UNIVERSITY *of York*

This is a repository copy of *On Sticky Bookmaking as a Learning Device in Horse-Racing Betting Markets*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/191268/

Version: Accepted Version

Article:

Zhang, Chi and Thijssen, Jacco orcid.org/0000-0001-6207-5647 (2022) On Sticky Bookmaking as a Learning Device in Horse-Racing Betting Markets. Journal of Economic Dynamics and Control. 104525. ISSN 0165-1889

https://doi.org/10.1016/j.jedc.2022.104525

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Journal Pre-proof

On Sticky Bookmaking as a Learning Device in Horse-Racing Betting Markets

Chi Zhang, Jacco Thijssen

PII: S0165-1889(22)00229-9
DOI: https://doi.org/10.1016/j.jedc.2022.104525
Reference: DYNCON 104525

To appear in: *Journal of Economic Dynamics & Control*

Received date: 24 June 2021
Revised date: 9 September 2022
Accepted date: 9 September 2022

Please cite this article as: Chi Zhang, Jacco Thijssen, On Sticky Bookmaking as a Learning Device in Horse-Racing Betting Markets, *Journal of Economic Dynamics & Control* (2022), doi: https://doi.org/10.1016/j.jedc.2022.104525

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier B.V.

Journal Pre-proof

On Sticky Bookmaking as a Learning Device in Horse-Racing Betting Markets

Chi Zhang[*] Jacco Thijssen[†]

Journal Pre-proof

Abstract

We present a stochastic dynamic model of the adjustment of betting odds by bookmakers in a horse-racing betting market. We use optimal stopping theory in a two-horse benchmark model with both informed and noise punters. A costly learning process discloses what information the informed traders possess and a risk-neutral bookmaker selects a stopping time at which the betting odds for each horse in a race are adjusted. Our main finding shows that an increased fraction of informed punters has a non-monotonic effect on the loss per trade to the bookmaker. We also find that as the fraction of noise traders goes up, the learning process is less informative, so that the decision to change the prices for each horse is taken sooner.

KEYWORDS: Horse racing; Betting markets; Dynamic pricing; Optimal stopping time; Sequential hypothesis testing

JEL: Z23, C61, D82

[*]Corresponding author. School of Economics, Shandong University, Jinan, Shandong Province, China; c.zhang@sdu.edu.cn.

[†]Management School, University of York, York, North Yorkshire, United Kingdom; jacco.thijssen@york.ac.uk.

Journal Pre-proof

1 Introduction

Horse-racing betting markets provide economists with a ready-made laboratory and large datasets. Therefore, it is no surprise that there is a large body of (empirical) literature examining these markets from different perspectives, e.g., by testing market efficiency theories (see Vaughan Williams (2005) and Hausch et al. (2008) for an overview), explaining the *favorite-longshot bias* (see Griffith (1949), McGlothlin (1956), Ali (1977), Hausch et al. (1981), Asch et al. (1984), Asch et al. (1982), Henery (1985), Brown et al. (1994), Shing and Koch (2008), Ottaviani and Sorensen (2010), Jeong et al. (2019), and references therein), or detecting the incidence of insider trading and decision-making under uncertainty. The literature documenting these problems covers both parimutuel markets, which prevail in North America, and bookmaking markets in the United Kingdom and other Commonwealth countries. The latter system can be traced back to the late 18th century and, on average, the bookmakers have been the winners in the *punter-versus-bookmaker* battle ever since the UK Parliament passed the Gaming Act in 1845.[1]

Journal Pre-proof

The markets for horse-racing betting have various structures. At North American racetracks, betting odds are determined by the parimutuel method, in which the win bet fractions (also known as subjective winning probabilities) are proportional to the amounts wagered. The public directly determines the odds in a parimutuel betting market—the more the public bets on a horse, the shorter its odds and the lower its return upon winning. In the UK and other Commonwealth countries, the dominant structure is the fixed-odds system in which bookmakers offer the odds. In the fixed-odds system, bookmakers set odds that may vary throughout the betting period. Bettors lock in their odds at the time they place a bet, even if the offered odds change later. Bets can also be placed at the "Starting Price", which is defined as the last available odds for the selection. Starting price bets can be placed at any time prior to the race and remain available up until the start of the race. Although far less popular, a parimutuel betting system using the Tote board is also available in the UK for betting both on and off track.

Each race convenes for 20–30 minutes, and is then followed by the next race. In this market,

[1] "All bets are off: Why bookmakers aren't playing fair," https://www.theguardian.com/global/2015/aug/02/betting-horses-gambling-bookmakers-accounts-closed.

Journal Pre-proof

the outcome of the race, and the payoffs occur at the end of each race. Bettors can choose from a variety of wagers. The simplest is a wager to win, which involves picking the horse that finishes first. Alternatively, "Each Way" (EW) bets can be placed, consisting of the *Win* part and the *Place* part.[2] These two betting methods are very popular in the UK.

When compared with the trading frequency in financial markets, Levitt (2004) mentioned that price adjustments are typically small and relatively infrequent in sports wagering and horse racing. While it is possibly optimal to continuously adjust betting odds to avoid any losses, this is not what is observed in practice. In our previous research[3] on on-course racetrack betting markets, the posted odds change, on average, 4–5 times during the betting period. Thus, prices initially remain fairly constant before changing. In this paper, we will analyze the bookmaking market from the point of view of a bookmaker who acts as a market intermediary by setting betting odds for each horse in any race. In particular, we are trying to find the first time (for example, at time t) that the bookmaker decides to deviate from the posted prices, as shown in Figure 1. We posit that bookmakers use given odds to learn about the information that market participants hold. Gabriel and Marsden (1990) argued that the existence of insiders would provide a market signal to bookmakers to cut the odds. Dowie (1976) illustrated that the starting price incorporates more information than any of the various odds generated early, and should include any inside or monopolistically held information. Therefore, we believe that bookmakers choose to change the odds when the losses made on trading by informed punters sufficiently outweigh the benefits of learning.

It is clear from Figure 1 that all bettors in the market can only wager at the opening odds from the opening to time t. During this betting period, the bookmaker learns from the trades and decides whether there are insiders and which horse is more favorable. The realized bets contain less information when punters are less informed (Ottaviani and Sorensen, 2010).

Journal Pre-proof

[2]Horse Racing Rules. https://support.skybet.com/s/article/Horse-Racing-Overall-Rules#Horse_Racing_Rules.

[3]Data in our research containing all price adjustments for each racehorse at different racetracks in Yorkshire were collected from *The Sporting Life* for other purposes. For more details, please refer to https://www.semanticscholar.org/paper/Economic-Analysis-of-Horseracing-Betting-Markets-Zhang/d3345bb01da11658e3a3467e9078935e14d6927b.

Journal Pre-proof

Figure 1: A toy example on the betting period in the bookmaking market. The betting market convenes for 30 minutes. Bookmakers usually post the opening odds to signal punters that they start to take bets. The posted odds remain fairly constant at the initial stage (e.g. from 0 to 10 minutes) and change relatively more frequently in the last few minutes. Note that bettors can only lock in the current posted odds before changing.

In this paper we shall not explore the reasons for price stickiness, but we will focus on an important side-effect. Namely, stickiness allows the bookmaker to learn about information that is held by informed punters. So, while bookmakers may lose money, in expectation, by not changing their odds, they do get some value in return.

We investigate this in the context of a simple stochastic dynamic model in which punters are either informed or are pure noise traders. Thus, the sticky bookmaking is that a rational bookmaker is willing to take initial losses to collect information about informed traders, because these punters are the main source of bookmakers' losses. That is, the presence of insiders imposes adverse selection costs on bookmakers and noise traders. We call this the *learning process*. Our benchmark model offers a simple rule for finding the optimal time at which the value of learning no longer outweighs the losses made on informed punters, and thus, the time at which it is optimal to change the betting odds. The main driver of our result is that, to learn about the private information held by informed punters, bookmakers must keep the betting odds constant for a while. In such a short period, bookmakers can extract the signal from the noise.

The main contribution of this study is the construction of a dynamic theory of price discovery in horse-racing markets. Additionally, we provide a methodological contribution by introducing optimal stopping techniques into the betting literature. In economic terms, our most important result is that the decision bounds become wider as the fraction of informed punters becomes larger, but the expected loss made on those punters is non-monotonic. This occurs because of two opposing effects. It is intuitively clear that, if there are more insiders in the market, the bookmakers' losses will be higher, because informed punters can exploit their information advantage. However, a larger

Journal Pre-proof

proportion of informed traders implies that the learning process is more informative. Thus, a choice between the two states is made sooner, which, in turn, lowers the bookmakers' losses, as expected.

In contrast with the standard literature on real options (see, e.g., Dixit and Pindyck, 1994), which suggests that higher uncertainty implies wider decision bounds and, therefore, the higher value of an investment project, we find that the posterior bounds become narrower due to volatility. This is because of the uncertainty introduced to the observations in the form of the trades made by noise punters. If there are more noise punters, then the bookmakers learn less quickly, which leads to narrower bounds of the continuation region. It turns out that the expected loss increases with the volatility, because a decision is taken at worse posterior odds, leading to a higher probability of an erroneous decision being taken.

This study presents a continuous-time two-state model for a two-horse benchmark framework in which the bookmaker changes the odds of horse A if one particular state is revealed or changes the odds of horse B if the evidence sufficiently points to the other state prevailing. No matter which odds are adjusted, the other odds change automatically. We assume that the cumulative sales of tickets on horse A, net of the cumulative sales of tickets on horse B, can be modeled by an arithmetic Brownian motion with either a positive or negative trend. These are the two states of the world: the *high state* implies that informed traders know horse A is going to win, whereas the *low state* means horse B is more favorable among insiders. The true state remains unknown to the bookmaker *ex ante*, so they have no information about which horse the informed traders are backing. These two states determine the direction in which the odds are changed. This information is gradually disclosed through a sequential observation of trades. The bookmaker then acts and will accept the hypothesis of the high state if the cumulative sales on horse A exceed the tickets sold on horse B by an (endogenously determined) sufficient margin. If cumulative sales point to the low state, the odds on horse B will be changed.

Our model is analyzed using the theory of sequential hypothesis testing, as described by, e.g., Peskir and Shiryayev (2006). However, where they analyze an optimal stopping problem framed purely in terms of (non-discounted) statistical losses, we extend the model using the discounted benefits and costs developed by Thijssen and Bregantini (2017) to account for the uncertain number

Journal Pre-proof

of bets that are observed. Our model allows for an analytical solution. We show that there is a unique solution to the problem from simultaneously (and implicitly) solving two nonlinear equations. This solution provides the upper and lower boundaries of the posterior belief of informed punters who know that horse A will win (the high state). These boundaries indicate the optimal point at which the odds should be changed. The upper boundary is the trigger for the bookmaker to change the odds of horse A, whereas the lower boundary is the trigger for horse B. The "continuation region" lies between these two boundaries. In this region, the bookmaker should keep observing trades without changing the odds.

The rest of the paper is structured as follows. We review related literature in Section 2. In Section 3 we build a continuous-time statistical-probabilistic framework of a two-horse betting market. In Section 4, we formulate our optimal stopping problem and its solution. A comparative statics analysis is reported in Section 5, and Section 6 concludes.

Journal Pre-proof

2 Related Literature

Racetrack betting has many parallels with traditional financial markets, and much of the analysis is similar. A betting market is a simple example of a financial market, having a large number of investors (bettors) with heterogeneous beliefs, extensive market information, and ease of entry. There are vast amounts of data in the form of prices and other information. This market also raises informational issues. The major difference is that there is a well-defined termination point at which the value of the asset is fixed at the end of the betting period. The added advantage is that we are no longer bounded by future cash flow and net present value. Although investments in betting markets have negative expected value for most investors, they have wide appeal for entertainment.

Insider trading and adverse selection costs

A central notion of the market microstructure literature is that trades are informative. Information affecting security prices is one of the main subjects in financial markets, and has been researched by numerous scholars. In a series of papers, Bagehot (1971), Copeland and Galai (1983), and Glosten and Milgrom (1985) studied the effect of informed traders on the spreads between bid and ask prices.

Our model is mainly related to the standard literature on insider trading. The idea of insider trading was first introduced by Bagehot in 1971, who mentioned that "every time one investor benefits from a trade [...] another loses" and showed that the market maker always loses to traders with special information and gains with liquidity-motivated transactors. The gains from noise traders must exceed the losses to informed traders. Since this study of Bagehot (1971), the presence of traders with superior information has received widespread attention in the finance literature. Due to the existence of insiders, a closely related topic is the estimation of the theoretical components of the bid–ask spread. Kyle (1985) developed a dynamic model of insider trading in which prices exhibit Brownian motion and the market maker can only see the order imbalance. With sequential auctions, informed traders make positive profits while noise traders provide camouflage, but eventually all private information is incorporated into the prices. He also argued that the presence of insiders imposes adverse selection costs on market makers and liquidity traders. A wider spread is set to compensate the market maker for losses from informed traders. In the Glosten and Milgrom (1985) model, the bid–ask spread can be explained by adverse selection as well as the exogenous arrival pattern of traders. The presence of insiders leads to a positive spread. Because the transaction prices are informative, the spread declines as trading continues. Easley and O'Hara (1987) showed that informed traders prefer to trade larger amounts at any given price, if they wish to trade at all. As a result, larger trade volumes indicate a greater likelihood that the market maker is trading with informed traders. Although the market maker faces uncertainty about whether any individual trader is informed, there is also uncertainty about whether any new information exists. This latter uncertainty dictates that both the size and the sequence of trades matter in determining the price–quantity relationship. Li and Song (2019) showed that market makers are willing to initiate trades with more-informed traders to obtain their private information and then set higher bid–ask spreads to recoup from the liquidity traders.

