θ_2 *for which the maximum is obtained.*

Proof (by induction). We prove the lemma true for $n = k + 1$, given that it holds for $n = k$. To prove that

$$R_{k+1}(a, b) = \frac{a + b}{2} + P_{k+1}(d),$$

we must maximise

$$\begin{aligned} & R_k(a - (1 - \theta_1), b + \theta_1) a_{10}(\theta_1, \theta_2) + R_k(a + \theta_2, b - (1 - \theta_2)) a_{01}(\theta_1, \theta_2) \\ & \quad + R_k(a, b) a_{00}(\theta_1, \theta_2) + R_k(a - (1 - \theta_1 - \theta_2), b - (1 - \theta_1 - \theta_2)) a_{11}(\theta_1, \theta_2). \end{aligned}$$

By substituting for R_k and regrouping, we find that this may be written as

$$\frac{a+b}{2}[a_{10}(\theta_1,\theta_2)+a_{01}(\theta_1,\theta_2)+a_{00}(\theta_1,\theta_2)+a_{11}(\theta_1,\theta_2)]$$
$$+\left(\theta_1-\frac{1}{2}\right)[a_{10}(\theta_1,\theta_2)+a_{11}(\theta_1,\theta_2)]+\left(\theta_2-\frac{1}{2}\right)[a_{01}(\theta_1,\theta_2)+a_{11}(\theta_1,\theta_2)]$$
$$+P_k(d)[a_{00}(\theta_1,\theta_2)+a_{11}(\theta_1,\theta_2)]+P_k(d-1)a_{10}(\theta_1,\theta_2)+P_k(d+1)a_{01}(\theta_1,\theta_2)$$
$$=\frac{a+b}{2}+\left(\theta_1-\frac{1}{2}+P_k(d-1)\right)a_{10}(\theta_1,\theta_2)+\left(\theta_2-\frac{1}{2}+P_k(d+1)\right)a_{01}(\theta_1,\theta_2)$$
$$+(\theta_1+\theta_2-1)a_{11}(\theta_1,\theta_2)+P_k(d)[1-a_{10}(\theta_1,\theta_2)-a_{01}(\theta_1,\theta_2)].$$

Reorganising, we find this to be equal to

$$\frac{a+b}{2}+P_k(d)+\left(\theta_1-\frac{1}{2}+P_k(d-1)-P_k(d)\right)a_{10}(\theta_1,\theta_2)$$
$$+\left(\theta_2-\frac{1}{2}+P_k(d+1)-P_k(d)\right)a_{01}(\theta_1,\theta_2)+(\theta_1+\theta_2-1)a_{11}(\theta_1,\theta_2).$$

Thus, our expression for $R_{k+1}(a,b)$ becomes

$$R_{k+1}(a,b)=\frac{a+b}{2}+P_k(d)+\max_{\theta_1,\theta_2}\left\{\left[\theta_1-\frac{1}{2}+P_k(d-1)-P_k(d)\right]a_{10}(\theta_1,\theta_2)\right.$$
$$+\left[\theta_2-\frac{1}{2}+P_k(d+1)-P_k(d)\right]a_{01}(\theta_1,\theta_2)$$
$$\left.+(\theta_1+\theta_2-1)a_{11}(\theta_1,\theta_2)\right\}$$
$$=\frac{a+b}{2}+P_{k+1}(d).$$

We have proved the lemma to hold for $n=k+1$ if it does for $n=k$. It is true for $n=0$ since, by definition, $R_0(a,b)=\min\{a,b\}=(a+b)/2+P_0(d)$, where $P_0(d)=-|d|/2$. Thus, it is true for all $n\in\mathbb{N}$, the set of natural numbers.

Lemma 2. $\hat{\theta}_1^n(d)+\hat{\theta}_2^n(d)\geq 1$.

Proof. By Lemma 1, the values $\hat{\theta}_1^n(d)$ and $\hat{\theta}_2^n(d)$ are chosen to maximise

$$(\theta_1-\tfrac{1}{2}+G_1)a_{10}(\theta_1,\theta_2)+(\theta_2-\tfrac{1}{2}+G_2)a_{01}(\theta_1,\theta_2)+(\theta_1+\theta_2-1)a_{11}(\theta_1,\theta_2),$$

where

$$G_1=P_{n-1}(d-1)-P_{n-1}(d),\qquad G_2=P_{n-1}(d+1)-P_{n-1}(d).$$

Let $\theta_1+\theta_2\geq 1$ and recall that p denotes the customer's probability of horse A winning the race; the customer's probability of horse B winning is thus $1-p$. The customer will bet only on horse A if $p>\theta_1$ and $1-p<\theta_2$ or, equivalently, $p>1-\theta_2$. Now, $1-\theta_2\leq\theta_1$, so the condition becomes simply $p>\theta_1$. This occurs with probability $a_{10}(\theta_1,\theta_2)=1-F(\theta_1)$. Similarly, the customer will bet only on horse B if $p<\theta_1$ and $1-p>\theta_2$ or, equivalently, $p<1-\theta_2$. Again, since $1-\theta_2\leq\theta_1$, this means that the necessary condition is $p<1-\theta_2$. This occurs with probability $a_{01}(\theta_1,\theta_2)=F(1-\theta_2)$.

If, however, we restrict ourselves to having $\theta_1 + \theta_2 \geq 1$, the customer cannot bet on both horses, for the following reason: in order for the customer to bet on both horses, we must have $p > \theta_1$ and $1 - p > \theta_2$, i.e. $p < 1 - \theta_2 \Rightarrow \theta_1 < 1 - \theta_2 \Rightarrow \theta_1 + \theta_2 < 1$. Our only remaining choice for the customer, therefore, is not to bet; the probability of this is

$$a_{00}(\theta_1, \theta_2) = 1 - [1 - F(\theta_1) + F(1 - \theta_2)] = F(\theta_1) - F(1 - \theta_2).$$

The objective function is now

$$(\theta_1 - \tfrac{1}{2} + G_1)[1 - F(\theta_1)] + (\theta_2 - \tfrac{1}{2} + G_2)F(1 - \theta_2).$$

Next, let us consider the case where $\theta_1 + \theta_2 = 1 - \varepsilon$ with $\varepsilon > 0$, so $\theta_1 + \theta_2 < 1$. Again, the customer will bet only on horse A if $p > \theta_1$ and $1 - p < \theta_2$. Now, $1 - \theta_2 > \theta_1$, so the condition becomes simply $p > 1 - \theta_2$. This occurs with probability $a_{10}(\theta_1, \theta_2) = 1 - F(1 - \theta_2)$. Similarly, the customer will bet only on horse B if $p < \theta_1$ and $1 - p > \theta_2$ or, equivalently, $p < 1 - \theta_2$. Again, since $1 - \theta_2 > \theta_1$, the necessary condition is $p < \theta_1$. This occurs with probability $a_{01}(\theta_1, \theta_2) = F(\theta_1)$.

