Author's Accepted Manuscript

www.elsevier.com/locate/jedc

Relative pricing of binary options in live soccer betting markets

Vera Hofer, Johannes Leitner

PII: S0165-1889(16)30211-1
DOI: http://dx.doi.org/10.1016/j.jedc.2016.12.007
Reference: DYNCON3381

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

Received date: 15 August 2016
Revised date: 21 December 2016
Accepted date: 23 December 2016

Cite this article as: Vera Hofer and Johannes Leitner, Relative pricing of binary options in live soccer betting markets, *Journal of Economic Dynamics and Control,* http://dx.doi.org/10.1016/j.jedc.2016.12.007

Relative pricing of binary options in live soccer betting markets

Vera Hofer, Johannes Leitner[a]

[a]*Department of Statistics and Operations Research, University of Graz, Universitaetsstrasse 15/E3, 8010 Graz, Austria,*

Abstract

Live soccer betting markets differ from other hinary options markets in that all fundamental information is observable, the options mature in less than two hours and the markets are highly liquid. This study presents a new method for the identification of hidden information in market prices. The method is based on two independent Poisson distributions and on a numerical algorithm for the aggregation of all market price information into one rational number. The method is applied to an empirical dataset of real time market prices in 29,413 soccer games. The results indicate that the method selects the most profitable markets and allows for a significant improvement in average investment returns.

Keywords:
Binary options, sports betting, gambling, sports betting analytics

1. Introduction

Sports bets are cash-or-nothing binary options whose payout depends on the results of sporting events. The total annual volume of the sports betting market is estimated to be one trillion US dollars, of which 70% is soccer betting (Keogh and Rose, 2013). Sports bets are traded before the sporting event takes place or during the course of the event. The latter are called live, "in running" or "real-time" bets. Live betting makes up 70% of all trading volume (Pantheon-Sorbonne and the International Centre for Sport Security, 2014). Information technologies have changed the sports betting market substantially over the past decade and have served to make live betting much more relevant. Traces of this evolution can be found in the literature. For instance, about a decade ago Levitt (2004) and Kuypers (2000) stated that once bookmakers announce market prices they hardly adjust them at all. Nowadays, many bookmakers and betting exchanges provide datafeeds of their market prices via their own application programming interfaces. Market participants have easy internet-based access to automated odds comparison services that publish highest odds and arbitrage opportunities for games across hundreds of providers. Companies such as Betradar supply information on events (red cards, fouls, etc.) directly from the soccer stadium to bookmakers within seconds. This information is published on the bookmakers' websites and prices adapt immediately to crucial events (see Croxson and Reade (2014)). Sports betting markets are affected by market inefficiencies. The

Email address: vera.hofer@uni-graz.at (Vera Hofer, Johannes Leitner)

most prominent inefficiency is the favorite longshot bias (Cain et al., 2003, Hwang and Kim, 2015) that refers to mispricings of unlikely events.

Sports bets are different from binary options traded on traditional financial markets in several aspects. Firstly, the final state of the underlying of a sports bet is known. This is not the case for traditional binary options where the underlying can theoretically have an infinite lifespan. Secondly, unlike traditional options live bets are very short-lived. A soccer game lasts for about 90 minutes and a bet can mature in a few minutes or even seconds. Thirdly, in live sports markets all price-relevant information, e.g. scored goals, injuries, red cards etc. is publicly available simultaneously to all market participants.

These unique characteristics of the options traded on betting markets allow powerful tests to be made of the market participants' ability to process the available information efficiently. Despite this fact most of the literature analyzes pre-game (non-live) market price data (betting odds). Franck et al. (2013) show that there are inter-market arbitrage opportunities when the market prices of bookmakers and sports betting exchanges are combined. Only pre-game markets are used and the dataset dates back to the years 2004 to 2006. Vlastakis et al. (2009) report that naive trading strategies generate abnormal returns in a pre-game dataset from 2002 to 2004. Vergin and Scriabin (1978) developed strategies for betting on National Football League games based on bookmakers' point spreads. They used 1969-1974 data to investigate the performance of the strategies. Using 1975-1981 data, Tryfos et al. (1984) also investigated different betting strategies. They suggested measuring the performance of a strategy in terms of profitability instead of forecasting accuracy.

Soccer betting markets offer a variety of bets on the final results of a game, on the final results of the first half, on the number of red cards or on the number of corners. The present study focuses on the analysis of the market prices[1] for live bets on final game results. With respect to the trading volume, the most important live betting markets for final results are Asian handicap (AHC) and Over Under (OU) because these markets define payouts for the bets that only depend on future goals.[2] A handicap $h \in \mathbb{Q}$, specific for each AHC market, is added to the number of goals scored by the home team until the end of the game. If this value exceeds the number of additional goals scored by the away team, the bet on the home team has won. For each Over Under market a specific total goal target line $l \in \mathbb{Q}^+$ is defined. If the target is exceeded, the Over bet wins. If the target is not reached, the Under bet wins.[3]

[1]Conventionally, the term 'odds' is used for the payouts that a bettor receives in the case of a win. In this study the term 'market price' is used instead of odds to indicate that the payouts result from the trading activity of market participants similar to prices on the stock exchange.

[2]Most bookmakers (especially those operating in Asia) do not publish data on their turnovers. It is estimated that the Asian Handicap market represents up to 90% of the total trading volume (Hawkins, 2014).

[3]Besides the Asian handicap and the Over Under markets several other live betting markets exist. The most relevant are 1X2 and correct score markets. The 1X2 market defines payouts for three possible outcomes, a win of the home team (1), a win of the away team (2) or a draw (X). Correct score markets provide the opportunity to bet on a specific final result of a game, e.g. 1:4 or 3:2. 1X2 and correct score markets have not been considered in this study. 1X2 market do not have significant trading volume at specific scores. In correct score markets the probabilities of certain events can easily be derived, but the bets entail high bookmaker fees which cannot be calculated exactly. Both types of markets are inappropriate for our purposes.

As an example, Table 1 presents market data during a soccer game with four Asian handicap markets and four Over Under markets. The market prices give information on the possible payouts in the winning case. Assume that the current score of the game is 0:0 and the game ends 1:0. A bet of \$ 1 on the home team in the Asian handicap market with handicap $h = -0.5$ will receive a payout of \$ 2.316 because the handicap added to the number of goals scored by the home team exceeds the number of goals scored by the away team, i.e. $1 - 0.5 > 0$. If the current score was 1:0 the bet would be lost because no additional goals were scored until the end of the game by either team and $0 - 0.5 < 0$. A \$ 1 bet on Under in an Over Under market with line 1.5 will yield a return of \$ 2.429 because the total number of goals is smaller than 1.5. Although goals can only be integers handicaps and lines are not restricted to integers. The reason for this will be explained in the subsequent sections.

Table 1: Sample data from live betting markets

Asian handicap			Over Under		
	Market prices			Market prices	
Handicap	Home	Away	Line	Over	Under
-0.75	2.724	1.500	1.50	1.600	2.429
-0.50	2.316	1.700	1.75	1.800	2.170
-0.25	1.940	2.000	2.00	2.149	1.800
0.00	1.570	2.530	2.25	2.429	1.600

The market price information in Table 1 is comparable to an option ladder of market prices for options traded on traditional financial markets. The handicap h and the line l for sports bets are analogous to an option strike price. Smaller handicaps relate to options that are more out of the money. The investment decision of a rational participant in any options market is based on risk adjusted returns. When several markets are available, the probabilities of final scores implied by the market prices need to be compared and the risk of the bets needs to be quantified in order to identify the most attractive bet.

A new method for solving this nontrivial problem is introduced in this paper. The method is called *Implied Symmetry*. It is based solely on publicly available market price data. It applies independent Poisson distributions with the parameters λ_1 and λ_2 modeling the team strengths. The method derives the information on the probabilities of final scores implied by the market prices. The two parameters λ_1 and λ_2 are used for the comparison of implied probabilities and for identifying relatively cheap and expensive bets. This method provides the foundation for optimal trading decisions.

In section 2 below the properties of Over Under and Asian handicap markets are presented in detail. The mechanisms of sports betting markets are less well-known than those of traditional financial markets. In order to help understand the motives behind the research problems, the characteristics of the markets and the payouts are presented. In section 3 it will be shown that market prices of sports bets do not allow the derivation of implied probabilities for final scores. This is the motivation behind the Implied Symmetry method that is presented in section 4. Based on this method a betting strategy is defined in section 5 that is supposed to maximize expected returns irrespective of the variance of the returns. A second strategy is proposed that minimizes variance of the returns irrespective of their expected value. Section 6 shows the ability of the Implied Symmetry method to identify differences in probability valuations between markets. Market prices do not indicate these differences directly. This hidden information is revealed by the Implied Symmetry method. In section 7 the Implied Symmetry method is applied to a dataset of market prices from almost 30,000 games. The results are completely in line with the betting strategies formulated in section 5 and show statistically significant higher returns and lower variances. In section 8 the results are summarized.

2. Mathematical Properties of Betting Markets

The mathematical model of soccer bets is based on a discrete probability distribution of final scores, $\mathbf{P}_t$, as shown in Table 2. The current score of the home team is denoted g_t^H and the current score of the away team g_t^A at time $t \in [0, T]$. Time T is the total duration of the game. A soccer game generally lasts for 90 minutes and is extended for several minutes by the referee to cover 'injury time'. The final scores are denoted g_T^H and g_T^A with $g_T^H \geq g_t^H \geq 0$ and $g_T^A \geq g_t^A \geq 0$. At time $t \in [0, T]$, the probability matrix $\mathbf{P}_t$ contains the values $p_{ij}^{(t)} \geq 0$ which are the true but unknown probabilities for the final results $g_T^H : g_T^A = i : j$, $i, j \in \mathbb{N}_0$. In order to simplify notation the superscript t is omitted from $p_{ij}^{(t)}$.

		Away goals				
		0	1	2	3	$\cdots$
Home goals	0	p_{00}	p_{01}	p_{02}	p_{03}	$\cdots$
	1	p_{10}	p_{11}	p_{12}	p_{13}	$\cdots$
	2	p_{20}	p_{21}	p_{22}	p_{23}	$\cdots$
	3	p_{30}	p_{31}	p_{32}	p_{33}	$\cdots$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Table 2: Probability matrix $\mathbf{P_t}$ of final scores in a soccer game

In live betting markets $\mathbf{P}_t$ changes during the game. At the beginning of the game $(t = 0)$ $\mathbf{P}_0$ reflects all the relevant information such as home team advantage, recent team performance or injuries of key players. When $t > 0$, $\mathbf{P}_t$ additionally reflects all information that is revealed during the game. This includes player disqualifications, weather conditions, or the unanticipated strength of a team. It also includes the score of the game, since scoring events automatically nullify certain probabilities. If the current score is 1:0, the first line of matrix $\mathbf{P}_t$ contains zeros, i.e. $p_{0j} = 0 \, \forall j$. In general, the probabilities p_{ij} are zero for scores $i : j$ where $i < g_t^H$ or $j < g_t^A$ holds.

Soccer bets are based on an estimate of the unknown true $\mathbf{P}_t$ and define a payout for a specific subset of all possible final scores. Such a subset might be a draw with scores $i : i$ where $i \geq 0$ and with probability $\sum_{i \geq 0} p_{ii}$. The payout is directly related to the probability of the subset. Based on $\mathbf{P}_t$ this relation is described in detail in the sections below for the Asian handicap market and the Over Under market.

