Economica

Economica (2013) **80**, 149–170
doi:10.1111/j.1468-0335.2012.00939.x

Optimal Setting of Point Spreads

By Jeremy Sandford and Paul Shea

University of Kentucky and Bates College

Final version received 20 December 2011.

We develop a model of competitive gambling markets addressing two empirical puzzles. First, why do bookmakers not set unbiased lines that try to equalize betting on both sides, and thus profit from commissions with minimal risk? Second, why is there little evidence of bookmakers competing through lower commissions? We show that the interaction between bookmakers' and gamblers' private information can induce biased lines even when all players are maximizing their chances of winning. We also offer an explanation for persistently high commissions charged by seemingly competitive bookmakers; these commissions are necessary to compensate books for assuming the disadvantage of moving first.

Introduction

In 1999, gamblers wagered (mostly illegally) between \$80 and \$380 billion on sporting events in the USA.[1] This figure has likely increased since that time. Much of this wagering was 'threshold' betting on football and basketball. In this game, the bookmaker (also known as the house, or bookie in the case of illegal gambling) acts as the first mover by choosing a betting line: either a threshold margin of victory (a point spread), or a total number of points scored between both teams (an over-under). The gambler moves second, by betting on a team in the case of point spread gambling, or whether more or fewer points are scored in the case of over-under betting. The house charges a fee (known as a *vigorish* or *vig*), often 10% of winning bets, in exchange for its services.[2]

Despite the widespread prevalence of sports gambling, there has been very little theoretical work on how bookmakers set lines and what determines the equilibrium vigorish in a competitive market. This paper attempts to fill this gap. Conventional wisdom suggests that bookmakers set lines in order to equalize the money bet on both sides. Empirical evidence, however, suggests that this is not true. Levitt (2004) finds that in only 20% of NFL games is at least 45% of the total money bet on each side, and that in nearly 10% of games, more than 80% of money is bet on one side. Paul and Weinbach (2007, 2009, 2011) provide further evidence from professional football, college football and professional basketball. They find that more bets are placed on favourites, especially those playing away from home, and on overs when making over-under bets.

The most common explanation for this behaviour is that gamblers are exogenously biased towards certain types of bets.[3] Levitt (2004), for example, finds that gamblers have a preference for road favourites when betting on college and professional football. It is then straightforward for bookmakers to profit from this bias by adjusting these point spreads upwards. Our paper offers an alternative explanation for biased lines.

The paper's first major contribution is the novel result that bookmakers may bias their lines even when bettors are maximizing the expected monetary payoff of their bets. It is thus possible to explain the empirical result of Levitt (2004) without having to assume that gamblers have exogenous biases. Depending on the distribution of the game's outcome, the bookmaker may tend to bias its line upwards, and thus exaggerate its information, or bias it downwards, dampening its information. While the paper discusses the incentives behind this bias at length, the next subsection provides a simple example with most of the intuition of our formal model.

The second empirical puzzle that we address is that while sports gambling appears to be a competitive market, there is little price competition on vigorishes. In the USA, legal Nevada casinos face extensive competition from underground bookies as well as online gambling sites of undetermined legality.[4] In the UK, legal bookmakers proliferate on busy streets.[5] Levitt (2004) notes: 'a major puzzle in this industry is the rarity of price competition, i.e. the vig is almost universally 10%. It is possible that the bearing of risk somehow supports this equilibrium.' He expresses scepticism that ordinary risk-aversion, along with operational costs, is sufficient to support this vigorish. Yet the persistence of 4.5% commissions suggests either the presence of market power or costs sufficiently high to sustain these prices. Since the former seems unlikely, we focus on the latter possibility, viewing the vigorish as compensation for the first mover disadvantage of the bookmaker.[6]

The lack of competition on vigorishes is especially puzzling given that at first glance, a bookmaker would seem to have low fixed costs and near-zero marginal costs. Furthermore, although small bookmakers may risk bankruptcy from a run of bad luck, larger operations should be able to aggregate their risk away through better access to capital and playing the game many times. Yet seemingly competitive books still take approximately 4.5% of every bet in vigorishes. Our paper provides an explanation for this puzzle within a competitive sports gambling market: if gamblers have information at the time when they place their bets that the bookmaker did not have at the time when it set its lines, then the bookmaker is at a first mover disadvantage, even if it has private information of its own. The paper's second main result is that, depending on the distribution of information, this first mover disadvantage alone may be enough to justify a vigorish at least as large as those observed in the data, even in a perfectly competitive market.

The paper's third major contribution is that even if the bookmaker has private information at the time when it sets the line, the advantage of such information may be negated by gamblers correctly inferring the information through the line. We consider two different information regimes. In both regimes, gamblers possess information that the bookmaker cannot observe. Under *public information*, the gambler also has access to the bookmaker's information. Under *private information*, the gambler cannot observe the bookmaker's information, but instead attempts to extract information based on the bookmaker's observed strategy. Surprisingly, under private information, a separating equilibrium exists in which gamblers are able to fully extract the bookmaker's information. The equilibrium is thus the same under both information regimes.

Example

This paper develops a model of sports gambling that works as follows. First, bookmakers decide whether to enter the market, and those that do compete over the vigorish. Second, bookmakers receive some information (which may be either public or private) that allows them to better forecast the result of a sporting event. Third, bookmakers set their lines.[7] Finally, gamblers receive a separate piece of information that allows them to better forecast the outcome, and they then decide which, if any, bet to make. In setting the line, bookmakers must therefore act without observing gamblers' private information.[8] The following simple scenario captures most of the model's intuition.

The Miami Heat and Los Angeles Lakers are playing a basketball game. The median outcome is a tie. Bookmakers then receive information that Luke Walton, the Lakers' best player, is hurt and will not play. Conditional on this information, the median outcome is now a 10-point Heat victory.[9] A bookmaker must now set a line without knowing gamblers' private information. Should it set the line at the conditional median (Heat by 10 points), bias the line upwards (Heat by more than 10 points), or bias the line downwards (Heat by less than 10 points)?

The answer depends on how gamblers' private information interacts with the bookmaker's information. Suppose that the remaining uncertainty about the basketball game results from Kobe Bryant, the Lakers' second best player, being hurt and possibly unable to play. Suppose that the bookmaker considers both (i) and (ii) below to be equally likely:

(i) Bryant will not play, and the median outcome becomes a 30-point Heat victory. Gamblers with this information will bet on the Heat for any line in the neighbourhood of 10 points. Because this information has a large effect on the expected result, outcomes around a 10-point Heat victory are relatively unlikely, and therefore increasing the line above 10 points to increase the bookmaker's probability of winning has only a small benefit.

(ii) Bryant will play, and the median outcome becomes a 5-point Heat victory. Now gamblers will bet on the Lakers for any line around 10 points. However, outcomes in the neighbourhood of a 10-point Heat victory are relatively likely, so lowering the line below 10 points to increase the bookmaker's probability of winning has a relatively large benefit.

Because the marginal effect of changing the line is larger under (ii) than (i), the bookmaker will choose to bias its line downwards when faced with this uncertainty. This slightly lowers the bookmaker's probability of winning in (i), but raises its probability of winning in (ii). Because (i) and (ii) are equally likely, at the time the bookmaker sets the line (before either (i) or (ii) is realized), it is better off setting a line below the median of 10 points.

In this example, gamblers are thus more likely to bet on favourites, overs on over-under bets with large totals, and unders on over-under bets with low totals. This happens because the injuries to Luke Walton and Kobe Bryant are complements; the marginal effect of Bryant's injury on top of Walton's injury is larger than Walton's injury alone. Another possibility is that the injuries to Walton and Bryant are substitutes. If, for example, Bryant needs Walton to play in order to be successful, then the marginal effect of the second injury is less than the first. The opposite argument applies and the bookmaker biases its line upwards. A third possibility is that the informational draws are neither substitutes nor compliments. In this case, lines are not biased. Section III provides numerical examples of all three cases.

Related literature

Despite the popularity of threshold betting on sports, almost no theoretical work examines how bookmakers optimally set lines. One exception is Cain *et al.* (2000), who focus on the risk attitudes of gamblers, assuming away private information and the resulting strategic behaviour that is the focus of our paper. We are not aware of a paper that argues that bias can emerge even when both bookmakers and gamblers rationally maximize the monetary payoff of bets.

Much empirical work on sports betting focuses on the efficiency of point spreads (cf. Gray and Gray 1997; Sauer *et al.* (1988). Several papers study the favourite–longshot bias in horse racing, i.e. the phenomenon of favourites generically performing better than longshots. Shin (1991, 1992) studies this bias in European horse betting markets, while

Hurley and McDonough (1995) examine North American parimutuel betting, in which odds automatically adjust based on the amount bet on each outcome. All three of these papers posit some gamblers as 'informed' about the true state of the world, while other 'uninformed' gamblers simply guess what will happen. Unlike in our model, this literature assumes that gamblers possess a direct informational advantage over the bookmaker. As a result, a 'favourite–longshot' bias emerges where it is optimal for uninformed gamblers to bet on favourites and avoid longshots. Our results are noticeably different. We show that depending on the exact distribution of the event's outcome, it may be optimal for the bookmaker to distort the line in favour of either the favourite or the longshot.