The customer will bet on both horses if $p > \theta_1$ and $1 - p > \theta_2$, i.e. $p < 1 - \theta_2$, so we have $a_{11}(\theta_1, \theta_2) = F(1 - \theta_2) - F(\theta_1)$. The probabilities of these outcomes sum to 1; therefore, they are the only outcomes possible in this case. We may now express the objective function representing the bookie's profit in this case as follows:

$$\begin{aligned}(\theta_1 - \tfrac{1}{2} + G_1)[1 - F(1 - \theta_2)] + (\theta_2 - \tfrac{1}{2} + G_2)F(\theta_1) + (\theta_1 + \theta_2 - 1)[F(1 - \theta_2) - F(\theta_1)] \\ = (1 - \theta_2 - \tfrac{1}{2} + G_1)[1 - F(1 - \theta_2)] + (1 - \theta_1 - \tfrac{1}{2} + G_2)F(\theta_1) - \varepsilon.\end{aligned}$$

This is less than

$$(1 - \theta_2 - \tfrac{1}{2} + G_1)[1 - F(1 - \theta_2)] + (1 - \theta_1 - \tfrac{1}{2} + G_2)F(\theta_1).$$

Since $(1 - \theta_1) + (1 - \theta_2) > 1$, this is less than or equal to

$$\max_{\theta_1, \theta_2}\{(\theta_1 - \tfrac{1}{2} + G_1)[1 - F(\theta_1)] + (\theta_2 - \tfrac{1}{2} + G_2)F(1 - \theta_2)\},$$

where the maximum is taken over values of θ_1 and θ_2 such that $\theta_1 + \theta_2 \geq 1$.

3. Alternative strategy

We now introduce an alternative strategy for determining the quoted probabilities, which aims to maximise

$$\min\{\mathrm{E}[A_0 \mid A_N = 0,\ B_N = 0], \mathrm{E}[B_0 \mid A_N = 0,\ B_N = 0]\}.$$

This greatly reduces the amount of calculation involved in the algorithm.

For $0 \leq n \leq N$, define $S_n(a, b)$ to be the maximum value achievable for

$$\min\{\mathrm{E}[A_0 \mid A_n = a,\ B_n = b], \mathrm{E}[B_0 \mid A_n = a,\ B_n = b]\},$$

and let $\{\theta_{1i}(d/n)\colon 1 \leq i \leq n\}$ and $\{\theta_{2i}(d/n)\colon 1 \leq i \leq n\}$ denote the optimal sets of quoted probabilities. We now prove a number of lemmas relating to this strategy, culminating in Lemma 6, which gives the most concise form of the algorithm for maximum profit.

https://doi.org/10.1017/S0001867800001385 Published online by Cambridge University Press

Definition 1. The function $F(\theta)$ is said to be of *concave character* if

(a) the functions $\theta[1 - F(\theta)]$ and $(\theta - 1)[1 - F(\theta)]$ are concave over the interval $[\theta_1^*, \theta_1^{\mathrm{R}}]$ and

(b) the functions $\theta F(1-\theta)$ and $(\theta - 1)F(1-\theta)$ are concave over the interval $[\theta_2^*, \theta_2^{\mathrm{R}}]$,

where θ_1^* is the maximum value that θ_1 takes, and is the value of θ that maximises $\theta[1 - F(\theta)]$, and θ_2^* is the maximum value that θ_2 takes, and is the value of θ that maximises $\theta F(1-\theta)$, as we will prove in Lemma 5.

In the remainder of the paper we shall assume that $F(\theta)$ is of concave character.

Lemma 3. *The maximum value $S_n(a, b)$ is attained by choosing θ_1 and θ_2 to maximise $g(\theta_1, \theta_2; d/n)$, where*

$$g(\theta_1, \theta_2; x) = a_{10}(\theta_1, \theta_2)\theta_1 + a_{01}(\theta_1, \theta_2)\theta_2 + a_{11}(\theta_1, \theta_2)(\theta_1 + \theta_2 - 1) + \min\left\{\frac{x}{2} - a_{10}(\theta_1, \theta_2), -\frac{x}{2} - a_{01}(\theta_1, \theta_2)\right\}.$$

Furthermore, we can write

$$S_n(a, b) = \frac{a+b}{2} + Q_n(d),$$

where $d = a - b$ and $Q_n(d) = nG(d/n)$ with $G(x) = g(\tilde{\theta}_1(x), \tilde{\theta}_2(x); x)$, $\tilde{\theta}_1(x)$ and $\tilde{\theta}_2(x)$ being chosen to maximise $g(\theta_1, \theta_2; x)$.

Proof. Barry and Hartigan (1996) proved that if F is of concave character then this alternative approach produces quoted probabilities which are independent of i, the number of bets on A purchased. Observe that we can write

$$\mathrm{E}[A_0 \mid A_n = a,\ B_n = b] = a + n[a_{10}(\theta_1, \theta_2)(\theta_1 - 1) + a_{01}(\theta_1, \theta_2)\theta_2 + a_{11}(\theta_1, \theta_2)(\theta_1 + \theta_2 - 1)]$$

and

$$\mathrm{E}[B_0 \mid A_n = a,\ B_n = b] = b + n[a_{10}(\theta_1, \theta_2)\theta_1 + a_{01}(\theta_1, \theta_2)(\theta_2 - 1) + a_{11}(\theta_1, \theta_2)(\theta_1 + \theta_2 - 1)].$$

Hence, we may write $\min\{\mathrm{E}[A_0 \mid A_n = a,\ B_n = b], \mathrm{E}[B_0 \mid A_n = a,\ B_n = b]\}$ as

$$n[a_{10}(\theta_1, \theta_2)\theta_1 + a_{01}(\theta_1, \theta_2)\theta_2 + a_{11}(\theta_1, \theta_2)(\theta_1 + \theta_2 - 1)] + \min\{a - na_{10}(\theta_1, \theta_2), b - na_{01}(\theta_1, \theta_2)\},$$

which may be written as

$$\frac{a+b}{2} + ng\left(\theta_1, \theta_2; \frac{d}{n}\right),$$

thus establishing the result.