In the horse-racing betting markets, especially under the fixed-odds system whereby bookmakers set the odds to balance the book, taking into account the adverse selection costs recoups the

Journal Pre-proof

Journal Pre-proof

loss from uninformed traders. While bookmakers face the adverse selection problem, the optimal pricing strategy is to contract the odds, which leads to a favorite-longshot bias.

Shin (1991, 1992, 1993) extended the seminal Glosten and Milgrom (1985) model to the horse-racing context and used it to provide a measure of the degree of insider trading in UK horse-racing betting markets. Fingleton and Waldron (1999) generalized the Shin model and showed that informed trading, the bookmaker's attitude towards risk, and the bookmaker margin are three factors that determine the optimal odds. Schnytzer and Shilony (1995) provided empirical evidence of the presence of inside information in Melbourne horse-betting markets. They showed that punters with inside information, even those exposed to "secondhand" information, change their behavior to increase their payoffs. Positive information can be a significant predictor of a race's outcome. Gabriel and Marsden (1990) provided evidence that Tote (Totalizator) returns on winning bets exceed the starting price returns paid out by bookmakers. Cain et al. (2001) showed that bookmakers pay more generously on favorites than the Tote, but less generously on high-odds bets. The difference between these two papers is that the latter considers the incidence of insider trading.

Favorite-longshot bias and market efficiency

A central theme of the betting literature is the occurrence of the favorite-longshot bias (FLB). That is, the expected return on longshot bets tends to be systematically lower than on favorite bets in betting markets. This FLB is a widely documented empirical fact, often perceived to be an important deviation from the market efficiency hypothesis. The first documentation of this bias is attributed to Griffith (1949), and it has been observed in racetrack betting data from around the world covering both bookmaker and parimutuel markets. Over recent decades, a number of theories have been developed to explain the FLB.

After Griffith first uncovered the FLB, Isaacs (1953) advanced an explanation based on the fact that, in the parimutuel system, bettors placing more money on a horse results in lower odds. Thus, an informed bettor would want to limit the amount bet in order to maximize the profits made. Ali (1977) assumed that bettors have heterogeneous prior beliefs. If bettors are risk neutral but not sophisticated, the parimutuel market mechanism may generate the FLB. Hurley and McDonough (1995) demonstrated that the FLB can be explained on the basis of limited arbitrage by bettors who possess superior information. As the amount of arbitrage is limited and bets placed on the favorite are unable to produce the expected return on a longshot to the same level, the presence of the track take induces an asymmetry in the rational bets, resulting in the FLB. Ottaviani and Sorensen (2010) proposed a purely informational explanation for the FLB in the context of parimutuel betting. To illustrate this explanation, they considered the simplest case in which two horses are ex-ante equally likely to win (1/2), outsiders bet equal amounts on the two horses, the track take is zero, and the number of privately informed insiders is large. They concluded that when there are many privately informed bettors, the FLB arises if they place last-minute bets without knowing the final market odds. Cain et al. (2003) found that sports with higher incidences of insider trading exhibit a strong FLB, while Vaughan Williams and Paton (1997) supported the conclusion that the adverse selection problem may be the reason for the observed FLB. Smith et al. (2006) concluded that the incidence of insider trading, and hence the FLB, is higher in the bookmaking setting than in betting exchanges.

For fixed-odds betting markets, Shin (1991, 1992, 1993) also explained the FLB as the response of an uninformed bookmaker to the private information possessed by insiders. Our paper is closely related to these papers, as we assume that a bookmaker faces different types of bettors—some are informed, while others are uninformed. The insiders are perfectly informed, meaning that they always pick the winning horse; the outsiders are assumed to have uniformly distributed beliefs. In the absence of outsiders, the bookmaker would make a loss at any finite price. In the bookmaking system, the bookmaker sets prices considering that outsiders are involved in the betting. This is a natural precondition for the bookmaker to run the market and make any positive profit at all. Intuitively, a lower market price encourages fewer outsiders to participate, and enhances the bias chosen by the bookmaker to protect against the adverse selection of bettors. Our model uses some assumptions proposed by Shin, but analyzes a different question. We are interested in the optimal time at which the bookmaker decides to adjust the odds of the horses. Up to this point, we observe the cumulative sales and bear some costs to learn whether there are insiders in the market and how

Journal Pre-proof

they place their bets.

Most of the literature on horse-racing betting concerns statistical tests of the extent to which these markets are efficient. The abovementioned FLB is often perceived to be an important deviation from the market efficiency hypothesis. Evidence of inefficiency also provides indirect support for the popular hypothesis that insiders operate in these markets.

Research on market efficiency theory was first conducted by Fama in the early 1960s. Market efficiency *per se* is not testable and must proceed on the models of market equilibrium, which can be described in terms of expected returns in theoretical and empirical work. Dowie (1976) found that if insider information does exist, it will not be available to the public and only the final Starting Prices can reveal it. Therefore, he concluded that there is no insider information. Crafts (1985) suggested that the way Dowie tests the "strong inefficient" market is not appropriate. To argue this, Crafts redesigned a test using the same data source as Dowie and gave evidence to prove that there exist profitable opportunities for insider trading. Gabriel and Marsden (1990, 1991) compared the returns to the bookmaker's starting prices and to the parimutuel Tote bets, based on data from the 1978 horse-racing season in the UK. Their analysis indicated that, on average, the Tote returns for winning bets are persistently higher than the odds set by bookmakers given that the risk and payoffs are widely available under both systems. This implies that the UK horse-racing betting market fails to satisfy the conditions of semi-strong and strong efficiency. Both the FLB and market inefficiency imply that informed traders do exist in betting markets.

Journal Pre-proof

3 Bookmaking as Sequential Hypothesis Testing

In this section, we describe the market in which bets are placed on a two-horse race. Consider a bookmaker who sets betting odds for the two horses. Unlike betting activity in other sports where bookmakers are allowed to change the odds during the game based mainly on the players' performance (football, tennis, etc.), we pay particular attention to horse racing because of the short racing period during which it is not suitable to adjust prices. Since a horse's in-game performance has little effect on the price-adjusting process, bookmakers must rely only on the demand for each horse when updating the betting odds. At any point in time, the bookmaker faces a *two-sided* decision: to change either the price of horse A or the price of horse B.

We assume that the bookmaker always tries to maximize their profits by setting prices[4] in which the summation of all the odds should be greater than or equal to 1, no matter which horse wins, otherwise there is an arbitrage opportunity. The over-roundness, which is a way of expressing to what extent the odds are in favor of the bookmaker, is the bookmaker's profit margins. That is, if there are two horses in a race, then $P_A + P_B = 1 + \epsilon$ where P_A denotes the price for horse A, P_B the price for horse B, and ϵ is the bookmaker's margin. Note that the cost per ticket is £1, which means the bookie pays out £1 for the winning horse per ticket sold. For the marginal noise trader who is indifferent between two horses, the fair price is one half each and there is also no need to change the price because the marginal noise punter goes with each horse with probability 1/2. There exist some traders who have privileged information, and those people are endowed with information about which horse is going to win. They are known as *informed traders*. Because of the existence of insiders, bookmakers face uncertainty when setting up the betting odds.

Since there is empirical evidence of the presence of informed traders in horse-betting markets,[5] we assume that a (known) fraction μ of traders is informed. We separate uninformed traders into liquidity traders and noise traders. Therefore, there are three types of traders in the market: informed traders, noise traders, and "liquidity" traders (Weston, 2001). Liquidity traders enter the market purely for fun[6]. They come to the racecourse for a picnic, a day out, etc. Trades made by such people are very small and do not influence the bookmaker's decision. So, we ignore this type of trader in the following analysis. The marginal noise trader, whose valuation is uniformly distributed on the interval $[0, 1]$, has his own subjective winning probabilities for each horse. The proportion of noise punters is denoted $\sigma \leq 1 - \mu$.

Our dynamics are modeled in continuous time, but to illustrate the main idea, we build one

[4]In this paper, we convert fractional odds to decimal prices. For example, odds of 2/1 correspond to the price £0.33 $(= 1/(1 + 2))$, which means you pay £0.33 for a ticket in return for £1 from the bookmaker, if that horse wins.
[5]See, for example, Shin (1991, 1992, 1993) and Schnytzer and Shilony (1995).
[6]See Ziemba (2019). Liquidity traders are defined as those people who "attend a day's card, particularly on Saturdays and Sundays, do not go to the track all that frequently. They go every so often to enjoy a good day's racing and to have an entertaining afternoon or evening. They do not expect to win but they certainly would enjoy a winning day and particularly a big exciting payoff."

Journal Pre-proof

Journal Pre-proof

in discrete time. Suppose $t \geq 0$ is the first time that the bookmaker stops the learning process and decides to adjust the prices. Consider a time interval $[0, t]$, which we partition into n parts of equal length $dt = t/n$. During this time, assume that we observe the net cumulative sales sequentially at n equally spaced occasions. Over a small time interval dt, say, from time t_{i-1} to t_i, we expect the net cumulative sales (of horse A over horse B) to go up by $u = \theta\mu dt + \sigma\sqrt{dt}$ and to go down by $d = \theta\mu dt - \sigma\sqrt{dt}$, where μ is the fraction of informed traders and $\sigma \leq 1 - \mu$ is the fraction of uninformed noise traders as defined above. The value of θ is unknown to the bookmaker and from the observed trades at given odds, she learns about its value. Note that, while the initial odds remain unchanged, the bookmaker loses on bets from informed traders, but these losses give her the information to change odds in the right direction later on. The net cumulative sales depend on what the informed traders know. If they know that horse A is going to win, their net cumulative sales will go up. Otherwise they will go down. So, θ is going to be either 1 or -1. $\sqrt{dt}$ here keeps the variance of the cumulative relative sales finite when we take the limit $dt \to 0$. In addition, the marginal noise trader's bet is such that each of the binomial steps occurs with probability $p = 1/2$.

The total net accumulated sales at time t after the n-th trade are defined as $S^{(n)}(t) := \sum_{i=1}^{n} X_i$. In order to obtain analytical results, we move to the continuous-time limit of this simple model. For fixed t, as $n \to \infty$, we define $S_t := \lim_{n\to\infty} S^{(n)}(t)$, where the limit is understood to be in distribution.[7] According to the central limit theorem, the conditional distribution of S_t is well-defined and is given by

$$S_t \mid \theta \sim N(\theta\mu t, \sigma^2 t).$$

Thus the continuous-time stochastic process $(S_t)_{t\geq 0}$ gives these distributions at each finite time t is the arithmetic Brownian motion (ABM)

$$dS_t = \theta\mu dt + \sigma dB_t,$$

where B_t is a standard Brownian motion.

Since the betting period will not last forever, this learning process will definitely cease at some

[7]Note that here we applied Donsker's invariance principle (see, e.g., Steele, 2012).

Journal Pre-proof

time. For analytical convenience, we propose a Poisson process with intensity λ that jumps from 0 to 1. Once it has jumped, our learning process will stop.

We use a jump process rather than a deterministic end time for the following reason. The *betting period* usually lasts for about 30 minutes in on-course horse-racing gambling markets. We know for sure that betting takes place over a fixed time interval. The race will start at a certain time and the market will then close at the starting prices. However, the number of bets placed during that period is unknown *ex ante*. Theoretically there is an almost infinite array of bets that can be made with a bookmaker (Gabriel and Marsden, 1990). To model this, we use an infinite horizon model, which stops after the first jump of the Poisson process with intensity λ. A jump process is just a mechanism to introduce a finite betting time and we can still model our problem with an infinite betting time horizon using all the analytical tools.

In the real world, punters come and go in no particular order. Trades do not happen at regular betting time intervals so that the bookmaker does not know exactly how many trades are going to be made and how fast the trade will come. The intensity λ plays the same role as a discount factor. Uncertainty over the number of trades that a bookmaker will make during the fixed and known betting period of time that bets can be made has the same effect as having a discount factor in an infinite horizon model. Therefore, our model, which has a random cut-off point, is discounted at λ. To get analytical results, we assume that there is a random process that determines when the decision should be made. There is no discounting *per se*, but since trades do not appear regularly and since trading will cease at some point, the cut-off occurs when the bookmaker decides to end the learning period and changes the price.

The way to turn λ into a discount rate is standard (see, e.g., Dixit and Pindyck, 1994). We assume that there is a Poisson process $(q_t)_{t\geq 0}$, independent of $(S_t)_{t\geq 0}$, with intensity λ and $q_0 = 0$. If $\bar{\tau} := \inf\{t \geq 0 \mid q_t = 1\}$, and $\bar{\tau}$ follows an exponential distribution with parameter λ, we know that $P(\bar{\tau} \geq t) = e^{-\lambda t}$, and the expected value of the project equals

$$F(s) = E_s\left[\int_0^\infty S_t P(\bar{\tau} \geq t)dt\right] = E_s\left[\int_0^\infty e^{-\lambda t} S_t dt\right].$$

By introducing the Poisson process with intensity λ, we model how many trades occur in the time interval $[0, t]$, and the time elapsed between any two trades is exponentially distributed with mean $1/\lambda$.

4 The Optimal Stopping Problem and Main Result

Recall that S_t denotes the cumulative sales of tickets on horse A, net of cumulative sales of tickets on horse B. If we assume that punters arrive sequentially in continuous time, then we observe

$$dS_t = \theta\mu dt + \sigma dB_t, \tag{1}$$

where $\mu \geq 0$ and $\sigma > 0$ are constant, and $(B_t)_{t\geq 0}$ is a standard Brownian motion.

The hypotheses that the bookmaker is testing are $H_0 : \theta = -1$ and $H_1 : \theta = 1$. The intuition behind this is that if insiders know that horse A (B) will win, we should have a positive (negative) trend. The decision depends on the state of the world, θ: conditional on the event $\{\theta = 1\}$ the bookmaker makes a decision to increase the price of horse A as all informed traders are backing horse A, thereby decreasing the price for horse B accordingly. However, if the state of the world turns out to be $\{\theta = -1\}$, the price of horse B should be changed. The true value of θ is not known, but the bookmaker has a (subjective) prior probability, ξ, that the state is $\{\theta = 1\}$. If the decision is correct, positive profits should be expected. There is also a chance that the decision is wrong.