2.1. Asian Handicap Markets

Asian handicap markets offer the opportunity to bet on a home team or an away team win. A handicap $h \in \mathbb{Q}$ specific for each market is added to the future goals of the home team. If $g_T^H - g_t^H + h$ exceeds the future goals of the away team, $g_T^A - g_t^A$, the bet on the home team has won. The handicap h is set in such a way that preferably equal probabilities of a home team win and an away team win are achieved and market prices approximately equal 2. This is the reason for the empirical observation that only about three to five liquid handicap markets exist for any game at the same time (see Table 1). The choice of h will depend on the current estimate of the probability matrix $\mathbf{P}_t$.

There are three kinds of handicaps each with a specific payout mechanism: Integer handicaps $h \in \mathbb{Z}$, handicaps $h \in \mathbb{Z} + 0.5$ and linear combinations of these two. For integer handicaps $h \in \mathbb{Z}$ the stake is returned, if $g_T^H - g_t^H + h$ equals $g_T^A - g_t^A$. Such neutral returns are impossible for handicaps of the form $h \in \mathbb{Z} + 0.5$ (for example $h = -0.5$ or $h = 1.5$). Further Asian handicap markets are obtained as a linear combination of the two basic Asian handicap markets. This results in markets with handicaps of the form $0.5^n(\mathbb{Z} + 0.5)$, where $n \in \mathbb{N}$. Examples of this type are $h = -0.125$ or $h = 0.75$. To ease exposition, the Asian handicap market with handicap h is denoted as AHC(h).

Figure 1 illustrates the partition of matrix $\mathbf{P}_t$ for three AHC(0) markets. In the left panel the current score is 0:0, in the middle panel it is 1:1, and in the right panel the current score is 1:0. White cells indicate where given the current score impossible final scores are rated with zero probability. Since in Asian handicap markets only future goals

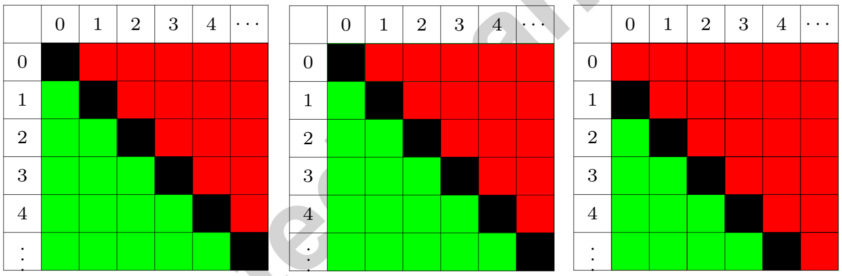


Figure 1: Three Asian handicap markets with handicap 0 at current scores 0:0, 1:1 and 1:0 (home team scores in the rows, away team scores in the columns)

are relevant for the payouts, the partitioning of the probability matrix in the right panel of Figure 1 changes depending on the current score.

2.1.1. Integer Handicaps:

The payout $\Pi_H^{(h)}$ when betting one monetary unit on the home team in a handicap market with integer handicaps at time t is

$$\Pi_H^{(h)} = \begin{cases} q_H^{(h)} & \dots \quad g_T^H - g_t^H + h > g_T^A - g_t^A \\ 1 & \dots \quad g_T^H - g_t^H + h = g_T^A - g_t^A \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

In the case of a draw the stake is returned to the bettor and in the winning case she receives the amount $q_H^{(h)}$. Otherwise the stake is lost. A corresponding definition of payouts applies to the bets on the away team.

The fair prices of bets on these markets are calculated under the assumption that there are no bookmaker fees. In this case the expected payout for a bet on the home team or the away team equals the stake, i.e. in this case one monetary unit.

$$q_H^{(h)} \cdot W_H^{(h)} + 1 \cdot W_D^{(h)} = 1 \tag{2}$$

$$q_A^{(h)} \cdot W_A^{(h)} + 1 \cdot W_D^{(h)} = 1 \tag{3}$$

The calculations require three probabilities. These are the probabilities of a draw $W_D^{(h)}$, of a home win $W_H^{(h)}$ and the probabilities of the away team winning the game, $W_A^{(h)}$. These events are defined such that the handicap is already taken into consideration with respect to the additional number of goals scored between the time t and end of the game at time T. For instance a draw (D) is defined as a result where the handicap added to the additional number of goals by the home team equals the additional number of goals by the away team, i.e. $g_T^H + h - g_t^H = g_T^A - g_t^A$. These three probabilities are defined as:

$$\begin{aligned}
W_D^{(h)} &= P(g_T^H - g_t^H + h = g_T^A - g_t^A) = \begin{cases} \sum_{i=0}^{\infty} p_{i+g_t^H-h,i+g_t^A} & \dots \quad h \leq 0 \\ \sum_{i=0}^{\infty} p_{i+g_t^H,i+g_t^A+h} & \dots \quad h > 0 \end{cases} \\
W_H^{(h)} &= P(g_T^H - g_t^H + h > g_T^A - g_t^A) = \begin{cases} \sum_{i=0}^{\infty} \sum_{j=0}^{i} p_{1+i+g_t^H-h,j+g_t^A} & \dots \quad h \leq 0 \\ \sum_{i=0}^{\infty} \sum_{j=0}^{i+h-1} p_{i+g_t^H,j+g_t^A} & \dots \quad h > 0 \end{cases} \\
W_A^{(h)} &= P(g_T^H - g_t^H + h < g_T^A - g_t^A) = 1 - W_D^{(h)} - W_H^{(h)}
\end{aligned}$$

For instance, the definition of $W_D^{(h)}$ refers to the black diagonals in Figure 1 for the handicap $h = 0$. By rearranging (2) and (3) the fair market prices are given by

$$\begin{aligned}
q_H^{(h)} &= \frac{1 - W_D^{(h)}}{W_H^{(h)}} \\
q_A^{(h)} &= \frac{1 - W_D^{(h)}}{W_A^{(h)}}.
\end{aligned}$$

The probabilities W_D^h, W_H^h and W_A^h and consequently the market prices $q_H^{(h)}$ and $q_A^{(h)}$ depend on t. However, for reasons of simplicity of notation, the time dependence is not made explicit.

2.1.2. Handicaps $h \in \mathbb{Z} + 0.5$:

For all bets with handicaps $h \in \mathbb{Z} + 0.5$ the payouts $\Pi_H^{(h)}$ for betting one monetary unit on the home team are defined as

$$\Pi_H^{(h)} = \begin{cases} q_H^{(h)} & \dots \quad g_T^H - g_t^H + h > g_T^A - g_t^A \\ 0 & \dots \quad \text{otherwise} \end{cases} \tag{4}$$

The definition of the payout for bets on the away team is analogous. In contrast to integer handicaps the payouts are based on only two conditions, i.e. the payouts are only based on two probabilities. In Figure 2, three AHC markets with $h = -0.5$ are illustrated.

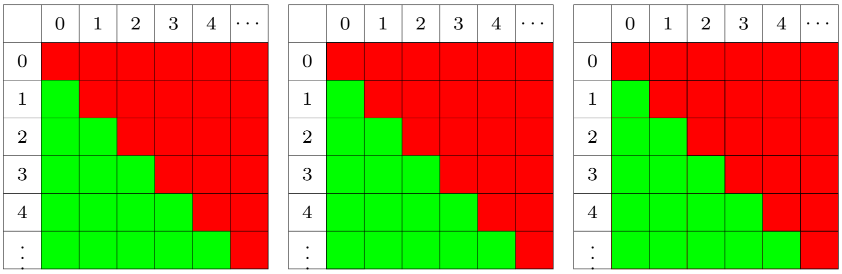


Figure 2: Three Asian handicap markets with handicap -0.5 at current scores 0:0, 1:1 and 1:0 (home team scores in the rows, away team scores in the columns)

In a fair market the expected payouts of the bets equal the stakes. In the case when bets of one monetary unit are placed on the home team or the away team respectively this means

$$\begin{aligned} q_H^{(h)} \cdot W_H^{(h)} &= 1\,, \\ q_A^{(h)} \cdot W_A^{(h)} &= 1\,. \end{aligned}$$

For the calculation of the fair prices two probabilities are used. These probabilities are defined as

$$\begin{aligned} W_H^{(h)} &= P(g_T^H - g_t^H + h > g_T^A - g_t^A) = \sum_{i=1+g_t^H}^{\infty} \sum_{j=g_t^A}^{i-1} p_{ij}\,, \\ W_A^{(h)} &= P(g_T^H - g_t^H + h < g_T^A - g_t^A) = 1 - W_H^{(h)}\,. \end{aligned}$$

The fair prices of an AHC(h) market with $h \in \mathbb{Z} + 0.5$ are given by

$$\begin{aligned} q_H^{(h)} &= \frac{1}{W_H^{(h)}}\,, \\ q_A^{(h)} &= \frac{1}{W_A^{(h)}}\,. \end{aligned}$$

Even though $W_H^{(-0.5)}$ and $W_H^{(0)}$ are defined using the same components p_{ij}, the prices of the bets are not necessarily the same, i.e. $q_H^{(-0.5)} \neq q_H^{(0)}$ because the AHC(0) market neutralises the draw. The buyer of an AHC(-0.5) home bet will lose her stake if the result is a draw, the buyer of an AHC(0) bet will get her stake back. As long as the probabilities of a draw in a football game are not zero, i.e. $W_D^{(0)} > 0$, this has the consequence that $q_H^{(-0.5)} > q_H^{(0)}$. Otherwise arbitrage would be possible.

2.1.3. Other Handicaps:

Markets with handicaps $h \in 0.5^n(\mathbb{Z}+0.5)$, where $n \in \mathbb{N}$ are created by linear combinations of bets of the form $h \in \mathbb{Z}$ and $h \in \mathbb{Z}+0.5$. For example, under the assumptions that all market prices are based on the same values p_{ij} and that prices are fair the $h = -0.25$ market is equivalent to betting 50% of the stake on an $h = -0.5$ market and 50% on an $h = 0$ market. The AHC(-0.25) fair market prices are then derived as

$$\frac{1}{2}\left(q_H^{(0)} \cdot W_H^{(0)} + W_D^{(0)}\right) + \frac{1}{2} q_H^{(-0.5)} \cdot W_H^{(-0.5)} = 1$$

Using $W_H^{(-0.5)} = W_H^{(0)}$ yields

$$q_H^{(-0.25)} \cdot W_H^{(0)} + \frac{1}{2} W_D^{(0)} = 1$$
$$q_H^{(-0.25)} = \frac{1 - \frac{1}{2} W_D^{(0)}}{W_H^{(0)}}$$

Similarly, for the away team bet

$$\frac{1}{2}\left(q_A^{(0)} \cdot W_A^{(0)} + W_D^{(0)}\right) + \frac{1}{2} q_A^{(-0.5)} \left(W_D^{(0)} + W_A^{(0)}\right) = 1$$
$$q_A^{(-0.25)} \left(W_A^{(0)} + \frac{1}{2} W_D^{(0)}\right) = 1 - \frac{1}{2} W_D^{(0)}$$
$$q_A^{(-0.25)} = \frac{1 - \frac{1}{2} W_D^{(0)}}{W_A^{(0)} + \frac{1}{2} W_D^{(0)}}$$

The Asian handicap markets for other handicaps such as 0.625 or -0.825 are defined analogously.

2.2. Over Under Markets

Over Under markets define bets on the total number of goals scored, $g_T^H + g_T^A$. Which team scores is irrelevant. Each market is characterized by a specific non-negative value l, referred to as the line that corresponds to the handicap h in Asian handicap markets. This line $l \in \mathbb{Q}^+$ denotes the total number of goals scored in a game that has to be exceeded such that the Over bet wins. Where this score is not met the Under bet wins. The line not only allows for a variety of bets. It also provides a tool for offering Over and Under bets with preferably equal probabilities. In contrast to lines $l \in \mathbb{N}_0$, a draw is impossible for lines of the form $\mathbb{N} - 0.5$. Analogously to Asian handicap markets, more complex Over Under markets can be derived from linear combinations of the two basic types of Over Under markets. The Over Under market for line l is denoted as OU(l).