Lemma 4. $\tilde{\theta}_1(x) + \tilde{\theta}_2(x) \geq 1$, *where the values $\tilde{\theta}_1(x)$ and $\tilde{\theta}_2(x)$ are chosen to maximise $g(\theta_1, \theta_2; x)$.*

https://doi.org/10.1017/S0001867800001385 Published online by Cambridge University Press

Proof. As previously shown in the dynamic programming case, if $\theta_1 + \theta_2 \geq 1$ then $a_{10}(\theta_1, \theta_2) = 1 - F(\theta_1)$, $a_{01}(\theta_1, \theta_2) = F(1 - \theta_2)$, and $a_{11}(\theta_1, \theta_2) = 0$; thus, in this case, we must maximise

$$\min\{a + n([1 - F(\theta_1)](\theta_1 - 1) + F(1 - \theta_2)\theta_2), b + n([1 - F(\theta_1)]\theta_1 + F(1 - \theta_2)(\theta_2 - 1))\},$$

which we may rewrite as

$$n\biggl([1 - F(\theta_1)]\theta_1 + F(1 - \theta_2)\theta_2 + \min\biggl\{\frac{a}{n} - [1 - F(\theta_1)], \frac{b}{n} - F(1 - \theta_2)\biggr\}\biggr).$$

Next, consider values of θ_1 and θ_2 which satisfy $\theta_1 + \theta_2 < 1$, i.e. $\theta_1 + \theta_2 = 1 - \varepsilon$ with $\varepsilon > 0$. In this case, as before, $a_{10}(\theta_1, \theta_2) = 1 - F(1 - \theta_2)$, $a_{01}(\theta_1, \theta_2) = F(\theta_1)$, and $a_{11}(\theta_1, \theta_2) = F(1 - \theta_2) - F(\theta_1)$. Hence, the objective function becomes

$$\begin{aligned}\min\{&a + n([1 - F(1 - \theta_2)](\theta_1 - 1) + F(\theta_1)\theta_2 + [F(1 - \theta_2) - F(\theta_1)](\theta_1 + \theta_2 - 1)),\\ &b + n([1 - F(1 - \theta_2)]\theta_1 + F(\theta_1)(\theta_2 - 1) + [F(1 - \theta_2) - F(\theta_1)](\theta_1 + \theta_2 - 1))\}\end{aligned}$$

or, alternatively,

$$\begin{aligned}n\biggl(&[1 - F(1 - \theta_2)]\theta_1 + F(\theta_1)\theta_2 + [F(1 - \theta_2) - F(\theta_1)](\theta_1 + \theta_2 - 1)\\ &+ \min\biggl\{\frac{a}{n} - [1 - F(1 - \theta_2)], \frac{b}{n} - F(\theta_1)\biggr\}\biggr).\end{aligned}$$

We may rewrite this as

$$n\biggl((1 - \theta_2)[1 - F(1 - \theta_2)] + (1 - \theta_1)F(\theta_1) + \min\biggl\{\frac{a}{n} - [1 - F(1 - \theta_2)], \frac{b}{n} - F(\theta_1)\biggr\} - \varepsilon\biggr),$$

which is strictly less than

$$n\biggl(u_1[1 - F(u_1)] + u_2F(1 - u_2) + \min\biggl\{\frac{a}{n} - [1 - F(u_1)], \frac{b}{n} - F(1 - u_2)\biggr\}\biggr),$$

where we let $u_1 = 1 - \theta_2$ and $u_2 = 1 - \theta_1$. We find that $u_1 + u_2 = 2 - \theta_1 - \theta_2 \geq 1$, so this, in turn, is less than or equal to the maximum over the set where $\theta_1 + \theta_2 \geq 1$; hence, $\theta_1 + \theta_2 \geq 1$ is optimal.

Lemma 5. *Let θ_1^* be the value of θ for which $\theta[1 - F(\theta)]$ is a maximum and let θ_2^* be the value of θ for which $\theta F(1 - \theta)$ is a maximum. Then $\theta_1 \geq \theta_1^*$ and $\theta_2 \geq \theta_2^*$.*

Proof. As previously described, we must maximise the minimum of

$$\begin{aligned}\mathrm{E}[A_0 \mid A_n = a,\ B_n = b] = a + \sum_{i=1}^{n}\{&a_{10}(\theta_{1i}, \theta_{2i})(\theta_{1i} - 1) + a_{01}(\theta_{1i}, \theta_{2i})\theta_{2i}\\ &+ a_{11}(\theta_{1i}, \theta_{2i})(\theta_{1i} + \theta_{2i} - 1)\}\end{aligned}$$

and

$$\begin{aligned}\mathrm{E}[B_0 \mid A_n = a,\ B_n = b] = b + \sum_{i=1}^{n}\{&a_{10}(\theta_{1i}, \theta_{2i})\theta_{1i} + a_{01}(\theta_{1i}, \theta_{2i})(\theta_{2i} - 1)\\ &+ a_{11}(\theta_{1i}, \theta_{2i})(\theta_{1i} + \theta_{2i} - 1)\}.\end{aligned}$$

Using Lemma 3 and the probabilities derived in consequence of Lemma 4, we may rewrite this as the minimum of

$$\mathrm{E}[A_0 \mid A_n = a,\ B_n = b] = a + n([1 - F(\theta_1)](\theta_1 - 1) + F(1 - \theta_2)\theta_2)$$

and

$$\mathrm{E}[B_0 \mid A_n = a,\ B_n = b] = b + n([1 - F(\theta_1)]\theta_1 + F(1 - \theta_2)(\theta_2 - 1)),$$

which minimum may be rewritten in turn as

$$n([1 - F(\theta_1)]\theta_1 + F(1 - \theta_2)\theta_2) + \min_{\theta_1,\theta_2}\{a - n[1 - F(\theta_1)], b - nF(1 - \theta_2)\}.$$

The first term is maximised for $\theta_1 = \theta_1^*$ and the second term is maximised for $\theta_2 = \theta_2^*$. Of the two terms within braces, the first is an increasing function of θ_1 and the second is an increasing function of θ_2. This shows that we want the values of θ_1 and θ_2 never to fall below θ_1^* and θ_2^*. As they are probabilities, neither may be greater than 1. This completes the proof.

Lemma 6. *Define* $x_{\mathrm{R}} = 1 - F(\theta_1^*)$ *and* $x_{\mathrm{L}} = -F(1 - \theta_2^*)$. *The following statements then hold.*

(a) *For* $x \geq x_{\mathrm{R}}$, *we have* $\tilde{\theta}_1(x) = \theta_1^*$, $\tilde{\theta}_2(x) = 1$, *and* $G(x) = -x/2 + \theta_1^*[1 - F(\theta_1^*)]$.

(b) *For* $x \leq x_{\mathrm{L}}$, *we have* $\tilde{\theta}_1(x) = 1$, $\tilde{\theta}_2(x) = \theta_2^*$, *and* $G(x) = x/2 + \theta_1^* F(1 - \theta_2^*)$.

(c) *For* $x_{\mathrm{L}} < x < x_{\mathrm{R}}$, $\tilde{\theta}_1(x)$ *and* $\tilde{\theta}_2(x)$ *are chosen to maximise*

$$g(\theta_1, \theta_2; x) = (\theta_1 - \tfrac{1}{2})[1 - F(\theta_1)] + (\theta_2 - \tfrac{1}{2})F(1 - \theta_2).$$

Hence,

$$G(x) = (\tilde{\theta}_1(x) - \tfrac{1}{2})[1 - F(\tilde{\theta}_1(x))] + (\tilde{\theta}_2(x) - \tfrac{1}{2})F(1 - \tilde{\theta}_2(x))$$

with

$$x = 1 - F(\tilde{\theta}_1(x)) - F(1 - \tilde{\theta}_2(x)).$$

Proof. (a) We have $x = d/n > 1 - F(\theta_1^*)$ or, equivalently, $d > n[1 - F(\theta_1^*)]$. We must maximise

$$Q_n(d) = \max_{\theta_1,\theta_2}\left\{n[1-F(\theta_1)]\theta_1 + nF(1-\theta_2)\theta_2 + \min\left\{\frac{d}{2} - n[1-F(\theta_1)], -\frac{d}{2} - nF(1-\theta_2)\right\}\right\}.$$

