The Impact of Search and Holding Costs on Persistent Mis-pricing

Ben R. Marshall
Department of Finance, Banking and Property
Massey University
Private Bag 11-222
Palmerston North
New Zealand
Tel: 64 6 350 5799 ext 5402
Fax 64 6 350 5651
Email: B.Marshall@Massey.ac.nz.

JEL Classification: *G14*

Keywords: *Search Costs, Holding Costs, Arbitrage*

I would like to thank Jason Thompson of zero-risk-arbitrage.com for kindly providing the data. The views expressed in this paper are not necessarily those of zero-risk-arbitrage.com. I would also like to thank Hamish Anderson, Jonathan Batten, Henk Berkman, Charles Corrado, Nont Dhiensiri, John Gandar, Pinghsun Huang, Ben Jacobsen, Chris Malone, John Powell, Martin Young, Eric Zitzewitz, Rick Zuber, and seminar participants at the 2005 Western Finance Association Annual Meeting and the 2006 Asian FA/FMA annual meeting for valuable comments. This paper won the International Journal of Finance Best Paper Prize at the 2006 Asian FA/FMA annual meeting.

The Impact of Search and Holding Costs on Persistent Mis-pricing

This paper uses data from internet sports betting markets to provide some of the first empirical evidence that search and holding costs are important determinants of persistent arbitrage opportunities. I find information uncertainty regarding sporting event outcomes leads to betting odds that are sufficiently divergent for arbitrage profits to be made. These arbitrage opportunities persist for a median time of 15 minutes due to search and holding costs. I show the arbitrage profits are not driven by unique characteristics of sports betting markets and propose they are not compensation for fundamental or noise trader risk or implementation costs.

JEL Classification: *G14*

Keywords: *Search Costs, Holding Costs, Arbitrage*

Arbitrage is one of the central tenets of financial economics, enforcing the law of one price and keeping markets efficient. In the traditional framework where there are no frictions and agents are rational, any deviation of price from fundamental value creates an attractive investment opportunity which is exploited immediately. In reality, arbitrage requires capital and there are costs and risks involved so there are limits to arbitrage, and pervasive mis-pricing exists.

This paper documents the impact of search and holding costs on mis-pricing. To the best of my knowledge, there is no other work on the link between search costs and arbitrage opportunities, despite a large number of papers (e.g. Hortacsu and Syverson (2004)) highlighting the importance of search costs in explaining sustained price dispersion. Rather, previous work on search costs has considered markets where arbitrage is not possible. There is also little literature on the link between arbitrage opportunities and holding costs. Tuckman and Vila (1992) develop a theoretical model that highlights the link between holding costs and arbitrage activity and test their model using Treasury Bill data. They find arbitrage opportunities that involve higher holding costs persist for longer. The lack of other research on the relationship between search and/or holding costs and arbitrage opportunities may be due to difficulties in disentangling these effects from the three well documented barriers to arbitrage (fundamental risk[1], noise trader risk, and implementation costs). Holding costs are distinct to implementation costs. The latter refer to "transaction costs such as commissions, bid-ask spreads, and price impact (costs) … (and) short-selling costs." (Barberis and Thaler (2002, p. 6)).

The setting for my analysis is the global internet sports betting market. As noted in Durham, Hertzel and Martin (2005), sports betting markets share many important features with stock markets (e.g. large volume and wide availability of information) but have several strengths as an empirical laboratory. The particular advantages this paper is interested in are: numerous price (or odd) quotations on the same event by different decentralised market makers (bookmakers), diverse price quotations generating arbitrage opportunities, these arbitrage opportunities being easily identified and valued based on odds that are perfect substitutes (i.e. no fundamental risk), the odds having a short life and a well defined termination point (i.e. no noise trader risk), and these arbitrage transaction being executed with minimal cost via the internet (i.e. low transactions costs). This paper shows the arbitrage opportunities do not persist due to unique aspects of the internet betting market, such as credit risk or home country betting biases.

I show information uncertainty results in odds that are sufficiently divergent that arbitrage opportunities are created. Sporting events with three possible outcomes (team A wins, team B wins, or there is a draw) have a higher level of uncertainty then events with two possible outcomes (team A wins or team B wins) and also have more divergent odds. This in turn leads to more profitable arbitrage opportunities being generated for three-outcome sporting events.

Arbitrage opportunities involving three-outcome sporting events also have higher search costs because the odds of three rather than two bookmakers need to be analysed to determine if an arbitrage opportunity exists. I find strong evidence that those arbitrage opportunities with higher search costs persist for longer. There is also strong evidence that holding costs are a

determinant of arbitrage activity. An arbitrageur does not have time to transfer money to bookmaker accounts after identifying an opportunity, but rather has to have money sitting in multiple accounts ready for use if an arbitrage opportunity is created. I find those bookmakers involved in few arbitrage opportunities, which are those that have the highest holding costs from an opportunity cost perspective, generate arbitrage opportunities that persist for longer than those generated by bookmakers who post odds that lead to numerous opportunities.

The remainder of the paper is organised as follows: The first section describes the market setting and provides further motivation for the paper. Section II describes the data set and the methodology used to determine if an arbitrage opportunity exists. Theoretical background and hypotheses are presented in Section III. The results are presented in Section IV. Section V concludes the paper.

I. The Market Setting

A. Betting Markets as a Useful Laboratory

There is both theoretical and empirical work on the impact of search costs on price dispersion within markets. Stahl (1989) shows prices for the same good can differ across stores if consumers are faced with search costs. Hortacsu and Syverson (2004) show search costs are a partial explanation for the proliferation of S&P 500 index funds, which offer the same product but charge different fees. However, to the best of my knowledge, no paper has considered the link between search costs and the persistence of arbitrage opportunities. This is likely to be

due to difficulties in controlling for other pervasive determinants of arbitrage opportunities such as fundamental risk, noise trader risk and implementation costs. This factor is also likely to explain the lack of research on the impact of holding costs on mis-pricing. The only empirical paper is by Tuckman and Vila (1992), who show holding costs explain differences in prices of equivalent U.S. Treasury Bills.