Let the `Type I error` be the situation where all the informed traders know that horse B is going to win but the bookmaker thinks they favor horse A, conditional on the event $\{\theta = 1\}$. A `Type II error` is the opposite situation. In this case, insiders are backing horse A but the bookmaker thinks horse B is their choice. If these two types of errors occur, for example, the true state is $\{\theta = -1\}$, but the bookmaker decides it is $\{\theta = 1\}$, or *vice versa*, a huge loss is on the way. Therefore, the loss is temporary and the bookmaker gets the profit in the long run conditional on making the pricing decisions on the correct state θ.

To derive a proper value function, it seems appropriate to exclude two extreme cases. The first

Journal Pre-proof

Journal Pre-proof

case is that a rational bookmaker will never set a price equal to £1 (£0) for either horse A or horse B because this price immediately reveals to the market which horse will win. On the one hand, from the punters' point of view, given that the odds are properly determined, if they purchase the winning horse, there is no profit to earn as the cost of the ticket now equals the return. If punters buy the other one, they will definitely lose. So, no punter would be willing to be involved in such a market, except for the informed traders, who may have different information about the winning horse. From the bookmaker's perspective, on the other hand, they are taking the risk of setting the wrong prices, which implies making tremendous losses due to the insiders. In reality, a rational bookmaker would never set a price equal to 1 just to break down the market. They are willing to take losses to collect information about the true state in which, theoretically, they will make a profit eventually.

The second case is that all the punters are informed, and the bookmaker will learn in the first trade what information the traders have. The bookmaker would make a loss if the price of the winning horse is other than £1. This situation is not allowed in our model either.

In the following, we derive the optimal price for horse A at time t if $\{\theta = 1\}$ and also derive the loss due to a `Type I error`. In the current study, we consider that the fraction μ is informed. The bookmaker does not know what information they have. There are two possibilities. A fraction μ of punters knows either that horse A is the winning horse or that it is horse B. Let $P_A^* \in (0, 1)$ denote the adjusted price for horse A at time t, conditional on $\{\theta = 1\}$. The loss the bookmaker makes on informed traders is $\mu(P_A^* - 1)$.

We assume that the valuation of a horse by the marginal noise trader, denoted by V, is uniformly distributed on the interval $[0, 1]$, where 1 indicates that he is sure that horse A is going to win and 0 indicates the opposite, that is, he is sure that horse B is going to win. Let $\pi \in (0, 1)$ denote a noise punter's belief in horse A winning. We assume that π is independent of the odds. Note that the noise punter puts a bet on horse A if and only if

$$\pi \times 1 - P_A^* \geq 0 \Longleftrightarrow \pi \geq P_A^*,$$

Journal Pre-proof

and a bet on horse B if and only if

$$(1-\pi)\times 1 - P_B^* \geq 0 \Longleftrightarrow \pi \leq 1 - P_B^*.$$

where P_A^* and P_B^* are the prices *after* the odds are changed.

Conditional on the event $\{\theta = 1\}$, the expected profit of the bookmaker is

$$\begin{aligned} E[\text{profit} \mid \theta = 1] &= \mu\,(P_A^* - 1) + (1-\mu)\left\{\int_0^{1-P_B^*}(P_B^* - 0)\,d\pi + \int_{P_A^*}^{1}(P_A^* - 1)\,d\pi\right\} \\ &= -\{\underbrace{\mu\,(1-P_A^*) + (1-\mu)\,(1-P_A^*)^2}_{>0}\} + (1-\mu)P_B^*\,(1-P_B^*)\,, \end{aligned}$$

which is maximized for $P_B^* = 1/2$. Now we find the value for P_A^* such that

$$E[\text{profit} \mid \theta = 1] = \eta,$$

for some margin $\eta \geq 0$.

This problem is equivalent to finding the value for P_A^* under the constraint $P_B^* = 1/2$, such that

$$-\{\mu\,(1-P_A^*) + (1-\mu)\,(1-P_A^*)^2\} + \frac{1}{4}(1-\mu) = \eta. \tag{2}$$

By solving Eq. (2), we immediately arrive at

$$P_A^* = \frac{2 - \mu - \sqrt{2\mu^2 + (4\eta - 2)\mu + 1 - 4\eta}}{2(1-\mu)}.$$

where $0 \leq \mu < 1$ and $0 \leq \eta < \frac{\sqrt{2}-1}{2}$.[8]

With the prices P_A^* and P_B^* in hand, we can compute the expected loss of `Type I error`

[8]See Appendix A for the proof.

conditional on $\{\theta = 1\}$, i.e.,

$$
\begin{aligned}
E[\text{loss} \mid \theta = 1] &= \mu\,(P_B^* - 1) + (1-\mu)\left\{\int_0^{1-P_B^*} (P_B^* - 1)\,d\pi + \int_{P_A^*}^1 (P_A^* - 0)d\pi\right\} \\
&= -\{\mu\,(1-P_B^*) + (1-\mu)\,(1-P_B^*)^2\} + (1-\mu)P_A^*\,(1-P_A^*) \\
&= -\frac{\mu}{2} - \frac{1-\mu}{4} + (1-\mu)P_A^*\,(1-P_A^*) \\
&\triangleq \ell.
\end{aligned}
$$

It is obvious that $P_A^*(1-P_A^*) < 1/4$ when $P_A^* > 1/2$, thus $E[\text{loss} \mid \theta = 1] < 0$. A similar analysis holds conditional on the event $\{\theta = -1\}$, the solutions to the prices P_A^{**} and P_B^{**} are

$$
P_A^{**} = 1/2, \quad P_B^{**} = \frac{2-\mu-\sqrt{2\mu^2+(4\eta-2)\mu+1-4\eta}}{2(1-\mu)},
$$

with the same constraints $0 \le \mu < 1$ and $0 \le \eta < \frac{\sqrt{2}-1}{2}$. The expected loss of `Type II error` conditional on the event $\{\theta = -1\}$ can be written as

$$
\begin{aligned}
E[\text{loss} \mid \theta = -1] &= \mu\,(P_A^{**} - 1) + (1-\mu)\left\{\int_{P_A^{**}}^1 (P_A^{**} - 1)\,d\pi + \int_{P_B^{**}}^1 (P_B^{**} - 0)d\pi\right\} \\
&= -\{\mu\,(1-P_A^{**}) + (1-\mu)\,(1-P_A^{**})^2\} + (1-\mu)P_B^{**}\,(1-P_B^{**}) \\
&= -\frac{\mu}{2} - \frac{1-\mu}{4} + (1-\mu)P_B^{**}\,(1-P_B^{**}) \\
&\triangleq \ell.
\end{aligned}
$$

Table 1 summarizes the expected profits and losses for different states of the world.

Table 1: Expected profits and losses.

		State of the world	
		$\theta = 1$	$\theta = -1$
Decision	Adjust the Price of Horse A	η	ℓ
	Adjust the Price of Horse B	ℓ	η

Suppose the posterior belief of the bookmaker in the event $\{\theta = 1\}$ at time t is $p_t \in (0, 1)$, the

Journal Pre-proof

expected value functions, denoted by $V_1(p_t)$ and $V_0(p_t)$ respectively, are

$$V_1(p_t) = p_t \times \eta + (1 - p_t) \times \ell, \tag{3}$$

$$V_0(p_t) = p_t \times \ell + (1 - p_t) \times \eta. \tag{4}$$

The margin is determined in equilibrium such that *if the bookmaker decides on the correct odds*, they make a profit to compensate for the loss they make during the learning phase. A natural question that arises is for what levels of η not all losses can be covered. By plugging ℓ into Eq. (3), we obtain that

$$\begin{aligned} & p_t \cdot \eta \leq (1 - p_t) \left[\frac{\mu}{2} + \frac{1-\mu}{4} - (1-\mu) P_A^* (1 - P_A^*) \right] \\ \Longleftrightarrow\ & \eta \leq \frac{1-p_t}{p_t} \left[\frac{\mu}{2} + \frac{1-\mu}{4} - (1-\mu) P_A^* (1 - P_A^*) \right] \\ \Longleftrightarrow\ & \eta \leq \inf \left\{ \frac{1-p_t}{p_t} \left[\frac{\mu}{2} + \frac{1-\mu}{4} - (1-\mu) P_A^* (1 - P_A^*) \right] \right\} = \frac{1-p_t}{p_t} \cdot \frac{\mu}{2}. \end{aligned}$$

This gives an upper bound on the level of η that one would expect to observe in real-life betting markets.

Since we are mainly interested in the learning effect of keeping odds constant for a while, we will focus on the extreme case where $\eta = 0$, which will maximize the loss of learning for the bookmaker. So, for simplicity, we focus on the problem with $P_A^* + P_B^* = 1$.

The expected profit of a trade conditional on $\{\theta = 1\}$ becomes

$$E(\text{PROFIT} \mid \theta = 1) = \mu(P_A^* - 1) + (1-\mu)\left[(1 - P_A^*)(P_A^* - 1) + P_A^*(1 - P_A^* - 0)\right]. \tag{5}$$

Setting (5) equal to 0, we obtain

$$P_A^* = \frac{1}{2(1-\mu)}, \tag{6}$$

where $0 \leq \mu < 1/2$. As can be seen from (6), as $\mu \to 1$, which implies all punters in the market are informed, then the price the bookmaker should set goes to infinity. In this case, no one is willing to participate in the market. If $\mu = 1/2$, then $P_A^* = 1$ and the market also breaks down. The market

Journal Pre-proof

functions only when the fraction μ of informed traders is less than half and the bookmaker gets a chance to break even on the noisy ones.

To sum up, the expected profit is 0 if the bookmaker sets the price P_A^* conditional on making the right decision that the true state is indeed $\{\theta = 1\}$. All insiders will back horse A. However, if the `Type I error` occurs, the expected loss is

$$\begin{aligned} E_A(\text{LOSS} \mid \theta = 1) &= \mu(1 - P_A^* - 1) + (1-\mu)[(1-P_A^*)(P_A^* - 0) + P_A^*(1 - P_A^* - 1)] \\ &= -\mu P_A^* + (1-\mu)P_A^*(1 - 2P_A^*) \\ &= -\frac{\mu}{2(1-\mu)} + \frac{1-\mu}{2(1-\mu)}\left[1 - \frac{2}{2(1-\mu)}\right] \\ &= -\frac{\mu}{1-\mu}. \end{aligned} \tag{7}$$

The expectation is that if $\mu = 0$, the price $P_A^* = 1/2$, and the punter is uniformly distributed, then the bookmaker makes zero loss. As $\mu \to 1/2$, the price goes to 1 and the loss per ticket goes to 1 as well. So, the loss is bounded between 0 and 1, which is consistent with our assumptions. The loss due to the `Type II error` is also $\mu/(1-\mu)$.[9] Table 2 summarizes the losses due to the two types of errors for different states of the world.

Table 2: Loss of the two types of error.

		State of the world	
		$\theta = 1$	$\theta = -1$
Decision	Adjust the Price of Horse A	0	$\frac{\mu}{1-\mu}$
	Adjust the Price of Horse B	$\frac{\mu}{1-\mu}$	0

If the posterior belief in the event $\{\theta = 1\}$ at time t is $p_t \in (0, 1)$, the expected loss due to the `Type I error`, denoted by L_1, is

$$\begin{aligned} L_1(p_t) &= p_t \times 0 + (1-p_t)\left(\frac{\mu}{1-\mu}\right) \\ &= (1-p_t)\left(\frac{\mu}{1-\mu}\right). \end{aligned} \tag{8}$$

[9]See Appendix B for more details.

Similarly, the expected loss due to the `Type II error`, denoted by L_0, is

$$\begin{aligned} L_0(p_t) &= p_t\left(\frac{\mu}{1-\mu}\right) + (1-p_t)\times 0 \\ &= p_t\left(\frac{\mu}{1-\mu}\right). \end{aligned} \tag{9}$$

Note that at time t, the bookmaker will choose to change the price of horse A if and only if

$$p_t > \bar{p} \equiv \frac{1}{2}. \tag{10}$$

We model the net cumulative sales with Brownian motion. Initially, the bookmaker is trying to ensure that the expected profit they make from the marginal noise trader is zero. So, if the opening odds for both horses are 1/2 and if there are no informed traders in the market, then the marginal noise trader buys horse A with probability 1/2 and horse B with probability 1/2. This gives us a binomial tree, and the net cumulative sales could either go up or down. In this case, the loss per trade is always going to be 1/2, no matter what information the marginal noise punter has or which horse he bets on because the bookmaker always gets 1/2 from selling both types of ticket. For informed traders, the bookmaker should charge either 1 or for horse A, but if they charge 1/2 for horse A instead, the loss is also always going to be 1/2 per trade. So the expected loss, if the bookmaker does not make a decision, is

$$\frac{\mu}{2} + (1-\mu)\left[\frac{1}{2}\left(\frac{1}{2}-1\right) + \frac{1}{2}\left(\frac{1}{2}-0\right)\right] = \frac{\mu}{2}.$$

Therefore, until a decision is reached, the bookmaker is assumed to take a loss during this learning process. From time 0 to τ, the loss function is

$$\begin{aligned} L^*(p) &= \inf_{\tau\in\mathcal{T}} E_p\left[\int_0^{\tau} e^{-\lambda t}\frac{\mu}{2}dt + e^{-\lambda\tau}\min\{L_1(p), L_0(p)\}\right] \\ &= E_p\left[\int_0^{\tau^*} e^{-\lambda t}\frac{\mu}{2}dt + e^{-\lambda\tau^*}\min\{L_1(p_{\tau^*}), L_0(p_{\tau^*})\}\right], \end{aligned} \tag{11}$$

where $\mathcal{T}$ is the set of all stopping times. This is our optimal stopping problem. Since the aim of

Journal Pre-proof

the bookmaker is to minimize the loss from either the informed or uninformed noise punters, we choose the minimum functions in Eq. (11).