A small literature on forecasting considers models in which two or more agents try to outguess each other. Ottaviani and Sorensen (2005, 2006) consider the case in which different forecasters simultaneously state their predictions, and find that forecasters shade their predictions away from their best guess under private information, as they have an incentive to differentiate themselves from the crowd. Steele and Zidek (1980), Pittenger (1980) and Hwang and Zidek (1982) consider the special case of two forecasters moving sequentially, finding that the second mover wins three out of four times under broad assumptions. However, their first movers act naively, simply forecasting their best guess based on their private information, regardless of what that information is. Our paper adds a strategic first mover, which generically biases its forecast (threshold) away from its best guess based on its private information.

I. The Model

We now develop a model to demonstrate the paper's main results. To most clearly present the model's mechanisms, we suppress features of actual gambling markets (e.g. risk-averse bookmakers, marginal costs, and exogenously biased gamblers) that do not affect our main conclusions. All proofs are in the Appendix.

Risk-neutral gamblers and bookmakers wager over the outcome of a random variable, denoted γ. Bookmakers set a line, or threshold value of γ, denoted y. Gamblers then choose either 'over' or 'under', and the random variable's value is realized. If $\gamma > y$, gamblers who chose over win $(1 - v)$ units from the house, while those who chose under lose 1 unit to the bookmaker, where v is the vigorish charged by the bookmaker. If $\gamma < y$, then under wins and over loses.

Assume that both bookmakers and gamblers have some information about the distribution of γ at the time when they make their decisions. If a bookmaker has information z_1 and a gambler has information z_2, then the distribution of γ is $G(\gamma | z_1, z_2)$.

Bookmakers can freely enter and exit the market. For simplicity, the gambler's side of the market is exogenously determined; there are a continuum of gamblers of measure 1, each betting one unit. The game proceeds in three stages:

1. Bookmakers compete on vigorishes, and entry/exit decisions are made.
2. Information z_1 is realized, and bookmakers set lines $y(z_1)$ for each bet.
3. Gamblers choose over, under, or not to wager at all based on z_2. γ is realized and payoffs are realized.

If we consider only pecuniary payoffs, the interaction between a bookmaker and a gambler is a zero-sum game. Because profit-motivated bookmakers won't operate at a loss, the wide extent of gambling makes it apparent that gamblers receive utility simply from playing the game. We therefore assume that gamblers get a utility of ζ from each wager placed.[10] Assume that each bookmaker pays a fixed cost of c on each line that it sets, while the marginal cost of taking a bet is 0.[11]

We make the following assumptions on the distributions of z_1, z_2 and γ:

(A) z_1 and z_2 are independently and uniformly distributed on the intervals $[-\alpha_1, \alpha_1]$ and $[-\alpha_2, \alpha_2]$, respectively.
(B) For any (z_1, z_2), G is strictly increasing and continuously differentiable in γ on compact support $\Delta_G \subset \mathbb{R}$, with density g.
(C) A higher z_i means that higher values of γ are more likely, i.e. $\partial G/\partial z_i < 0$ for $i = 1, 2$.

Assumptions (A) and (B) are for simplicity; more general distributions for z_1, z_2 and γ can be considered, but they complicate the analysis without adding much generality. Assumption (C) says that the distribution of γ for large z_1 and z_2 stochastically dominates the distribution of γ for low z_1 and z_2; this assumption is why bookmakers and gamblers view z_1 and z_2 as meaningful information.

We assume that all bookmakers are identical and thus receive the same information, and that all gamblers likewise receive the same private information. Although actual gambling markets certainly include additional heterogeneity, the model's main mechanism can be communicated most clearly using representative agents.[12] Further, we assume for simplicity that bookmakers compete on vigs, but not lines (i.e. a bookmaker can steal customers from his competitors by offering a lower vig but not by setting an intentionally unfavourable line).[13]

In our application to sports betting, γ is the result of a sporting event, most likely a margin of victory or sum of points scored. Using a football game as an example, γ may be the actual margin of victory by the home team, while y is the point spread set by the bookmaker. In the context of a sporting event, z_1 and z_2 represent facts that are not universally known but allow agents to better forecast the sporting event's outcome. Examples may include information about injured players, how the teams match up against each other, a team's motivation, or differences in how the house and gambler interpret the same piece of information.

Public information

We first consider the case of *public information* where z_1 is known to both bookmakers and gamblers, while z_2 is known only to gamblers. This informational assumption may apply if a delay exists between bookmakers setting y and gamblers choosing over or under, so that z_2 represents the additional public information that becomes known during the delay. Likewise, it may apply if bookmakers must rely on public information while gamblers possesses additional private information, or if bookmakers must reveal z_1 when setting y. In the next subsection, we show that the public information equilibrium coincides with an equilibrium under *private information* where gamblers extract information about z_1 from a bookmaker's choice of y.

We characterize the subgame-perfect equilibria of the three-stage game between bookmakers and gamblers, beginning with stage 3.

Stage 3: Gamblers choose over or under We solve the model backwards, beginning with a gambler's problem for a fixed vig v, a line y, publicly available information z_1, and private information z_2. A gambler's probability of winning equals $G(y|z_1, z_2)$ if she plays under and $1 - G(y|z_1, z_2)$ if she plays over. Her utility is given by

$$(1)\qquad \max\begin{cases}(1-v)G(y|z_1,z_2)-(1-G(y|z_1,z_2))+\zeta & \text{if she plays under,}\\ (1-v)(1-G(y|z_1,z_2))-G(y|z_1,z_2)+\zeta & \text{if she plays over,}\\ 0 & \text{if she does not play.}\end{cases}$$

Because a gambler can ensure that she has at least a $\frac{1}{2}$ probability of winning any bet by choosing over if $G(y|z_1,z_2)<\frac{1}{2}$ and under otherwise, a sufficient condition for a gambler to place a bet is

$$(2)\qquad v<2\zeta.$$

For a sufficiently low v, the solution to (1) is described by a threshold value of z_2, denoted $x^*(y, z_1)$, above which she chooses over and below which she chooses under. She optimally chooses $x^*(y, z_1)$ so that $G(y|z_1,x^*(y,z_1))=\frac{1}{2}$.[14] Assumptions (B) and (C) directly imply that

$$\frac{\partial}{\partial y}x^*(y,z_1)>0 \quad\text{and}\quad \frac{\partial}{\partial z_1}x^*(y,z_1)<0,$$

and that $x^*(y, z_1)$ is continuous.

Stage 2: The bookmaker's line-setting problem We next consider the bookmaker's problem of setting an optimal line for a fixed vig v. A bookmaker sets his line $y(z_1)$ aware of z_1 but not z_2. By assumption, a bookmaker ignores competitive effects in setting lines; all books with the same z_1 will set the same line.

A bookmaker chooses a line y to maximize its revenue from vigs and losing bets subject to $x=x^*(y,z_1)$. A bookmaker gets revenue 1 from a bet that a gambler loses, and revenue $-(1-v)$ from a bet that a gambler wins. Denote the probability that a bookmaker wins a bet by

$$(3)\qquad \psi(z_1,y)=\frac{1}{2\alpha_2}\left(\int_x^{\alpha_2}G(y|z_1,z_2)\mathrm{d}z_2+\int_{-\alpha_2}^{x}1-G(y|z_1,z_2)\mathrm{d}z_2\right).$$

The first term of $\psi(z_1, y)$ is the probability that under wins and the gambler chooses over. The second term is the probability that over wins and the gambler chooses under. A bookmaker's optimal line $y^*(z_1)$ maximizes its profit:

$$(4)\qquad y^*(z_1)\quad\text{maximizes}\quad \psi(z_1,y)-(1-\psi(z_1,y))(1-v).$$

Lemma 1 characterizes the bookmaker's optimal strategy $y^*(z_1)$.

Lemma 1

1. A solution to the bookmaker's problem (4) exists and is continuous in z_1.
2. The bookmaker's public information equilibrium strategy $y^*(z_1)$ is described by

$$(5)\qquad \frac{1}{2\alpha_2}\left(\int_{x^*(y^*(z_1),z_1)}^{\alpha_2}g(y^*(z_1)|z_1,z_2)\mathrm{d}z_2-\int_{-\alpha_2}^{x^*(y^*(z_1),z_1)}g(y|z_1,z_2)\mathrm{d}z_2\right)=0.$$

3. A sufficient condition for $y^*(z_1)$ to be a strictly increasing function is that $\frac{\partial}{\partial z_1} g(y|z_1, z_2)$ and $\frac{\partial}{\partial y} g(y|z_1, z_2)$ be sufficiently small.

$y^*(z_1)$ is the line that maximizes the bookmaker's probability of winning a bet under public information, and minimizes each gambler's win probability. The bookmaker loses money on the average bet; it wins a bet with probability below $\frac{1}{2}$, as gamblers play over when they have a high z_2 and under when they have a low z_2, whereas a bookmaker sets the line knowing only the distribution of z_2. Therefore the vig that a bookmaker charges must be high enough to compensate it for its lower probability of winning.

The next subsection and Section II further characterize $y^*(z_1)$. The next subsection shows that when z_1 is a bookmaker's private information, there is an equilibrium in which the book sets line $y^*(z_1)$, as if z_1 is public information, and gamblers correctly infer the bookmaker's information from the line set. Section II shows that the lines that bookmakers set under either informational regime are generally biased (different from the median outcome conditional on z_1), and gives conditions for when lines are biased upwards and downwards.