Empirical tests of theoretical models within the financial economics literature utilise data from four broad categories: product/service markets, financial markets, experimental markets, and sports betting markets. Each setting has its own advantages and disadvantages. Product/service and financial markets have the obvious advantage of being the setting in which the theory relates to. The disadvantage of using data from such a market is that it is often impossible to find conclusive evidence of a relationship due to confounding effects of other variables and the inability to control for them.

This issue is particularly relevant when it comes to quantifying the impact of search and holding costs on the persistence of arbitrage opportunities. Evidence of these two effects is only clear if well known determinants of arbitrage activity, such as fundamental risk, noise trader risk, and implementation costs do not apply or can be controlled for. This is virtually impossible in financial markets as two assets are rarely perfect substitutes for each other. In addition, there is almost never a well defined termination point at which time the true value of the two assets is determined. This leads to the possibility of noise traders irrationally causing prices to diverge in the short-term forcing an arbitrageur to liquidate their positions at a loss. Finally, implementation costs in the form of short-selling constraints are difficult to control for.

Experimental markets enable researchers to control for factors such as the main arbitrage determinants and consider the impact of other factors such as search and holding costs. However, Durham, Hertzel, and Martin (2005) point out it is sometimes difficult to extrapolate these inferences back to the real world because in the real world, individuals have their own wealth at stake.

Sports betting markets can be a useful middle ground, which has lead to these data being widely used in the financial economics literature. Gandar, Dare, Brown, and Zuber (1998) use National Basketball Association (NBA) data to document evidence consistent with market efficiency. Avery and Chevalier (1999) use National Football League (NFL) to show investor sentiment can affect prices. Durham, Hertzel, and Martin (2005) use college football betting market data to test the regime shifting model of Barberis, Shleifer, and Vishny (1998).

Compared to experimental markets, betting markets offer the advantage of vast volumes of real money on the line over lengthy time periods. Compared to financial markets, betting markets have the advantage of providing a market where arbitrage opportunities are easily identified and valued and based on odds that are perfect substitutes (i.e. no fundamental risk), a well defined termination point to each arbitrage opportunity (i.e. no noise trader risk), and minimal transaction costs. In addition, sports betting markets have no unique characteristics that lead to other impediments to arbitrage activity that are not prevalent in financial markets.[2] For instance, I show credit risk is not a likely explanation for these opportunities.

B. Institutional Details: Internet Betting Markets

It is difficult to estimate the size of the internet sports betting market, but the evidence suggests it is in well in excess of U.S.$20 billion per annum. Sinclair (1998) estimates the size of the highly regulated Nevada market was US$2.7 billion by 1997 and Strumpf (2003) estimates internet betting on U.S. sports alone is now up to ten times bigger than the Nevada market. There are now hundreds of internet bookmakers simultaneously offering odds on the same sporting event.

The sports betting market system is quite different from the pari-mutuel structure used by many horse race tracks. In pari-mutuel markets, money is returned to winning bettors, following the deduction of the market makers' profits, in proportion to their individual stake. The actual payoffs to winning bettors are therefore not known until the close of the betting period. In contrast, sports bookmakers offer odds a bettor is guaranteed to receive regardless of how odds change subsequent to their placing the bet.

Prior to 2006 (i.e. the period of this study) non-U.S. based companies offering internet betting services to U.S. based individuals were not violating any U.S. federal laws. This changed with the adoption of the Unlawful Internet Gambling Enforcement Act (2006). This legislation criminalizes the acceptance of funds from bettors by operators of most online gambling websites. U.S. gambling law still permits Nevada, Oregon, and Delaware to offer state regulated sports betting, but offering sports betting services in other states is illegal. This means that U.S. residents are no longer able to pursue the strategies outlined in this paper, but there opportunities are still available to non-U.S. residents.

Internet sports bookmakers in non-US countries face varying levels of regulation. Countries such as Great Britain, Canada, and Australia conduct stringent checks before issuing licenses to ensure each bookmaker is a legitimate business. In addition, most U.K. bookmakers are regulated by the Independent Betting Arbitration Service which is an authoritative, totally independent third party offering adjudication for customers who have an unresolved betting dispute with their bookmaker. Moreover, many European and U.K. bookmakers are stock-exchange listed. However, in other countries such as the Caribbean, a bookmaker can gain a license by simply paying a fee to the Government. Very few, if any, checks are carried out in these countries (Rose (1999)).

II. Data

A. Source

The data are sourced from zero-risk-arbitrage.com, a British company that identifies sports betting arbitrage opportunities in the odds quoted by internet sports bookmakers. Zero-risk-arbitrage.com uses proprietary software to scan bookmaker websites in real-time. It then alerts subscribers to an arbitrage opportunity as soon as it is created via an SMS message or "ArbAlarm", which is proprietary software that transmits to subscribers' desktops in real-time. Zero-risk-arbitrage.com data are unique in both their depth and breadth. Zero-risk-arbitrage.com started identifying and recording arbitrage opportunities electronically in late 2002. Data for the full 2003 and 2004 years are unavailable due to the loss of September 2003 – March 2004 data by zero-risk-arbitrage.com, so the data set spans the January 2003 – December 2005 period, with the exception of the period for which data are missing.[3]

The original data set contains 529,561 arbitrage opportunities - however, many of these are on the same sporting event involving the same bookmaker. This occurs because when one bookmaker (say bookmaker A) posts an odd that is sufficiently different from the other bookmakers to create an arbitrage opportunity, zero-risk-arbitrage.com records arbitrage opportunities between bookmaker A and every other bookmaker. In reality, an individual would be able to pursue only one of these opportunities due to time and/or bet maximum constraints with bookmaker A. Therefore this paper examines only one of such bookmaker duplicate arbitrage opportunities and those that do not involve one common bookmaker on the same sporting event. The approach outlined above resulted in all but 19,882 arbitrage opportunities being removed from the data set.

Zero-risk-arbitrage.com focuses on arbitrage opportunities in sporting events that involve two or three-outcomes. The sports included in the zero-risk-arbitrage.com database are Cricket, Darts, Formula One Motor Racing, Golf, Major League Baseball (MLB), Nascar Motor Racing, National Basketball Association (NBA), National Football League (NFL), National Hockey League (NHL), Rugby, Snooker, Soccer, and Tennis.