Since $(p_t)_{t\geq 0}$ is Markovian, the optimal stopping problem (11) can be written as

$$\begin{aligned} L^*(p) &= \frac{1}{\lambda}\cdot\frac{\mu}{2} + \inf_{\tau} E_p\left[e^{-\lambda\tau}\min\left\{L_0(p_\tau) - \frac{\mu}{2\lambda}, L_1(p_\tau) - \frac{\mu}{2\lambda}\right\}\right] \\ &= \frac{1}{\lambda}\cdot\frac{\mu}{2} + \inf_{\tau} E_p\left[e^{-\lambda\tau}\min\left\{p_\tau\left(\frac{\mu}{1-\mu}\right) - \frac{\mu}{2\lambda}, (1-p_\tau)\left(\frac{\mu}{1-\mu}\right) - \frac{\mu}{2\lambda}\right\}\right] \\ &= \frac{1}{\lambda}\cdot\frac{\mu}{2} + \inf_{\tau} E_p\left[e^{-\lambda\tau}\min\left\{G_0(p_\tau), G_1(p_\tau)\right\}\right]. \end{aligned} \tag{12}$$

Note that since we assume $\mu \in [0, 1/2)$, $G_0(\cdot)$ and $G_1(\cdot)$ are increasing and decreasing, respectively. That condition also ensures that $G_0(0) < G_1(0)$ and $G_0(1) > G_1(1)$. Note that $\bar{p} = 1/2$ is the unique point where $G_0(\bar{p}) = G_1(\bar{p})$.

This learning process reveals information about the true state of the world. We model the optimal stopping problem (11) as one of Bayesian sequential testing of two simple hypotheses in continuous time. Following Peskir and Shiryayev (2006), uncertainty is modeled on a *probability* space $(\Omega, \mathcal{F}, P_\xi)$ for a family of probability measures $(P_\xi)_{\xi\in[0,1]}$. For fixed $\xi \in [0, 1]$, we use a random variable $\tilde{\theta}$ to represent the true state of the world, which has values in $\{-1, 1\}$. The parameter θ is assumed to be the realization of a random variable $\tilde{\theta}$. For $p \in (0, 1)$, the probability measure P_ξ is obtained as follows:

$$P_\xi = \xi P_1 + (1-\xi)P_0,$$

where P_1 and P_0 denote degenerate distributions of the observed process under H_1 and H_0, respectively, with $P_1(\tilde{\theta} = 1) = P_0(\tilde{\theta} = -1) = 1$. So, for the Bayes formulation, in a measurable space $(\Omega, \mathcal{F})$, a probability measure P_ξ is given such that

$$P_\xi\{\theta = 1\} = \xi, \qquad P_\xi\{\theta = -1\} = 1 - \xi.$$

In other words, we assume that $\tilde{\theta}$ has the two values 1 and -1 with *a priori* probabilities ξ and $1 - \xi$.

Since we are sequentially testing the hypotheses H_1 and H_0, we observe the continuous process

Journal Pre-proof

Journal Pre-proof

$(S_t)_{t\geq 0}$ defined in Eq. (1). This process generates the natural filtration $\mathcal{F}_t^S = \sigma(S_s : 0 \leq s \leq t)$, which is augmented with the P_ξ-null set.

Let $p_t = P_\xi(\tilde{\theta} = 1 \mid \mathcal{F}_t^S)$ with $t \geq 0$. Thus,

$$\begin{aligned} p_t &= \frac{\xi \cdot \exp\left\{-(S_t - \mu t)^2/2\sigma^2 t\right\}}{\xi \cdot \exp\left\{-(S_t - \mu t)^2/2\sigma^2 t\right\} + (1-\xi) \cdot \exp\left\{-(S_t + \mu t)^2/2\sigma^2 t\right\}} \\ &= \left[1 + \frac{1-\xi}{\xi} \exp\left\{-\frac{2\mu}{\sigma^2} S_t\right\}\right]^{-1}. \end{aligned} \tag{13}$$

The Radon–Nikodym derivative

$$\varphi_t = \frac{d\left(P_1 \mid \mathcal{F}_t^S\right)}{d\left(P_0 \mid \mathcal{F}_t^S\right)}, \tag{14}$$

defines the *likelihood ratio process* between the two hypotheses $(\varphi_t)_{t\geq 0}$. It is well known that (Shiryayev, 1978)

$$\varphi_t = \exp\left\{\frac{2\mu}{\sigma^2} S_t\right\}. \tag{15}$$

The posterior process $(p_t)_{t\geq 0}$ can be shown in terms of the likelihood ratio process using Bayes' rule

$$p_t = \frac{[\xi/(1-\xi)]\varphi_t}{1 + [\xi/(1-\xi)]\varphi_t}. \tag{16}$$

From Ito's lemma, the right-hand side of (16) becomes

$$dp_t = \frac{2\mu^2}{\sigma^2} p_t(1-p_t)(1-2p_t)dt + \frac{2\mu}{\sigma^2} p_t(1-p_t)dS_t.\text{[10]} \tag{17}$$

Note that in our study $p_t = P_\xi(\tilde{\theta} = 1 \mid \mathcal{F}_t^S) = E_\xi[\tilde{\theta} \mid \mathcal{F}_t^S]$. We consider the process

$$\overline{B}_t = \sigma^{-1}\left(X_t + \mu t - 2\mu \int_0^t p_s ds\right). \tag{18}$$

Combining (17) and (18), we obtain that $(p_t)_{t\geq 0}$ follows the stochastic differential equation

$$dp_t = \frac{2\mu}{\sigma} p_t(1-p_t)d\overline{B}_t, \quad \text{with } p_0 = \xi, \tag{19}$$

[10] See Appendix C for derivations.

Journal Pre-proof

Journal Pre-proof

where $(\overline{B}_t)_{t\geq 0}$ is also a standard Brownian motion, called the *innovation process* (see, e.g., Poor and Hadjiliadis, 2009). Using (15) and (16), it can be verified that the process $(p_t)_{t\geq 0}$ is time-homogeneous and strongly Markovian under P_ξ with respect to the natural filtration. Note that if $\theta = 1$, then $p_t \xrightarrow{a.s.} 1$ and if $\theta = -1$, $p_t \xrightarrow{a.s.} 0$ as $t \to \infty$. So, as $t \to \infty$, $\mathrm{Var}(dp_t) \to 0$ holds in either case since $\mathrm{Var}(dp_t) = (2\mu/\sigma)^2\, p_t^2(1-p_t)^2 dt$.

In the following, we adapt Thijssen and Bregantini's model (2017) as the starting point. The closer $(p_t)_{t\geq 0}$ gets to either 0 or 1, the less likely it is that the loss will decrease upon continuation. This suggests that there are two points $p_B \in (0, 1/2)$ and $p_A \in (1/2, 1)$ that distinguish the adjusting regions. Thus, the state space $(0, 1)$ can be divided into three regions. The first is called the *continuation region*. It is the region around $\overline{p}$ where keeping the price unchanged is optimal. It is denoted by

$$\mathcal{C} = \{p \in (0,1) \mid L^*(p) < \min(L_0(p), L_1(p))\} = (p_B, p_A),$$

where p_B and p_A, with $0 < p_B < \overline{p} < p_A < 1$, are the boundaries for changing the price of horse B and changing the price of horse A, respectively. When p is small enough, we enter the *adjusting horse B region*, where the cumulative sales of horse B are larger than those of horse A, so that we change the price of horse B. This region is denoted by

$$\mathcal{D}_B = \{p \in (0,1) \mid L^*(p) = L_0(p)\} = (0, p_B].$$

Conversely, when p is large enough, we enter the *adjusting horse A region*, where we should increase the price of horse A. This region is denoted by

$$\mathcal{D}_A = \{p \in (0,1) \mid L^*(p) = L_1(p)\} = [p_A, 1).$$

It follows that the stopping time

$$\tau^* = \inf\{t \geq 0 : p_t \notin (p_B, p_A)\} \tag{20}$$

is optimal in (11).

We give a proof of the following proposition in the appendix.

Proposition 4.1 *Suppose that*

1. $0 \leq \mu < \frac{1}{2}$ *and*
2. $\frac{1-\mu}{2\lambda} > \frac{1}{2} + \frac{1}{4\gamma}$.

In testing the two simple hypotheses $H_0 : \theta = -1$ and $H_1 : \theta = 1$ on the observations of the process given by $dS_t = \theta\mu dt + \sigma dB_t$, there are unique thresholds p_B^* and p_A^*, $p_B^* < \bar{p} < p_A^*$, such that the optimal stopping time is the first exit time of (p_B^*, p_A^*). That is,

$$\tau_p^* = \inf\{t \geq 0 \mid p_t^p \notin (p_B^*, p_A^*)\}$$

and the loss function L is explicitly given by

$$L^*(p) = \begin{cases} p \cdot \left(\frac{\mu}{1-\mu}\right), & \text{if } p \in (0, p_B^*], \\ \frac{\mu}{2\lambda} + \hat{v}_{p_B^*, p_A^*}(p) G_1(p_A^*) + \check{v}_{p_B^*, p_A^*}(p) G_0(p_B^*), & \text{if } p \in (p_B^*, p_A^*), \\ (1-p) \cdot \left(\frac{\mu}{1-\mu}\right), & \text{if } p \in [p_A^*, 1), \end{cases} \tag{21}$$

where

$$\hat{v}_{p_B^*, p_A^*}(p) := \sqrt{\frac{p(1-p)}{p_A^*(1-p_A^*)}} \frac{\left(\frac{1-p_B^*}{p_B^*}\frac{p}{1-p}\right)^\gamma - \left(\frac{p_B^*}{1-p_B^*}\frac{1-p}{p}\right)^\gamma}{\left(\frac{1-p_B^*}{p_B^*}\frac{p_A^*}{1-p_A^*}\right)^\gamma - \left(\frac{p_B^*}{1-p_B^*}\frac{1-p_A^*}{p_A^*}\right)^\gamma}, \tag{22}$$

and

$$\check{v}_{p_B^*, p_A^*}(p) := \sqrt{\frac{p(1-p)}{p_B^*(1-p_B^*)}} \frac{\left(\frac{1-p}{p}\frac{p_A^*}{1-p_A^*}\right)^\gamma - \left(\frac{p}{1-p}\frac{1-p_A^*}{p_A^*}\right)^\gamma}{\left(\frac{1-p_B^*}{p_B^*}\frac{p_A^*}{1-p_A^*}\right)^\gamma - \left(\frac{p_B^*}{1-p_B^*}\frac{1-p_A^*}{p_A^*}\right)^\gamma}, \tag{23}$$

are the expected discount factors of first reaching p_B^* and p_A^*, respectively, given the current posterior probability $p \in (0, 1)$.

The first condition in 4.1 states that the fraction of informed bettors is less than one half, which is consistence with the empirical evidence in Shin (1993). A large number of informed traders will break down the markets. The second condition, to some extent, may be too strong. The proof of 4.1 gives more details.

In the region $\mathcal{C} = (p_B, p_A)$ the loss incurred is equal to the loss of not changing the odds, plus the expected loss of making two types of errors at p_B^* and p_A^*, respectively, discounted back to the current time using the expected discount factor, i.e., the parameter of the jump process λ.

5 Comparative Statics

In this section, we obtain analytical results for the posterior bounds p_A^* and p_B^*, as well as the loss function given different parameters. To assess the quantitative effects of the bounds, we consider the base-case scenario described in Table 3.

Table 3: Parameters for a base-case numerical example.

Base-case scenario
$\mu = 0.1$
$\sigma = 0.1$
$\lambda = 0.1$

For this particular case, we obtain $p_B^* = 0.2527$ and $p_A^* = 0.7473$.[11] The loss is shown in Figure 2 for different values of the posterior belief in the event $\{\tilde{\theta} = 1\}$. Note that the curved line between the thresholds p_B^* and p_A^* shows the loss while waiting for belief. The continuation region is between the thresholds p_B and p_A, as the loss is higher if the bookmaker decides to stop learning and takes action by adjusting the price. In the continuation region, to guarantee the minimum loss for the bookmaker, instead of making a decision, it is optimal to continue learning. The maximum loss is at the point 1/2 because under this circumstance the bookmaker is unable to distinguish which horse the informed traders believe is the winning horse. The odds are kept at 1/2, so that the loss on the winning horse is one half, regardless of what information the punters possess. Note that this figure includes the loss from both sides, whether the bookmaker decides to change the

[11] All calculations were done in MatLab.