2.2.1. Lines $l \in \mathbb{N}$:

Over under markets with lines $l \in \mathbb{N}$ are very similar to AHC markets with integer handicaps. The payout for one monetary unit placed on Over is

$$\Pi_O^{(l)} = \begin{cases} q_O^{(l)} & \dots & g_T^H + g_T^A > l \\ 1 & \dots & g_T^H + g_T^A = l \\ 0 & \dots & \text{otherwise} \end{cases} \tag{5}$$

If the total number of goals exceeds the line l the bettor is paid $q_O^{(l)}$. If the total number of goals scored equals the line l the stake is returned. Otherwise the bet is lost. Figure 3 illustrates the three probabilities of the Over Under market with the line $l = 2$. The matrix of final scores is split by secondary diagonals.

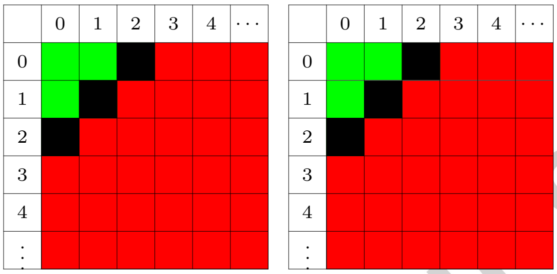


Figure 3: Over Under Markets with line $l = 2$ for the current scores $g_t^H : g_t^A = 0 : 0$ and 1:0 (home team scores in the rows, away team scores in the columns)

The fair market prices for given l are defined as follows:

$$\begin{aligned} q_U^{(l)} &= \frac{1 - V_X^{(l)}}{V_U^{(l)}} \\ q_O^{(l)} &= \frac{1 - V_X^{(l)}}{V_O^{(l)}} \end{aligned}$$

The three probabilities required for the calculation are defined as:

$$\begin{aligned} V_X^{(l)} &= P(g_T^H + g_T^A = l) = \sum_{i=0}^{l} p_{i,l-i} \\ V_U^{(l)} &= P(g_T^H + g_T^A < l) = \sum_{i=0}^{l-1} \sum_{j=0}^{l-1-i} p_{ij} \\ V_O^{(l)} &= P(g_T^H + g_T^A > l) = 1 - V_X^{(l)} - V_U^{(l)} \end{aligned}$$

The probability $V_X^{(l)}$ refers to the black diagonals in Figure 3. It is analogous to the draw in the Asian handicap markets but in order to avoid confusion with the term draw the index X is used. Unlike in the Asian handicap markets the definitions do not depend on the current scores g_t^H and g_t^A. In the situation depicted in the right panel of Figure

3 the probabilities $p_{0j} = 0 \; \forall j \geq 0$. Using these probabilities in the summation does not change the results.

2.2.2. Lines $l \in \mathbb{N} - 0.5$:

Over Under markets with lines $l \in \mathbb{N} - 0.5$ are comparable to AHC markets with handicaps $h \in \mathbb{Z} + 0.5$. The payout for one monetary unit placed on Over is

$$\Pi_O^{(l)} = \begin{cases} q_O^{(l)} & \dots \quad g_T^H + g_T^A > l \\ 0 & \dots \quad \text{otherwise} \end{cases} \tag{6}$$

The fair market prices are

$$\begin{aligned} q_U^{(l)} &= \frac{1}{V_U^{(l)}} \\ q_O^{(l)} &= \frac{1}{V_O^{(l)}} \end{aligned}$$

In Figure 4 the OU(2.5) market is illustrated for three different current scores.

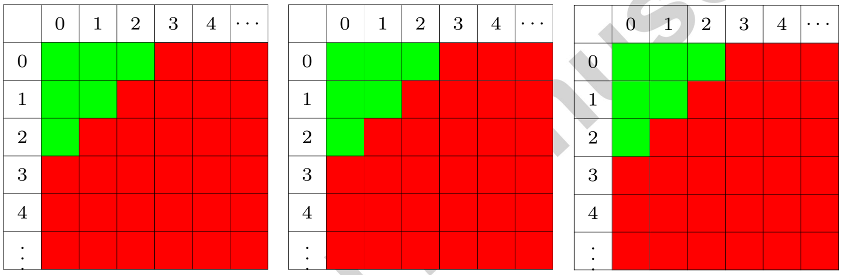


Figure 4: Over Under Markets with line $l = 2.5$ for the current scores $g_t^H : g_t^A = 0:0$, 1:0 and 1:1 (home team scores in the rows, away team scores in the columns)

In these markets only two partitions of the probability matrix exist

$$\begin{aligned} V_U^{(l)} &= \sum_{i=0}^{\lfloor l \rfloor} \sum_{j=0}^{\lfloor l \rfloor - i} p_{ij} = P(g_T^H + g_T^A < l) \\ V_O^{(l)} &= 1 - V_U^{(l)} = P(g_T^H + g_T^A > l)\,. \end{aligned}$$

All other Over Under markets with lines 2.25, 1.75 or 1.125 (or more generally: lines of the form $0.5^n(\mathbb{N} - 0.5)$ where $n \in \mathbb{N}$) are linear combinations of these two markets.

2.3. Extensions to Real Market Conditions

Due to betting exchanges such as Betfair.com the traditional one-sided role allocation of a bookmaker who offers bets and a gambler who accepts bets is no longer valid. On an exchange each market participant can both accept and offer bets and therefore take the role of a bookmaker. Bookmakers themselves use exchanges to hedge their positions and oftentimes avoid taking risky positions but offer prices in a way to balance bets on the outcome of an event in proportions so that they make a profit regardless of which outcome prevails (Cortis, 2015).

The absence of long term arbitrage opportunities justifies the assumption that prices published by various bookmakers and exchanges are very similar and that these prices are the result of the trading activities of thousands of market participants. Each market participant i has an individual estimate $\widehat{\mathbf{P}}_t^i$ of the probabilities of the final scores. Their collective trading activities on each market m_k, $k = 1, 2, \ldots$ influence the prices and thus represent an estimate $\widehat{\mathbf{P}}_t^m$ of the true but unknown probabilities. Not all markets need to imply the same probabilities, i.e. in general $\widehat{\mathbf{P}}_t^{m_1} \neq \widehat{\mathbf{P}}_t^{m_2}$. It will be assumed below that market particpants exploit the differences between bookmakers or exchanges and always choose the maximum of all available prices.

The definitions of fair market prices as described in the sections above are based on the assumption that $E(\Pi) = 1$.[4] Under real market conditions market participants are not offered the fair prices $q_H^{(h)}$ and $q_A^{(h)}$, or $q_O^{(l)}$ and $q_U^{(l)}$ but these are reduced by a transaction fee called 'vigorish', v, or 'juice' that usually lies between 2% and 10%. Bookmakers only publish these unfair prices $\widetilde{q}$. On betting exchanges prices are usually higher than the prices offered by bookmakers, but winners needs to pay 2% to 5% of their turnover to the betting exchange. The effective prices, i.e. implied returns, on exchanges can be considered equal to the bookmaker prices $\widetilde{q}$. The observed unfair market prices have to be adjusted for the vigorish for the analysis of the market assessments described in the following sections. These are called adjusted prices subsequently.

The unfair AHC(h) market prices offered by a bookmaker or betting exchange j are $\widetilde{q}_{j,H}^{(h)} = q_{j,H}^{(h)}\,(1 - v_j)$ and $\widetilde{q}_{j,A}^{(h)} = q_{j,A}^{(h)}\,(1 - v_j)$. Here $q_{j,H}^{(h)}$ and $q_{j,A}^{(h)}$ denote the fair prices for the home team and away team bets, respectively.[5] These fair prices are not $q_H^{(h)}$ and $q_A^{(h)}$ since these are based on the true and unknown probabilities p_{ij}. Instead, each market at each bookmaker is based on a different assessment $\widehat{\mathbf{P}}_t$ of $\mathbf{P}_t$ caused by trading activities. These prices imply

$$(1 - v_j) = \frac{1}{\frac{1}{\widetilde{q}_{j,H}^{(h)}} + \frac{1}{\widetilde{q}_{j,A}^{(h)}}}$$

Rational gamblers are assumed to choose the highest available prices for an AHC(h) bet, $\max_j \widetilde{q}_{j,H}^{(h)}$ and $\max_j \widetilde{q}_{j,A}^{(h)}$. Arbitrage opportunities exist if the following inequality holds (see Appendix A):

$$\frac{1}{\frac{1}{\max_j \widetilde{q}_{j,H}^{(h)}} + \frac{1}{\max_j \widetilde{q}_{j,A}^{(h)}}} > 1 \tag{7}$$

In general this inequality is not satisfied because market participants quickly exploit arbitrage opportunities. Selecting the maximum available market prices minimizes the total deviation from the fair prices. In the text below, the prices are assumed to be the maximum available prices $\max_j \widetilde{q}_{j,H}^{(h)}$ and $\max_j \widetilde{q}_{j,A}^{(h)}$.

[4]For simplicity the index and the superscript of the payout Π is omitted here.

[5]that the definition of unfair market prices only holds if the unfair market price $\widetilde{q}$ is larger than 1. If the fair price was 1.01 the bookmaker could not charge a vigorish of 5%. The market price $\widetilde{q} = 1.01 \cdot 0.95 = 0.9595$ implies a sure financial loss for the bettor even if the bet is won.

Assume that the market prices are $\widetilde{q}_{1,H} = \widetilde{q}_{1,A} = 1.99$, $\widetilde{q}_{2,H} = 1.97$ and $\widetilde{q}_{2,A} = 2.01$. The fair market prices for the two bookmakers would be $q_{1,H} = q_{1,A} = 2$, $q_{2,H} = 1.98$ and $q_{2,A} = 2.02$. These two markets assess the relative team strengths differently but they do not provide an arbitrage opportunity. In this example the maximum available prices would be $\max_j \widetilde{q}^{(h)}_{j,H} = 1.99$ and $\max_j \widetilde{q}^{(h)}_{j,A} = 2.01$.

All the above statements also hold for Over Under markets.

3. Deriving Implied Probabilities

Neither the probability matrix $\mathbf{P}_t$ nor its estimates $\widehat{\mathbf{P}}_t$ and its partitions can be observed on the betting markets. Only (unfair) market prices are observable. The question is, how might probabilities be derived from market prices? Such probabilities are denoted as implied probabilities in the text below.

Only two prices for betting on the home team $q^{(h)}_H$ and the away team $q^{(h)}_A$ exist on AHC(h) markets. If the prices are fair, i.e. the bookmakers do not charge any fees, the following is true:

$$\frac{1}{q^{(h)}_H} + \frac{1}{q^{(h)}_A} = 1 \Leftrightarrow q^{(h)}_H = \frac{q^{(h)}_A}{q^{(h)}_A - 1} \tag{8}$$

The AHC(h) market is completely defined by one fair price and a handicap h. The second price can easily be derived by (8). For integer handicaps $h \in \mathbb{Z}$ this implies that the probability of a draw $W^{(h)}_D$ (note that 'draw' only refers to future goals considering the handicap) is unknown with given $q^{(h)}_H$ and $q^{(h)}_A$. For all integer values of h, only the ratios of probabilities can be derived. In an Asian handicap market with the handicap $h = 0$ and the market prices $q^{(0)}_H = q^{(0)}_A = 2$, the prices imply that the probabilities of a home team win and an away team win are estimated to be equal, i.e. $W^{(0)}_H = W^{(0)}_A$, but their values are unknown. Only handicaps of the form $h \in \mathbb{Z} + 0.5$ allow the derivation of probabilities implied by the market instead of only ratios of these implied probabilities.