Zero-risk-arbitrage.com has added more bookmakers to the universe it scans each year. To ensure the results are comparable across years, this research includes only arbitrage opportunities that stem from the fifty bookmakers that zero-risk-arbitrage.com tracked in each of the three years. These bookmakers were based in Costa Rica, Netherlands Antilles, West Indies (Curacao), Canada, Australia, Austria, Isle of Man, Gibraltar, UK, Ireland, Malta, Cyprus.

B. Market Conventions and Mechanics

A simple formula can be applied to determine if an arbitrage opportunity exists. Assuming there are n outcomes in a contest and that a successful bet on outcome i returns X_i units (inclusive of the original bet) the bettor must wager $\frac{1}{X_i}$ on outcome i for all $1 \leq i \leq n$. This requires a total outlay (TO) of:

$$TO = \sum_{i=1}^{n} \frac{1}{X_i} \tag{1}$$

For an arbitrage opportunity to exist the total outlay that guarantees a payoff of 1 unit (equation 1) must be less than one.

The total revenue earned (*REVENUE*) per unit bet is:

$$REVEUNE = \frac{1 - \sum_{i=1}^{n} \frac{1}{X_i}}{\sum_{i=1}^{n} \frac{1}{X_i}} \text{ subject to } \sum_{i=1}^{n} \frac{1}{X_i} < 1 \tag{2}$$

I use the term *REVENUE* rather than "profit" to describe the payback because arbitrageurs do face minor costs such as internet service provider costs. I discuss these in Section III and show they are small compared to the revenues on offer.

The proportion of the total bet the arbitrageur should bet on each outcome (i = 1,2, …, n). to achieve certain revenue regardless of the outcome is given by:

$$\text{Proportion to bet on outcome } i = \frac{\frac{1}{X_i}}{\frac{1}{X_1}+\frac{1}{X_2}+\ldots+\frac{1}{X_n}} \tag{3}$$

By way of example, on the 28th of August 2003 the following odds were on offer for a Major League Baseball (MLB) match between the Boston Red Sox and Oakland Athletics:

- Boston Red Sox to Win at Stan James at 1.83
- Oakland Athletics to Win at ToteXpress at 2.38

To determine if an arbitrage opportunity exists, the arbitrageur should calculate:

$$TO = \sum_{i=1}^{n} \frac{1}{X_i} = \frac{1}{1.83} + \frac{1}{2.38}$$

$TO = 0.9666$

AP is less than 1 so an arbitrage opportunity does exist.

To receive guaranteed revenue regardless of the outcome the arbitrageur must bet (1/1.83 / (1/1.83+1/2.38)) or 56.53% his/her total bet on Boston Red Sox and (1/1.83 / (1/1.83+1/2.38)) or 43.47% on Oakland Athletics. Assuming the arbitrageur has $966.60 to bet, s/he should bet $546.40 on Boston Red Sox to win at odds of $1.83 at Stan James and $420.20 on Oakland Athletics to win at odds of $2.38 at ToteXpress.

Based on these bets the return per \$1 bet is $R = \frac{1 - 0.9666}{0.9666} = 0.0346$.

In this situation, the arbitrageur earns \$0.0346 for every \$1 bet regardless of the outcome of the event. The arbitrageur commits \$966.60 to the event so s/he receives \$33.44 *REVENUE*.

C. Summary Statistics

The summary statistics displayed in Table I indicate arbitrage *REVENUEs* range from 0.91 percent to 11.10 percent. The mean is 2.03 percent and the median is 1.51 percent, indicating skewness. *DURATION* is the number of minutes it takes for an arbitrage opportunity to be removed. The shortest length of time an arbitrage opportunity lasts for is 0.02 of a minute or 1.2 seconds. Clearly, it would be very difficult to exploit this opportunity, however, 75 percent of opportunities last for 6.3 minutes or longer. The longest duration of an opportunity is just under a day. As with the *REVENUE* results, there is positive skewness in the durations, with a mean of 59.4 minutes versus a median of 15.4 minutes.

[Insert Table I About Here]

III. Theoretical Background and Hypotheses

The structures of the Nasdaq stock market and the internet sports betting market are similar in that both involve dealers posting two-way quotes. However, unlike sports betting markets the quotes of all Nasdaq market makers are displayed for participants to view. Investors wishing

to purchase or sell a stock trade at the inside quotes which consist of the best (highest) bid and (lowest) ask price. Although the trading system opens at 9:30 a.m., the quotation reporting system of the Nasdaq opens much earlier. During the preopening period from 8:00 a.m. to 9:30 a.m., market makers transmit their bid-ask quotes, observe other dealers' quotes, and revise their own quotes in response to the quotes of others.

Cao, Ghysels, and Hatheway (2000) find information uncertainty and diversity of opinion among dealers lead to 23.6% of the inside quotes being crossed (the highest bid being greater than the lowest ask) during the pre-opening session. Quotes remain crossed for an average duration of 29.2 minutes. Nasdaq regulations discourage crossed trades during trading hours. The market maker who creates a cross is contacted by Nasdaq officials to restore it quickly.

I propose information uncertainty regarding the true probabilities of each sporting event outcome lead to divergent odds being posted by different bookmakers. The quotes of bookmakers are not displayed in a centralised "exchange" and there is no coordination of quotes so there is nothing forcing bookmakers to remove crossed quotes. Thus the first hypothesis is:

H_1:

Larger arbitrage opportunities are created when there is more information uncertainty.

Bookmakers could remove crossed odds themselves after viewing the odds of their competitors but I suggest they may not do this for two possible reasons. Either they believe their odds better reflect the outcome probabilities than do those of their competitors.

Alternatively, they may purposely quote some odds at a level above those of their competitors as a marketing ploy to attract customers to their website. This is consistent with the theory of spatial price dispersion discussed by Salop and Stiglitz (1977) and Varian (1980).

Stahl (1989) shows, via a theoretical model, that costly search can sustain price dispersion in homogenous product markets. Further evidence of this relationship is provided by Zhao (2006) in grocery markets. Sports betting arbitrage is a particularly time-consuming process. The market for an information aggregator like zero-risk-arbitrage.com supports this notion. Anyone who does not subscribe to an alert service like zero-risk-arbitrage.com has to manually compare odds on different websites. Even zero-risk-arbitrage.com subscribers need to devote several hours each day to implementing the alerts they receive if they are to maximise their profit.