Journal Pre-proof

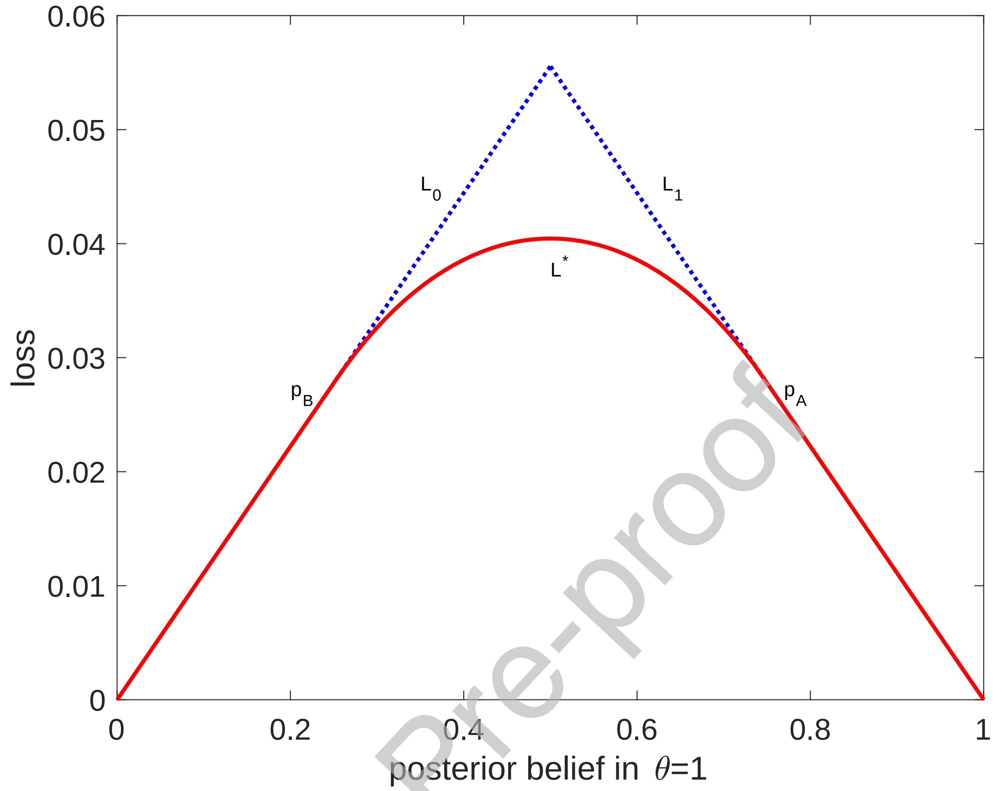

Figure 2: The loss function.

price of horse A or horse B. As expected, the curve is symmetric because we assumed that for the bookmaker to achieve zero profits, the sum of the prices of the two horses is 1.

In the following, we check the sensitivity of the solution to (11) with respect to μ, σ, and λ.

Proposition 5.1 *Suppose that the conditions of Proposition 1 hold. Let* $\tau^* = \inf\{t \geq 0 \mid p_t \notin (p_B, p_A)\}$ *be the unique solution to (11). It then holds that*

- p_B *is decreasing and* p_A *is increasing in the proportion of insiders* μ*, conditional on* $\theta = 1$*, provided that* $p_B < 1/2 < p_A$*.*
- p_B *is increasing and* p_A *is decreasing with the volatility* σ*, provided that* $p_B < 1/2 < p_A$*.*
- p_B *is decreasing and* p_A *is increasing as the parameter of the jump process* λ *increases.*

Journal Pre-proof

Journal Pre-proof

A proof of this proposition can be found in Appendix E.

We can also derive the expected time to make a decision. The procedure is standard and follows Poor and Hadjiliadis (2009). We then obtain the following equation:

$$\begin{aligned} E_p[\tau^*] &= (1-p)E_0[\tau^*] + pE_1[\tau^*] \\ &= \frac{2\sigma^2}{(2\mu)^2} \cdot \log\left[\left(\frac{p}{1-p}\right)^{1-2p}\left(\frac{1-p_B}{p_B}\right)^{1-2p_B}\right] \\ &\quad + \frac{2\sigma^2}{(2\mu)^2} \cdot \frac{p-p_B}{p_A-p_B} \cdot \log\left[\left(\frac{p_B}{1-p_B}\right)^{1-2p_B}\left(\frac{1-p_A}{p_A}\right)^{1-2p_A}\right]. \end{aligned} \tag{24}$$

The comparative statics for the parameter μ are shown in Figure 3. Recall that the parameter μ denotes the fraction of the informed traders in the learning process in unit time. It is easily seen that p_B is decreasing and p_A is increasing with μ. The posterior bounds become wider because the higher value of μ implies the learning process is more informative. The non-monotonicity in the loss function can be explained from two perspectives. On the one hand, having a large proportion of insiders in the market would lead to a huge loss for the bookmaker. On the other hand, since the bookmaker can get more information in each time period from the insiders, the decision can be taken sooner, which means the loss declines. The base-case scenario explicitly shows that the maximum loss occurs when the fraction $\mu = 16\%$.

The comparative statics for the parameter σ are shown in Figure 4. As p_B is increasing and p_A is decreasing, the posterior bounds narrow to the point 0.5, which implies that a decision can be reached earlier. This is partially different from the standard literature on real options, in which more uncertainty widens the decision bounds. In this paper, more uncertainty lowers the bounds but increases the loss. Since this volatility also represents the fraction of noise traders in the market, the larger the value of σ, the more noise traders there are and the fewer informed traders there are. As having more noise traders makes the learning process less informative as the signal is noisier, the loss is incurred during learning, but the costs are the same, so we expect the decision would be taken sooner, thereby lowering the loss accordingly.

The comparative statics for the parameter λ are shown in Figure 5. As p_B is decreasing and

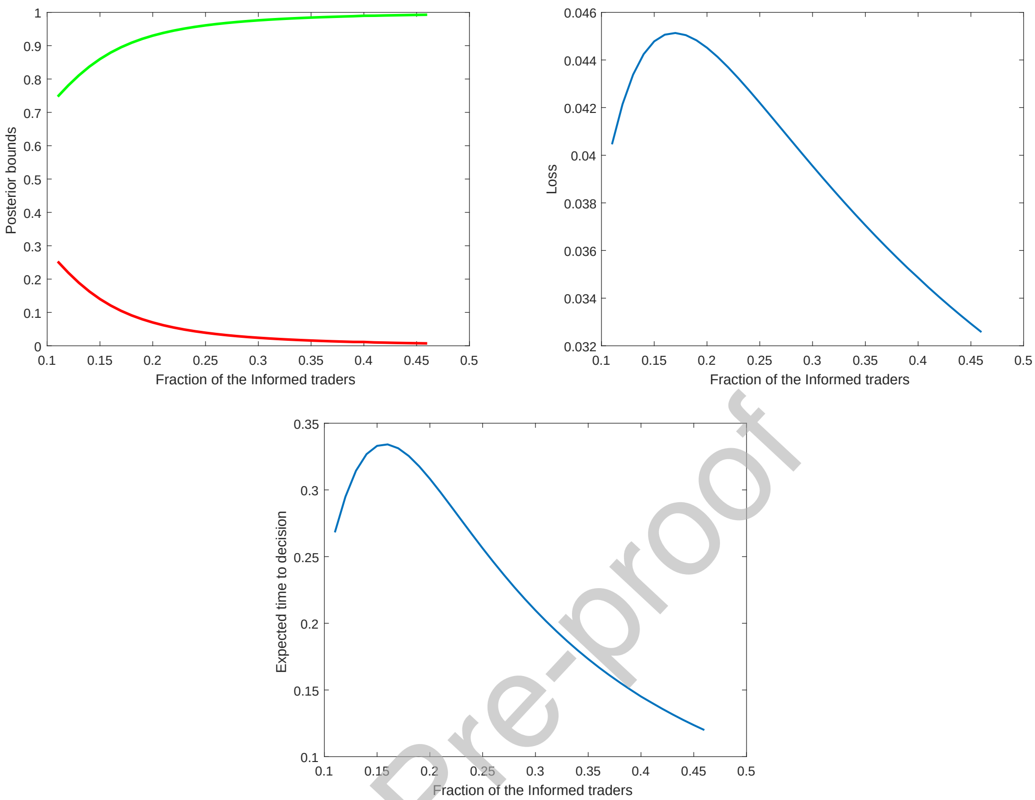

Figure 3: Comparative statics for μ.

p_A is increasing, the posterior bounds widen, which implies the learning process becomes more informative. A higher λ reduces the expected cost of keeping the current odds and makes waiting longer more attractive. However, the higher λ acts as a discount factor. The less we care about the future, the faster a decision is reached. These are two opposing effects, one of which apparently dominates. In our setting, the running loss is important. The loss is decreasing and the expected time to make a decision is increasing, as the expectation is that it takes longer as the bounds widen.

6 Conclusions

The Bayesian model presented in this paper focuses on the optimal stopping problem for a bookmaker who has to decide when to change the odds for each horse in a race. There are informed traders in the market, but which horse the informed traders prefer is unknown *ex ante*. The book-

Journal Pre-proof

Journal Pre-proof

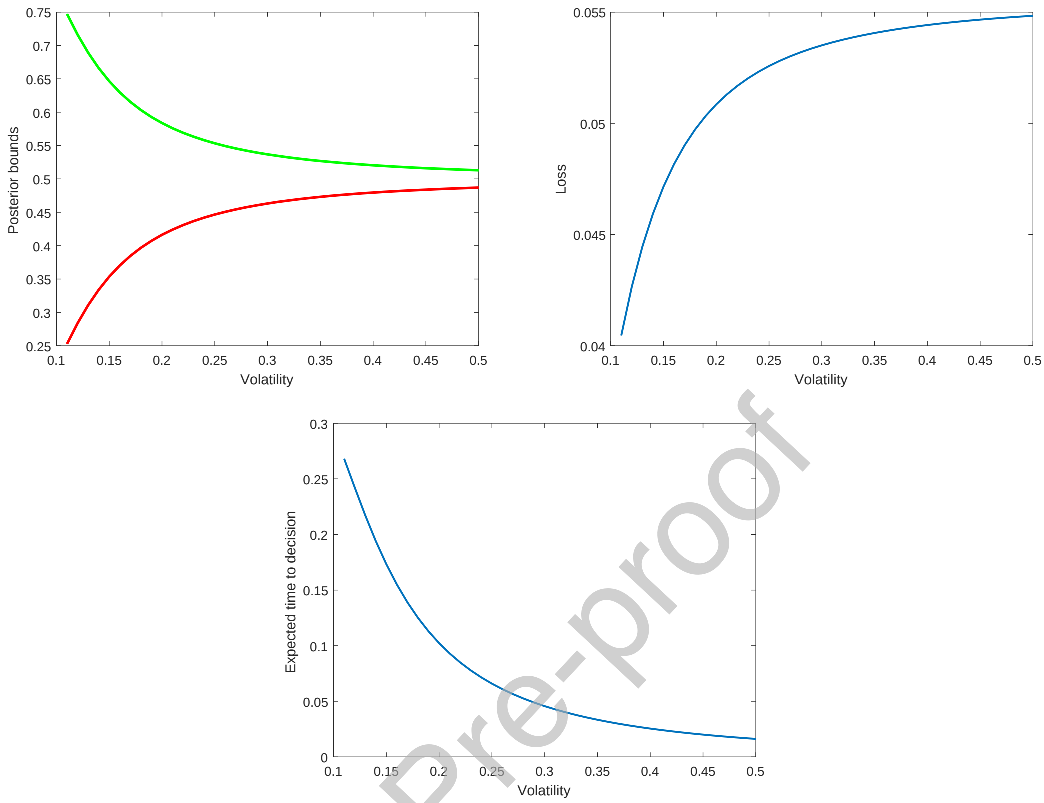

Figure 4: Comparative statics for σ.

maker goes through a learning process to assess whether they have gathered enough statistical evidence to determine which is the favorite horse of the insiders at each point in time by applying a sequential hypothesis test.

For this two-horse benchmark model, we have provided analytical results that give posterior bounds for adjusting the price of horse A or the price of horse B when the learning stops. Once a decision is taken, no matter what it is, it is guaranteed that the loss to the bookmaker is minimized.

We also investigated the sensitivity of the model parameters to the proposed solution to (11), which is consistent with our intuition. We found that: (i) As the fraction of informed punters gets bigger, the posterior bounds widen and the bookmaker receives more information per time period. Non-monotonicity occurs because of two opposing effects. On the one hand, having a large proportion of insiders leads to a huge loss to the bookmaker. On the other hand, since the

Journal Pre-proof

Journal Pre-proof

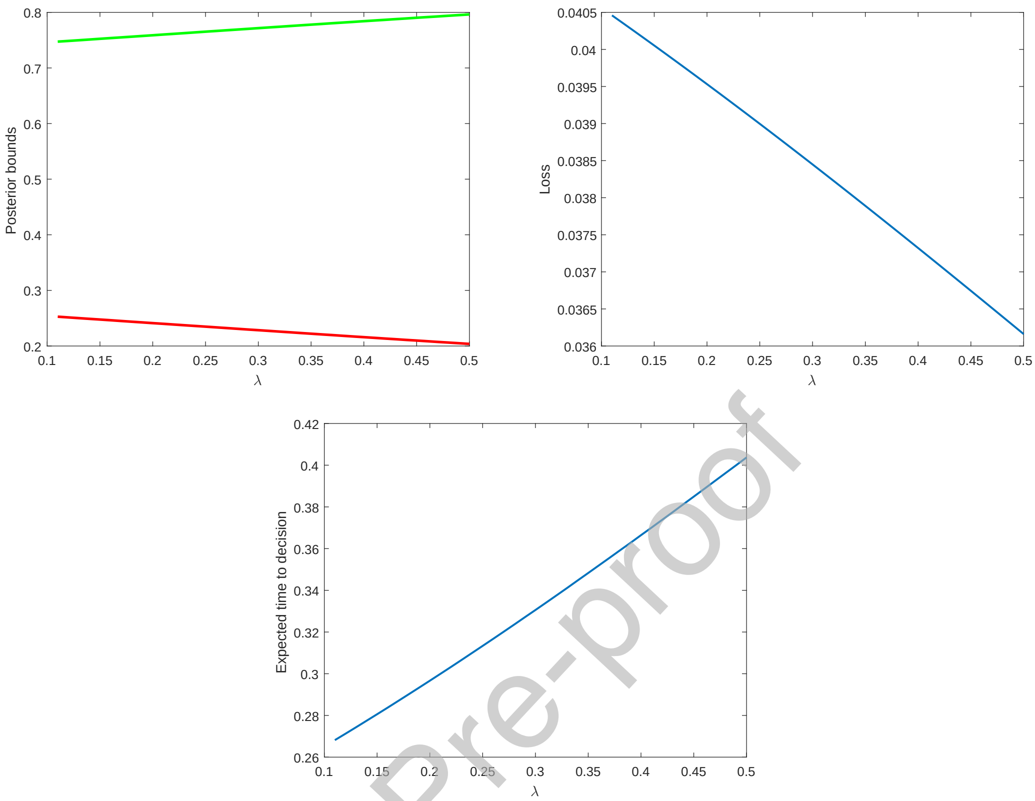

Figure 5: Comparative statics for λ.

bookmaker can get more information per time period, decisions will be taken sooner, thus reducing the loss. (ii) For higher σ, there are more noise traders in the market so that the signal the bookmaker observes is noisier. The bookmaker has to pay to keep the learning process going as less information is provided. A loss is incurred in learning, but the process is not informative and the costs are the same. We, therefore, expect a decision will be reached sooner as the posterior bounds get narrower. (iii) The bounds widen as λ increases, and as expected, it takes longer to reach a decision. A higher λ reduces the expected cost of keeping the current odds and makes waiting longer more attractive. On the other hand, the higher λ, the less we care about the future and the faster a decision is reached. These are the two opposing effects, one of which apparently dominates.