Since Asian handicap markets with $h \in 0.5^n(\mathbb{Z} + 0.5)$, where $n \in \mathbb{N}$, are linear combinations of the payments for integer handicaps and markets with handicaps $h \in \mathbb{Z} + 0.5$, these markets inherit the characteristics of the integer Asian handicap markets such that the market prices only represent ratios of probabilities. The numeric values of the probabilities implied by the market prices remain unknown.

The implications for the derivation of the Over Under markets' assessment of the probabilities are analogous. Only the Over Under markets with lines $0.5, 1.5, \ldots$ give exact values for the implied probabilities. For the other markets only ratios of probabilities can be calculated.

If there are more than two Over Under or Asian handicap markets available, it needs to be checked whether the market prices imply different estimates of $\mathbf{P}_t$. The market price data from Table 3 is used for a numerical illustration of the problem.

The market prices of these three Over Under markets represent the markets' estimates

Market m	Line l	Market prices $\widetilde{q}$		Adjusted prices q	
		Over	Under	Over	Under
1	2.25	1.6	2.4	1.667	2.5
2	2.5	1.72	2.21	1.778	2.285
3	2.75	2	1.9	2.053	1.950

Table 3: Over Under market price data adjusted for vigorish

for four partitions of the matrix $\widehat{\mathbf{P}}_t$:

$$V_U^{(2)} = P(g_T^H + g_T^A < 2) \tag{9}$$
$$V_X^{(2)} = P(g_T^H + g_T^A = 2) \tag{10}$$
$$V_X^{(3)} = P(g_T^H + g_T^A = 3) \tag{11}$$
$$V_O^{(3)} = P(g_T^H + g_T^A > 3) \tag{12}$$

The market prices $q_O^{(2.25)}$, $q_U^{(2.5)}$ and $q_U^{(2.75)}$ are defined as:

$$q_O^{(2.25)} = \frac{1-\frac{1}{2}V_X^{(2)}}{V_X^{(3)}+V_O^{(3)}} = \frac{1-\frac{1}{2}V_X^{(2)}}{1-V_U^{(2)}-V_X^{(2)}} \tag{13}$$
$$q_U^{(2.5)} = \frac{1}{V_U^{(2)}+V_X^{(2)}} \tag{14}$$
$$q_U^{(2.75)} = \frac{1-\frac{1}{2}V_X^{(3)}}{V_U^{(2)}+V_X^{(2)}} \tag{15}$$

In order to be able to compare the markets OU(2.25), OU(2.5) and OU(2.75) they are analysed in pairs. For this reason the probability assessment for a combination of two markets is analysed, e.g. $\widehat{\mathbf{P}}_t^{m_1,m_2}$, and not the assessment of one market, e.g. $\widehat{\mathbf{P}}_t^{m_1}$. It is assumed that each pair is based on identical assessments of the partitions (9) - (12) of the probability matrix of final scores $\widehat{\mathbf{P}}_t$. The first pair OU(2.25) and OU(2.5) requires solving the equations (13) and (14) with the price data from Table 3 for $V_U^{(2)}$ and $V_X^{(2)}$. This gives the results presented in the first row in Table 4. From a comparison of the markets OU(2.25) and OU(2.5), $V_X^{(3)}$ and $V_O^{(3)}$ cannot be derived individually, but only the sum $V_X^{(3)} + V_O^{(3)}$. The pair OU(2.25) and OU(2.75) gives a system of two equations with three variables. From these two markets a common estimate $\widehat{\mathbf{P}}_t^{m_1,m_3}$ cannot be derived. The combination of OU(2.5) and OU(2.75) gives the third row in Table 4. This pair of markets only gives an assessment of the sum of $V_U^{(2)}$ and $V_X^{(2)}$. If all markets

Market pairs			$V_U^{(2)}$	$V_X^{(2)}$	$V_X^{(3)}$	$V_O^{(3)}$
OU(2.25)	and	OU(2.5)	0.313	0.125	0.562	
OU(2.25)	and	OU(2.75)	cannot be solved			
OU(2.5)	and	OU(2.75)	0.398		0.271	0.331

Table 4: Implied probabilities by pairwise comparison of markets

were based on the same estimation of $\mathbf{P}_t$, the probabilities in Table 4 would be identical.

However, there are substantial differences. The first row of Table 4 gives a value of $V_U^{(2)}+V_X^{(2)} = 0.313+0.125 = 0.438$. The comparison of the markets OU(2.5) and OU(2.75) in the third row of Table 4 gives a value of 0.398. The probabilities are different, which shows that $\widehat{\mathbf{P}}_t$ differs between markets.

These differences cannot be exploited by a betting strategy. If a market participant estimates that the probability of less than 2.5 goals equals 50%, she wants to place a bet if any market m_k estimates this probability to be smaller than 50%. This is the case in all three OU markets of Table 4, but no particular market can be chosen. The third row of Table 4 contains the two Over Under markets with the lowest assessment of $P(g_T^H + g_T^A < 2.5)= 0.398$. If a bet is placed on Under on both OU(2.5) and OU(2.75) markets then it needs to be considered that the OU(2.5) market is also a component of the pair in the first row of Table 4. This would put a bettor in the unfortunate position of having to bet on the market with the highest and the lowest assessment of $P(g_T^H + g_T^A < 2.5)$ simultaneously.

A further drawback of this method is that there are always at least three markets needed to find differences in implied probabilities. If there are only two markets available there would only be one line of Table 4 and no other pairs of market to compare.

This method only serves to illustrate that the markets do not have identical assessments of probabilities. It does not provide any information on which bets maximize expected returns.

4. Implied Symmetry – A Method for Relative Pricing

4.1. Algorithm

Owing to the drawbacks associated with implied probabilities in betting, we introduce the concept of Implied Symmetry. An Asian handicap h and a market price $q_H^{(h)}$ or $q_A^{(h)}$ describe the markets' assessment of implied probabilities for a partition of $\mathbf{P}_t$. We want to represent a market not by two numbers but by only one rational number. This number is defined as a handicap or an Over Under line for which both respective market prices are $q_H = q_A = 2$ or $q_O = q_U = 2$ and therefore can be ignored. This number is called the symmetrical Asian handicap (SymmAHC) or the symmetrical Over Under line (SymmOU). This number is useful for comparing several markets at once. It provides information on differences between markets when assessing the probabilities of events. In contrast to implied probabilities, symmetrical Asian handicaps and Over Under lines can be used as a relative pricing criterion in betting strategies.

A fair market without vigorish ($v = 0$) can be represented by SymmAHC or SymmOU without loss of information. When an unfair market is condensed into a symmetrical handicap or into a symmetrical Over Under line, information is lost. The amount of lost information depends on the size of the vigorish. In the following simple example the symmetrical handicap can be derived from the market prices without the need for further calculations. Assume bookmaker 1 offers an AHC(0) market with the (fair) prices $q_{1,H}^{(0)} = q_{1,A}^{(0)} = 2$. The handicap $h = 0$ causes both fair prices to be equal and therefore $h =$ SymmAHC $=0$. Bookmaker 2 also offers an AHC(0) market with the prices $\widetilde{q}_{2,H}^{(0)} = \widetilde{q}_{2,A}^{(0)} = 1.99$. The symmetrical handicap for the AHC(0) market of bookmaker 2 also equals 0 because both markets have the same relative assessments of a home team win and an away team win. Judging by the symmetrical handicaps alone, both markets are equivalent. However, rational gamblers would only place bets on the market of bookmaker 1. Note that for all the analyses presented below the unfair market prices $\widetilde{q}$ are divided by the vigorish to generate the fair prices $q_H^{(0)} = q_A^{(0)} = 1.99/(1-0.005) = 2$. Obviously the markets are not equivalent just because they have the same symmetrical handicap. SymmAHC and SymmOU are calculated under the assumption that the vigorish is low enough such that it can be neglected. This assumption is assured by selecting the maximum prices from all available bookmakers. The plausibility of this assumption will be tested on the empirical dataset in Section 7.2.

The calculation of SymmAHC and SymmOU follows a two step procedure. First, pairs of markets are compared. Two markets m_1 and m_2 that form a pair for the comparison are based on probability matrices $\widehat{\mathbf{P}}_t^{m_1}$ and $\widehat{\mathbf{P}}_t^{m_2}$. Comparing the markets yields a matrix $\widehat{\mathbf{P}}_t^{m_1,m_2}$ that shares the probabilities in particular partitions with $\widehat{\mathbf{P}}_t^{m_1}$ and $\widehat{\mathbf{P}}_t^{m_2}$. Second, the probability matrix $\widehat{\mathbf{P}}_t^{m_1,m_2}$ obtained in the first step is partitioned such that the prices are identical. Derivation of a symmetrical handicap requires knowledge of the individual components of $\widehat{\mathbf{P}}_t^{m_1,m_2}$. For this purpose, the parameters λ_1, λ_2 of two independent Poisson distributions are introduced. Thus, the game results are described as

$$p_{g_T^H g_T^A} = P(X = g_T^H - g_t^H, Y = g_T^A - g_t^A | \lambda_1, \lambda_2) = \frac{e^{-\lambda_1} \cdot \lambda_1^{g_T^H}}{g_T^H!} \cdot \frac{e^{-\lambda_2} \cdot \lambda_2^{g_T^A}}{g_T^A!} \tag{16}$$

The Poisson assumption is commonly used in the literature for modeling the final scores of soccer games (Dixon and Robinson, 1998; Koning, 2003; Karlis and Ntzoufras, 2005; Maher, 1982). The parameters of the Poisson distribution represent the strengths of each team. These parameters are estimated using pre-game data (e.g. Maher 1982; Dixon and Coles 1997) or live data (e.g. Dixon and Robinson, 1998). Maher (1982) found that the independent Poisson model gives an accurate description of the final scores (see also Koning et al. 2003). A pair of markets m_1 = AHC(h) and m_2 =OU(l) is represented by four pieces of information $q_H^{(h)}, h, q_U^{(l)}$ and l. These are used for the calculation of the implied parameters, $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$, of the Poisson distribution. The four numbers $(q_H^{(h)}, h, q_U^{(l)}, l)$ are used for the construction of a system of two nonlinear equations. The handicap h and the line l determine which components $p_{g_T^H g_T^A}$ of $\widehat{\mathbf{P}}_t^{m_1,m_2}$ are used for the estimation of $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$. For instance, the handicap $h = -0.5$ with the fair price $q_H^{(-0.5)}$, and the line $l = 0.5$ with the fair price $q_U^{(0.5)}$, result in two nonlinear equations:

$$\sum_{i=1}^{15} \sum_{j=0}^{i-1} \frac{e^{-\widehat{\lambda}_1} \cdot \widehat{\lambda}_1^i}{i!} \cdot \frac{e^{-\widehat{\lambda}_2} \cdot \widehat{\lambda}_2^j}{j!} = \frac{1}{q_H^{(-0.5)}} \tag{17}$$

$$e^{-\widehat{\lambda}_1} \cdot e^{-\widehat{\lambda}_2} = \frac{1}{q_U^{(0.5)}} \tag{18}$$

The left sides of the equations are the components $\widehat{p}_{ij}$ addressed by the market. The right sides of the equations are the probabilities implied by the fair market prices. The AHC market implies probabilities for an infinite number of components (see for example the definition of $W_H^{(0)}$). For the calculation, the matrix is restricted to 15 rows and columns. The loss of numerical accuracy is small enough to be ignored.

Each pair consists of one AHC and one OU market. Pairs consisting of two Over

Under markets or two Asian handicap markets often cause numerical problems because the nonlinear equations use very similar components p_{ij}.