Arbitrage opportunities involve sports with two possible outcomes (team A wins, team B wins) and three possible outcomes (team A wins, team B wins, and there is a draw). Three-outcome arbitrage opportunities are more difficult to find because the odds of three bookmakers need to be compared. This leads to the second hypothesis:

H_2:

Arbitrage opportunities last for longer when search costs are higher.

Tuckman and Vila (1992) develop a theoretical model that highlights the link between holding costs and arbitrage activity and test their model using Treasury Bill data. They find arbitrage opportunities that involve higher holding costs persist for longer. More recently,

Abreu and Brunnermeier (2002) develop a model where agents delay acting on arbitrage opportunities due to the holding costs they incur. To take advantage of opportunities as they arise an arbitrageur has to have money in an account with the bookmaker prior to the opportunity being created. Bookmakers do not pay interest on funds deposited with them so an arbitrageur incurs a holding cost (in the form of foregone interest) on money deposited with each bookmaker. There is a large variation in the number of arbitrage opportunities the fifty bookmakers studied by this paper are involved in so there is also a lot of variation in the holding costs an arbitrageur incurs by leaving money on deposit with each bookmaker. Bookmakers who are involved in few (many) arbitrage opportunities have relatively higher (lower) holding costs. Thus the third hypothesis:

H_3:

Arbitrage opportunities last for longer when holding costs are higher.

IV. Results

In this section I present the results and discuss whether the implications they have for each hypothesis in Section III. Part A contains the core results while Part B contains results for the robustness checks.

A. Core Results

The mean and median *REVENUE* for each two and three-outcome event are presented in Table II. Panel A contains the results for the entire 2003-2005 period, while Panels B-D

contains results for each year separately. I am interested in determining whether there are differences between the mean and median *REVENUE* for two and three-outcome events so I calculate a standard difference of mean *t*-test and the Wilcoxon (1945) signed ranked test.

Turning to the results for all two and three bookmaker opportunities (Panel A), it is clear two bookmaker opportunities are far more common (17,468 versus 2,414). The mean *REVENUE* is lower for two-outcome (1.98 percent) than three-outcome (2.18 percent) opportunities, and this difference is highly statistically significant (p-value <0.0001). The fact two-outcome opportunities tend to generate less profitable arbitrage opportunities is also evident when the medians are analysed. The median *REVENUE* is considerably lower for two-outcome opportunities (1.49 percent versus 1.72 percent) and this difference is also highly statistically significant (p-value <0.0001).

The results displayed in Panels B-D indicate this finding is robust to each sub-period. Mean and median *REVENUEs* are lower for two-outcome opportunities in both 2003 and 2005. In 2004 the mean *REVENUE* is higher for two-outcome opportunities but this difference is not statistically significant. However, the 2004 median difference follows the trend observed in all other periods and is statistically significant. The results presented in Table II are supportive of Hypothesis One. There is strong evidence those sporting events with three possible outcomes rather than two (those with greater uncertainty about the likely outcome) generate larger arbitrage opportunities.

[Insert Table II About Here]

I now consider the results generated to test Hypothesis Two, which is "arbitrage opportunities last for longer when search costs are higher." Two-outcome event arbitrage opportunities are easier to locate than their three-outcome counterparts because they require an arbitrageur to only compare two sets of odds rather than three. This suggests two arbitrage opportunities should be found and removed more quickly.

The results displayed in Table III suggest this is indeed the case. The mean duration for all two-outcome arbitrage opportunities is 50.93 minutes compared to 101.14 minutes for their three-outcome counterparts. This difference is high statistically significant (p-value <0.0001). The medians are considerable lower, but the same trend is however evident in these results. The median duration for two-outcome events (13.83 minutes) is statistically significantly less than the median outcome for three-outcome events (32.49 minutes). This result is robust across each of the three sub-periods. Median and mean *DURATION*s are consistently lower for two-outcome sporting events. This is consistent with Hypothesis Two. When search costs are lower arbitrage opportunities are removed more quickly.

The other notable result apparent in Panels B-D is the decline in *DURATION*s between 2003 and 2005. Median *DURATION*s have declined from 20.60 minutes and 51.00 minutes for two and three-outcome events respectively in 2003 to 9.88 minutes and 21.38 minutes for two and three-outcome events respectively in 2005. This indicates search costs have declined over this period as arbitrageurs develop new technologies to take advantage of these opportunities more efficiently.

[Insert Table III About Here]

Fifty bookmakers are consistently tracked by zero-risk-arbitrage.com over the three-year period of this study. However, there is wide variation in the number of times one of these bookmakers posted an odd that is sufficiently different from the rest of the market to create an arbitrage opportunity. One bookmaker posted only one such odd, while another posted 784 such odds. To determine if there are different characteristics across the odds quoted by bookmakers involved in a few and many arbitrage opportunities in this paper the fifty bookmakers are sorted according to the number of arbitrage opportunities they created. The top and bottom twenty percent (ten) bookmakers are then identified. Table IV contains the characteristics of the arbitrage opportunities created by these bookmakers. *Low* refers to the bottom ten bookmakers while *high* refers to the top ten.

Differences in the length of time arbitrage opportunities last at low and high bookmakers can be used to give an insight into the relationship between holding costs and arbitrage. Arbitrageurs do not have time to transfer money to a bookmaker after they identify an arbitrage opportunity. Rather, they have to have money sitting in accounts at bookmakers in anticipation of arbitrage opportunities being created. Since bookmakers do not pay interest on deposited funds the holding costs, represented by the interest lost, is related to the frequency with which money at a bookmaker can be applied to an arbitrage opportunity. This leads to Hypothesis Three which states "arbitrage profits last for longer when holding costs are higher."

The results displayed in Table IV indicate the bottom ten bookmakers create just 201 opportunities while the top ten bookmakers create 4,928 opportunities. The mean *DURATION* for the low number of opportunity bookmakers is 79.05 minutes versus 49.22

minutes for their high number bookmaker counterparts. This difference is statistically significant at the one percent level. A similar trend is evident in the median *DURATION* results. Median *REVENUEs* are 43.83 minutes and 13.29 minutes for low and high number of opportunity bookmakers respectively. This relationship is clear in both the mean and median results for the three-outcome events (Panel C) and in the median results for the two-outcome events (Panel B). Mean *REVENUEs* are higher for low number of opportunity bookmakers in the two-outcome events but this difference is not statistically significant. Overall, I find strong evidence in support of Hypothesis Three. Arbitrage profits do appear to last for longer when holding costs are higher.