Now, by Lemma 5, θ_1^* is the minimum value which θ_1 may take, so $1 - F(\theta_1^*) \geq 1 - F(\theta_1)$ and

$$d > n[1 - F(\theta_1^*)] \geq n[1 - F(\theta_1)] > n[1 - F(\theta_1) - F(1 - \theta_2)].$$

By inspection, we find that

$$\min\left\{\frac{d}{2} - n[1 - F(\theta_1)], -\frac{d}{2} - nF(1 - \theta_2)\right\} = -\frac{d}{2} - nF(1 - \theta_2)$$

if and only if $d \geq n[1 - F(\theta_1) - F(1 - \theta_2)]$, as is the case here. Thus, we must now maximise

$$\begin{aligned} Q_n(d) &= \max_{\theta_1,\theta_2}\left\{n[1 - F(\theta_1)]\theta_1 + nF(1 - \theta_2)\theta_2 - \frac{d}{2} - nF(1 - \theta_2)\right\} \\ &= \max_{\theta_1,\theta_2}\left\{n[1 - F(\theta_1)]\theta_1 + nF(1 - \theta_2)(\theta_2 - 1) - \frac{d}{2}\right\}. \end{aligned}$$

The maximum occurs for $\theta_1 = \theta_1^*$ and $\theta_2 = 1$, and we obtain

$$Q_n(d) = -\frac{d}{2} + n\theta_1^*[1 - F(\theta_1^*)] = n\left\{-\frac{d}{2n} + \theta_1^*[1 - F(\theta_1^*)]\right\} = nG(x),$$

where $G(x) = -x/2 + \theta_1^*[1 - F(\theta_1^*)]$.

(b) We have $x = d/n < -F(1 - \theta_2^*)$ or, equivalently, $d < n[-F(1 - \theta_2^*)]$. Again, we must maximise

$$Q_n(d) = \max_{\theta_1,\theta_2}\left\{n[1-F(\theta_1)]\theta_1 + nF(1-\theta_2)\theta_2 + \min\left\{\frac{d}{2} - n[1-F(\theta_1)], -\frac{d}{2} - nF(1-\theta_2)\right\}\right\}.$$

By Lemma 5, θ_2^* is the minimum value which θ_2 may take, so $-F(1-\theta_2^*) \leq -F(1-\theta_2)$ and

$$d < n[-F(1-\theta_2^*)] \leq n[-F(1-\theta_2)] < n[1 - F(\theta_1) - F(1-\theta_2)].$$

By inspection, we find that

$$\min\left\{\frac{d}{2} - n[1 - F(\theta_1)], -\frac{d}{2} - nF(1-\theta_2)\right\} = \frac{d}{2} - n[1 - F(\theta_1)]$$

if and only if $d \leq n[1 - F(\theta_1) - F(1-\theta_2)]$, as is the case here. Thus, in this case we must now maximise

$$\begin{aligned} Q_n(d) &= \max_{\theta_1,\theta_2}\left\{n[1 - F(\theta_1)]\theta_1 + nF(1-\theta_2)\theta_2 + \frac{d}{2} - n[1 - F(\theta_1)]\right\} \\ &= \max_{\theta_1,\theta_2}\left\{n[1 - F(\theta_1)](\theta_1 - 1) + nF(1-\theta_2)\theta_2 + \frac{d}{2}\right\}. \end{aligned}$$

The maximum occurs for $\theta_1 = 1$ and $\theta_2 = \theta_2^*$, and we obtain

$$Q_n(d) = \frac{d}{2} + n\theta_2^*[F(1-\theta_2^*)] = n\left\{\frac{d}{2n} + \theta_2^*[F(1-\theta_2^*)]\right\} = nG(x),$$

where $G(x) = x/2 + \theta_2^*[F(1-\theta_2^*)]$.

(c) As (a) corresponds to the case $d > n[1 - F(\theta_1^*)]$ and (b) corresponds to the case $d < n[-F(1-\theta_2^*)]$, we now have $-nF(1-\theta_2^*) \leq d \leq n[1 - F(\theta_1^*)]$. We may write the objective function to be maximised as follows:

$$\min\left\{\frac{d}{2} - n(1-\theta_1)[1-F(\theta_1)] + n\theta_2 F(1-\theta_2), -\frac{d}{2} + n\theta_1[1-F(\theta_1)] - n(1-\theta_2)F(1-\theta_2)\right\}.$$

This is a minimum of two functions, the first of which is an increasing function of θ_1 and a decreasing function of θ_2 for $\theta_2 \geq \theta_2^*$, and the second of which is an increasing function of θ_2 and a decreasing function of θ_1 for $\theta_1 \geq \theta_1^*$. Hence, the maximum occurs at values of θ_1 and θ_2 for which the two functions are equal.

We may express θ_2 in terms of θ_1 because, by equating the two functions, we find that

$$\frac{d}{n} = 1 - F(\theta_1) - F(1-\theta_2) \quad \Longrightarrow \quad \theta_2 = 1 - F^{-1}\left[1 - F(\theta_1) - \frac{d}{n}\right],$$

https://doi.org/10.1017/S0001867800001385 Published online by Cambridge University Press

which gives us

$$Q_n(d) = \max_\theta \left\{ \frac{d}{2} - n(1-\theta)[1-F(\theta)] + n\left[1 - F^{-1}\left(1 - F(\theta) - \frac{d}{n}\right)\right]\left[1 - F(\theta) - \frac{d}{n}\right]\right\}.$$

We may rewrite this as $Q_n(d) = nG(d/n)$, where

$$G\left(\frac{d}{n}\right) = \max_\theta \left\{ \frac{d}{2n} - (1-\theta)[1-F(\theta)] + \left[1 - F^{-1}\left(1 - F(\theta) - \frac{d}{n}\right)\right]\left[1 - F(\theta) - \frac{d}{n}\right]\right\}.$$

This completes the proof.

4. The 'changing expectations' strategy

In this section we use the algorithm associated with the strategy introduced and examined in the previous section. We now calculate the quoted probabilities based on the expectation that they will remain fixed for all remaining customers, thus reducing the complexity of the algorithm; in reality, we will recalculate them after each customer. We then place a bound on the difference between profits obtained using the different methods.