Stage 1: Entry/exit and the bookmaker's vig-setting problem In a market with free entry and exit, the equilibrium vig will be such that bookmakers make zero expected profit.[15] By assumption, all gamblers will patronize only the bookmaker with the lowest vig, and if there are two or more such bookmakers, gamblers split evenly across all such books. Each bookmaker's profit from setting a line is equal to the measure of gamblers multiplied by the expected revenue from each bet placed, minus the fixed cost c. Given that each bookmaker has a fixed cost of operating, it is straightforward that the competitive outcome will leave a single firm, whose vig is just large enough to break even. Proposition 2 describes the competitive equilibrium of the stage 1 game and, by extension, the subgame-perfect equilibrium of the overall public information game.

Proposition 2 In a competitive equilibrium, one bookmaker operates, sets a line $y^*(z_1)$ according to (5), and charges a vig equal to

$$v^* = \frac{c + 1 - 2E_{z_1}\psi(z_1, y^*(z_1))}{1 - E_{z_1}\psi(z_1, y^*(z_1))}. \tag{6}$$

Only one bookmaker operates in a competitive equilibrium because of our assumption of a constant marginal cost of taking a bet. If bookmakers had increasing marginal costs, then the number of bookmakers operating in equilibrium would be increasing in the concavity of the bookmaker cost function. Our approach is for simplicity. The bookmaker's line-setting problem does not depend on this assumption.

Private information

We now consider private information, where z_1 is a bookmaker's private information and gamblers observe only y and z_2. We might expect that this assumption would increase the bookmaker's chances of winning and thus reduce the equilibrium vig. Surprisingly, should the bookmaker's public information strategy $y^*(z_1)$ be strictly increasing, there is a private information equilibrium in which a bookmaker behaves as though its information is public, and gamblers correctly infer a bookmaker's private information from the line. A bookmaker is thus no better off by having private information.

If a gambler believes that a bookmaker is playing its public information strategy, it would seem that a bookmaker has an incentive to manipulate its line in order to convey false information to gamblers. For example, by setting a line $y > y^*(z_1)$, a bookmaker may convince gamblers that its private information suggests a higher outcome of γ than it actually does, and thus may lead gamblers to interpret their own information differently. But bookmakers do not do this because a gambler's public information strategy is chosen to minimize the probability of a bookmaker winning the bet. The marginal effect of a small change in a gambler's strategy on a bookmaker's win probability is therefore zero. A bookmaker thus sets the line thinking only of the direct effect of a marginal change in the line on its probability of winning, and ignores the indirect effect of a line change on a gambler's strategy, as in equation (5). At the margin, the bookmaker's problem is therefore identical to the public information case, where a change in the line set has zero impact on a gambler's strategy.[16]

We begin our formal analysis in stage 3. For any $y \in \Delta_G$, a gambler has beliefs $\pi_y(z_1)$, assigning a probability to each $z_1 \in [-\alpha_1, \alpha_1]$ such that

$$\frac{1}{2\alpha_1}\int_{-\alpha_1}^{\alpha_1} \pi_y(z_1)\mathrm{d}z_1 = 1.$$

A gambler's strategy under private information is a function $\widetilde{x} : \Delta_G \to [-\alpha_2, \alpha_2]$ mapping the threshold y into a choice of $\widetilde{x}$ such that she plays 'over' if and only if $z_2 \geq \widetilde{x}(y)$. A gambler optimally sets $\widetilde{x}^*$ such that $E_\pi G(y|z_1, \widetilde{x}^*(y)) = \frac{1}{2}$.[17]

Suppose that gamblers hold the following beliefs:

$$z_1^*(y) = \begin{cases} z_1^{-1}(y^*) & \text{with probability 1 if } y \in [y^*(-\alpha_1), y^*(\alpha_1)], \\ -\alpha_1 & \text{with probability 1 if } y < y^*(-\alpha_1), \\ \alpha_1 & \text{with probability 1 if } y > y^*(\alpha_1). \end{cases} \tag{7}$$

$z_1^*(y)$ is the inverse of a bookmaker's strategy $y^*(z_1)$ from the public information case, with out-of-equilibrium beliefs defined so that a bookmaker has no incentive to set a very low or very high line.

If a gambler's beliefs are given by (7), then $\widetilde{x}^*(y) = x^*(y, z_1^*(y))$, that is, she plays the same threshold as in the public information case. Now consider stage 2. A bookmaker can do no better than playing its public information strategy $y^*(z_1)$, thus negating the value of any private information that it has. Proposition 3 demonstrates that these strategies and beliefs comprise an equilibrium in stage 2, so long as $y^*(z_1)$ is a strictly increasing function.

Proposition 3 Under private information, provided that $y^*(z_1)$ is a strictly increasing function, there is a perfect Bayesian equilibrium in which the following hold:

1. Gamblers play strategy $\widetilde{x}^*(y) = x^*(y, z_1^*(y))$, where $x^*(y, z_1)$ is the optimal public information strategy.
2. Gamblers hold beliefs $z_1^*(y)$ as described in (7).
3. Bookmakers play $y^*(z_1)$, i.e. their optimal public information strategy from (5).

A bookmaker therefore plays its public information strategy, even though doing so allows gamblers to indirectly observe z_1 without error. Given that behaviour in stages 2 and 3 is identical to the public information case, the equilibrium vig set by bookmakers in stage 1 will match the public information v^* from Proposition 2.

We now more fully characterize the bookmaker's equilibrium strategy $y^*(z_1)$, and examine the determinants of the direction of bias in the line.

II. Biased Lines

Returning to the example of the Introduction, suppose that bookmakers and gamblers are wagering over Miami's margin of victory over Los Angeles. Interpret the Luke Walton injury—the first piece of information revealed—as a high z_1. Certainty that Kobe Bryant will not play is a high z_2, while certainty that Bryant will play is a low z_2. Given the result of Proposition 3, it does not matter if information about Walton's injury is known only to bookmakers or is publicly known.

Let $m(z_1)$ be the median of γ conditional only on z_1, that is,

$$\frac{1}{2\alpha_2}\int_{-\alpha_2}^{\alpha_2} G(m(z_1)|z_1, z_2)\mathrm{d}z_2 = \frac{1}{2}. \tag{8}$$

In this section, we characterize $y^*(z_1)$ in relation to $m(z_1)$. If a bookmaker sets $y^*(z_1) > m(z_1)$ when z_1 is large (upward bias), it will also set $y^*(z_1) < m(z_1)$ for low z_1 (downward bias), as the cases of high and low z_1 are mirror images of one another. For simplicity, we focus only on high values of z_1. Define m^* to be the median of γ before z_1 and z_2 are realized, that is,

$$\frac{1}{4\alpha_1\alpha_2}\int_{-\alpha_1}^{\alpha_1}\int_{-\alpha_2}^{\alpha_2} G(m^*|z_1, z_2)\mathrm{d}z_1\mathrm{d}z_2 = \frac{1}{2}. \tag{9}$$

That G is continuously decreasing in z_1 implies that there exists an uninformative draw of z_1 such that the bookmaker's conditional median equals m^*. We define this draw as z_1^*:

$$\frac{1}{2\alpha_2}\int_{-\alpha_2}^{\alpha_2} G(m^*|z_1^*, z_2)\mathrm{d}z_2 = \frac{1}{2}.$$

Consider a z_1 such that $z_1 > z_1^*$. There are three possibilities for the optimal line, $y^*(z_1)$:

1. $y^*(z_1) > m(z_1)$ (bookmaker biases line upwards);
2. $y^*(z_1) < m(z_1)$ (bookmaker biases line downwards);
3. $y^*(z_1) = m(z_1)$ (bookmaker does not bias line).

The direction of bias in the bookmaker's line depends on the complementarity or substitutability of the two pieces of information. Specifically, if a high z_2 does not increase the median outcome much (z_1 and z_2 are substitutes), then the density of γ will be high in the neighbourhood of $m(z_1)$. As a gambler with a high z_2 will choose over, a bookmaker has a strong incentive to set $y(z_1) > m(z_1)$ to increase its chances of winning with under in the event that z_2 is large. If a small z_2 decreases the median outcome far below $m(z_1)$, then the density of γ around $m(z_1)$ will be low, and the benefit to the bookmaker of setting $y(z_1) < m(z_1)$ will be small (but positive, as gamblers with low z_1 will play under). The bookmaker, being relatively more concerned with a high z_2 than a low z_2, sets $y(z_1) > m(z_1)$. The opposite story holds if z_1 and z_2 are complements.

Proposition 4 formalizes the above intuition, giving sufficient conditions for upward and downward bias. For any given distribution, and for $z_1 > z_1^*$ evaluate the bookmaker's first-order condition (5) at $m(z_1)$. If the result is positive, a bookmaker biases the line upwards; if the first-order condition at $m(z_1)$ is negative, a bookmaker sets the line with a downward bias. If the first-order condition is zero, he does not bias the line. Proposition 4 gives tractable sufficient conditions for upward and downward bias.[18]

Proposition 4

A (upward bias). If $g(m(z_1)|z_1, \varepsilon) \geq g(m(z_1)|z_1, -\varepsilon)$ for all $z_1 > z_1^*$ and for all $\varepsilon > 0$, and $G(m(z_1)|z_1, z_2)$ is concave in z_2, then a bookmaker will bias the line upwards so that $y(z_1) > m(z_1)$ for all $z_1 > z_1^*$.