[Insert Table IV About Here]

The data do not allow a conclusion to be drawn on the annual percentage *REVENUEs* available to an arbitrageur. Sporting events are typically concluded within one week of the odds first being offered (and the arbitrage opportunity potentially being created). Moreover, bookmakers return money to winning bettors on the day after an event's conclusion. However, a reasonable estimate of annual arbitrage *REVENUEs* is substantially less than the 216% implied by annualizing the average *REVENUE* of 1.51% by 52 weeks. Arbitrageurs are unlikely to have accounts with all 50 bookmakers in the zero-risk-arbitrage.com database or the time required to pursue each opportunity so no one arbitrageur would be in a position to take advantage of all the opportunities documented.

B. Robustness Checks

In this section I show that unique characteristics of sports betting markets are not the driver of the core results. The likelihood of suffering major loss of capital through bookmaker default does not appear to be high. No bookmaker tracked by zero-risk-arbitrage.com defaulted in the period I study. In addition, an arbitrageur's capital is typically distributed among 25 or more bookmakers so the effect of the default of one bookmaker on the arbitrageur's capital is not extreme. Despite the above I investigate whether bookmaker default or the risk of a bookmaker not acting honorably in a transaction is so high that these arbitrage opportunities are fair compensation for the chance of losing one's money. I also consider whether the larger arbitrage profits and longer *DURATION*s are explained by more-risky bookmakers being involved, rather than the search and holding cost arguments I put forward. I investigate this by including three dummy variables for bookmaker credit / reputation risk.

One measure of bookmaker risk is related to the country they are domiciled in. Bookmakers Review, an independent organisation, considers "each country's economical and political stability, the presence of government's control over the betting industry and/or the existence of a gaming authority or commission, the procedures and requirements for a betting company to be licensed, the international reputation as a betting jurisdiction, the ease of information access and the availability of modern communication infrastructures, and classifies each jurisdiction into one of five tiers." (http://www.bookmakersreview.com/Jurisdictions/):

The first dummy variable *CTRY_RISK* is based on the risk attached to bookmakers in different countries by Bookmakers Review. The most distinct increase in risk occurs between

Tiers 2 and Tier 3[4] so *CTRY_RISK* equals one if all bookmakers are from a strongly regulated country (Tiers 1 or 2) and zero if one or all bookmakers in an arbitrage opportunity are from a weakly regulated country (tiers 3-4). Strongly regulated countries include Canada, Australia, Isle of Man. Malta and the U.K. Weakly regulated countries include Costa Rica, Netherlands Antilles, West Indies, Austria, Gibraltar, Ireland, and Cyprus.

A second way of measuring the credit risk inherent in dealing with a bookmaker relates to the level of disclosure of the bookmaker's financial position. Of the 50 bookmakers included in my sample 19 are stock exchange listed public companies. These companies have a considerably higher level of disclosure than their private counterparts. This information can be used by arbitragers to reduce the risk of having money with a bookmaker close to collapsing so the second dummy variable to proxy for risk *PUBLIC* equals one if all bookmakers in an arbitrage opportunity are public companies and zero if one or all are not.

Bookmakers can choose to be members of the Independent Betting Arbitration Service (IBAS), which is an authoritative, totally independent third party offering adjudication for customers who have an unresolved betting dispute with their bookmaker. Bookmakers who register with IBAS undertake to be bound by the rulings of its arbitrators. Failure to comply with an IBAS ruling would mean that a bookmaker had not honoured his obligation under the scheme - which would result in the bookmaker being publicly de-registered from the service. Being an IBAS member is therefore a signal that a bookmaker is likely to act honourable with clients. The third variable I introduce as a risk proxy, *OTHER_REG*, equals one if all bookmakers are members of IBAS and zero if one or all bookmakers are not members.

It is also possible that some "investor biases" are prevalent in betting markets to a larger degree than they are in stock markets and that these are driving the results. The most plausible bias is a "Home Country Bias" where individuals bet on their teams from their own country even when it is not rational to do so based on the odds on offer and the probability of their team winning. Kang and Stulz (1997) find that stock market investors favour investments in their own country. This could be explained to the Heath and Tversky's (1991) competence hypothesis which suggests that individuals prefer to operate in an environment where they feel knowledgeable and competent rather than one where they feel ignorant. This same hypothesis could be expected to apply to sports betting markets, and it is reasonable to assume that it might be exacerbated by bettors favouring their home country team due to some emotional attachment that would not exist with stocks. I test for a home country bias using a dummy variable *HOME_CTRY* which equals one if all teams / players are from the same country and zero otherwise.

I investigate the influence that these bookmaker characteristic variables have on my results via regression analysis. I introduce a new bookmaker dummy variable, *BM_NO*, which equals one if there are two bookmakers involved in an opportunity and zero if there are three bookmakers are involved. I compute a correlation matrix to ensure this analysis is not being driven by multicolineratiy in the dummy variables. These results, which are displayed in Table V, indicate the largest correlation, which is between *BM_NO* and *HOME_CTRY*, is 0.4665. This suggests that arbitrage opportunities involving two bookmakers are more likely to come from the same country. The next largest correlation, which is between the *PUBLIC* dummy variable and *OTHER_REG*, is 0.2948. This indicates that publicly listed bookmakers are more likely to be members of IBAS. I re-run my regressions omitting one of these

variables (*HOME_CTRY*, *PUBLIC*, and *OTHER_REG*) each time, and verify that none of the reported results have a material affect on the statistical and / or economic significance of the results I report.

Other correlations are very low. There is a weak negative relationship between bookmakers being IBAS members and bookmakers being from countries with strong regulation, and a weak negative relationship between bookmakers being public listed and being from strongly regulated countries. The relationship between the *HOME_CTRY* variable and the credit risk variables is inconsistent.