Journal Pre-proof

Journal Pre-proof

Acknowledgments

We are grateful to the editor and two anonymous referees for their helpful and excellent comments. We also thank Zaifu Yang, Michael Thornton, Takashi Yamagata, Bredan McCabe, Yang Li, Alexander Kovalenkov, and participants at various conferences for useful comments and suggestions. Chi Zhang gratefully acknowledged the financial support from the Fundamental Research Funds of Shandong University and the Taishan Scholar Program of Shandong Province. This research was also supported by the National Natural Science Foundation of China (NSFC) (No.11901351). We remain responsible for any errors.

Journal Pre-proof

Journal Pre-proof

Appendices

Appendix A. Proof of restrictions on μ and η.

Re-arranging Eq. (2), we get

$$(1-\mu)(1-P_A^*)^2 + \mu\,(1-P_A^*) + \eta - \frac{1}{4}(1-\mu) = 0.$$

By solving the above equation, we immediately arrive at

$$P_A^* = \frac{2-\mu-\sqrt{2\mu^2+(4\eta-2)\mu+1-4\eta}}{2(1-\mu)}.$$

Note that conditional on the event $\{\theta = 1\}$, the price P_A^* needs to satisfy $P_A^* > P_B^* = 1/2$, but $P_A^* < 1$, i.e.,

$$1/2 < \frac{2-\mu-\sqrt{2\mu^2+(4\eta-2)\mu+1-4\eta}}{2(1-\mu)} < 1, \tag{A.1}$$

and

$$2\mu^2 + (4\eta-2)\mu + 1 - 4\eta > 0, \tag{A.2}$$

To guarantee (A.2) holds, $\Delta = (4\eta-2)^2 - 8(1-4\eta) = 4(4\eta^2+4\eta-1) < 0$. We obtain $0 \le \eta < \frac{\sqrt{2}-1}{2}$. By solving (A.1), we can obtain $0 \le \mu < 1$.

Journal Pre-proof

Appendix B. Loss of the `Type II error`

Similarly, let P_B^{**} denote the adjusted price for horse B at time t given the event $\{\theta = -1\}$. The valuation of the marginal noise trader, V, is also uniformly distributed on the interval [0, 1]. For the informed trader, the bookmaker loses $\mu(P_B^{**} - 1)$. For the marginal noise trader, he buys a horse-A ticket if and only if his own valuation is higher than P_B^{**}. With probability $1 - P_B^{**}$, the marginal noise trader buys horse B and the bookmaker loses $P_B^{**} - 1$ on this kind. With probability P_B^{**}, a horse-A ticket is bought and the loss of this kind is 0. So the expected profit of a trade is

$$E(\text{PROFIT} \mid \theta = -1) = \mu(P_B^{**} - 1) + (1-\mu)[(1 - P_B^{**})(P_B^{**} - 1) + P_B^{**}(1 - P_B^{**} - 0)].$$

Setting the above equation equal to 0, we can get

$$P_B^{**} = \frac{1}{2(1-\mu)}.$$

So the loss to the `Type II error` follows the same analysis.

$$\begin{aligned} E_B(\text{LOSS} \mid \theta = -1) &= \mu(1 - P_B^{**} - 1) + (1-\mu)[(1 - P_B^{**})(P_B^{**} - 0) + P_B^{**}(1 - P_B^{**} - 1)] \\ &= -\mu P_B^{**} + (1-\mu)P_B^{**}(1 - 2P_B^{**}) \\ &= -\frac{\mu}{2(1-\mu)} + \frac{1-\mu}{2(1-\mu)}[1 - \frac{2}{2(1-\mu)}] \\ &= -\frac{\mu}{(1-\mu)}, \end{aligned}$$

where $0 \leq \mu < 1/2$.

Journal Pre-proof

Appendix C. Derivations of Eq. (17)

Let $p_t = P_\xi(\tilde{\theta} = 1 \mid \mathcal{F}_t^S)$ with $t \geq 0$. Thus,

$$p_t = \frac{\xi \cdot \exp\left\{-(S_t - \mu t)^2/2\sigma^2 t\right\}}{\xi \cdot \exp\left\{-(S_t - \mu t)^2/2\sigma^2 t\right\} + (1-\xi)\cdot \exp\left\{-(S_t + \mu t)^2/2\sigma^2 t\right\}}$$
$$= \left[1 + \frac{1-\xi}{\xi}\exp\left\{-\frac{2\mu}{\sigma^2}S_t\right\}\right]^{-1}.$$

The Radon–Nikodym derivative

$$\varphi_t = \frac{d\left(P_1 \mid \mathcal{F}_t^S\right)}{d\left(P_0 \mid \mathcal{F}_t^S\right)},$$

defines the *likelihood ratio process* between the two hypotheses $(\varphi_t)_{t\geq 0}$. It is well known that (Shiryayev, 1978)

$$\varphi_t = \exp\left\{\frac{2\mu}{\sigma^2}S_t\right\}, \tag{C.1}$$

and since for $\xi < 1$

$$p_t = \frac{\xi d(P_1 \mid \mathcal{F}_t^S)}{\xi d(P_1 \mid \mathcal{F}_t^S) + (1-\xi) d(P_0 \mid \mathcal{F}_t^S)},$$

it follows that

$$p_t = \frac{[\xi/(1-\xi)]\varphi_t}{1 + [\xi/(1-\xi)]\varphi_t}. \tag{C.2}$$

Applying Ito's lemma to equation (C.1) yields

$$\begin{aligned} d\varphi_t &= \frac{2\mu^2}{\sigma^2}\varphi_t dS_t + \frac{1}{2}(\frac{2\mu}{\sigma^2})^2\varphi_t (dS_t)^2 \\ &= \frac{2\mu^2}{\sigma^2}\varphi_t dS_t + \frac{1}{2}(\frac{2\mu}{\sigma^2})^2\varphi_t \sigma^2 dt \\ &= \frac{2\mu}{\sigma^2}\varphi_t (dS_t + \mu dt). \end{aligned}$$

which follows the geometric Brownian motion:

$$\frac{d\varphi}{\varphi} = \left(\theta\frac{2\mu^2}{\sigma^2} + \mu\right)dt + \frac{2\mu}{\sigma}dB_t.$$

Journal Pre-proof

Journal Pre-proof

Similarly applying Ito's lemma to equation (C.2) yields

$$
\begin{aligned}
dp_t &= -\frac{\xi(\xi-1)}{(\xi\varphi_t-\xi+1)^2}d\varphi_t + \frac{1}{2}\frac{2\xi^2(\xi-1)}{(\xi\varphi_t-\xi+1)^3}d\varphi_t^2 \\
&= -\frac{\xi(\xi-1)}{(\xi\varphi_t-\xi+1)^2}\frac{2\mu}{\sigma^2}\varphi_t(dS_t+\mu dt) \\
&\quad +\frac{1}{2}\frac{2\xi^2(\xi-1)}{(\xi\varphi_t-\xi+1)^3}(\frac{2\mu}{\sigma^2}\varphi_t(dS_t+\mu dt))^2 \\
&= \frac{2\mu}{\sigma^2}(\frac{\xi\varphi_t}{\xi\varphi_t-\xi+1})(\frac{1-\xi}{\xi\varphi_t-\xi+1})(dS_t+\mu dt) \\
&\quad -(\frac{2\mu}{\sigma^2})^2\sigma^2(\frac{\xi\varphi_t}{\xi\varphi_t-\xi+1})^2(\frac{1-\xi}{\xi\varphi_t-\xi+1})dt \\
&= \frac{2\mu}{\sigma^2}p_t(1-p_t)dS_t + \frac{2\mu^2}{\sigma^2}p_t(1-p_t)(1-2p_t)dt.
\end{aligned}
$$

Journal Pre-proof

Appendix D. Proof of Proposition 4.1

We need to find a function $L^* \in \mathcal{C}^2$ that solves the following *free-boundary problem*:

$$\begin{cases} \mathcal{L}L^* - \lambda L^* = 0, \quad \text{for } p \in (p_B, p_A), \\ L^*(p_B; p_A) = G_0(p_B), \\ L^*(p_A; p_B) = G_1(p_A), \\ L^{*\prime}(p_B; p_A) = G_0'(p_B), \quad \text{(smooth fit)} \\ L^{*\prime}(p_A; p_B) = G_1'(p_A). \quad \text{(smooth fit)} \end{cases} \tag{D.1}$$

Here, $\mathcal{L}$ defines the characteristic operator of $(p_t)_{t\geq 0}$ (Øksendal, 2003). That is, for any $\psi \in \mathcal{C}^2$,

$$\mathcal{L}\psi(p) = \frac{1}{2}\left(\frac{2\mu}{\sigma}\right)^2 p^2(1-p)^2\psi''(p).$$

To derive the loss function of (11), we introduce the parameter:

$$\gamma := \frac{1}{2}\sqrt{1 + 2\lambda\left(\frac{\sigma}{\mu}\right)^2} > \frac{1}{2}.$$

There are two *fundamental solutions* to the differential equation $\mathcal{L}\psi - \lambda\psi = 0$:

$$\hat{\psi}(p) = \sqrt{p(1-p)}\left(\frac{p}{1-p}\right)^\gamma,$$
$$\breve{\psi}(p) = \sqrt{p(1-p)}\left(\frac{1-p}{p}\right)^\gamma.$$

Note that $\hat{\psi}(\cdot)$ is increasing and $\breve{\psi}(\cdot)$ is decreasing. The general solution to $\mathcal{L}\psi - \lambda\psi = 0$ is of the form

$$\psi(p) = \hat{A}\hat{\psi}(p) + \breve{A}\breve{\psi}(p),$$

Journal Pre-proof

where $\hat{A}$ and $\breve{A}$ are arbitrary constants. Furthermore, it is easily obtained that

$$\hat{\psi}'(p) = \hat{\psi}(p)\frac{1/2+\gamma-p}{p(1-p)} > 0,$$
$$\breve{\psi}'(p) = \breve{\psi}(p)\frac{1/2-\gamma-p}{p(1-p)} < 0,$$
$$\hat{\psi}''(p) = \hat{\psi}(p)\frac{\gamma^2-1/4}{p^2(1-p)^2} > 0,$$

and

$$\breve{\psi}''(p) = \breve{\psi}(p)\frac{\gamma^2-1/4}{p^2(1-p)^2} > 0.$$

1. For $p_L \le \bar{p}$, we define a mapping in the $p \mapsto \breve{L}(p;p_L)$ by

$$\breve{L}(p;p_L) = \hat{A}(p_L)\hat{\psi}(p) + \breve{A}(p_L)\breve{\psi}(p), \tag{D.2}$$

where the constants $\hat{A}(p_L)$ and $\breve{A}(p_L)$ are given by

$$\hat{A}(p_L) = \frac{\breve{\psi}(p_L)}{2\gamma}\left[G_0'(p_L) - \frac{1/2-\gamma-p_L}{p_L(1-p_L)}G_0(p_L)\right], \tag{D.3}$$

and

$$\breve{A}(p_L) = \frac{\hat{\psi}(p_L)}{2\gamma}\left[\frac{1/2+\gamma-p_L}{p_L(1-p_L)}G_0(p_L) - G_0'(p_L)\right]. \tag{D.4}$$

Notice that $\mathcal{L}\breve{L}(p;p_L) - \lambda\breve{L}(p;p_L) = 0$ for all $p \in (0,1)$. Furthermore, we show that $\breve{L}(p_L;p_L) = G_0(p_L)$ and $\breve{L}'(p_L;p_L) = G_0'(p_L) > 0$. With Assumption 2, it is ensured that $G_0(\bar{p}) < 0$ and since $G_0'(p) = \frac{\mu}{1-\mu} > 0$, it implies $G_0(\cdot)$ is monotonically increasing. With $p_L \in (0,\bar{p})$, it follows $G_0(p_L) < 0$ and

$$\frac{\partial\hat{A}(p_L)}{\partial p_L} = -\frac{\breve{\psi}(p_L)}{2\gamma}\cdot\frac{\gamma^2-\frac{1}{4}}{p_L^2(1-p_L)^2}\cdot G_0(p_L) > 0,$$

and

$$\frac{\partial\breve{A}(p_L)}{\partial p_L} = \frac{\hat{\psi}(p_L)}{2\gamma}\cdot\frac{\gamma^2-\frac{1}{4}}{p_L^2(1-p_L)^2}\cdot G_0(p_L) < 0.$$

Journal Pre-proof

Therefore, it is easily verified that $\breve{A}(p_L) < 0$ for all p_L. Since $\breve{A}(p_L)$ is decreasing in p_L, and $\breve{A}(\cdot)$ is strictly monotone between 0 and $\bar{p}$, it is seen that $\breve{A}(\bar{p}) < 0$. Assumption 2 also ensures that $\hat{A}(\bar{p}) < 0$. Since $\hat{A}(\cdot)$ is monotonically increasing on $(0, \bar{p})$, $\hat{A}(p_L) < 0$. Thus, the function $p \mapsto \breve{L}(p; \bar{p})$ is concave on $(0, 1)$[12] and it satisfies $\breve{L}(0+; \bar{p}) = \breve{L}(1-; \bar{p}) = -\infty$.