B (downward bias). If $g(m(z_1)|z_1, \varepsilon) \leq g(m(z_1)|z_1, -\varepsilon)$ for all $z_1 > z_1^*$ and for all $\varepsilon > 0$ and $G(m(z_1)|z_1, z_2)$ is convex in z_2 then a bookmaker will bias the line downwards so that $y(z_1) < m(z_1)$ for all $z_1 > z_1^*$

C (no bias). If $g(m(z_1)|z_1, \varepsilon) = g(m(z_1)|z_1, -\varepsilon)$ for all $\varepsilon > 0$ and $G(m(z_1)|z_1, 0) = \frac{1}{2}$ for all z_1 then a bookmaker will set an unbiased line.

It follows immediately that if a bookmaker biases his line upwards, then the magnitude of bias in the line is increasing in z_1. Thus a bookmaker actually sets y so that he exaggerates his information more the stronger it is. Of course, the conditions of Proposition 4 are not exhaustive. It is also entirely possible there exists some distribution such that a bookmaker biases the line upwards for some $z_1 > z_1^*$ and downwards for other $z_1 > z_1^*$

Biased lines have empirical support. $y^*(z_1) = m(z_1)$ represents the conventional wisdom that books set lines that result in equal wagering on each type of bet. The outcomes in A and B of Proposition 4, however, follow the empirical results of Levitt (2004) and Paul and Weinbach (2007, 2009, 2011), which suggest that lines are often biased. Consider the outcome in B. Here, the bookmaker dampens its information and sets lines between the unconditional median and the median conditional on z_1 As a result, gamblers on observing z_2 are more likely to play over when the line is above the unconditional median, and under when it is below the unconditional median. If the line represents a point spread, then bettors are biased towards favourites. If the line is an over-under, then gamblers will be biased towards over the total when the game is expected to be high scoring, and under the total when it is expected to be low scoring. Importantly, these biases do not result from irrationality or bettors maximizing something besides their probability of winning. They are a result of both sides maximizing their chances of winning the game. We can therefore interpret the results of Levitt (2004) and Paul and Weinbach (2007, 2009, 2011) on point spread football as being consistent with substitutability of information in the context of football. It would be interesting for future work to study complementarity versus substitutability in other sports.

Section III gives examples for all three types of bias. Of course, for $z_1 < z_1^*$ the inequalities in possibilities 1–3 reverse. Note that $y^*(z_1^*) = m^* = m(z_1^*)$ regardless of whether a bookmaker otherwise biases the line upwards or downwards.

III. Numerical Examples

We now discuss three examples that illustrate the three types of bias discussed in Proposition 4, as well as an example that allows one agent to have *ex ante* better information than the other.

A numerical example of upward bias

Suppose that z_1 and z_2 are uniformly distributed on $[-1,\ 1]$. Suppose further that

$$G(\gamma|z_1, z_2) = \gamma\left(1 - \tfrac{1}{2}(z_1 + z_2)\right) + (z_1 + z_2)\frac{\gamma^2}{2},$$

with $\Delta_G = [0,\ 1]$. The density of γ is then linear in $z_1 + z_2$, with slope $z_1 + z_2$. It is direct that z_1, z_2 and G satisfy assumptions (A)–(C), and

$$(10) \qquad x^*(y, z_1) = \begin{cases} 1 & \text{if } \frac{1-2y}{y^2-y} - z_1 > 1, \\ -1 & \text{if } \frac{1-2y}{y^2-y} - z_1 < -1, \\ \frac{1-2y}{y^2-y} - z_1 & \text{otherwise.} \end{cases}$$

Note that $z_1^* = 0$. Evaluating the first-order condition (5) at $y(z_1) = m(z_1)$ yields

$$(11) \qquad \left.\frac{\partial\psi(y, z_1)}{\partial y}\right|_{y=m(z_1)} = m(z_1) - \frac{1}{2}$$

as $x^*(m(z_1),\ z_1) = 0$. Because $m(0) = 0$ and m is increasing in its argument, (11) is positive for $z_1 > 0$ and negative for $z_1 < 0$, and a bookmaker biases the line upwards for positive z_1 and downwards for negative z_1. It is also direct to verify that the assumed G meets the conditions for Proposition 4A.

Figure 1 plots the difference between a bookmaker's equilibrium strategy and the conditional median, $m(z_1)$ (i.e. the degree of bias in the line set by a bookmaker).[19] Figure 2 displays a bookmaker's probability of winning for each $z_1 \in [-1,\ 1]$. Integrating over all z_1, a bookmaker's *ex ante* probability of winning in this example is 43.9%. For this probability, as $c \rightarrow 0$, the equilibrium vigorish in a competitive market approaches 22%. With our model, it is thus possible to generate equilibrium vigorishes greater than the 10% observed in gambling markets.[20]

Given that, in equilibrium, gamblers have better information than bookmakers, it is unsurprising that bookmakers win with probability less than 0.5. To eliminate this informational advantage, Figures 1 and 2 also consider the special case of *naive expectations*. Under naive expectations, we assume that a gambler ignores a bookmaker's information by acting as if $z_1 = z_1^* = \frac{1}{2}$. As shown in Figure 2, under naive expectations, a bookmaker's *ex ante* probability of winning is 45.9%, resulting in a first mover disadvantage equal to 4.1%. Public or private information (where a bookmaker's overall win probability is 43.9%) imposes the maximum possible informational disadvantage on a bookmaker. These results roughly quantify this informational disadvantage at 2.0%, surprisingly less than half of a bookmaker's first mover disadvantage.

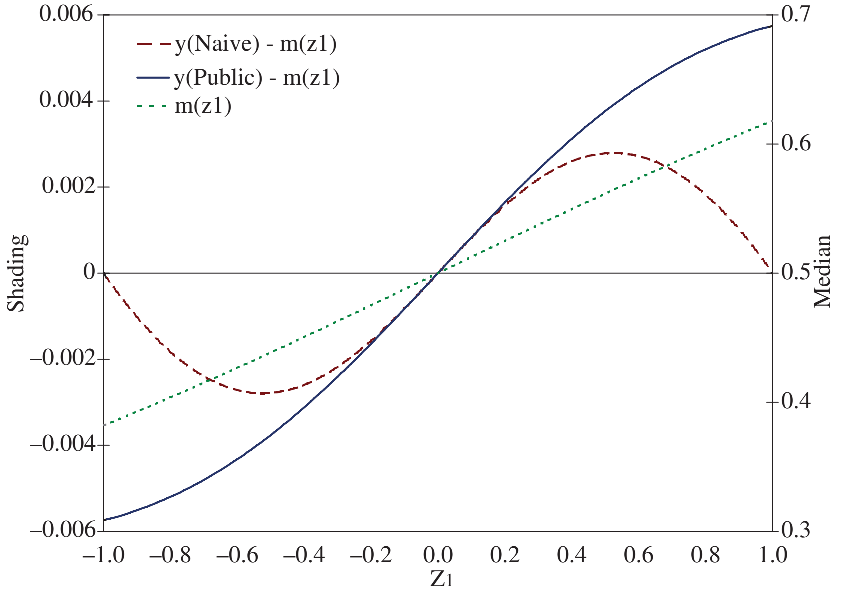

FIGURE 1. Bookmaker's equilibrium bias.

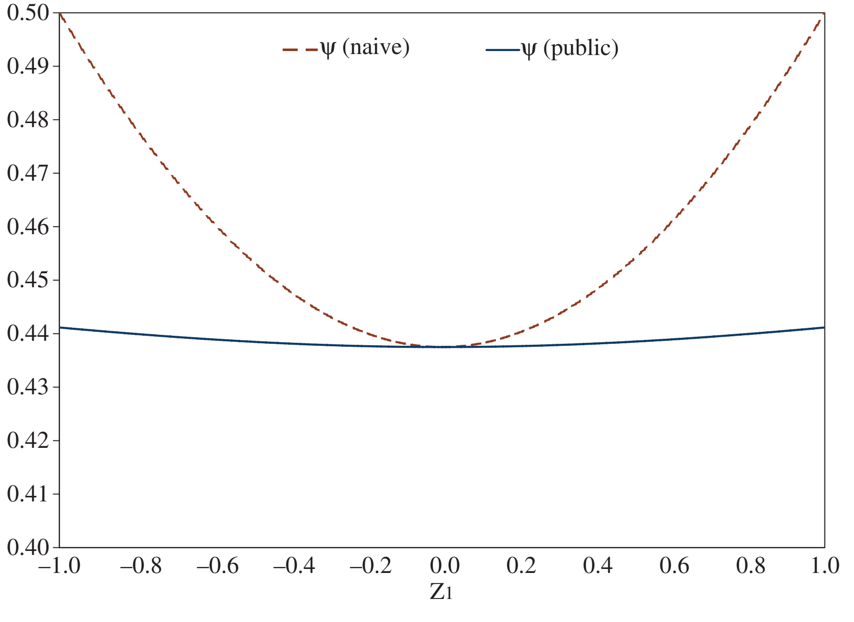

FIGURE 2. Bookmaker's probability of winning.