[Insert Table V About Here]

As mentioned earlier, there is skewness in the *REVENUE* and *DURATION* variables. To ensure that this is not affecting my regression results I follow the approach adopted by Allayannis and Weston (2001) and many other researchers, and take the natural logarithm of these variables before using them in my regression analysis. To see if the earlier conclusions made about H_1 "Larger arbitrage opportunities are created when there is more information uncertainty" stand up to the robustness checked described above I then run the regression specified in equation 4:

$$LN(REVENUE) = \alpha + \beta_1 BM\_NO + \beta_2 CTRY\_RISK + \beta_3 OTHER\_REG + \beta_4 PUBLIC + \beta_5 HOME\_CTRY + \varepsilon \quad (4)$$

To determine if my conclusions made about H_2 "Arbitrage opportunities last for longer when search costs are higher" are robust to the variables described above I run the regression specified in equation 5.

$$LN(DURATION) = \alpha + \beta_1 BM\_NO + \beta_2 CTRY\_RISK + \beta_3 OTHER\_REG + \beta_4 PUBLIC + \beta_5 HOME\_CTRY + \varepsilon \quad (5)$$

Testing the robustness of the third hypothesis requires the creation of another dummy variable, which is based on the whether the bookmaker generating the arbitrage opportunity is involved in many or few other opportunities. *LOW_HIGH* is a dummy variable that equals one if the bookmaker is in the top ten bookmakers (20%) based on the total number of arbitrage opportunities created and zero if there are in the bottom ten (20%). I then test the robustness of the third hypothesis "arbitrage profits last for longer last for longer when holding costs are higher" with the regression specified in equation 6.

$$LN(DURATION) = \alpha + \beta_1 LOW\_HIGH + \beta_2 CTRY\_RISK + \beta_3 OTHER\_REG + \beta_4 PUBLIC \beta_5 HOME\_CTRY + \varepsilon \quad (6)$$

The robustness test results presented in Table VI indicate that the core results are robust to these new variables. The results displayed in Panel A indicate that more profitable arbitrage opportunities are created when information uncertainty is higher. The *BM_NO* coefficient is negative and statistically significant, which indicates *REVENUE* is lower when there are two bookmakers even when the control variables are included.

All three credit risk variables are positive and statistically significant. This suggests that *REVENUE* is higher for arbitrage opportunities involving bookmakers from strongly regulated countries, bookmakers who are members of IBAS, and bookmakers who are publicly listed. This indicates that credit risk is not an issue. *REVENUE* is, on average, higher when credit risk is lower. The coefficient of *HOME_CTRY* is negative and statistically significant. This suggests that arbitrage opportunities involving players / teams from the same country generate lower *REVENUE* than those involving players / teams from different countries. A home country betting bias does appear to exist.

The results displayed in Panel B relate to the earlier conclusion that arbitrage opportunities with higher search costs last for longer. This also withstands the inclusion of control variables. Arbitrage opportunities involving two bookmakers have lower durations on average and this results is statistically significant at the 1% level. The relationship between *DURATION* and bookmaker credit risk is mixed. Arbitrage opportunities involving bookmakers from strong regulation countries last longer, while arbitrage opportunities involving listed bookmakers and those belonging to IBAS are removed more quickly. Arbitrage opportunities involving teams from the same country last longer.

The robustness of the conclusion that arbitrage opportunities with higher holding costs last for longer is tested in Panel C. The *HIGH_LOW* variable is negative and statistically significant at the 1% level indicating that durations are shorter for bookmakers involved in a high number of opportunities. The conclusion about holding costs being an important determinant of the persistence of arbitrage opportunities is therefore robust. There is some evidence of a negative risk premium in results, with *CTRY_RISK* and *OTHER_REG* both having positive,

statistically significant coefficients. This suggests that durations of the *HIGH_LOW* subset tend to be higher for bookmakers who are better regulated.

A key difference between internet sports betting markets and financial markets is the lack of depth in sports betting markets. Internet bookmakers have bet maximums which preclude placing very large bets on any one contest. These maximums are typically equivalent to US$2,000. While this is obviously quite different to financial markets, I propose that this does not affect my conclusions about the importance of search and holding costs in explaining persistent mis-pricing. Firstly, maximum bet sizes are an important element of bookmaker marketing so they tend to be very similar across all bookmakers in my sample. Secondly, my key conclusions are based on differences in *REVENUE* and *DURATION* between two and three outcome events, which are both offered by all bookmakers in my sample.

While it is not possible to directly test, I propose that the traditional explanations for arbitrage opportunities are unlikely to be driving the opportunities in these data. Fundamental risk, as described by Barberis and Thaler (2002), is non-existent in sports betting arbitrage as the odds offered by different bookmakers are perfect substitutes for each other. The short time frame between placing an arbitrage and the sports event taking place means that noise trader risk (risk that the mis-pricing being exploited by the arbitrageur worsens in the short run, forcing arbitrageurs to liquidate their positions early resulting in losses) is also not a big factor in sports betting arbitrage.

The fact sport betting arbitrage opportunities are removed with a median time of 15.4 minutes rather than persisting for the entire duration leading up the game commencement suggests

transactions costs are not the explanation. Anyone wishing to pursue sports betting arbitrage needs to open accounts with a range of bookmakers (zero-risk-arbitrage.com advise that 25 is a good number). Money can be transferred to and from bookmakers using credit cards and online banking services, such as NETeller. Credit card deposits incur a fee of between 1% and 3% from bookmakers. Deposits via NETeller incur no fees. Bookmakers typically offer a deposit bonus of up to 10% of the funds deposited, so an arbitrageur's account generally has a higher balance (after the deposit fees have been subtracted and deposit bonus has been added) than the amount they deposited. Bookmakers generally offer accounts denominated in a wide range of currencies so it is possible to remove currency risk.

V. Conclusions

This paper uses a unique data set from the internet sports betting market to document some of the first empirical evidence on the importance of search and holding costs in explaining persistent mis-pricing. Sports betting markets share many important features with stock markets (e.g. large volume and wide availability of information) but have the advantage of numerous price (or odd) quotations on the same event by different decentralised market makers (bookmakers), diverse price quotations generating arbitrage opportunities. Fundamental risk, noise trader risk and implementation costs have been shown to be pervasive influences of arbitrage activity in financial markets, which makes it difficult to investigate the independent influences of other arbitrage determinants. In the internet sports betting market arbitrage opportunities are easily identified and valued and based on odds are perfect substitutes (i.e. no fundamental risk), the odds having a short life and a well defined

termination point (i.e. no noise trader risk), and these arbitrage transaction being executed with minimal cost via the internet (i.e. low transactions costs).