The parameters $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$ are used as inputs for the algorithm to find the symmetrical Asian handicap (SymmAHC) and the symmetrical Over Under line (SymmOU). The algorithm for the calculation of the symmetrical Asian handicap is described in Algorihm 1.

```
Data: λ̂1, λ̂2
generate matrix P̂^{m1,m2} with the probabilities from the Poisson distribution (16)
and parameters λ̂1, λ̂2;
initialize the symmetrical handicap as h ← ⌈λ̂2 − λ̂1⌉;
initialize the interval of the symmetrical handicap as interval[1] ← -8 and
interval[2] ← 8;
initialize stop ← FALSE;
while (stop==FALSE) do
    price ← fair price of bet given (P̂^{m1,m2} ,h);
    diff ← price - 2;
    if |diff| < 2^(−8) then
        stop ← TRUE;
    else
        if diff< 0 then
            interval[2] ← h;
            h ←(h + interval[1])/2;
        else
            interval[1] ← h;
            h ← (h + interval[2])/2;
        end
    end
end
Result: h
```

Algorithm 1: Algorithm for the calculation of the symmetrical Asian handicap

The choice of $2^{(-8)}$ is based on a tradeoff between accuracy and processing time. The algorithm for the calculation of the symmetrical Over Under line is analogous.

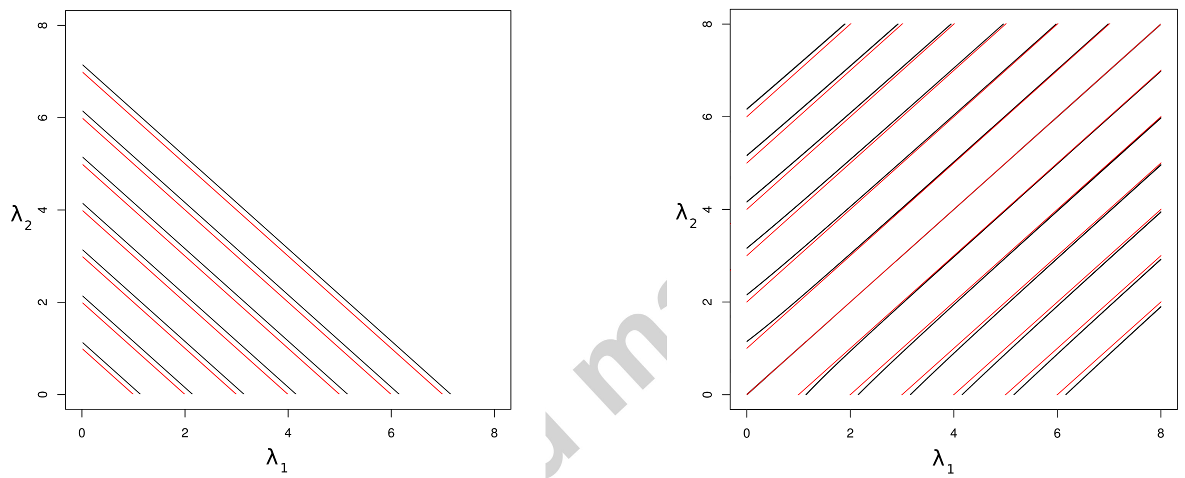
4.2. Uniqueness of Symmetrical Handicaps/Lines

Symmetrical handicaps and lines fully represent the market's assessment of the outcome probabilities. However, the information provided by the market, e.g. an Over Under line and two corresponding prices will not give a unique combination of the parameters $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$ that are used for the calculation of SymmOU.

Symmetrical handicaps and lines are non-injective functions of $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$. A single Over Under market only provides information on the total number of expected goals but not on the strength of each team. The symmetrical Over Under line inherits this characteristic. The sum of two independent Poisson distributions with respective parameters λ_1 and λ_2, is Poisson-distributed with parameter $\lambda_1 + \lambda_2$. All pairs (λ_1, λ_2) with $\lambda_1 + \lambda_2 = c$, yield the same univariate Poisson distribution. The points (1,0), (0.5,0.5) and (0,1) in the (λ_1, λ_2) plane will have the same Over Under market prices q_O and q_U. Therefore the values of λ_1 and λ_2 cannot be derived from the market prices of one market.

SymmOU and SymmAHC are not the medians of the Poisson distribution. When the SymmOU line is derived as the median of the Poisson distribution (see for example Adell and Jodrá, 2005 and the references therein) it is expected that SymmOU(1,0) = 1. However, the symmetrical Over Under line of $(\lambda_1 = 1, \lambda_2 = 0)$ does not equal 1, but 0.859 because SymmOU is derived such that the prices, but not necessarily the probabilities, are identical. The median of a discrete distribution such as the Poisson distribution does not meet this requirement of identical prices.

In Figure 5(a) the black straight lines represent all points SymmOU$(\lambda_1, \lambda_2) \in \{1, 2, \ldots, 7\}$. Red straight lines $\lambda_2 = -\lambda_1 + d$ with $d \in \{1, 2, \ldots, 7\}$ are plotted for reference purposes. The Figure illustrates that SymmOU$(\lambda_1, \lambda_2) < \lambda_1 + \lambda_2 = c$.



(a) Isoquants for SymmOU$(\lambda_1, \lambda_2) \in \{1, 2, \ldots, 7\}$ in black. Straight lines $\lambda_2 = -\lambda_1 + d$ with $d \in \{1, 2, \ldots, 7\}$ in red

(b) Isoquants for SymAHC$(\lambda_1, \lambda_2) \in \{-6, -5, \ldots, 6\}$ in black. Straight lines $\lambda_2 = \lambda_1 + d$ with $d \in \{-6, -5, \ldots, 6\}$ in red

Figure 5: SymmAHC and SymmOU isoquants

The above statements hold for the symmetrical handicap in terms of ratios of team strengths. From an Asian handicap market alone, the total number of expected goals cannot be derived, only the relative strengths of the teams. The symmetrical Asian handicaps are plotted in Figure 5(b) for SymAHC$(\lambda_1, \lambda_2) \in \{-6, -5, \ldots, 6\}$ in black. For comparison purposes straight lines with a slope of 1 and intercepts of $-6, -5, \ldots, 6$ are plotted in red. Figure 5(b) contains the obvious result that SymmAHC$(\lambda_1, \lambda_2) = 0$ for $\lambda_1 = \lambda_2$. If the home team is exactly as strong as the away team then the AHC(0) market has identical prices for both teams. The isoquants of order $\neq 0$ are not straight lines but they converge to the straight lines with slope 1 plotted in red in Figure 5(b). This means that SymmAHC(1,1) = SymmAHC(2,2)= 0 but SymmAHC(1,2) = 0.943 and SymmAHC(2,3) = 0.965.

For our purposes it is only relevant to see that isoquants are strictly monotonic. Thus, the symmetrical handicaps and symmetrical Over Under lines are unique and contain the same information as the Over Under and Asian handicap markets from which they were derived, under the assumption of Poisson marginal distributions and fair market prices. Information from two markets is always needed in order to find a unique solution for these parameters. Thus, the exact notation for two markets i and j would be $\widehat{\lambda}_1^{m_i,m_j}$ and $\widehat{\lambda}_2^{m_i,m_j}$ but this excessive notation is avoided whenever possible.

5. Betting Strategies

In this study the Implied Symmetry method is investigated with respect to its ability to identify bets with the highest expected return. Questions of how to generate an assessment $\widehat{\mathbf{P}}_t^i$ or how to bet optimally given $\widehat{\mathbf{P}}_t^i$ are not discussed here, since this would confuse the empirical evaluation of Implied Symmetry. It is assumed that a market participant's decision to bet on Home, Away, Over or Under is given.

If the market prices are based on the same probability matrix $\mathbf{P}_t$ and all prices are fair, all bets provide the same expected payout $E(\Pi_H^{(h)}) = 1$. In this case we assume that a rational gambler will choose the bets with the lowest variance. It can be shown that the variance of the home team bet decreases when the handicap increases, i.e. $Var(\Pi_H^{(h_1)}) \geq Var(\Pi_H^{(h_2)})\ \forall\ h_1 > h_2$. For the away team bets the opposite is true, i.e. the variance of payouts for the away team bets decreases with decreasing handicap h. The variance minimizing strategy in a fair Asian handicap market can be summarized as follows: Choose the market where the handicap is a maximum, when placing a bet on the home team, and choose the market where the handicap is a minimum, when placing a bet on the away team. In the Over Under market the variance is minimized by choosing the market where the line is a minimum, when placing an Over bet, and by choosing the market where the line is a maximum, when placing an Under bet.

If markets deviate from each other in terms of their assessments of the event probabilities they will also have different expected returns. In this case the market participants want to select the bet with the highest expected return. The betting decision itself is based on an individual assessment of the event probabilities $\widehat{\mathbf{P}}_t^i$. If market participant i thinks that the game will end with a final score 2:0, i.e. $\widehat{p}_{20}^i = 1$, she will place a bet only if there is a market m_j with $\widehat{p}_{20}^{(m_j)} < 1$.

The betting strategy can be demonstrated by a stylized example. Assume that the market prices on an OU(2) and on an OU(2.5) market are given as $q_O^{(2)} = q_U^{(2)} = q_O^{(2.5)} = q_U^{(2.5)} = 2$. These market prices have the unrealistic implication that the total number of goals in the game equals 2 is $P(g_T^H + g_T^A = 2) = V_X^{(2)} = 0$. In this case the symmetrical Over Under line of the OU(2) market equals 2 and of the OU(2.5) market equals 2.5. These can be derived from the price data without the need for calculations.

A gambler who thinks $P(g_T^H + g_T^A < 2) > 0.5$ and wants to bet on Under would place a bet on OU(2.5), not on OU(2). All other factors equal, more events (in this case $g_T^H + g_T^A = 2$) are covered by the OU(2.5) market. If the market participant wants to bet on Over, the OU(2) market strictly dominates OU(2.5) for the same reason.

The example demonstrates that buying "more events" at the same price will increase the expected return. This is the idea of the Implied Symmetry method. The market prices hardly ever present obvious ineffciencies as in the example. The Implied Symmetry method helps discover the markets with highest expected return when the deviations are not obvious.

According to Implied Symmetry, the optimal decision with respect to the expected return is to place an Over bet on the market with the minimum SymmOU or an Under bet on the market with the maximum SymmOU. Analogous considerations hold for the Asian handicap market. The expected return is maximized by choosing the market with the maximum SymmAHC for a Home bet, and choosing the market with the minimum SymmAHC for an Away bet. The strategy does not take variances into account.

6. Relevance of Implied Symmetry

The relevance of the Implied Symmetry method for the identification of hidden information in the market prices is demonstrated by the pairwise analysis of real market data from twelve soccer games. Table 5 summarizes general information on the games. In all these games three Asian handicap markets and three Over Under markets are available. The very special feature of the data analysed in this section is that for each pair of soccer games the three Asian handicap markets have identical handicaps and market prices. This would suggest that the handicap markets believe the relative strengths of the teams to be exactly identical within these six pairs of games. However, it can be shown that the symmetrical handicaps reveal differing assessments of team strengths which cannot be seen from the identical AHC market prices alone. These differences are derived from the analysis of the Over Under market prices using the Implied Symmetry method.