Why does a bookmaker bias the line in this way? Consider Figure 3. Suppose that its information suggests that γ is likely to be high (say $z_1 = 1$). The expected median outcome conditional on this information is then above the *ex ante* median of $\frac{1}{2}$; it equals 0.62. The density of γ is becoming increasingly steep in z_2, being completely flat for $z_2 = -1$ (together, z_1 and z_2 provide no information) and steepest for $z_2 = 1$ (very likely that γ will be high). A bookmaker thus gains by biasing the line upwards from $m(z_1) = 0.62$, as the rate at which its probability of winning with over decreases (the slope of the density for low z_2) is less than the rate at which its probability of winning with under increases (the slope of the density for high z_2). A bookmaker can increase its win probability on the margin by biasing the line upwards from $m(z_1)$, until the two marginal effects discussed above are equalized. Of course, if z_1 is negative, then it will hedge against negative draws of z_2 by setting $y < m(z_1)$.

A numerical example of downward bias

Now suppose that the density of γ is equal to 2 at $\frac{1}{2}+\frac{1}{4}(z_1+z_2)$, equal to 0 for all $\gamma \notin [-1, 1]$, and linear otherwise. Here, the density of γ is triangle-shaped: unimodal, with a peak at $\frac{1}{2}+\frac{1}{4}(z_1+z_2)$, and linear otherwise. Formally,

$$(12) \qquad g(\gamma) = \begin{cases} \frac{8\gamma}{2+z_1+z_2} & \text{if } \gamma \in [0, \frac{1}{2}+\frac{1}{4}(z_1+z_2)], \\ \frac{8(1-\gamma)}{2-(z_1+z_2)} & \text{if } \gamma \in [\frac{1}{2}+\frac{1}{4}(z_1+z_2), 1], \\ 0 & \text{otherwise.} \end{cases}$$

As in the previous subsection, continue to assume that z_1 and z_2 are uniform on $[-1, 1]$. It is straightforward that

$$m(z_1) = \tfrac{1}{2} - \tfrac{1}{2}z_1 - \tfrac{1}{24}(6-2z_1)^{3/2} + \tfrac{1}{3}\left(\tfrac{3}{2}+\tfrac{1}{2}z_1\right)^{3/2}$$

and that

$$(13) \qquad x(y, z_1) = \begin{cases} -1 & \text{if } y < 1 - \frac{\sqrt{6+2z_1}}{4}, \\ 16y - 8y^2 - 6 - z_1 & \text{if } y \in \left[1 - \frac{\sqrt{6+2z_1}}{4}, \frac{1}{2}\right], \\ 8y^2 - 2 - z_1 & \text{if } y \in \left[\frac{1}{2}, \sqrt{\frac{3+z_1}{8}}\right], \\ 1 & \text{if } y > \sqrt{\frac{3+z_1}{8}}. \end{cases}$$

We numerically compare the maximizer of (5) with $m(z_1)$ in Figure 4. A bookmaker sets a downward-biased line for any $z_1 > 0$, with more bias the higher z_1 is. Figure 5 explains why. If a bookmaker's information is that γ is likely to be high (say $z_1 = 1$), val-

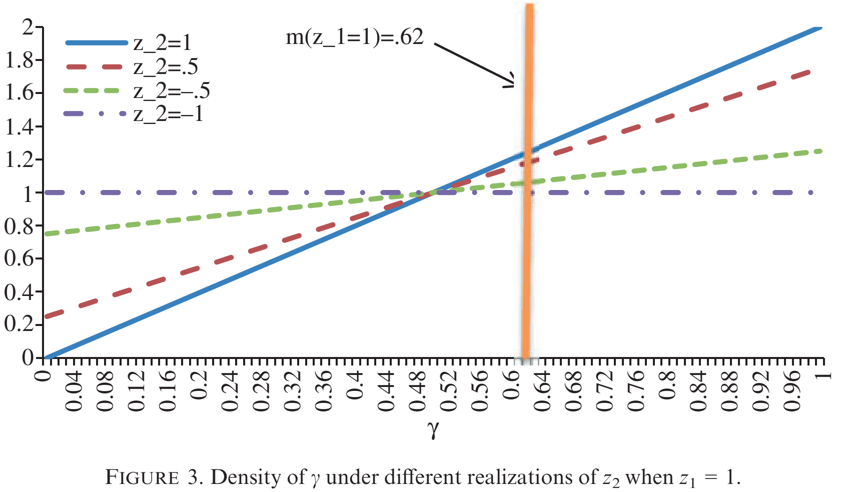

FIGURE 3. Density of γ under different realizations of z_2 when $z_1 = 1$.

ues of γ around the conditional median $m(z_1) = 0.599$ are more likely with a low draw of z_2 than with a high draw. Therefore what a bookmaker gains from lowering the threshold y below $m(z_1)$—an increased probability of winning should z_1 be low—outweighs what it loses—a decreased probability in the event of a high z_2. Because a low z_2 and a high z_2 are equally likely, a bookmaker optimally sets $y < m(z_1)$. The equilibrium vig charged by the house in this example goes to 30.6% as fixed cost c approaches 0.

A numerical example of no bias

In the previous two subsections, a bookmaker has an incentive to bias the line away from or towards the unconditional median because more of the density of γ is concentrated around $m(z_1)$ for high or low z_2, respectively. If, however, a high z_2 and a low z_2 affect the distribution of γ symmetrically, then a bookmaker may set an unbiased line. Suppose that γ is distributed normally with mean $A + B(z_1 + z_2)$ and variance σ^2, where A, B and σ are exogenous parameters. Consider Figure 6. Here, for a given z_1, a high z_2 moves the density of γ to the right by exactly as much as a draw of $-z_2$ would move it to the left. As a result, $g(m(z_1)|z_1, z_2) = g(m(z_1)|z_1, -z_2)$ for any z_1 and z_2. Suppose that $z_1 = 1$. Were a bookmaker to bias the line towards the median, the increased probability that it would win given a low z_2 is exactly equal to the decreased probability of winning from a high z_2. There is thus no reason to bias the line downwards or upwards.

A numerical example where one agent has better information

The previous three examples assume that the bookmaker and gambler have *ex ante* identical information. It can reasonably be argued, however, that either the bookmaker's or the gambler's information is likely to be *ex ante* superior to the other's. We therefore examine the effects of z_1 and z_2 having different supports in the context of our illustrative example at the start of the section, the case of upward bias. We assume that z_1 is distributed uniformly on $[-(2-\kappa), 2-\kappa]$, while z_2 is uniformly distributed on $[-\kappa, \kappa]$. All other assumptions are as at the start of the section.

Increasing κ increases the strength of the gambler's information relative to the house, and increases the line's bias.[21] By the explanation in Section II, for example, a bookmaker sets an upward-biased line because it is relatively more concerned with high values of z_2

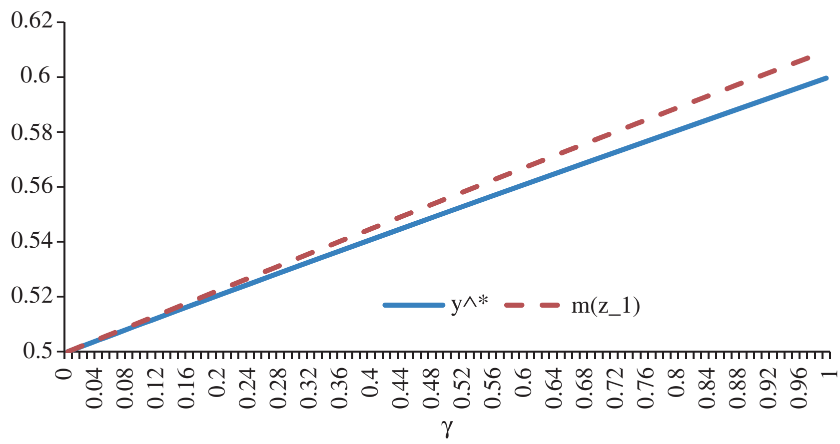

FIGURE 4. Bookmaker biases the line downwards from the conditional median.

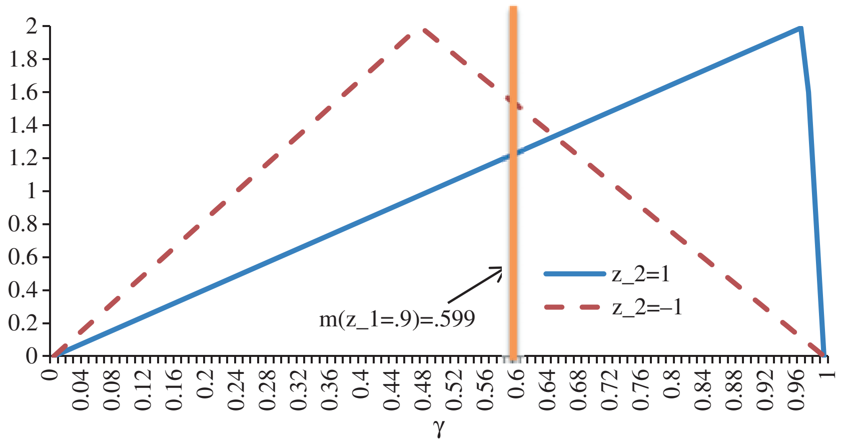

FIGURE 5. Bookmaker biases the line towards the median with $z_1 = 0.9$.