Let $\hat{R}_n(a,b)$ be the value of $\mathrm{E}[\min\{A_0, B_0\} \mid A_n = a,\ B_n = b]$ achieved if the bookie quotes probabilities $\tilde{\theta}_1(d/n)$ and $\tilde{\theta}_2(d/n)$ instead of the optimal values $\hat{\theta}_1^n(d)$ and $\hat{\theta}_2^n(d)$. Then clearly, by definition,

$$\hat{R}_n(a,b) \le R_n(a,b).$$

How much worse (less) than $R_n(a,b)$ is $\hat{R}_n(a,b)$? Using the formula derived in Lemma 1, we can write

$$\hat{R}_n(a,b) = \frac{a+b}{2} + \hat{P}_n(d),$$

where $\hat{P}_0(d) = -|d|/2$ and, for $n \ge 1$,

$$\begin{aligned}\hat{P}_n(d) = \hat{P}_{n-1}(d) &+ \left[1 - F\left(\tilde{\theta}_1\left(\frac{d}{n}\right)\right)\right]\left[\tilde{\theta}_1\left(\frac{d}{n}\right) - \frac{1}{2} + \hat{P}_{n-1}(d-1) - \hat{P}_{n-1}(d)\right] \\ &+ F\left(1 - \tilde{\theta}_2\left(\frac{d}{n}\right)\right)\left[\tilde{\theta}_2\left(\frac{d}{n}\right) - \frac{1}{2} + \hat{P}_{n-1}(d+1) - \hat{P}_{n-1}(d)\right].\end{aligned}$$

Theorem 1. *Assume that F is of concave character. Define $\gamma_0 = 0$ and, for $n \ge 1$,*

$$\gamma_n = \gamma_{n-1} + \max_d\{V_n(d)\},$$

where

$$\begin{aligned}V_n(d) &= Q_n(d) - Q_{n-1}(d) \\ &\quad - \left[1 - F\left(\tilde{\theta}_1\left(\frac{d}{n}\right)\right)\right]\left[\tilde{\theta}_1\left(\frac{d}{n}\right) - \frac{1}{2} + Q_{n-1}(d-1) - Q_{n-1}(d)\right] \\ &\quad - F\left(1 - \tilde{\theta}_2\left(\frac{d}{n}\right)\right)\left[\tilde{\theta}_2\left(\frac{d}{n}\right) - \frac{1}{2} + Q_{n-1}(d+1) - Q_{n-1}(d)\right].\end{aligned}$$

Then

$$Q_n(d) - \gamma_n \le \hat{P}_n(d) \le P_n(d) \le Q_n(d)$$

for all d.

https://doi.org/10.1017/S0001867800001385 Published online by Cambridge University Press

Proof. It follows from Theorem 1 of Barry and Hartigan (1996) that $P_n(d) \le Q_n(d)$. It remains to prove that $\hat{P}_n(d) \ge Q_n(d) - \gamma_n$.

The theorem is true for $n = 0$ since $Q_0(d) = \hat{P}_0(d) = -|d|/2$. We now proceed by induction. Assume that

$$\hat{P}_{n-1}(d) \ge Q_{n-1}(d) - \gamma_{n-1}$$

for all d. Then, since the function

$$x + \left[1 - F\left(\tilde{\theta}_1\left(\frac{d}{n}\right)\right)\right]\left[\tilde{\theta}_1\left(\frac{d}{n}\right) - \frac{1}{2} + y - x\right] + F\left(1 - \tilde{\theta}_2\left(\frac{d}{n}\right)\right)\left[\tilde{\theta}_2\left(\frac{d}{n}\right) - \frac{1}{2} + z - x\right]$$

is an increasing function of y and z, and an increasing function of x provided that $F(\tilde{\theta}_1(d/n)) - F(1 - \tilde{\theta}_2(d/n)) \ge 0$ (which is true by Lemma 4), we have

$$\begin{aligned}
\hat{P}_n(d) &\ge Q_{n-1}(d) - \gamma_{n-1} + \left[1 - F\left(\tilde{\theta}_1\left(\frac{d}{n}\right)\right)\right]\left[\tilde{\theta}_1\left(\frac{d}{n}\right) - \frac{1}{2} + Q_{n-1}(d-1) - Q_{n-1}(d)\right] \\
&\quad + F\left(1 - \tilde{\theta}_2\left(\frac{d}{n}\right)\right)\left[\tilde{\theta}_2\left(\frac{d}{n}\right) - \frac{1}{2} + Q_{n-1}(d+1) - Q_{n-1}(d)\right] \\
&\ge Q_n(d) - \gamma_n,
\end{aligned}$$

provided that $\gamma_n \ge \gamma_{n-1} + V_n(d)$, which is true by definition. The result follows by induction.

Lemma 7. *Assume that the function $G(x)$ has a first derivative, denoted $G'(x)$, and that there exists a constant $M < \infty$ such that*

$$|G'(x) - G'(y)| \le M|x - y|$$

for all x and y. Then the sequence $\{\gamma_n\}$ converges to ∞ at the same rate as $\log n$.

Proof. Recall that we can write $Q_n(d) = nG(d/n)$ and that, hence,

$$\begin{aligned}
V_n(d) &= nG\left(\frac{d}{n}\right) - (n-1)G\left(\frac{d}{n-1}\right) \\
&\quad - \left[1 - F\left(\tilde{\theta}_1\left(\frac{d}{n}\right)\right)\right]\left[\tilde{\theta}_1\left(\frac{d}{n}\right) - \frac{1}{2} + (n-1)\left(G\left(\frac{d-1}{n-1}\right) - G\left(\frac{d}{n-1}\right)\right)\right] \\
&\quad - F\left(1 - \tilde{\theta}_2\left(\frac{d}{n}\right)\right)\left[\tilde{\theta}_2\left(\frac{d}{n}\right) - \frac{1}{2} + (n-1)\left(G\left(\frac{d+1}{n-1}\right) - G\left(\frac{d}{n-1}\right)\right)\right].
\end{aligned}$$

The theorem will follow if we can prove that $\max_d\{V_n(d)\}$ is of order $1/n$ for sufficiently large n.

We must consider a number of cases.

Case (a): $d/n \ge x_{\mathrm{R}}$. Here $\tilde{\theta}_1(d/n) = \theta_1^*$ and $\tilde{\theta}_2(d/n) = 1$. Hence,

$$\begin{aligned}
V_n(d) &= nG\left(\frac{d}{n}\right) - (n-1)G\left(\frac{d}{n-1}\right) \\
&\quad - [1 - F(\theta_1^*)]\left[\theta_1^* - \frac{1}{2} + (n-1)\left(G\left(\frac{d-1}{n-1}\right) - G\left(\frac{d}{n-1}\right)\right)\right].
\end{aligned}$$

Since $d/(n-1) \ge d/n \ge x_{\mathrm{R}}$, we have

$$G\left(\frac{d}{n}\right) = -\frac{d}{2n} + \theta_1^*[1 - F(\theta_1^*)]$$

https://doi.org/10.1017/S0001867800001385 Published online by Cambridge University Press

and

$$G\left(\frac{d}{n-1}\right) = -\frac{d}{2(n-1)} + \theta_1^*[1 - F(\theta_1^*)].$$