Since $\hat{A}(p_L)$ increases and $\breve{A}(p_L)$ decreases in p_L, it holds that $\frac{\partial \breve{L}(p;p_L)}{\partial p_L} = \frac{\partial \hat{A}(p_L)}{\partial p_L}\hat{\psi}(p) + \frac{\partial \breve{A}(p_L)}{\partial p_L}\breve{\psi}(p) > 0$ for all $p > p_L$.[13]

So there is a unique point $p_H \in (\bar{p}, 1)$ such that $\breve{L}'(p_H; p_L) = G_1'(p_H)$, ensuring that $\breve{L}(p_H; p_L)$ increases in p_L, but decreases in p_H. Then we get the existence of a unique point $p_B \in (0, \bar{p})$ for which there is $p_A \in (\bar{p}, 1)$ such that $\breve{L}(p_A; p_B) = G_1(p_A)$ and $\breve{L}'(p_A; p_B) = G_1'(p_A)$.

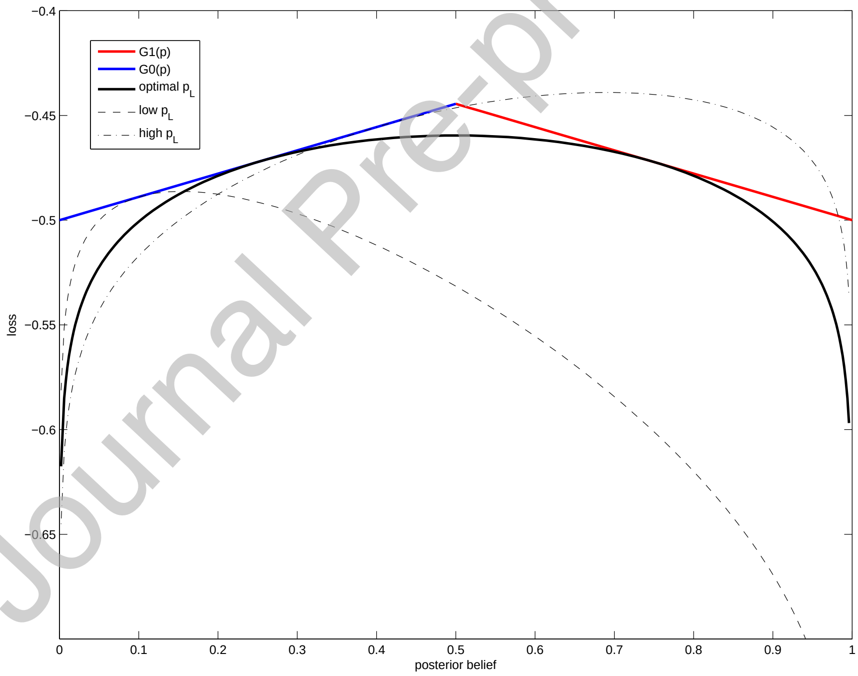

Figure D.1: The Loss Function for Different Values of p_L.

[12] $\breve{L}''(p; \bar{p}) = \hat{A}(\bar{p})\hat{\psi}''(p) + \breve{A}(\bar{p})\breve{\psi}''(p) < 0$.

[13] For every $p > p_L$, $\hat{\psi}(p) > \hat{\psi}(p_L)$ and $\breve{\psi}(p) < \breve{\psi}(p_L)$, then $\frac{\partial \hat{A}(p_L)}{\partial p_L}\hat{\psi}(p) > \frac{\partial \hat{A}(p_L)}{\partial p_L}\hat{\psi}(p_L)$ and $\frac{\partial \breve{A}(p_L)}{\partial p_L}\breve{\psi}(p) > \frac{\partial \breve{A}(p_L)}{\partial p_L}\breve{\psi}(p_L)$, therefore $\frac{\partial \breve{L}(p;p_L)}{\partial p_L} = \frac{\partial \hat{A}(p_L)}{\partial p_L}\hat{\psi}(p) + \frac{\partial \breve{A}(p_L)}{\partial p_L}\breve{\psi}(p) > \frac{\partial \hat{A}(p_L)}{\partial p_L}\hat{\psi}(p_L) + \frac{\partial \breve{A}(p_L)}{\partial p_L}\breve{\psi}(p_L) = 0$.

2. For $p_H \geq \bar{p}$, we can also define the mapping in the $p \to \hat{L}(p; p_H)$ by

$$\hat{L}(p; p_H) = \hat{B}(p_H)\hat{\psi}(p) + \breve{B}(p_H)\breve{\psi}(p), \tag{D.5}$$

where the constants $\hat{B}(p_H)$ and $\breve{B}(p_H)$ are given by

$$\hat{B}(p_H) = \frac{\breve{\psi}(p_H)}{2\gamma}\left[G_1'(p_H) - \frac{1/2 - \gamma - p_H}{p_H(1-p_H)}G_1(p_H)\right], \tag{D.6}$$

and

$$\breve{B}(p_H) = \frac{\hat{\psi}(p_H)}{2\gamma}\left[\frac{1/2 + \gamma - p_H}{p_H(1-p_H)}G_1(p_H) - G_1'(p_H)\right]. \tag{D.7}$$

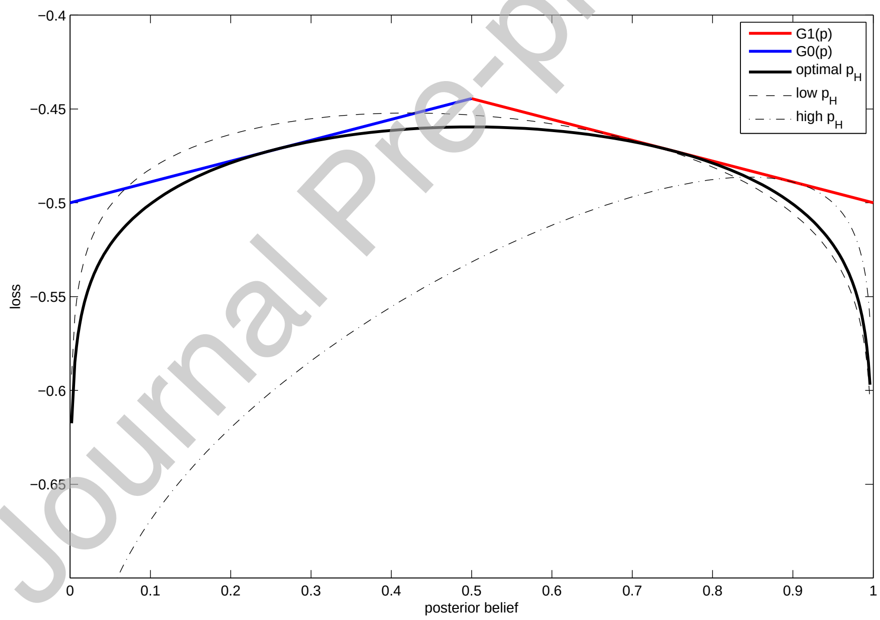

Figure D.2: The Loss Function for Different Values of p_H.

It is easily verified that $\hat{L}(p_H; p_H) = G_1(p_H)$ and $\hat{L}'(p_H; p_H) = G_1'(p_H)$. Using a similar method it holds that $\hat{B} < 0$, $\breve{B} < 0$, and $\hat{L}$ is concave on $(0, 1)$.

3. It is easy to prove that a solution to the free-boundary problem (D.1) is also a solution to

Journal Pre-proof

the optimal stopping problem (11). That is,

$$G^*(p) := L^*(p) - \frac{\mu}{2\lambda} = \inf_\tau E_p\left[e^{-\lambda\tau}G(p_\tau)\right] =: \inf_\tau J^\tau(p).$$

Obviously, it holds that $G^*(p) \geq \inf_\tau J^\tau(p)$. To prove the reverse inequality, take any stopping time τ. It now holds that

$$\begin{aligned} G^*(p) &= E_p\left[e^{-\lambda\tau}G^*(p_\tau)\right] - E_p\left[\int_0^\tau e^{-rt}\left(\mathcal{L}-\lambda\right)G^*(p_t)dt\right] \\ &\leq E_p\left[e^{-\lambda\tau}G^*(p_\tau)\right] \leq E_p\left[e^{-\lambda\tau}G(p_\tau)\right] = J^\tau(p), \end{aligned}$$

where the first equality follows from Dynkin's formula (see Øksendal (2003)), the first inequality follows from $\mathcal{L}G^* - \lambda G^* \geq 0$, and the second inequality follows from $G^* \leq G$. Hence, $G^*(p) \leq \inf_\tau J^\tau(p)$. So $G^*(p) = \inf_\tau J^\tau(p)$.

Journal Pre-proof

Journal Pre-proof

Appendix E. Proof of Proposition 5.1

Let $\tau^* = \inf\{t \geq 0 \mid p_t \notin (p_B, p_A)\}$ be the unique solution to (12). Consider a change in γ. Such a change affects $\hat{A}$, $\breve{A}$, $\hat{B}$, $\breve{B}$, $\hat{\psi}$ and $\breve{\psi}$. The constants $\hat{A}$ and $\breve{A}$ in (D.3, D.4) can be rewritten as

$$\hat{A}(p_B) = \frac{p_B(1-p_B)}{2\gamma\hat{\psi}(p_B)}\left(G_0'(p_B) - \frac{\frac{1}{2}-\gamma-p_B}{p_B(1-p_B)}G_0(p_B)\right),$$

and

$$\breve{A}(p_B) = \frac{p_B(1-p_B)}{2\gamma\breve{\psi}(p_B)}\left(\frac{\frac{1}{2}+\gamma-p_B}{p_B(1-p_B)}G_0(p_B) - G_0'(p_B)\right).$$

Similarly, the constants $\hat{B}$ and $\breve{B}$ in (D.6, D.7) can be rewritten as

$$\hat{B}(p_A) = \frac{p_A(1-p_A)}{2\gamma\hat{\psi}(p_A)}\left(G_1'(p_A) - \frac{\frac{1}{2}-\gamma-p_A}{p_A(1-p_A)}G_1(p_A)\right),$$

and

$$\breve{B}(p_A) = \frac{p_A(1-p_A)}{2\gamma\breve{\psi}(p_A)}\left(\frac{\frac{1}{2}+\gamma-p_A}{p_A(1-p_A)}G_1(p_A) - G_1'(p_A)\right).$$

The proof of Proposition 4.1 was based on the fact that the mapping $p \mapsto \breve{L}(p;p_L)$ is such that $\breve{L}(p_L;p_L) = G_0(p_L)$ and $\breve{L}'(p_L;p_L) = G_0'(p_L)$ for all $p_L \in (0,\bar{p})$; and the mapping $p \to \hat{L}(p;p_H)$ is such that $\hat{L}(p_H;p_H) = G_1(p_H)$ and $\hat{L}'(p_H;p_H) = G_1'(p_H)$ for all $p_H \in (\bar{p},1)$.

We have also shown that $\hat{A} < 0$, $\breve{A} < 0$, $\hat{B} < 0$ and $\breve{B} < 0$ in the proof of Proposition 4.1. Then we prove that

$$\begin{aligned}
\frac{\partial\gamma}{\partial\lambda} &= \frac{1}{2}\cdot\frac{1}{2}\cdot\left[1+2\lambda\left(\frac{\sigma}{\mu}\right)^2\right]^{-\frac{1}{2}}\cdot 2\cdot\left(\frac{\sigma}{\mu}\right)^2 > 0,\\
\frac{\partial\gamma}{\partial\mu} &= \frac{1}{2}\cdot\frac{1}{2}\cdot\left[1+2\lambda\left(\frac{\sigma}{\mu}\right)^2\right]^{-\frac{1}{2}}\cdot 2\lambda\cdot 2\left(\frac{\sigma}{\mu}\right)\cdot\left(-\frac{\sigma}{\mu^2}\right) < 0,\\
\frac{\partial\gamma}{\partial\sigma} &= \frac{1}{2}\cdot\frac{1}{2}\cdot\left[1+2\lambda\left(\frac{\sigma}{\mu}\right)^2\right]^{-\frac{1}{2}}\cdot 2\lambda\cdot 2\left(\frac{\sigma}{\mu}\right)\cdot\frac{1}{\mu} > 0.
\end{aligned}$$

So γ is increasing in λ and σ, but decreasing in μ.