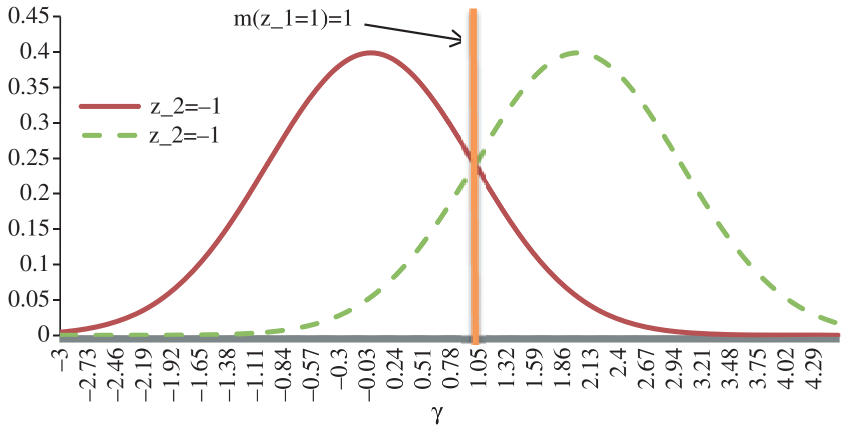

FIGURE 6. Bookmaker sets an unbiased line with $z_1 = 1$.

than with low values. Increasing κ increases the relative importance of z_2 relative to z_1, so it is unsurprising that this amplifies the effect.

Figure 7 plots the bookmaker's probability of winning, averaged across draws of z_1, under both public information and naive expectations. Unsurprisingly, these odds decrease as κ and the informative power of z_2 increase. Under public information, a bookmaker's odds of winning are maximized at $\frac{1}{2}$ when $\kappa = 0$ and z_2 is completely uninformative. The competitive vig (for c close to zero) ranges from 0 when $\kappa = 0$ to 40% when $\kappa = 2$. It equals 10% when $\kappa = 0.44$, a value where the bookmaker enjoys a significant *ex ante* informational advantage over gamblers.

Figure 8 plots the average absolute value of bias in the line. Under public information, low values of κ eliminate the bookmaker's incentive to hedge against high draws of z_2 and little bias occurs. As κ increases, the degree of bias increases for any given z_1. At the same time, higher magnitudes of z_1, which exhibit the most bias, are eliminated from the support of z_1. Initially, the former effect dominates the latter and average bias increases with κ. Above $\kappa = 1.45$, however, the latter effect dominates and average bias begins to converge towards 0 as κ approaches 2.

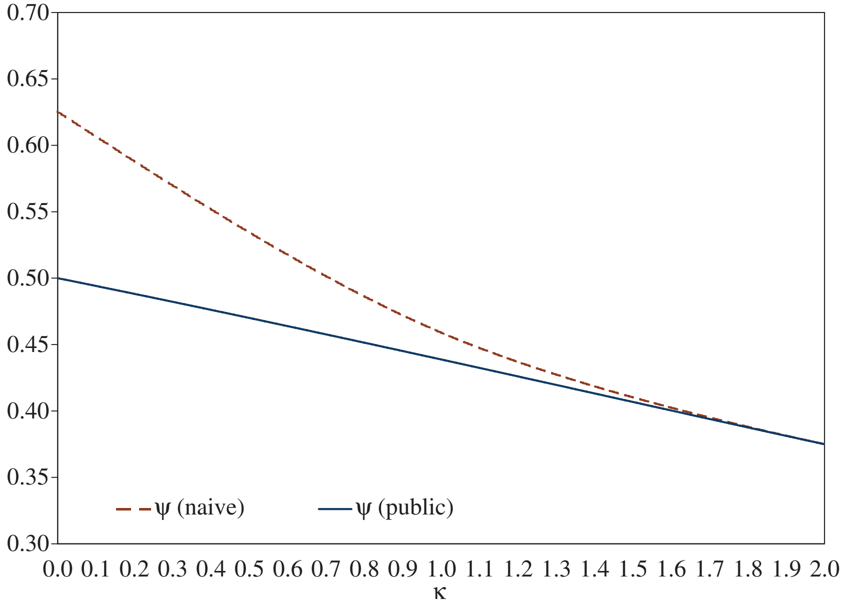

FIGURE 7. Bookmaker's probability of winning.

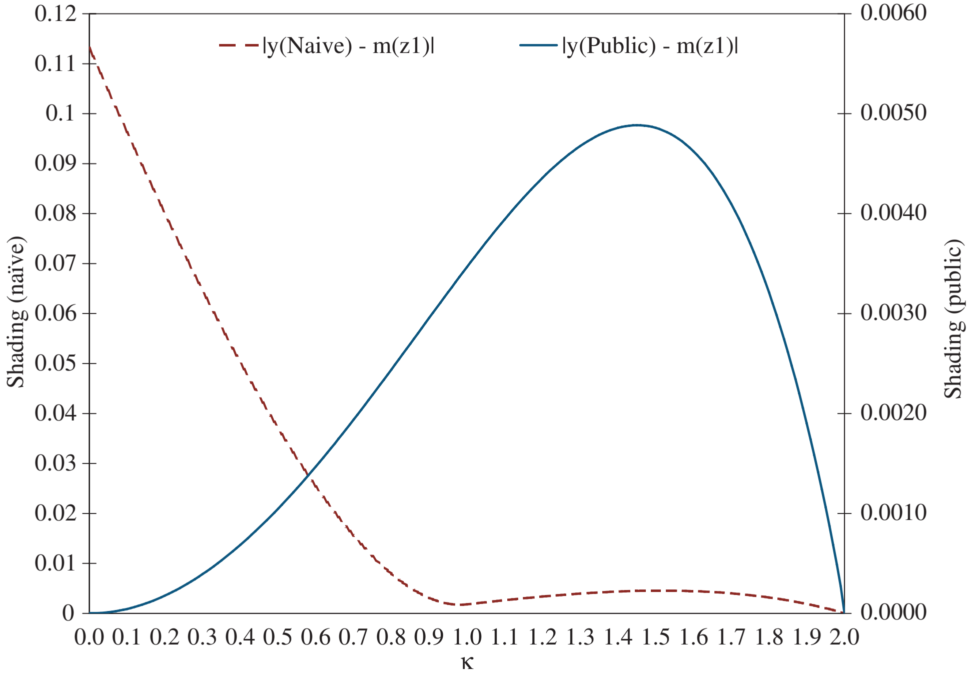

FIGURE 8. Average magnitude of bias.

IV. Line Movements

Bookmakers may revise their initial lines in response to new information or the betting patterns of the public. There is some dispute regarding the significance of actual line revisions. Levitt (2004), for example, argues that changes are small, infrequent, and not significant in his dataset of NFL lines.[22] Kuypers (2000) finds a similar result for UK soccer betting. Gandar *et al.* (1998, 2000), however, present evidence that line changes effectively eliminate the bias that often exists in opening lines in their dataset of NBA games.

So far, we have assumed that the bookmaker's initial line remains unchanged throughout the game. Our model is therefore best suited for explaining the behaviour of the initial pool of bettors. In this section, we introduce another stage to the game where the house may revise its line after z_2 is revealed. Unsurprisingly, additional information induces movements in the line.

When both sides know z_1 and z_2, there is little room for strategic behaviour. The bookmaker sets its revised line at the median, conditional on both z_1 and z_2, and agents are indifferent between over and under. Each side wins the game with probability $\frac{1}{2}$. Proposition 5 formalizes this result.

Proposition 5 In the additional stage of the game, on observing z_1 and z_2, the bookmaker revises its line to the median, conditional on both z_1 and z_2, and both players win with probability $\frac{1}{2}$.

The line therefore moves in the direction of bettors' private information. Like the major result of Gandar *et al.* (1998), the movement of the line eliminates the bias present in the opening line. The bookmaker's odds of winning improve, and now equals a weighted (by the relative fractions of early and late bettors) average of the results from Section I and $\frac{1}{2}$. It is straightforward that if the bookmaker is able to mitigate its first mover disadvantage by adjusting its line in response to new information, the equilibrium vig will be lower.

A different behavioural assumption, however, allows the final line to continue to exhibit bias. Suppose that gamblers form naive expectations where they ignore the house's private information by acting as if $z_1 = 0$. In this case, gamblers play over if and only if $G(y|0, z_2) \leq \frac{1}{2}$. If $z_1 > 0$, then the bookmaker optimally chooses y such that $G(y|0, z_2)$ is arbitrarily greater than $\frac{1}{2}$. In this case, gamblers play under, but are likely to lose. If $z_1 < 0$, however, then the bookmaker sets y such that $G(y|0, z_2)$ is arbitrarily less than $\frac{1}{2}$. Gamblers then play over and are likely to lose.[23] In setting the final line, the bookmaker therefore largely ignores its own information and instead moves the line closer to the line suggested by viewing the gambler's private information in isolation.

V. Conclusion

While common sense suggests that bookmakers should set lines in order to equate betting on each side of an event in order to profit from commissions risk-free, we argue that bookmakers can do better than this if they have information that interacts with private information that is unknown at the time when the line is set. Depending on whether or not a bookmaker's information is a complement or a substitute of information learned by gamblers, a bookmaker can set a biased line either above or below the median outcome, and gamblers will prefer certain types of bets. This theoretical result can help to explain empirical results that overwhelmingly suggest that bookmakers do not set lines so as to equate money placed on both sides of a bet.