Pair	Home Team	Away Team	League	Country	Minute	Score
1	SC Paderborn 07	TSV 1860 Muenchen	2. Bundesliga	Germany	30	0:1
	RKC Waalwijk	Roda JC Kerkrade	Eredivisie	Netherlands	32	0:0
2	FC Dordrecht	PEC Zwolle	Eredivisie	Netherlands	37	0:0
	Queens Park Rangers	Arsenal FC	Premier League	England	33	1:0
3	TSV 1860 Muenchen	SpVgg Greuther Fuerth	2. Bundesliga	Germany	29	1:2
	Preston North End	Huddersfield Town FC	League Cup	England	42	2:0
4	Granada CF	Rayo Vallecano	Primera Division	Spain	40	0:1
	1. FC Union Berlin	SG Dynamo Dresden	2. Bundesliga	Germany	13	0:0
5	Parma FC	SSC Napoli	Serie A	Italy	43	0:1
	TSG 1899 Hoffenheim	FC Schalke 04	Bundesliga	Germany	16	0:0
6	AZ Alkmaar	SC Heerenveen	Eredivisie	Netherlands	15	0:0
	SS Lazio	US Citta di Palermo	Serie A	Italy	11	0:0

Table 5: Games from the season 2012 with pairwise identical Asian handicap markets

In Table 6, data from the first pair of soccer games are presented. The first nine lines of the table show market data from the game SC Paderborn against TSV 1860 Muenchen. The lines 10-18 are market data from the game RKC Waalwijk against Roda JC Kerkrade. In both games the three Asian handicap markets AHC(-0.75), AHC(-0.5) and AHC(-0.25) have the identical market prices for the home team $\widetilde{q}_H^{(-0.75)} = 2.36, \widetilde{q}_H^{(-0.5)} = 1.99$ and $\widetilde{q}_H^{(-0.25)} = 1.7$. The market prices for the away team are also identical. The Over Under markets are not identical, but very similar. In the first game the price for betting on over 3 goals at a current score of 0:1 equals 2.13. In the second game the price for betting on over 2 goals at a current score of 0:0 equals 2.11. The number of expected additional goals is almost equal.

The similarity of all market prices means that the markets have very similar assessments of the probabilities of future goals in both games. Also, the implied parameters $\widehat{\lambda}_1^{m_i,m_j}$ and $\widehat{\lambda}_2^{m_i,m_j}$ are similar. These are plotted with the isoquant corresponding to the symmetrical Over Under lines and the 45° line in Figure 6(a). The 45° line represents the

$\widehat{\lambda}_1$	$\widehat{\lambda}_2$	Minute	SymmAHC	SymmOU	$\widetilde{q}_H^{(h)}$	$\widetilde{q}_A^{(h)}$	h	$\widetilde{q}_O^{(l)}$	$\widetilde{q}_U^{(l)}$	l
1.277	0.753	30	-0.461	1.875	2.36	1.65	-0.75	1.60	2.42	2.5
1.276	0.753	30	-0.461	1.875	2.36	1.65	-0.75	1.80	2.11	2.75
1.278	0.755	30	-0.461	1.881	2.36	1.65	-0.75	2.13	1.78	3
1.283	0.747	30	-0.477	1.875	1.99	1.93	-0.5	1.60	2.42	2.5
1.283	0.746	30	-0.477	1.875	1.99	1.93	-0.5	1.80	2.11	2.75
1.284	0.748	30	-0.477	1.881	1.99	1.93	-0.5	2.13	1.78	3
1.278	0.752	30	-0.461	1.875	1.70	2.28	-0.25	1.60	2.42	2.5
1.278	0.752	30	-0.461	1.875	1.70	2.28	-0.25	1.80	2.11	2.75
1.279	0.753	30	-0.461	1.881	1.70	2.28	-0.25	2.13	1.78	3
1.284	0.762	32	-0.461	1.893	2.36	1.65	-0.75	1.78	2.13	1.75
1.284	0.762	32	-0.461	1.893	2.36	1.65	-0.75	2.11	1.80	2
1.286	0.765	32	-0.461	1.898	2.36	1.65	-0.75	2.47	1.58	2.25
1.291	0.755	32	-0.477	1.893	1.99	1.93	-0.5	1.78	2.13	1.75
1.291	0.755	32	-0.477	1.893	1.99	1.93	-0.5	2.11	1.80	2
1.294	0.758	32	-0.477	1.898	1.99	1.93	-0.5	2.47	1.58	2.25
1.286	0.760	32	-0.465	1.893	1.70	2.28	-0.25	1.78	2.13	1.75
1.286	0.760	32	-0.465	1.893	1.70	2.28	-0.25	2.11	1.80	2
1.289	0.763	32	-0.465	1.898	1.70	2.28	-0.25	2.47	1.58	2.25

Table 6: Market prices of Pair 1: Minute 30 of SC Paderborn vs. TSV 1860 Muenchen and minute 32 of RKC Waalwijk vs. Roda JC Kerkrade

isoquant for SymmAHC = 0 and is plotted for reference purposes in all charts. Different colors refer to the respective Asian handicap. Red denotes the AHC(-0.75) market, green the AHC(-0.5) market and black the AHC(-0.25) market. The points are very close to each other. This means that there are hardly any differences between the implied probabilities, between the markets within each game and the markets between the two games.

Figure 6(b) plots the parameters $\widehat{\lambda}_1^{m_i,m_j}$ and $\widehat{\lambda}_2^{m_i,m_j}$ from pair 2. As was the case for pair 1 in Figure 6(a), within each of the two games the markets' assessments of the implied probabilities are very similar. The nine tuples of $(\widehat{\lambda}_1^{m_i,m_j}, \widehat{\lambda}_2^{m_i,m_j})$ of each game are very close to each other. However, unlike pair 1, the markets' assessments between the games are quite different. Since the Asian handicap markets have exactly the same prices and handicaps, these differences are again caused by the Over Under markets. The Over Under markets reflect differences in the number of expected goals. Thus, the values of $\widehat{\lambda}_1^{m_i,m_j}$ and $\widehat{\lambda}_2^{m_i,m_j}$ are higher in the first game than in the second game for pair 2.

The situations in the top row of Figure 6 are consistent with the assumption of a common market expectation $\widehat{\mathbf{P}}_t$. In reality, this is hardly ever the case, as the analysis of pairs 3-6 reveals. Figure 6(c) plots the data for pair 3. As before, the handicaps and market prices on three Asian handicap markets are identical. As is the case in Figure 6(b), market assessments differ in terms of the expected total number of goals. The symmetrical Over Under lines are different. More importantly, the market assessments of the probabilities mainly match in one game, and differ much more in the other game. The nine points in the $(\widehat{\lambda}_1,\widehat{\lambda}_2)$ plane diverge significantly. This contradicts the assumption of a common market expectation $\widehat{\mathbf{P}}_t$ in the second game for pair 3.

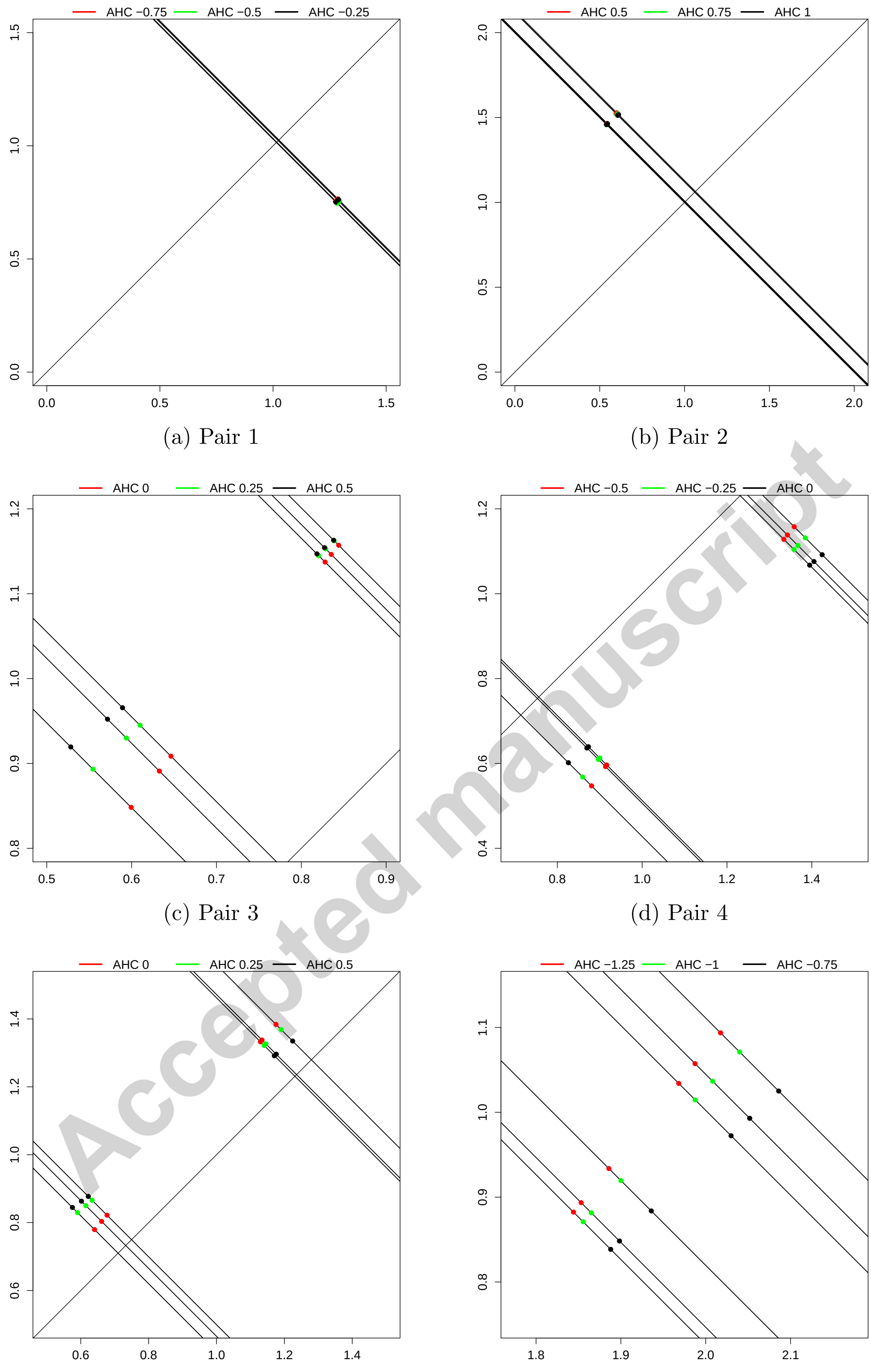
Figures 6(d)-6(f) show pairs where the markets in both games significantly disagree in terms of implied probabilities. The top right corner of Figure 6(d) shows the second game of pair 4, see rows 10-18 in Table 8 . The market with the smallest handicap is $h = -0.5$. The corresponding values of $(\widehat{\lambda}_1, \widehat{\lambda}_2)$ are plotted in red. This is also the market with the largest symmetrical handicap (SymmAHC$_{max} = -0.18$). The red dots are closer to the diagonal than the green and the black dots. The data for the second game is plotted in the lower left corner of Figure 6(d). Since this game has identical handicaps the smallest handicap is also $h = -0.5$. However, this is the market with the smallest symmetrical handicap (SymmAHC$_{min} = -0.266$). Figure 6(e) can be considered analogously.

In Figures 6(c)-6(f) the Implied Symmetry method indicates that markets deviate from each other. In these cases the betting strategies discussed in section 5 may produce different recommendations than the variance minimizing strategies. Table 7 compares information on four strategies. These are:

1. Choose the market with the maximum handicap h_{max},
2. choose the market with the maximum symmetrical handicap SymmAHC$_{max}$,
3. choose the market with the minimum handicap h_{min} or
4. choose the market with the minimum symmetrical handicap SymmAHC$_{min}$.