Consistent with the propositions of several theoretical models, I find strong evidence arbitrage opportunities with higher search costs last for longer. The arbitrage opportunities I study span both two- and three-outcome sports. Arbitrage opportunities on three-outcome sports are more difficult to find because they involve comparing three sets of odds. I find these opportunities last for considerably longer than two-outcome sports opportunities.

Some bookmakers are involved in numerous arbitrage opportunities while others are involved in few opportunities. Arbitrageurs must have money on deposit with bookmakers before an opportunity is identified to be in a position to take advantage of it. Such deposits earn no interest so arbitrageurs incur a holding cost in the form of lost interest. I show arbitrage opportunities last for longer at bookmakers who generate fewer opportunities which suggests that arbitrage opportunities with higher holding costs last for longer. I show that my conclusions are not being driven by unique characteristics of sports betting markets (e.g. credit risk and home team betting biases). My key results are robust to inclusion of a number of variables that act as proxies for these effects.

References

Allayannis, George and James P. Weston, 2001, The Use of Foreign Currency Derivatives and Firm Market Value, The Review of Financial Studies 14(1), 243-276.

Abreu, Dilip and Markus K. Brunnermeier, 2002, Synchronization Risk and Delayed Arbitrage, *Journal of Financial Economics* 66, 341-360.

Avery, Christopher, and Judith Chevalier, 1999, Identifying investor sentiment from price paths: The case of football betting, *Journal of Business* 72, 607–36.

Barberis, Nicholas. and Richard Thaler, 2002, A survey of behavioural finance, Handbook of the economics of finance, edited by George Constantinides, Milt Harris and Rene Stulz.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–43.

Cao, Charles Q., Ghysels, Eric and Frank Hatheway, 2000, Price Discovery Without Trading: Evidence from the Nasdaq Preopening, *Journal of Finance* 55(3), 1339-1365.

Durham, Gregory R., Hertzel, Michael J., and J. Spenser Martin, 2005, The Market Impact of Trends and Sequences in Performance: New Evidence, *Journal of Finance* 60(5), 2551-2569.

Gandar, John M., Dare, William H., Brown, Craig R. and Zuber, Richard A., 1998, Informed Traders and Price Variations in the Betting Market for Professional Basketball Games, *Journal of Finance* 53(1), 385-386.

Heath, Chip and Amos. Tversky, 1991, Preference and Belief: Ambiguity and Competence in Choice under Uncertainty, *Journal of Risk and Uncertainty* 4, 5-28.

Hortacsu, Ali and Chad Syverson, 2004, Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds, *Quarterly Journal of Economics* 119(2), 403-456.

Kang, Jun-Koo and Rene M. Stulz, 1997, Why Is There a Home Bias? An Analysis of Foreign Portfolio Equity Ownership in Japan, *Journal of Financial Economics* 43, 3-28.

Pontiff, Jeffrey, 2006, Costly arbitrage and the myth of idiosyncratic risk, *Journal of Accounting and Economics* 42, 35-52.

Rose, I.Nelson., 1999, *The Law of Internet Gambling.* I Nelson Rose.

Salop, Steven C. and Joseph E. Stiglitz, 1977, Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion, *Review of Economic Studies* 44, 493-510.

Sinclair, Sebastian, 1998, *Casino Gambling and the Internet*, Christiansen/Cummings Associates, New York

Stahl, Dale O., 1989, Oligopolistic Pricing with Sequential Consumer Search, *American Economic Review* 79(4), 700-712.

Strumpf, Koleman S., 2003, *Illegal sports bookmakers*, Unpublished Working Paper.

Tuckman, Bruce., and Jean Luc. Vila, 1992, Arbitrage with Holding Costs: A Utility-Based Approach, *Journal of Finance* 47(4), 1283-1302.

Varian, Hal R., 1980, A Model of Sales, *American Economic Review* 70(4), 651-659.

Wilcoxon, Frank, 1945, Individual Comparisons by Ranking Methods, *Biometrics*, 1, 80-83.

Zhao, Ying, 2006, Price Dispersion in the Grocery Market, *Journal of Business*, 79(3), 1175-1192

Footnotes

1. Pontiff (2006) uses the term "idiosyncratic risk" to describe the risk that an asset price moves for some reason specific to that asset. Pontiff (2006) refers to this non-diversifable risk as being a component of holding costs, but I follow the more standard treatment in the literature (e.g. Barberis and Thaler 2002) and refer to this risk as "fundamental risk".
2. One limitation of the data set is the lack of depth data. As a result, I am unable to quantify exactly how much an arbitrageur could make by pursuing these strategies. It is reasonable to expect there to be considerably less depth than most large financial markets, but this does not detract from my conclusions on the importance of search and holding costs to persistent arbitrage opportunities, as I explain in Section V.
3. All results were re-run for the common non-missing months of each year (April – August) and no major differences were found so I present the entire year results in this paper.
4. Tier 2 is "assigned to jurisdictions offering a good combination between a stable economical and political environment and a regulatory framework setting out high standards to protect bettors' interests. Tier 3 is assigned to the jurisdictions that, even in presence of a stable economical and political environment, are lacking one or more of the following characteristics: proper governments' controls, strong gaming authorities or commissions, stringent licensing requirements, therefore transferring too many risks to bettors."(http://www.bookmakersreview.com/Jurisdictions/).

Table I
Summary Statistics

REVENUE is the gross payback arbitrages receive for conducting an arbitrage transaction. *DURATION* is the length of time each arbitrage transaction lasts for before being totally removed.

	Number	Mean	Std Dev	Min	LQ	Median	UQ	Max
REVENUE	19882	0.0203	0.0149	0.0091	0.0114	0.0151	0.0229	0.1110
DURATION	19882	59.424	127.219	0.017	6.333	15.358	50.046	1177.667

Table II
Arbitrage Revenues by Number of Event Outcomes

The results in this table compare *REVENUE* by the number of event outcomes. Two refers to those sporting events with two possible outcomes. Three refers to those sporting events with three possible outcomes. The p-value in column D is a generated from the standard difference of means parametric *t*-test. The p-value in column F is a generated from the Wilcoxon (1945) non-parametric *t*-test.