Journal Pre-proof

Differentiating $\hat{\psi}$ and $\breve{\psi}$ with respect to γ get

$$\frac{\partial\hat{\psi}(p)}{\partial\gamma} = \hat{\psi}(p)\log\left(\frac{p}{1-p}\right), \frac{\partial\breve{\psi}(p)}{\partial\gamma} = \breve{\psi}(p)\log\left(\frac{1-p}{p}\right).$$

Then we differentiate $\hat{A} < 0$ and $\breve{A} < 0$ with respect to γ,

$$\frac{\partial\hat{A}}{\partial\gamma} = \hat{A}\left[\log\left(\frac{1-p_B}{p_B}\right) - \frac{1}{\gamma}\right] + \frac{\breve{\psi}\left(p_B\right)}{2\gamma}\frac{G_0\left(p_B\right)}{p_B\left(1-p_B\right)},$$

$$\frac{\partial\breve{A}}{\partial\gamma} = \breve{A}\left[\log\left(\frac{p_B}{1-p_B}\right) - \frac{1}{\gamma}\right] + \frac{\hat{\psi}\left(p_B\right)}{2\gamma}\frac{G_0\left(p_B\right)}{p_B\left(1-p_B\right)}.$$

It is then easily obtained that

$$\begin{aligned}\frac{\partial\breve{L}\left(p_A;p_B\right)}{\partial\gamma} &= \left(\hat{A}\hat{\psi}\left(p_A\right) - \breve{A}\breve{\psi}\left(p_A\right)\right)\log\left(\frac{1-p_B}{p_B}\frac{p_A}{1-p_A}\right)\\ &+\frac{1}{2\gamma}\left[\frac{\hat{\psi}\left(p_A\right)\breve{\psi}\left(p_B\right)+\breve{\psi}(p_A)\hat{\psi}\left(p_B\right)}{p_B\left(1-p_B\right)}G_0\left(p_B\right) - 2G_1\left(p_A\right)\right].\end{aligned} \tag{E.1}$$

Note that the logarithmic term in (E.1) is positive since $0 < p_B < 1/2 < p_A < 1$ and the term in round brackets is negative. If we differentiate at p_A, we obtain

$$\begin{aligned}&\hat{A}\hat{\psi}'\left(p_A\right) + \breve{A}\breve{\psi}'\left(p_A\right) = G_1'\left(p_A\right) < 0,\\ \Longleftrightarrow&\frac{\gamma+\left(\frac{1}{2}-p_A\right)}{p_A\left(1-p_A\right)}\hat{A}\hat{\psi}\left(p_A\right) + \frac{-\gamma+\left(\frac{1}{2}-p_A\right)}{p_A\left(1-p_A\right)}\breve{A}\breve{\psi}\left(p_A\right) < 0,\\ \Longleftrightarrow&\hat{A}\hat{\psi}(p_A) < \frac{\gamma-\left(\frac{1}{2}-p_A\right)}{\gamma+\left(\frac{1}{2}-p_A\right)}\breve{A}\breve{\psi}\left(p_A\right) < \breve{A}\breve{\psi}\left(p_A\right).\end{aligned} \tag{E.2}$$

Journal Pre-proof

So the first term on the right-hand side is negative and the second term is also negative because

$$\begin{aligned}
&\tfrac{\hat{\psi}(p_A)\breve{\psi}(p_B)+\breve{\psi}(p_A)\hat{\psi}(p_B)}{p_B(1-p_B)}G_0\left(p_B\right)-2G_1\left(p_A\right)\\
&=\left(\tfrac{\hat{\psi}(p_A)}{\hat{\psi}(p_B)}+\tfrac{\breve{\psi}(p_A)}{\breve{\psi}(p_B)}\right)G_0\left(p_B\right)-2G_1\left(p_A\right)\\
&=\left(\tfrac{\hat{\psi}(p_A)}{\hat{\psi}(p_B)}+\tfrac{\breve{\psi}(p_A)}{\breve{\psi}(p_B)}\right)\left(\hat{A}\hat{\psi}\left(p_B\right)+\breve{A}\breve{\psi}\left(p_B\right)\right)-2\left(\hat{A}\hat{\psi}\left(p_A\right)+\breve{A}\breve{\psi}\left(p_A\right)\right)\\
&=\hat{A}\left[\hat{\psi}\left(p_A\right)+\tfrac{\breve{\psi}(p_A)\hat{\psi}(p_B)}{\breve{\psi}(p_B)}-2\hat{\psi}\left(p_A\right)\right]+\breve{A}\left[\tfrac{\hat{\psi}(p_A)\breve{\psi}(p_B)}{\hat{\psi}(p_B)}+\breve{\psi}\left(p_A\right)-2\breve{\psi}(p_A)\right]\\
&=\hat{A}\left[\tfrac{\breve{\psi}(p_A)\hat{\psi}(p_B)-\hat{\psi}(p_A)\breve{\psi}(p_B)}{\breve{\psi}(p_B)}\right]+\breve{A}\left[\tfrac{\hat{\psi}(p_A)\breve{\psi}(p_B)-\breve{\psi}(p_A)\hat{\psi}(p_B)}{\hat{\psi}(p_B)}\right]\\
&=\left(\breve{\psi}\left(p_A\right)\hat{\psi}\left(p_B\right)-\hat{\psi}\left(p_A\right)\breve{\psi}\left(p_B\right)\right)\left(\tfrac{\hat{A}}{\breve{\psi}(p_B)}-\tfrac{\breve{A}}{\hat{\psi}(p_B)}\right)\\
&=\underbrace{\frac{\breve{\psi}\left(p_A\right)\hat{\psi}\left(p_B\right)-\hat{\psi}\left(p_A\right)\breve{\psi}\left(p_B\right)}{\breve{\psi}\left(p_B\right)\hat{\psi}\left(p_B\right)}}_{<0}\underbrace{\left(\hat{A}\hat{\psi}\left(p_B\right)-\breve{A}\breve{\psi}\left(p_B\right)\right)}_{>0}.
\end{aligned} \tag{E.3}$$

Combining (E.2) and (E.3) shows that $\frac{\partial\breve{L}(p_A;p_B)}{\partial\gamma}<0$. To restore optimality after an increase in γ, the trigger p_B must be increased. A similar analysis shows that $\frac{\partial\hat{L}(p_B;p_A)}{\partial\gamma}<0$ so that p_A must be decreased after an increase in γ.

Journal Pre-proof

Appendix F. Expected Discount Factors

The general solution to the second order stochastic differential equation $\mathcal{L}\psi - \lambda\psi = 0$ is

$$\psi(p) = \hat{A}\hat{\psi}(p) + \breve{A}\breve{\psi}(p).$$

We need that $\psi(p_B) = 0$ and $\psi(p_A) = 0$, so

$$\hat{A}\hat{\psi}(p_B) + \breve{A}\breve{\psi}(p_B) = 0, \tag{F.1}$$

which means

$$\hat{A} = -\frac{\breve{\psi}(p_B)}{\hat{\psi}(p_B)}\breve{A}. \tag{F.2}$$

Substituting (F.2) into (F.1), we obtain

$$\psi(p) = -\frac{\breve{\psi}(p_B)}{\hat{\psi}(p_B)}\breve{A}\hat{\psi}(p) + \breve{A}\breve{\psi}(p).$$

Therefore the discount factor is

$$\begin{aligned}
\hat{v}_{p_B,p_A}(p) &= \frac{\psi(p)}{\breve{\psi}(p_A)} \\
&= \frac{-\frac{\breve{\psi}(p_B)}{\hat{\psi}(p_B)}\breve{A}\hat{\psi}(p) + \breve{A}\breve{\psi}(p)}{-\frac{\breve{\psi}(p_B)}{\hat{\psi}(p_B)}\breve{A}\hat{\psi}(p_A) + \breve{A}\breve{\psi}(p_A)} \\
&= \frac{-\breve{\psi}(p_B)\hat{\psi}(p) + \hat{\psi}(p_B)\breve{\psi}(p)}{-\breve{\psi}(p_B)\hat{\psi}(p_A) + \hat{\psi}(p_B)\breve{\psi}(p_A)} \\
&= \sqrt{\frac{p(1-p)}{p_A(1-p_A)}}\,\frac{(\frac{1-p_B}{p_B}\frac{p}{1-p})^\gamma - (\frac{p_B}{1-p_B}\frac{1-p}{p})^\gamma}{(\frac{1-p_B}{p_B}\frac{p_A}{1-p_A})^\gamma - (\frac{p_B}{1-p_B}\frac{1-p_A}{p_A})^\gamma}.
\end{aligned}$$

Journal Pre-proof

Journal Pre-proof

References

Ali, M. M., 1977. Probability and utility estimates for racetrack bettors. J. Political Econ. 85 (4), 803–815.

Asch, P., Malkiel, B. G., Quandt, R. E., 1982. Racetrack betting and informed behavior. J. Financial Econ. 10 (2), 187–194.

Asch, P., Malkiel, B. G., Quandt, R. E., 1984. Market efficiency in racetrack betting. J. Bus. 57 (2), 165–175.

Bagehot, W., 1971. The only game in town. Financial Anal. J. 27 (2), 12–14, 22.

Brown, L. D., D'Amato, R., Gertner, R., 1994. Racetrack betting: Do bettors understand the odds? CHANCE 7 (3), 17–23.

Cain, M., Law, D., Peel, D., 2003. The favouritelongshot bias, bookmaker margins and insider trading in a variety of betting markets. Bull. Econ. Res. 55 (3), 263–273.

Cain, M., Law, D., Peel, D. A., 2001. The relationship between two indicators of insider trading in British racetrack betting. Economica 68 (269), 97–104.

Copeland, T. E., Galai, D., 1983. Information effects on the bid-ask spread. J. Finance 38 (5), 1457–1469.

Crafts, N. F., 1985. Some evidence of insider knowledge in horse race betting in Britain. Economica 52 (207), 295–304.

Dixit, A. K., Pindyck, R. S., 1994. Investment under Uncertainty. Princeton University Press.

Dowie, J., 1976. On the efficiency and equity of betting markets. Economica 43 (170), 139–150.

Easley, D., O'Hara, M., 1987. Price, trade size, and information in securities markets. J. Financial Econ. 19 (1), 69–90.

Fama, E. F., 1970. Efficient capital markets: A review of theory and empirical work. J. Finance 25 (2), 383–417.

Journal Pre-proof

Fingleton, J., Waldron, P., 1999. Optimal determination of bookmakers' betting odds: Theory and tests, trinity Economic Paper Series, Technical Paper No. 96/9, Trinity College, Dublin.

Gabriel, P., Marsden, J. R., 1991. An examination of efficiency in British racetrack betting: Errata and corrections. J. Political Econ. 99 (3), 657–659.

Gabriel, P. E., Marsden, J. R., 1990. An examination of market efficiency in British racetrack betting. J. Political Econ. 98 (4), 874–885.

Glosten, L. R., Milgrom, P. R., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. J. Financial Econ. 14 (1), 71–100.

Griffith, R. M., 1949. Odds adjustments by American horse-race bettors. Am. J. Psychol. 62 (2), 290–294.

Hausch, D. B., Lo, V. S., Ziemba, W. T., 2008. Efficiency of racetrack betting markets. Vol. 2. World Scientific.

Hausch, D. B., Ziemba, W. T., Rubinstein, M., 1981. Efficiency of the market for racetrack betting. Manag. Sci. 27 (12), 1435–1452.

Henery, R. J., 1985. On the average probability of losing bets on horses with given starting price odds. J. R. Stat. Soc. 148 (4), 342–349.

Hurley, W., McDonough, L., 1995. A note on the hayek hypothesis and the favorite-longshot bias in parimutuel betting. Am. Econ. Rev., 949–955.

Isaacs, R., 1953. Optimal horse race bets. Am. Math. Mon. 60 (5), 310–315.

Jeong, J., Kim, J. Y., Ro, Y. J., 2019. On the efficiency of racetrack betting market: a new test for the favourite-longshot bias. Appl. Econ. 51 (54), 5817–5828.

Kyle, A. S., 1985. Continuous auctions and insider trading. Econometrica 53 (6), 1315–1335.

Levitt, S. D., 2004. Why are gambling markets organised so differently from financial markets? Econ. J. 114 (495), 223–246.

Journal Pre-proof

Li, W., Song, Z., 2019. Dealers as information intermediaries in over-the-counter markets. Available at SSRN 3351331.

McGlothlin, W. H., 1956. Stability of choices among uncertain alternatives. Am. J. Psychol. 69 (4), 604–615.

Øksendal, B., 2003. Stochastic differential equations. In: Stochastic differential equations. Springer, pp. 65–84.

Ottaviani, M., Sorensen, P. N., 2010. Noise, information, and the favorite-longshot bias in parimutuel predictions. Am. Econ. J.: Microecon. 2 (1), 58–85.

Peskir, G., Shiryayev, A., 2006. Optimal Stopping and Free Boundary Problems. Basel: Birkhäuser Verlag.

Poor, H. V., Hadjiliadis, O., 2009. Quickest Detection. Cambridge University Press Cambridge.

Schnytzer, A., Shilony, Y., 1995. Inside information in a betting market. Econ. J. 105 (431), 963–971.

Shin, H. S., 1991. Optimal betting odds against insider traders. Econ. J. 101 (408), 1179–1185.

Shin, H. S., 1992. Prices of state contingent claims with insider traders, and the favourite-longshot bias. Econ. J. 102 (411), 426–435.

Shin, H. S., 1993. Measuring the incidence of insider trading in a market for state-contingent claims. Econ. J. 103 (420), 1141–1153.

Shing, H.-F., Koch, A., 2008. Bookmaker and parimutuel betting: Is a (reverse) favourite-longshot bias built-in? J. Predict. Mark. 2 (2), 29–50.

Shiryayev, A. N., 1978. Optimal Stopping Rules. Springer-Verlag, New York.

Smith, M. A., Paton, D., Vaughan Williams, L., 2006. Market efficiency in person-to-person betting. Economica 73 (292), 673–689.

Journal Pre-proof

Steele, J. M., 2012. Stochastic Calculus and Financial Applications. Springer Science & Business Media.

Thijssen, J., Bregantini, D., 2017. Costly sequential experimentation and project valuation with an application to health technology assessment. J. Econ. Dyn. Control 77, 202–229.

Vaughan Williams, L., 2005. Information Efficiency in Financial and Betting Markets. Cambridge University Press.

Vaughan Williams, L., Paton, D., 1997. Why is there a favourite-longshot bias in british racetrack betting markets? Econ. J. 107 (440), 150–58.

Weston, J. P., 2001. Information, liquidity, and noise.
URL http://www.ruf.rice.edu/~westonj/pub/journals/weston_iln.pdf

Ziemba, W. T., 2019. Accuracy of the win odds. In: *Exotic Betting at the Racetrack*. World Scientific Publishing Co. Pte. Ltd., pp. 1–20.

Journal Pre-proof