Hence,

$$\begin{aligned} V_n(d) &= [1 - F(\theta_1^*)]\left[\frac{1}{2} - (n-1)\left(G\left(\frac{d-1}{n-1}\right) - G\left(\frac{d}{n-1}\right)\right)\right] \\ &= [1 - F(\theta_1^*)]\left[\frac{1}{2} + G'(a)\right] \end{aligned}$$

for some $a \in ((d-1)/(n-1), d/(n-1))$. However,

$$|\tfrac{1}{2} + G'(a)| = |-G'(x_{\mathrm{R}}) + G'(a)|.$$

If $a \geq x_{\mathrm{R}}$ then $|-G'(x_{\mathrm{R}}) + G'(a)| = 0$, and if $a \leq x_{\mathrm{R}}$ then $|-G'(x_{\mathrm{R}}) + G'(a)| \leq M|a - x_{\mathrm{R}}| \leq M/(n-1)$. In either case, we have

$$V_n(d) \leq \frac{M}{n-1}.$$

Case (b): $d/n \leq x_{\mathrm{L}}$. Here $\tilde{\theta}_2(d/n) = \theta_2^*$ and $\tilde{\theta}_1(d/n) = 1$. Hence,

$$\begin{aligned} V_n(d) &= nG\left(\frac{d}{n}\right) - (n-1)G\left(\frac{d}{n-1}\right) \\ &\quad - [F(1-\theta_2^*)]\left[\theta_2^* - \frac{1}{2} + (n-1)\left(G\left(\frac{d+1}{n-1}\right) - G\left(\frac{d}{n-1}\right)\right)\right]. \end{aligned}$$

Since $d/(n-1) \leq d/n \leq x_{\mathrm{L}}$, we have

$$G\left(\frac{d}{n}\right) = \frac{d}{2n} + \theta_2^* F(1-\theta_2^*)$$

and

$$G\left(\frac{d}{n-1}\right) = \frac{d}{2(n-1)} + \theta_2^* F(1-\theta_2^*).$$

Hence,

$$\begin{aligned} V_n(d) &= F(1-\theta_2^*)\left[\frac{1}{2} - (n-1)\left(G\left(\frac{d+1}{n-1}\right) - G\left(\frac{d}{n-1}\right)\right)\right] \\ &= F(1-\theta_2^*)\left[\frac{1}{2} - G'(b)\right] \end{aligned}$$

for some $b \in ((d+1)/(n-1), d/(n-1))$. However,

$$|\tfrac{1}{2} - G'(b)| = |G'(x_{\mathrm{L}}) - G'(b)|.$$

If $b \leq x_{\mathrm{L}}$ then $|G'(x_{\mathrm{L}}) - G'(b)| = 0$, and if $b \geq x_{\mathrm{L}}$ then $|G'(x_{\mathrm{L}}) - G'(b)| \leq M|b - x_{\mathrm{L}}| \leq M/(n-1)$. In either case, we have

$$V_n(d) \leq \frac{M}{n-1}.$$

Case (c): $x_{\mathrm{L}} < d/n < x_{\mathrm{R}}$. Here, by virtue of Lemma 5(c), we can write

$$\begin{aligned} V_n(d) = (n-1)\Big[G\Big(\frac{d}{n}\Big) - G\Big(\frac{d}{n-1}\Big)\Big] \\ - (n-1)\Big[1 - F\Big(\tilde{\theta}_1\Big(\frac{d}{n}\Big)\Big)\Big]\Big[G\Big(\frac{d-1}{n-1}\Big) - G\Big(\frac{d}{n-1}\Big)\Big] \\ - (n-1)F\Big(1 - \tilde{\theta}_2\Big(\frac{d}{n}\Big)\Big)\Big[G\Big(\frac{d+1}{n-1}\Big) - G\Big(\frac{d}{n-1}\Big)\Big]. \end{aligned}$$

Define

$$\Gamma = \Big\{\frac{d}{n}, \frac{d}{n-1}, \frac{d-1}{n-1}, \frac{d+1}{n-1}\Big\} \quad \text{and} \quad \hat{G}(x) = G\Big(\frac{d}{n-1}\Big) + \Big(x - \frac{d}{n-1}\Big)G'\Big(\frac{d}{n-1}\Big).$$

Since

$$G(x) = G\Big(\frac{d}{n-1}\Big) + G'(c)\Big(x - \frac{d}{n-1}\Big)$$

for some $c \in (x, d/(n-1))$, we have

$$|\hat{G}(x) - G(x)| = \Big|\Big[G'\Big(\frac{d}{n-1}\Big) - G'(c)\Big]\Big[x - \frac{d}{n-1}\Big]\Big|.$$

Since all points in Γ are within a distance $2/(n-1)$ of $d/(n-1)$, we can bound the difference between $V_n(d)$ and $\hat{V}_n(d)$ (i.e. $V_n(d)$ with G replaced by $\hat{G}$) by a multiple of $1/n$. This just leaves the term

$$\begin{aligned} (n-1)G'\Big(\frac{d}{n-1}\Big)\Big[\frac{d}{n} - \frac{d}{n-1} - \Big[1 - F\Big(\tilde{\theta}_1\Big(\frac{d}{n}\Big)\Big)\Big]\Big(\frac{d-1}{n-1} - \frac{d}{n-1}\Big) \\ - F\Big(1 - \tilde{\theta}_2\Big(\frac{d}{n}\Big)\Big)\Big(\frac{d+1}{n-1} - \frac{d}{n-1}\Big)\Big] \\ = G'\Big(\frac{d}{n-1}\Big)\Big[-\frac{d}{n} + 1 - F\Big(\tilde{\theta}_1\Big(\frac{d}{n}\Big)\Big) - F\Big(1 - \tilde{\theta}_2\Big(\frac{d}{n}\Big)\Big)\Big], \end{aligned}$$

which is equal to 0 since

$$1 - F\Big(\tilde{\theta}_1\Big(\frac{d}{n}\Big)\Big) - F\Big(1 - \tilde{\theta}_2\Big(\frac{d}{n}\Big)\Big) = \frac{d}{n}.$$

This completes the proof.

5. The effect of F

The expected average gain per customer when dynamic programming is used is $r_N = R_N(0, 0)/N$. In many instances, we have

$$r_N = s_N + O\Big(\frac{\log N}{N}\Big),$$

where $s_N = S_N(0, 0)/N = G(0)$. Therefore, we can examine $G(0)$ to investigate the effects of different choices of $F(\theta)$.

https://doi.org/10.1017/S0001867800001385 Published online by Cambridge University Press

Example 1. Suppose that

$$F(\theta) = \begin{cases} \dfrac{\theta}{2a}, & 0 \le \theta \le a, \\ \dfrac{1}{2}, & a \le \theta \le 1-a, \\ 1 - \dfrac{1-\theta}{2a}, & 1-a \le \theta \le 1, \end{cases}$$

for some value of $a \in [0, \frac{1}{2}]$. Observe that $a = \frac{1}{2}$ corresponds to the uniform distribution. It is easy to check that $\theta_1^* = \theta_2^* = 1 - a$, $x_{\mathrm{L}} = -\frac{1}{2}$, and $x_{\mathrm{R}} = \frac{1}{2}$. It follows in a straightforward manner that the functions $\theta[1 - F(\theta)] = \theta F(1-\theta)$ and $(\theta - 1)[1 - F(\theta)] = (\theta - 1)F(1-\theta)$ are concave on the interval $[1 - a, 1]$, whence F is of concave character.