	Number	Mean	p-Value	Median	p-Value
		Panel A: All			
Two	17468	0.0198		0.0149	
Three	2414	0.0218		0.0172	
Difference		-0.0020	<0.0001	-0.0023	<0.0001
		Panel B: 2003			
Two	4215	0.0227		0.0156	
Three	602	0.0255		0.0181	
Difference		-0.0028	<0.0001	-0.0025	<0.0001
		Panel C: 2004			
Two	3990	0.0184		0.0151	
Three	351	0.0177		0.0162	
Difference		0.0007	0.7533	-0.0011	0.0184
		Panel D: 2005			
Two	9263	0.0187		0.0142	
Three	1461	0.0206		0.0173	
Difference		-0.0019	<0.0001	-0.0031	<0.0001

Table III
Arbitrage Durations by Number of Event Outcomes

The results in this table compare *DURATION* by the number of event outcomes. Two refers to those sporting events with two possible outcomes. Three refers to those sporting events with three possible outcomes. The p-value in column D is a generated from the standard difference of means parametric *t*-test. The p-value in column F is a generated from the Wilcoxon (1945) non-parametric *t*-test.

	Number	Mean	p-Value	Median	p-Value
		Panel A: All			
Two	17468	50.93		13.83	
Three	2414	101.14		32.49	
Difference		-50.21	<0.0001	-18.67	<0.0001
		Panel B: 2003			
Two	4215	69.44		20.60	
Three	602	140.87		51.00	
Difference		-71.43	<0.0001	-30.40	<0.0001
		Panel C: 2004			
Two	3990	72.76		26.88	
Three	351	119.36		76.15	
Difference		-46.60	<0.0001	-49.28	<0.0001
		Panel D: 2005			
Two	9263	30.85		9.88	
Three	1461	72.32		21.38	
Difference		-41.47	<0.0001	-11.50	<0.0001

Table IV
Arbitrage Durations by Number of Opportunities Created by Bookmakers

The results in this table compare *DURATION* by the number of arbitrage opportunities created by a bookmaker. Low refers to the twenty percent of bookmakers (ten) who are involved in the fewest arbitrage opportunities. High refers to the twenty percent of bookmakers (ten) who are involved in the most arbitrage opportunities. The p-value in column D is a generated from the standard difference of means parametric *t*-test. The p-value in column F is a generated from the Wilcoxon (1945) non-parametric *t*-test.

	Number	Mean	p-Value	Median	p-Value
Panel A: All Events					
Low	201	79.05		43.83	
High	4928	49.22		13.29	
Difference		29.83	<0.0001	30.54	<0.0001
Panel B: Two-outcome Events					
Low	107	37.57		20.93	
High	4389	43.34		12.08	
Difference		-5.77	0.2325	8.85	0.0005
Panel C: Three-outcome Events					
Low	94	109.69		122.35	
High	539	86.01		28.72	
Difference		23.68	0.0276	93.63	<0.0001

Table V
Dummy Variable Correlations

This table contains the correlations between the dummy variables. *BM_NO* is a dummy variable that equals one if there are two bookmakers involved in an opportunity and zero if there are three bookmakers are involved. *CTRY_RISK* is a dummy variable that equals one if all bookmakers are from a strongly regulated country and zero if one or all bookmakers are from a weakly regulated country. *PUBLIC* is a dummy variable that equals one if all bookmakers are public companies and zero if one or all are not. *OTHER_REG* is a dummy variable that equals one if all bookmakers are members of IBAS and zero if one or all bookmakers are not members. *HOME_CTRY* is a dummy variable that equals one if all teams / players are from the same country and zero otherwise.

	BM_NO	*CTRY_RISK*	*PUBLIC*	*OTHER_REG*	*HOME_CTRY*
BM_NO	1				
CTRY_RISK	0.0060	1			
PUBLIC	-0.0220	0.0134	1		
OTHER_REG	0.0366	-0.0135	0.2948	1	
HOME_CTRY	0.4665	0.0295	-0.0963	-0.0078	1

Table VI
Robustness Checks

This table contains the regression results I employ to determine the robustness of my results. *REVENUE* is the payback arbitrageurs receive per transaction. *DURATION* is the number of minutes it takes for an arbitrage opportunity to be removed. *BM_NO* is a dummy variable that equals one if there are two bookmakers involved in an opportunity and zero if there are three bookmakers are involved. *CTRY_RISK* is a dummy variable that equals one if all bookmakers are from a strongly regulated country and zero if one or all bookmakers are from a weakly regulated country. *PUBLIC* is a dummy variable that equals one if all bookmakers are public companies and zero if one or all are not. *OTHER_REG* is a dummy variable that equals one if all bookmakers are members of IBAS and zero if one or all bookmakers are not members. *HOME_CTRY* is a dummy variable that equals one if all teams / players are from the same country and zero otherwise.

	α	β_1	β_2	β_3	β_4	β_5	Adj R^2
Panel A: LN(REVENUE) = α + β₁BM_NO + β₂CTRY_RISK + β₃OTHER_REG + β₄PUBLIC + β₅ HOME_CTRY + ε							
Coefficient	-3.9443	-0.0902	0.0477	0.1055	0.0359	-0.0786	0.0140
t-value	-246.62***	-7.03***	5.26***	9.50***	3.66***	-4.39***	
Panel B: LN(DURATION) = α + β₁BM_NO + β₂CTRY_RISK + β₃OTHER_REG + β₄PUBLIC + β₅ HOME_CTRY + ε							
Coefficient	3.592	-0.8579	0.1979	-0.1837	-0.0845	0.1130	0.0409
t-value	82.41***	-24.58***	8.01***	-6.08***	-3.18**	2.32**	
Panel C: LN(DURATION) = α + β₁HIGH_LOW + β₂CTRY_RISK + β₃OTHER_REG + β₄PUBLIC + β₅ HOME_CTRY + ε							
Coefficient	3.0790	-0.0398	0.3356	0.0240	-0.0053	-0.2709	0.0095
t-value	33.38***	-2.82***	6.18***	0.35	-0.12	-2.93***	