The strategies are analogous for Over Under. As indicated by the theoretical considerations outlined in the previous sections, each strategy is likely to affect return variances and means. The numerical examples in Table 7 refer to the second game in Table 8. The market with the maximum symmetrical handicap (i.e. the AHC(-0.5) market in the second game) is according to the Implied Symmetry method the market which provides the highest expected return. In this case it is also the market with the smallest available handicap ($h_{min} = -0.5$). This means that in this case the betting strategy selects the market with the maximum variance. In the typical Over Under data in Table 7 the market with the maximum symmetrical Over Under line and maximum l are both $l = 2.5$. The market with the SymmOU$_{min}$ and the minimum line l_{min} are both $l = 2$. These characteristics of the Over Under markets imply that the strategy with the maximum expected return also minimizes the variances.

All the statements are correct when the symmetrical Over Under lines are chosen for the selection of the optimal Over Under bet.



(a) Pair 1 (b) Pair 2

(c) Pair 3 (d) Pair 4

(e) Pair 5 (f) Pair 6

Figure 6: Market assessments of $\widehat{\lambda}_1^{m_i,m_j}$ (on the x-axes) and $\widehat{\lambda}_2^{m_i,m_j}$ (on the y-axes) in a pairwise comparison of two soccer games.

Bet on	Choose $\widetilde{q}$ from markets with	Example from Table 8			Expected effect on:	
		Criterion	Market	Price $\widetilde{q}$	Variance	Return
Home	h_{max}	0	0	$\widetilde{q}_H^{(0)} = 1.6$	min	
	SymmAHC$_{max}$	-0.18	-0.5	$\widetilde{q}_H^{(-0.5)} = 2.35$		max
	h_{min}	-0.5	-0.5	$\widetilde{q}_H^{(-0.5)} = 2.35$	max	
	SymmAHC$_{min}$	-0.297	0	$\widetilde{q}_H^{(0)} = 1.6$		min
Away	h_{max}	0	0	$\widetilde{q}_A^{(0)} = 2.47$	max	
	SymmAHC$_{max}$	-0.18	-0.5	$\widetilde{q}_A^{(-0.5)} = 1.66$		min
	h_{min}	-0.5	-0.5	$\widetilde{q}_A^{(-0.5)} = 1.66$	min	
	SymmAHC$_{min}$	-0.297	0	$\widetilde{q}_A^{(0)} = 2.47$		max
Under	l_{max}	2.5	2.5	$\widetilde{q}_U^{(2.5)} = 1.8$	min	
	SymmOU$_{max}$	2.344	2.5	$\widetilde{q}_U^{(2.5)} = 1.8$		max
	l_{min}	2	2	$\widetilde{q}_U^{(2)} = 2.42$	max	
	SymmOU$_{min}$	2.291	2	$\widetilde{q}_U^{(2)} = 2.42$		min
Over	l_{max}	2.5	2.5	$\widetilde{q}_O^{(2.5)} = 2.11$	max	
	SymmOU$_{max}$	2.344	2.5	$\widetilde{q}_O^{(2.5)} = 2.11$		min
	l_{min}	2	2	$\widetilde{q}_O^{(2)} = 1.6$	min	
	SymmOU$_{min}$	2.291	2	$\widetilde{q}_O^{(2)} = 1.6$		max

Table 7: The effects of four betting strategies on the returns for betting on Home, Away, Over or Under

7. Application to Real Market Data

In this section a large empirical dataset of market prices is used in order to test whether the Implied Symmetry method really does help to select the bets with the highest return.

7.1. Description of the Empirical Dataset

The dataset contains data from 29,413 soccer games from 470 international leagues. The data covers games from August 2011 until the end of June 2013. Table 9 reports details on the 20 leagues with the highest number of games in the dataset.

$\widehat{\lambda}_1$	$\widehat{\lambda}_2$	Minute	SymmAHC	SymmOU	$\widetilde{q}_H^{(h)}$	$\widetilde{q}_A^{(h)}$	h	$\widetilde{q}_O^{(l)}$	$\widetilde{q}_U^{(l)}$	l
0.881	0.547	40	-0.266	1.258	2.35	1.66	-0.5	1.51	2.63	2
0.914	0.593	40	-0.262	1.332	2.35	1.66	-0.5	1.82	2.08	2.25
0.916	0.596	40	-0.258	1.340	2.35	1.66	-0.5	2.17	1.75	2.5
0.860	0.568	40	-0.234	1.258	1.99	1.93	-0.25	1.51	2.63	2
0.897	0.609	40	-0.234	1.332	1.99	1.93	-0.25	1.82	2.08	2.25
0.900	0.613	40	-0.234	1.340	1.99	1.93	-0.25	2.17	1.75	2.5
0.826	0.602	40	-0.180	1.258	1.60	2.47	0	1.51	2.63	2
0.870	0.636	40	-0.188	1.332	1.60	2.47	0	1.82	2.08	2.25
0.873	0.639	40	-0.188	1.340	1.60	2.47	0	2.17	1.75	2.5
1.334	1.128	13	-0.180	2.291	2.35	1.66	-0.5	1.60	2.42	2
1.343	1.138	13	-0.180	2.309	2.35	1.66	-0.5	1.88	2.02	2.25
1.359	1.158	13	-0.180	2.344	2.35	1.66	-0.5	2.11	1.80	2.5
1.358	1.104	13	-0.227	2.291	1.99	1.93	-0.25	1.60	2.42	2
1.367	1.114	13	-0.227	2.309	1.99	1.93	-0.25	1.88	2.02	2.25
1.385	1.132	13	-0.227	2.344	1.99	1.93	-0.25	2.11	1.80	2.5
1.395	1.067	13	-0.289	2.291	1.60	2.47	0	1.60	2.42	2
1.405	1.076	13	-0.297	2.309	1.60	2.47	0	1.88	2.02	2.25
1.424	1.092	13	-0.297	2.344	1.60	2.47	0	2.11	1.80	2.5

Table 8: Market prices of Pair 4: Minute 40 of Granada vs. Rayo Vallecano and minute 13 of 1. FC Union Berlin vs. SG Dynamo Dresden

Country	League	Number of Games
-	Friendly matches	899
England	Championship	693
Italy	Serie B	570
Spain	Segunda Division	564
England	League One	518
Germany	3. Liga	495
France	Ligue 2	484
Argentina	Primera Division	454
Japan	J2 League	418
Italy	Serie A	418
Germany	2. Bundesliga	403
Romania	Liga I	398
Spain	Primera Division	388
Netherlands	Eerste Divisie	386
Brazil	Serie A	374
Belgium	2e Klasse	365
England	Premier League	357
Colombia	Primera Division	354
Mexico	Liga MX	349

Table 9: Leagues with the highest number of games in the empirical dataset

The market prices were collected from the websites of four of the largest bookmakers: SBObet, Pinnacle Sports, IBCbet and Singbet. For the analysis, only market prices for full minutes of the game $t \in \{1, \ldots, 90\}$ are considered. Market prices do not change frequently enough to make an analysis of seconds, i.e. $t \in \{1, \ldots, 5400\}$, sufficiently worthwhile. Market prices in the halftime break and in injury time are not considered. Only data from Asian handicap and Over Under markets are used.

Our method for the relative valuation of market prices cannot be applied when only one SymmAHC or one SymmOU value is available. Even in leagues with high trading volumes, several Over Under markets might be available in the beginning of the game, but only one market remains open in the final 10 minutes. This led to a large reduction of the data. The analysis below only includes data with at least two Asian handicap markets or two Over Under markets. Of the $29,413 \times 90$ rows about 700,000 rows remain for the analysis.

7.2. Results

Table 10 contains the average returns of the four betting strategies applied to betting on the home team and the away team. The Table includes data where the Implied Symmetry method indicates a deviation from probability assessments between markets. For the Asian handicap markets this means that $\text{SymmAHC}_{max} - \text{SymmAHC}_{min} > 0$. Each row of the Tables 10 and 11 reports the results when these differences exceed thresholds $0, 0.01 \ldots, 0.14$. The sample sizes reduce from 656,131 minutes of live game play to 222 minutes when the threshold equals 0.14.

Threshold	Number of Bets	Home				Away			
		maximum		minimum		maximum		minimum	
		h	SymmAHC	h	SymmAHC	h	SymmAHC	h	SymmAHC
0.00	656131	-2.593	**-2.413**	-4.508	-4.640	-1.530	-2.089	-0.861	**-0.279**
0.01	556906	-2.467	**-2.169**	-4.408	-4.643	-1.770	-2.382	-0.959	**-0.329**
0.02	402598	-2.320	**-1.868**	-4.289	-4.673	-2.119	-2.840	-1.074	**-0.338**
0.03	275743	-2.448	**-1.878**	-4.186	-4.689	-2.141	-3.055	-1.112	**-0.178**
0.04	157978	-2.450	**-1.680**	-3.909	-4.576	-2.270	-3.540	-1.394	**-0.119**
0.05	100649	-2.726	**-1.758**	-4.012	-4.900	-2.066	-3.774	-1.360	**0.375**
0.06	58692	-3.297	**-2.143**	-4.550	-5.695	-1.495	-3.668	-1.085	**1.184**
0.07	35754	-3.465	**-2.250**	-5.013	-6.292	-1.429	-3.914	-0.801	**1.791**
0.08	17215	-2.571	**-1.214**	-4.210	-5.673	-2.428	-5.216	-1.596	**1.342**
0.09	8800	-2.435	**-0.626**	-3.204	-5.157	-2.717	-6.264	-2.565	**1.186**
0.10	5033	-2.888	**-0.998**	-3.569	-5.658	-2.244	-6.612	-2.708	**1.854**
0.11	2246	-2.538	**0.908**	-1.945	-5.418	-2.880	-7.716	-3.982	**0.986**
0.12	1232	-2.126	**1.675**	-1.494	-5.246	-3.802	-8.280	-4.620	**-0.142**
0.13	532	0.880	**7.158**	6.195	0.159	-6.938	-12.588	-10.739	**-5.589**
0.14	222	-7.520	-3.149	-3.063	-7.234	3.581	-3.883	-3.773	3.586

Table 10: Returns (%) of betting strategies on home and away teams (Asian handicap)

The expectations on variances and returns for AHC and OU as derived from the theoretical considerations above, and summarized in Table 7, are completely supported by the empirical dataset. The results of the four betting strategies are compared for bets on the home team in columns 3-6. The variances of returns in column 3 of Table 11 are about 50% lower than the variances reported in column 5. This result is expected from our theory. More importantly, betting on the home team and choosing the h_{max} market not only has a lower variance, it also results in a higher average return, than betting on the home team and choosing the h_{min} market. The average return of this strategy is -2.593% and placing bets on the minimum available handicap has a return of -4.508% when $\text{SymmAHC}_{max} - \text{SymmAHC}_{min} > 0$. The average returns are higher in column 3 than in column 5 for the first 11 rows of Table 10. This means that the h_{min} markets offer betting opportunities that have both a higher expected value and a lower variance.

Our model suggests that betting on the home team and choosing the market corresponding to the maximum symmetrical handicap SymmAHC_{max} has significantly higher average returns than on the market with the maximum handicap h_{max}. The average returns of the model reported in column 4 of Table 10 are the highest of all columns. Based on a paired 2-sample t-test with $p < 0.01$ for rows 1-14 of Table 10 the differences between the strategies are significant. A comparison of columns 3 and 4 in Table 11 reveals that the significantly higher returns of the symmetrical handicap strategy come with higher variances.

Columns 7-10 of Table 10 report the average returns of the four strategies for betting on the away team. The variance minimizing strategy is to choose the h_{min} market. As the theory would lead us to expect, the variances of the returns in column 9 of Table 11 are lower than the variances of the h_{max} strategy in column 7. As we found for the home team, the variance minimizing strategy also has the highest average returns. For the 656,131 bets the average return is -0.861%. The variance maximizing strategy has a return of -1.53%.