Our model abstracts from two notable aspects of point spread gambling. First, we assume that the outcome's distribution is continuous. Most sports outcomes, however, are discrete, and certain margins of victory may be especially likely. If the median outcome occurs with significant probability, then the bookmaker's disadvantage is likely to be compounded.[24] Second, we assume that if there are multiple bookmakers, then they all share the same information at the time when they set their line. It might be interesting for future studies to examine the strategic interaction between early line-setters and late line-setters.

This model may be relevant for settings other than sports gambling. Suppose that the owner and a prospective buyer of a plot of land both commission engineer's reports on how much oil the plot contains; both then have private information. Suppose that one party makes a take-it-or-leave-it offer to the other. Should the offering party set a price equal to the expected value of the land based on his private information, or should he shade his offer away from this expected value, and, if so, in which direction? Future work will consider this question.

APPENDIX

Proof of Lemma 1

1. This follows from the continuity of G, $x^*(y,\ z_1)$ and ψ, and the fact that Δ_G is a compact set.
2. The derivative of (3) is given by

$$\begin{aligned}\frac{\partial\psi(y,z_1)}{\partial y} &= \frac{\partial x^*(y,z_1)}{\partial y}(1-2G(y|z_1,x^*(y,z_1)))f(x^*(y,z_1))\\ &\quad+\frac{1}{2\alpha_2}\left(\int_{x^*(y,z_1)}^{\alpha_2} g(y|z_1,z_2)\mathrm{d}z_2-\int_{-\alpha_2}^{x^*(y,z_1)} g(y|z_1,z_2)\mathrm{d}z_2\right)\\ &=0,\end{aligned}$$

   but the first term is 0, as $G(y|z_1,x^*(y,z_1))=\frac{1}{2}$ for any y. This reflects the fact that as a gambler chooses x^* to minimize ψ, the marginal effect on ψ of changing x^* is 0.
3. These conditions follow from examination of (5). Applying the Leibniz integral rule, the left-hand side of (5) is increasing in z_1 if $\partial\, g(y|z_1,\ z_2)/\ \partial\, z_1$ is sufficiently small and is decreasing in y if $\partial\, g(y|z_1,\ z_2)/\ \partial\, y$ is sufficiently small. These facts together imply that $y(z_1)$ is increasing in z_1.

Proof of Proposition 2

Suppose that there is an equilibrium with $n \geq 2$ firms operating. Let $v(n)$ describe the zero-profit vig if customers are split across the n firms. Book 1's expected profit is

$$\text{(A1)}\qquad \frac{v(n)(1-E_{z_1}\psi(z_1,y^*(z_1)))+2E_{z_1}\psi(z_1,y^*(z_1))-1}{n}-c=0.$$

Consider a deviation to vig $v' = v(n) - \varepsilon$. Then book 1 would capture the entire market, and despite earning a slightly lower vig per customer, would earn a higher total profit:

$$v'(1 - E_{z_1}\psi(z_1, y^*(z_1))) + 2E_{z_1}\psi(z_1, y^*(z_1)) - 1 - c > 0. \tag{A2}$$

The expression for v^* in the proposition follows from setting $n = 1$ in (A1) and solving for v.

Proof of Proposition 3

Under private information, the first-order condition for a maximizer to a bookmaker's win probability located in $y^*(-\alpha_1)$, $y^*(\alpha_1)]$ is given by

$$\begin{aligned}\frac{\partial\widetilde{\psi}(y,z_1)}{\partial y} &= \frac{\partial\widetilde{x}^*(y)}{\partial y}(1 - 2G(y|z_1, \widetilde{x}^*(y)))f(\widetilde{x}^*(y))\\ &\quad + \frac{1}{2\alpha_2}\left(\int_{\widetilde{x}^*(y)}^{\alpha_2} g(y|z_1,z_2)\mathrm{d}z_2 - \int_{-\alpha_2}^{\widetilde{x}^*(y)} g(y|z_1,z_2)\mathrm{d}z_2\right)\\ &= 0.\end{aligned} \tag{A3}$$

For any z_1, $G(\widetilde{y}^*(z_1)|z_1, \widetilde{x}^*(y)) = \frac{1}{2}$, so the first term is 0 at $y^*(z_1)$ and $\widetilde{x}^*(y)$. Then, by virtue of (5) holding at $y^*(z_1)$, (A3) holds when evaluated at $y^*(z_1)$, as $\widetilde{x}(y^*(z_1)) = x^*(y^*(z_1), z_1)$. Under $y^*(z_1)$, a gambler's beliefs are correct and $\widetilde{x}^*(y)$ is optimal for a gambler, as she is correctly inverting the bookmaker's public information strategy $y^*(z_1)$. Finally, given the gambler's stated beliefs for out-of-equilibrium lines, any line $y < y^*(-\alpha_1)$ is dominated by $y^*(\alpha_1)$, and $y > y^*(\alpha_1)$ is dominated by $y^*(\alpha_1)$.

Proof of Proposition 4

Consider the first-order condition characterizing the bookmaker's optimal strategy $y^*(z_1)$, equation (5). For $z_1 > z_1^*$, evaluate the left-hand side at $y = m(z_1)$.

First, if $G(m(z_1)|z_1, z_2)$ is concave in z_2, then $x^*(m(z_1), z_1) \leq 0$, while if $G(m(z_1)|z_1, z_2)$ is convex in z_2, then $x^*(m(z_1), z_1) \geq 0$. To see why, recall that $m(z_1)$ is defined by

$$\frac{1}{2\alpha_2}\int_{-\alpha_2}^{\alpha_2} G(m(z_1)|z_1 z_2)\mathrm{d}z_2 = \frac{1}{2}.$$

That is, $G(m(z_1)|z_1, z_2)$ is above $\frac{1}{2}$ for low z_2 and below $\frac{1}{2}$ for high z_2, with an average value of $\frac{1}{2}$ (recall by Assumption (C) that G is decreasing in z_2).

If $G(m(z_1)|z_1, z_2)$ is convex in z_2, then $x(m(z_1), z_1) \leq 0$.[25] Therefore for any $z_1 > z_1^*$, if the conditions of part A of the proposition are satisfied, then the left-hand side of (5) is positive when evaluated at $y = m(z_1)$, and so is satisfied only for $y > m(z_1)$. Similarly, if the conditions of part B are satisfied, then the left-hand side of (5) is negative at $y = m(z_1)$, and so satisfied only for $y < m(z_1)$. Finally, if the conditions of part C are satisfied, then the first-order condition (5) is satisfied for $y = m(z_1)$ and $y^*(z_1) = m(z_1)$.

Proof of Proposition 5

A gambler chooses over if and only if $G(y|z_1, z_2) \leq \frac{1}{2}$ and chooses under if and only if $G(y|z_1, z_2) > \frac{1}{2}$. Her probability of winning is therefore max $[G(y|z_1, z_2), 1 - G(y|z_1, z_2)]$. A bookmaker clearly minimizes this probability by setting y such that $G(y|z_1, z_2) = \frac{1}{2}$. A gambler thus plays over and wins with probability $\frac{1}{2}$.