There are two different solutions depending on whether or not $a \ge \frac{1}{4}$. Since $G(x)$ is symmetric, we need only consider the case $x \ge 0$. For $a \ge \frac{1}{4}$, we have

$$\tilde{\theta}_1(x) = \begin{cases} \dfrac{3}{4} - ax, & 0 \le x \le 1 - \dfrac{1}{4a}, \\ 1 - a, & x \ge 1 - \dfrac{1}{4a}, \end{cases}$$

and

$$\tilde{\theta}_2(x) = \begin{cases} \dfrac{3}{4} + ax, & 0 \le x \le 1 - \dfrac{1}{4a}, \\ 1 - a + 2ax, & 1 - \dfrac{1}{4a} \le x \le \dfrac{1}{2}, \\ 1, & x \ge \dfrac{1}{2}. \end{cases}$$

Hence, using symmetry, we have

$$G(x) = \begin{cases} \dfrac{x}{2} + \dfrac{1-a}{2}, & x \le -\dfrac{1}{2}, \\ \left(\dfrac{1}{2} - a\right) + \left(\dfrac{1}{2} - 2a\right)x - 2ax^2, & -\dfrac{1}{2} \le x \le \dfrac{1}{4a} - 1, \\ \dfrac{1}{16a} - ax^2, & \dfrac{1}{4a} - 1 \le x \le 1 - \dfrac{1}{4a}, \\ \left(\dfrac{1}{2} - a\right) - \left(\dfrac{1}{2} - 2a\right)x - 2ax^2, & 1 - \dfrac{1}{4a} \le x \le \dfrac{1}{2}, \\ -\dfrac{x}{2} + \dfrac{1-a}{2}, & x \ge \dfrac{1}{2}. \end{cases}$$

It is easy to check that $G'(x)$ is continuous and piecewise linear and, therefore, that there exists an M such that $|G'(x) - G'(y)| \le M|x - y|$. Observe that $G(0) = 1/16a$ and we would thus like a to be as small as possible, i.e. the uniform distribution is worst.

For $a \le \frac{1}{4}$, we have

$$\tilde{\theta}_1(x) = 1 - a, \qquad x \ge 0,$$

https://doi.org/10.1017/S0001867800001385 Published online by Cambridge University Press

and

$$\tilde{\theta}_2(x) = \begin{cases} 1 - a + 2ax, & 0 \le x \le \dfrac{1}{2}, \\ 1, & x \ge \dfrac{1}{2}. \end{cases}$$

Hence, using symmetry, we have

$$G(x) = \begin{cases} \dfrac{x}{2} + \dfrac{1-a}{2}, & x \le -\dfrac{1}{2}, \\ \left(\dfrac{1}{2} - a\right) - \left(\dfrac{1}{2} - 2a\right)|x| - 2ax^2, & -\dfrac{1}{2} \le x \le \dfrac{1}{2}, \\ -\dfrac{x}{2} + \dfrac{1-a}{2}, & x \ge \dfrac{1}{2}. \end{cases}$$

It is easy to check that $G'(x)$ is continuous and piecewise linear for $x > 0$. Observe that $G(0) = \frac{1}{2} - a$ and, so, we would again like a to be as small as possible.

Example 2. Suppose that

$$F(\theta) = \begin{cases} 0, & 0 \le \theta \le a, \\ \dfrac{\theta - a}{1 - 2a}, & a \le \theta \le 1 - a, \\ 1, & 1 - a \le \theta \le 1, \end{cases}$$

for some value of $a \in [0, \frac{1}{2}]$. Observe that $a = 0$ corresponds to the uniform distribution. It is easy to check that

$$\theta_1^* = \theta_2^* = \frac{1-a}{2}, \quad x_{\mathrm{L}} = -\frac{1-a}{2(1-2a)}, \quad x_{\mathrm{R}} = \frac{1-a}{2(1-2a)}, \qquad a \le \frac{1}{3},$$
$$\theta_1^* = \theta_2^* = a, \qquad x_{\mathrm{L}} = -1, \qquad x_{\mathrm{R}} = 1, \qquad a > \frac{1}{3}.$$

It follows in a straightforward manner that the functions $\theta[1 - F(\theta)] = \theta F(1 - \theta)$ and $(\theta - 1)[1 - F(\theta)] = (\theta - 1)F(1 - \theta)$ are concave on the interval $[a, 1 - a]$ and, thus, that F is of concave character.

There are two different solutions depending on whether or not $a \le \frac{1}{3}$. Since $G(x)$ is again symmetric, we need only consider the case $x \ge 0$. For $a \le \frac{1}{3}$, we have

$$\tilde{\theta}_1(x) = \begin{cases} \dfrac{1}{2}\left[\dfrac{3}{2} - x(1 - 2a) - a\right], & 0 \le x \le \dfrac{1}{2}, \\ 1 - a - x + 2ax, & \dfrac{1}{2} \le x \le \dfrac{1-a}{2(1-2a)}, \\ \dfrac{1-a}{2}, & x \ge \dfrac{1-a}{2(1-2a)}, \end{cases}$$

and

$$\tilde{\theta}_2(x) = \begin{cases} \dfrac{1}{2}\left[\dfrac{3}{2} + x(1 - 2a) - a\right], & 0 \le x \le \dfrac{1}{2}, \\ 1 - a, & x > \dfrac{1}{2}. \end{cases}$$

https://doi.org/10.1017/S0001867800001385 Published online by Cambridge University Press

Hence, using symmetry, we have

$$G(x) = \begin{cases} \dfrac{x}{2} + \dfrac{(1-a)^2}{4(1-2a)}, & x \le -\dfrac{1-a}{2(1-2a)}, \\[2ex] \dfrac{-x(1-2a)(1+2x)}{2}, & -\dfrac{1-a}{2(1-2a)} \le x \le -\dfrac{1}{2}, \\[2ex] \left(a - \dfrac{1}{2}\right)\left(x^2 - \dfrac{1}{4}\right), & -\dfrac{1}{2} \le x \le \dfrac{1}{2}, \\[2ex] \dfrac{x(1-2a)(1-2x)}{2}, & \dfrac{1}{2} \le x \le \dfrac{1-a}{2(1-2a)}, \\[2ex] -\dfrac{x}{2} + \dfrac{(1-a)^2}{4(1-2a)}, & x \ge \dfrac{1-a}{2(1-2a)}. \end{cases}$$

It is easy to check that $G'(x)$ is continuous and piecewise linear and, therefore, that there exists an M such that $|G'(x) - G'(y)| \le M|x - y|$.