The symmetrical handicap provides significantly valuable information. According to a paired 2-sample t-test with $p < 0.01$ for rows 1-14 of Table 10 it has the highest average return. These higher returns have a higher variance than the h_{min} strategy.

Threshold	Number of Bets	Home				Away			
		maximum		minimum		maximum		minimum	
		h	SymmAHC	h	SymmAHC	h	SymmAHC	h	SymmAHC
0.00	656131	0.501	0.654	0.956	0.802	0.890	0.761	0.531	0.659
0.01	556906	0.503	0.654	0.959	0.807	0.895	0.766	0.534	0.661
0.02	402598	0.517	0.666	0.963	0.813	0.906	0.775	0.546	0.676
0.03	275743	0.543	0.701	0.974	0.816	0.929	0.787	0.572	0.713
0.04	157978	0.562	0.736	0.992	0.821	0.953	0.795	0.591	0.750
0.05	100649	0.575	0.763	1.007	0.825	0.971	0.798	0.603	0.777
0.06	58692	0.582	0.787	1.014	0.818	0.985	0.793	0.607	0.801
0.07	35754	0.582	0.798	1.020	0.814	0.989	0.786	0.609	0.815
0.08	17215	0.577	0.808	1.027	0.804	0.993	0.775	0.614	0.838
0.09	8800	0.588	0.833	1.022	0.784	1.014	0.765	0.618	0.872
0.10	5033	0.590	0.842	1.014	0.767	1.031	0.756	0.622	0.900
0.11	2246	0.588	0.879	1.026	0.737	1.043	0.737	0.630	0.937
0.12	1232	0.580	0.902	1.049	0.729	1.022	0.723	0.641	0.941
0.13	532	0.594	0.967	1.095	0.725	1.072	0.746	0.658	0.981
0.14	222	0.620	0.939	1.095	0.764	1.127	0.762	0.665	1.032

Table 11: Variances of betting strategies on home and away teams (Asian handicap)

In Tables 12 and 13 the average returns and variances for the four strategies are reported for the Over Under markets. The conclusions derived from the AHC markets are the same as those from the OU markets. Variance minimizing strategies have higher average returns than variance maximizing strategies, see columns 3 and 5 of Table 12 for Over markets and columns 7 and 9 for Under markets. For both Over and Under taking the symmetrical Over Under lines into consideration when placing bets has significantly higher returns. See columns 6 and 8 of Table 12. Again, these higher returns are associated with higher variances when compared to the variance minimizing strategies.

Threshold	Number of Bets	Over				Under			
		maximum		minimum		maximum		minimum	
		l	SymmOU	l	SymmOU	l	SymmOU	l	SymmOU
0.00	606532	-2.161	-2.920	-1.505	**-0.635**	-4.159	**-3.744**	-5.257	-5.643
0.01	502047	-2.012	-2.899	-1.312	**-0.298**	-4.127	**-3.604**	-5.322	-5.814
0.02	356742	-1.868	-2.876	-1.034	**0.132**	-4.196	**-3.560**	-5.580	-6.183
0.03	255119	-1.802	-2.879	-0.906	**0.353**	-4.154	**-3.420**	-5.665	-6.367
0.04	168063	-1.654	-2.844	-0.777	**0.619**	-4.217	**-3.318**	-5.733	-6.634
0.05	109819	-1.246	-2.498	-0.263	**1.228**	-4.441	**-3.490**	-6.313	-7.271
0.06	63577	-1.915	-3.250	-0.632	**0.917**	-3.888	**-2.701**	-5.837	-7.036
0.07	43450	-2.854	-4.180	-1.518	**0.048**	-3.115	**-1.703**	-4.746	-6.131
0.08	22559	-2.452	-4.290	-1.453	**0.574**	-3.437	**-1.687**	-4.856	-6.618
0.09	11737	-2.809	-4.745	-1.591	**0.629**	-2.998	**-0.894**	-4.465	-6.662
0.10	6273	-0.546	-3.001	-0.304	**2.327**	-4.837	**-2.662**	-6.229	-8.308
0.11	3127	1.116	-2.023	-0.067	**3.219**	-6.077	**-3.027**	-6.494	-9.533
0.12	1437	0.239	-1.762	0.509	**2.998**	-5.289	**-2.688**	-7.126	-9.799
0.13	623	-1.612	-4.714	-0.385	**2.486**	-3.624	**0.522**	-5.064	-8.731
0.14	297	-2.830	-3.960	1.067	1.978	-2.423	-0.207	-6.212	-8.066

Table 12: Returns (%) of betting strategies on Over and Under

Threshold	Number of Bets	Over				Under			
		maximum		minimum		maximum		minimum	
		l	SymmOU	l	SymmOU	l	SymmOU	l	SymmOU
0.00	606532	1.048	0.945	0.807	0.907	0.556	0.703	0.899	0.746
0.01	502047	1.027	0.922	0.773	0.875	0.572	0.716	0.916	0.764
0.02	356742	1.018	0.911	0.748	0.853	0.583	0.726	0.928	0.776
0.03	255119	1.003	0.896	0.715	0.819	0.596	0.738	0.945	0.792
0.04	168063	0.998	0.884	0.688	0.801	0.609	0.751	0.953	0.800
0.05	109819	0.991	0.883	0.664	0.773	0.614	0.756	0.964	0.811
0.06	63577	0.984	0.881	0.650	0.758	0.617	0.763	0.974	0.817
0.07	43450	0.985	0.879	0.641	0.757	0.627	0.775	0.983	0.825
0.08	22559	0.991	0.877	0.645	0.768	0.632	0.785	0.988	0.823
0.09	11737	0.984	0.870	0.648	0.772	0.636	0.794	0.998	0.827
0.10	6273	0.985	0.874	0.649	0.777	0.638	0.796	1.003	0.832
0.11	3127	0.990	0.868	0.654	0.786	0.656	0.844	1.016	0.824
0.12	1437	0.980	0.869	0.651	0.780	0.650	0.828	1.001	0.812
0.13	623	0.975	0.845	0.650	0.789	0.654	0.864	1.010	0.798
0.14	297	0.929	0.799	0.643	0.781	0.638	0.867	1.008	0.776

Table 13: Variances of betting strategies on Over and Under

8. Summary

The market prices of soccer bets represent the market participants' assessments of event probabilities. The most important bet types for live soccer betting are Asian handicap bets for bets on teams and Over Under bets for bets on the total number of goals. The mathematical properties of these markets do not allow the derivation of probability assessments for exact final scores, e.g. the probability of the game ending 3:1. The market prices only provide information on sums of probabilities, e.g. that the total number of goals is smaller than 3. As a result, it is only possible to compare the implied probabilities of two markets under certain conditions and it remains unclear how differences between markets can be exploited.

To overcome these restrictions we introduce the Implied Symmetry method. This method derives the implied probabilities of a market under the assumption of Poisson-distributed team strengths. The implied probabilities are condensed into one number by a numerical algorithm. The symmetrical handicaps and symmetrical Over Under lines allow for the identification of deviations of probability assessments across markets.

Using a dataset of market prices from 29,413 soccer games we demonstrate that the Implied Symmetry method selects the markets with the highest average return. It is consistently shown that the opposite strategy has the lowest average returns.

A noteworthy empirical observation is the untypical relation between risk and return. Normally, markets with lower variances, i.e. lower risks, have lower average returns than the markets with higher variances. Here, however, market participants do not demand higher returns for higher risks, as is normally expected in traditional financial markets. The results thus indicate that market participants are risk seeking and not risk averse.

Appendix A. Arbitrage Condition

Let $j = 1, \ldots, J$ denote the J bookmakers offering an Asian handicap market for a given handicap h at prices $\widetilde{q}_{j,H}$ and $\widetilde{q}_{j,A}$. Assume that the condition

$$\frac{1}{\frac{1}{\max_j \widetilde{q}_{j,H}} + \frac{1}{\max_j \widetilde{q}_{j,A}}} > 1$$

holds, and $\widetilde{q}_{k,H} = \max_j(\widetilde{q}_{j,H})$ and $\widetilde{q}_{n,A} = \max_j(\widetilde{q}_{j,A})$. If the amount

$$a_1 = \frac{\frac{1}{\widetilde{q}_{k,H}}}{\frac{1}{\widetilde{q}_{k,H}} + \frac{1}{\widetilde{q}_{n,A}}}$$

is bet on the home team at bookmaker k, and an amount of $a_2 = 1 - a_1$ is bet on the away team at bookmaker n, the payout $a_1\Pi_H + a_2\Pi_A \geq 1$ at least equals the investment $a_1 + a_2 = 1$ irrespective of the final score. The market offers an arbitrage opportunity and market participants realise a risk-free return.

Adell, J. A., Jodrá, P., 2005. The median of the Poisson distribution. Metrika 61, 337–346.

Cain, M., Law, D., Peel, D., 2003. The favourite-longshot bias, bookmaker margins and insider trading in a variety of betting markets. Bulletin of Economic Research 55, 263–273.

Cortis, D., 2015. Expected values and variance in bookmaker payouts: A theoretical approach towards setting limits on odds. Journal of Prediction Markets 9, 1–14.

Croxson, K., Reade, J. J., 2014. Information and efficiency: goal arrival in soccer betting. The Economic Journal 124, 62–91.

Dixon, M. J., Coles, S. G., 1997. Modelling association football scores and inefficiencies in the football betting market. Applied Statistics 2 (2), 265–280.

Dixon, M. J., Robinson, M. E., 1998. A birth process model for association football matches. The Statistician 47 (3), 523–538.

Franck, E., Verbeek, E., Nueesch, S., 2013. Inter-market arbitrage in betting. Economica 80, 300–325.

Hawkins, E., September 2014. Mugs and millionaires: Inside the murky world of professional football gambling. http://bleacherreport.com/articles/2200795-mugs-and-millionaires-inside-the-murky-w accessed: 2016-11-12.

Hwang, J. H., Kim, M.-S., 2015. Misunderstanding of the binomial distribution, market inefficiency, and learning behavior: Evidence from an exotic sports betting market. European Journal of Operational Research 243, 333–344.

Karlis, D., Ntzoufras, I., 2005. Bivariate Poisson and diagonal-inflated bivariate Poisson regression models in r. Journal of Statistical Software 14, 1–36.

Keogh, F., Rose, G., 2013. Football betting - the global gambling industry worth billions. http://www.bbc.com/sport/0/football/24354124, accessed: 2015-02-07.

Koning, R. H., 2003. An econometric evaluation of the firing of a coach on team performance. Applied Economics 35, 555–564.

Koning, R. H., Koolhaas, M., Renes, G., Ridder, G., 2003. A simulation model for football championships. European Journal of Operational Research 148, 268–276.

Kuypers, T., 2000. Information and efficiency: an empirical study of a fixed odds betting market. Applied Economics 32, 1353–1363.

Levitt, S. D., 2004. Why are gambling markets organised so differently from financial markets. The Economic Journal 114, 223–246.

Maher, M. J., 1982. Modelling association football scores. Statistica Neerlandica 36, 109–118.

Pantheon-Sorbonne, U. P., the International Centre for Sport Security, May 2014. Protecting the integrity of sport competition the last bet for modern sport an executive summary of the sorbonne-icss integrity report sport integrity research programme, 2012-14.

Tryfos, P., Casey, S., Cook, S., Leger, G., Pylypiak, B., 1984. The profitability of wagering on NFL games. Management Science 30, 123–132.

Vergin, R. C., Scriabin, M., 1978. Winning strategies for wagering on National Football League games. Management Science 24, 809–818.

Vlastakis, N., Dotsis, G., Markellos, R. N., 2009. How efficient is the European football betting market? Evidence from arbitrage and trading strategies. Journal of Forecasting 28, 426–444.