NOTES

1. See National Gambling Impact Study Commission (1999).
2. Typically, bettors wager \$11 to win \$10 at a traditional bookmaker. Underground bookies often charge the commission on losses instead of wins. In both cases, the average commission is about 4.54% if bettors bet on both sides equally. For a detailed discussion of the rules followed by bookies and gamblers, see Jeffries and Oliver (2000).
3. Throughout the paper, we use bias to refer to instances where gamblers are more likely to make one type of bet (e.g. favourites as opposed to underdogs, or over as opposed to under) than another. We refer to a line as biased if it induces such behaviour on average. Betting behaviour exhibits *exogenous bias* if bettors simply get more utility from betting on a certain team or outcome regardless of monetary outcome. *Endogenous bias* refers to bettors maximizing their monetary outcome by making some types of bets more often than others. This paper is concerned only with the latter.
4. The legality of online sports betting in the USA through offshore websites remains unclear. In 2006, Congress passed the Unlawful Internet Gambling Enforcement Act, which makes it difficult for potential gamblers to conduct financial transactions with online books. Gamblers therefore face added transaction costs and potential expropriation risk, relative to legal books, when gambling online. Online books hosted in the USA are unambiguously illegal.
5. A Google Maps search shows nearly 100 bookmakers in London and its suburbs alone; while some bookmakers such as William Hill have multiple storefronts, there are at least a dozen distinct bookmakers operating in London.
6. Certainly, part of the vigorish is needed to compensate the bookmaker for its conventional costs, including both operating costs such as overheads and labour, and the risk that is assumed by the bookmaker if it is credit constrained or if the game is not played often enough for it to aggregate its risk away. Anecdotal reports of how books operate, such as Jeffries and Oliver (2000), offer compelling evidence that a bookie's costs for setting up an office and taking bets are quite small relative to other similar-sized businesses.
7. Allowing bookmakers to simultaneously compete over the vigorish and line does not affect our major results, subject to a minor regularity condition.
8. Section IV extends the game so that bookmakers may adjust their lines after observing gamblers' private information.
9. Here the conditional median implies integrating over all possible realizations of the gamblers' private information.
10. Multiple surveys (e.g. Neighbors *et al.* 2002; Gupta and Derevensky 1998) find that gamblers wager for 'fun', 'social reasons', 'excitement' and 'enjoyment'. Other papers making a similar assumption include Levitt (2004) and Farrell and Walker (1999). Peel and Law (2009) provide a more detailed discussion of this modelling approach. We assume that the utilities from winning and losing bets are the same, for simplicity.
11. If the fixed cost c is too large compared to ζ, then there may be no bookmakers in equilibrium. This is consistent with the observation that there is no wagering on many obscure sporting events.
12. An obvious generalization would be that individual gamblers receive a noisy signal around a 'true' z_2. The effect of this would be that the majority of gamblers, instead of all of them, make a certain wager for a given z_2. If each bookmaker receives a noisy signal of the true z_1, then only the bookmakers with the highest and lowest signals receive any business in equilibrium, and the equilibrium vigorish is higher. In both cases, the model's major conclusions are unaffected.
13. Some bookies offer clients the chance to 'purchase' a half-point movement in the line by agreeing to a double vig. One bookie describes this as a sucker bet, and the uniformity of vigs described in the Introduction suggests that there is little demand for a high-vig, bad-line sports book (see Jeffries and Oliver 2000). We also solved the model for the case where bookmakers simultaneously compete on the vig and the line. Subject to a mild regularity condition, the only notable change to our results is that the equilibrium vig may depend on z_1 in this case.
14. If no such x exists, $x^*(y, z_1) = -\alpha_2$ (gambler always plays over) if $G(y|z_1, z_2) < \frac{1}{2}$ for all $z_2 \in [-\alpha_2, \alpha_2]$, and $x^*(y, z_1) = \alpha_2$ (always plays under) if $G(y|z_1, z_2) > \frac{1}{2}$ for all $z_2 \in [-\alpha_2, \alpha_2]$.
15. Note that different values of z_1 may generate different profits for a bookmaker.
16. There is anecdotal evidence that gamblers attempt to extract the bookmaker's private information based on the line. In September 2000, the unranked Oregon Ducks were strangely about a 3.5-point favourite over the sixth-ranked UCLA Bruins in a college football game. Popular prognosticator Lee

Corso, however, picked Oregon to win the game, because the line suggested that 'somebody knows something' about Oregon being likely to win (*ESPN College Gameday*, 23 September 2000). Oregon won the game 29–10, and finished the season ranked seventh with a 10–2 record. UCLA finished the year 6–6 and unranked.

17. If no such $\tilde{x}$ exists, $\tilde{x}^*(y) = -\alpha_2$ if $E_\pi G(y|z_1, z_2) < \frac{1}{2}$ for all z_2, and $\tilde{x}^*(y) = \alpha_2$ if $E_\pi G(y|z_1, z_2) > \frac{1}{2}$ for all z_2.
18. Proposition 4 gives conditions for upward or downward bias for $z_1 > z_1^*$. The conditions for $z_1 < z_1^*$ are analogous.
19. As the public and private information equilibria coincide, both are represented in Figure 1.
20. If instead $G(\gamma|z_1, z_2) = \gamma(1 - \frac{1}{2}\tau(z_1 + z_2)) + \tau(z_1 + z_2)\gamma^2/2$, where $\tau \in (0, 1)$ is a scaling parameter, then the equilibrium vigorish will be reduced for low τ. As $\tau \to 0$, z_1 and z_2 are less informative, $\psi(z_1, y^*(z_1)) \to \frac{1}{2}$, and $v^* \to 0$. Thus for an appropriately chosen τ, the equilibrium vigorish will equal 10%.
21. This claim follows from direct calculation of the first-order condition (5).
22. One explanation for the lack of line movements is a bookmaker's fear of being 'middled', where it loses the bulk of bets placed on both teams. For example, in 1979, the Pittsburgh Steelers opened as 3.5-point favourites over the Dallas Cowboys in Super Bowl XIII. The majority of gamblers bet on Pittsburgh at this opening line, and the final line drifted up to 4.5 points, where the majority of bettors chose Dallas. Trailing 35–24 with 22 seconds remaining, Dallas scored a touchdown to make the final score 35–31 in favour of Pittsburgh, heroically beating the final point spread. Bookmakers, however, lost or tied (thus foregoing the vigorish) the majority of bets placed on both teams. The event is known as 'Black Sunday' among bookies.
23. This result is sensitive to our assumption that all bettors have the same information. If each gambler receives a noisy signal of z_2, then this strategy will ensure that approximately half of the gamblers play over. The optimal strategy in this case is to set y closer to the median, conditional on both z_1 and z_2.
24. In college football, for example, about 16% of games are decided by either 3 or 7 points.
25. Suppose that $G(m(z_1)|z_1, z_2)$ is convex in z_2 and $x(m(z_1), z_1) > 0$. Then, by convexity of G,

$$\frac{1}{2} - \int_{x(m(z_1,z_1))}^{\alpha_2} G(m(z_1)|z_1, z_2)\mathrm{d}z_2 \leq \int_{2x(m(z_1,z_1))-\alpha_2}^{x(m(z_1,z_1))} G(m(z_1)|z_1, z_2)\mathrm{d}z_2 - \frac{1}{2}.$$

But the fact that $x(m(z_1), z_1) > 0$ then contradicts equation (8). A similar proof shows that if $G(m(z_1)|z_1, z_2)$ is concave in z_2, then $x(m(z_1), z_2) \geq 0$.

REFERENCES

Cain, M., Law, D. and Lindley, D. (2000). The construction of a simple book. *Journal of Risk and Uncertainty*, **20**, 119–40.

Farrell, L. and Walker, I. (1999). The welfare effects of Lotto: evidence from the UK. *Journal of Public Economics*, **72**, 99–120.

Gandar, J., Dare, W., Brown, C. and Zuber, R. (1998). Informed traders and price variations in the betting market for professional basketball. *Journal of Finance*, **53**, 385–401.

———, Zuber, R. and Dare, W. (2000). The search for informed traders in the totals betting market for National Basketball Association games. *Journal of Sports Economics*, **1**, 177–86.

Gray, P. and Gray, S. (1997). Testing market efficiency: evidence from the NFL sports betting market. *Journal of Finance*, **52**, 1725–37.

Gupta, R. and Derevensky, J. (1998). Adolescent gambling behavior: a prevalence study and examination of the correlates associated with problem gambling. *Journal of Gambling Studies*, **14**, 319–45.

Hurley, W. and McDonough, L. (1995). A note on the Hayek hypothesis and the favorite–longshot bias in parimutuel betting. *American Economic Review*, **85**, 949–55.

Hwang, J. T. and Zidek, J. (1982). Limit theorems for out-guesses with mean-guided second guessing. *Journal of Applied Probability*, **19**, 321–31.

Jeffries, J. and Oliver, C. (2000). *The Book on Bookies: An Inside Look at a Successful Sports Gambling Operation*. Boulder, CO: Paladin Press.

Kuypers, T. (2000). Information and efficiency: an empirical study of a fixed odds betting market. *Applied Economics*, **32**, 1353–63.

Levitt, S. (2004). Why are gambling markets organised so differently from financial markets? *Economic Journal*, **114**, 223–46.

National Gambling Impact Study Commission (1999). National Gambling Impact Study Commission Final Report. Available at http://govinfo.library.unt.edu/ngisc/reports/fullrpt.html (accessed 23 March 2012).

NEIGHBORS, C., LOSTUTTER, T., CRONCE, J. and LARIMER, M. (2002). Exploring college student gambling motivation. *Journal of Gambling Studies*, **18**, 361–70.

OTTAVIANI, M. and SORENSEN, P. (2005). *Forecasting and rank-order contests*. Mimeo, Northwestern University.

——— and ———. (2006). The strategy of professional forecasting. *Journal of Financial Economics*, **81**, 441–66.

PAUL, R. and WEINBACH, A. (2007). Does sportsbook.com set pointspreads to maximize profits? Tests of the Levitt model of sportsbook behavior. *Journal of Prediction Markets*, **1**, 209–18.

——— and ——— (2009). Sportsbook behavior in the NCAA football betting market: tests of the traditional and Levitt models of sportsbook behavior. *Journal of Prediction Markets*, **3**, 21–37.

——— and ——— (2011). NFL bettor biases and price setting: further tests of the Levitt hypothesis of sportsbook behavior. *Applied Economic Letters*, **18**, 193–7.

PEEL, D. and LAW, D. (2009). A more general non-expected utility model as an explanation of gambling outcomes for individuals and markets. *Economica*, **76**, 251–63.

PITTENGER, A. O. (1980). Success probabilities for second guessers. *Journal of Applied Probability*, **17**, 1133–7.

SAUER, R., BRAJER, V., FERRIS, S. and MARR, M. (1988). Hold your bets: another look at the efficiency of the gambling market for national football league games. *Journal of Political Economy*, **96**, 206–13.

SHIN, H. S. (1991). Optimal betting odds against insider traders. *Economic Journal*, **101**, 1179–85.

——— (1992). Prices of state contingent claims with insider traders, and the favourite long shot bias. *Economic Journal*, **102**, 426–35.

STEELE, J. M. and ZIDEK, J. (1980). Optimal strategies for second guessers. *Journal of the American Statistical Association*, **75**, 596–601.