For $a \ge \frac{1}{3}$, again taking only values of $x \ge 0$, we have

$$\tilde{\theta}_1(x) = \begin{cases} \dfrac{1}{2}\left[\dfrac{3}{2} - x(1-2a) - a\right], & 0 \le x \le \dfrac{1}{2}, \\[2ex] 1 - a - x + 2ax, & \dfrac{1}{2} \le x \le 1, \\[2ex] a, & x \ge 1, \end{cases}$$

and

$$\tilde{\theta}_2(x) = \begin{cases} \dfrac{1}{2}\left[\dfrac{3}{2} + x(1-2a) - a\right], & 0 \le x \le \dfrac{1}{2}, \\[2ex] 1 - a, & x \ge \dfrac{1}{2}. \end{cases}$$

Hence, using symmetry, we have

$$G(x) = \begin{cases} \dfrac{x}{2} + \dfrac{(1-a)^2}{4(1-2a)}, & x \le -1, \\[2ex] \dfrac{-x(1-2a)(1+2x)}{2}, & -1 \le x \le -\dfrac{1}{2}, \\[2ex] \left(a - \dfrac{1}{2}\right)\left(x^2 - \dfrac{1}{4}\right), & -\dfrac{1}{2} \le x \le \dfrac{1}{2}, \\[2ex] \dfrac{x(1-2a)(1-2x)}{2}, & \dfrac{1}{2} \le x \le 1, \\[2ex] -\dfrac{x}{2} + \dfrac{(1-a)^2}{4(1-2a)}, & x \ge 1. \end{cases}$$

It is easy to check that $G'(x)$ is continuous and piecewise linear and, therefore, that there exists an M such that $|G'(x) - G'(y)| \le M|x - y|$.

https://doi.org/10.1017/S0001867800001385 Published online by Cambridge University Press

Observe that, in each case, $G(0) = \frac{1}{8} - a/4$, so we would like a to be as small as possible, i.e. the uniform distribution is best.

Example 3. Suppose that

$$F(\theta) = \begin{cases} 0, & 0 \leq \theta \leq a, \\ \dfrac{\theta - a}{b - a}, & a \leq \theta \leq b, \\ 1, & b \leq \theta \leq 1, \end{cases}$$

for some values of $a, b \in [0, 1]$ with $b \geq \max\{(a+1)/2, 2a\}$. Observe that having $a = 0$ and $b = 1$ corresponds to having the uniform distribution and that having $b = 1 - a$ corresponds to the situation in Example 2 in the case $a \leq \frac{1}{3}$. Generally, in this case we must have $\theta_1 \in [a, b]$ and $\theta_2 \in [1 - b, 1 - a]$. We observe that $\theta_1^* = b/2$ and $\theta_2^* = (1 - a)/2$ and that

$$x_{\mathrm{L}} = -\frac{1 - a}{2(b - a)} \quad \text{and} \quad x_{\mathrm{R}} = \frac{b}{2(b - a)}.$$

It follows in a straightforward manner that the functions

$$\theta[1 - F(\theta)] = \theta F(1 - \theta)$$

and

$$(\theta - 1)[1 - F(\theta)] = (\theta - 1)F(1 - \theta)$$

are concave on the interval $[a, b]$ and, so, that F is of concave character.

We have

$$\tilde{\theta}_1(x) = \begin{cases} b, & x \leq -\dfrac{1}{2}, \\ \dfrac{a}{4} + \dfrac{ax}{2} + \dfrac{3b}{4} - \dfrac{bx}{2}, & -\dfrac{1}{2} \leq x \leq \dfrac{1}{2}, \\ b - (b - a)x, & \dfrac{1}{2} \leq x \leq \dfrac{b}{2(b - a)}, \\ \dfrac{b}{2}, & x \geq \dfrac{b}{2(b - a)}, \end{cases}$$

and

$$\tilde{\theta}_2(x) = \begin{cases} \dfrac{1 - a}{2}, & x \leq -\dfrac{1 - a}{2(b - a)}, \\ (b - a)x + 1 - a, & -\dfrac{1 - a}{2(b - a)} \leq x \leq -\dfrac{1}{2}, \\ -\dfrac{3a + b}{4} + 1 + \dfrac{(b - a)x}{2}, & -\dfrac{1}{2} \leq x \leq \dfrac{1}{2}, \\ 1 - a, & x \geq \dfrac{1}{2}. \end{cases}$$

Hence, we have

$$G(x) = \begin{cases} \dfrac{x}{2} + \dfrac{(1-a)^2}{4(b-a)}, & x \le -\dfrac{1-a}{2(b-a)}, \\ \left(a - \dfrac{1}{2}\right)x - (b-a)x^2, & -\dfrac{1-a}{2(b-a)} \le x \le -\dfrac{1}{2}, \\ -\dfrac{(b-a)x^2}{2} + \dfrac{(a+b)x}{2} + \dfrac{b-a}{8} - \dfrac{x}{2}, & -\dfrac{1}{2} \le x \le \dfrac{1}{2}, \\ \left(b - \dfrac{1}{2}\right)x - (b-a)x^2, & \dfrac{1}{2} \le x \le \dfrac{b}{2(b-a)}, \\ -\dfrac{x}{2} + \dfrac{b}{2}\left(1 - \dfrac{b/2-a}{b-a}\right), & x \ge \dfrac{b}{2(b-a)}. \end{cases}$$

It is easy to check that $G'(x)$ is continuous and piecewise linear and, therefore, that there exists an M such that $|G'(x) - G'(y)| \le M|x - y|$.

Observe that $G(0) = (b-a)/8$ and, so, we would like $b - a$ to be as large as possible, i.e. the uniform distribution, for which $a = 0$ and $b = 1$, is best.

6. Final comments

In summary, in this paper we have presented a useful alternative strategy which the bookie may utilise to set odds. A bound has been placed on the penalty which the bookie may expect to incur by using this strategy, and it has been demonstrated using different examples of customers' betting behaviour. It is interesting to note from these examples that, of the cases where the customers' probabilities are distributed 'in the middle' of the range [0, 1] (i.e. Examples 2 and 3), the closer the range of probabilities is to [0, 1], the better it is for the bookie. On the other hand, as we saw in Example 1, if the customers' probabilities are extreme, i.e. clustered around 0 and 1 with few in the middle, this works out even better from the bookie's perspective. This would seem to correspond to the real world, in the sense that the bookie can better predict the behaviour of customers who have strong preferences for one horse or the other.

Throughout this paper, F, the distribution of customers' betting probabilities, was assumed to be known. In reality, of course, this is not generally the case, and the question of what to do when F is unknown provides scope for further investigation.

An important consideration involves the assumptions which were made, in particular those associated with Lemma 6. Future research could include the examination of the necessity of these assumptions and the possibility of relaxing them. We also assumed that the number of customers is known in advance, and placed a restriction on the amount which customers are allowed to bet. We could examine the effects of removing these restrictions.

Our bookie is not assumed to have any opinion on the outcome of the race. We could in future investigate how any such opinion might affect the bookie's strategy for setting odds. We have also restricted the number of horses in the race to only two. An obvious course for potential future research involves increasing the number of horses. This would involve an increase in the number of betting options available to customers, the possibility of betting on multiple horses, and the introduction of different distributions to represent the probabilities customers associate with the different horses.

References

Barry, D. and Hartigan, J. A. (1996). The minimax bookie. *J. Appl. Prob.* **33,** 1093–1107.