IIMB तेजस्वि नावधीतमस्तु
भारतीय प्रबंध संस्थान बंगलूर
INDIAN INSTITUTE OF MANAGEMENT BANGALORE

IMB 401

U DINESH KUMAR, V SANDEEP AND H SATYABALA

A GAME OF TWO HALVES: IN-PLAY BETTING IN FOOTBALL

Peter Smith has been a football (soccer) enthusiast since his schooldays. In 2012, he developed a keen interest in betting, particularly in-play betting, in which a bettor can place bets while the match is in progress. According to Peter:

> Unlike other gambling events and other forms of sports betting, in-play betting is a level player. Most gambling events set odds highly in favour of the gambling houses, whereas in in-play betting, the house edge is not completely biased towards the betting houses. A bettor can statistically analyze the situation before placing the bet.

Football (soccer) attracts not just fans, television audience, and sponsorships, but also betting. In 2012, betting was an important contributor to the economic value of any sports. Bwin, a sports betting company, offered 30,000 bets daily on 90 different sports in 22 different languages.[1] In 2011, Bwin's online betting revenue was USD 320 million. Bet365, a gambling company based in the United Kingdom, has been a leading online gambling house with an estimated customer base of 7 million across 200 countries. In 2012, Bet365 listed more than 30 sports across different continents in its betting options. According to the Fédération International de Football Association (FIFA), the sports industry itself generated revenue totalling USD 300 billion in 2011, while sports betting was worth USD 350–400 billion, with 50% of the bets placed in the gray market.[2] Sports betting has been an important service industry where every week, millions of people bet on various sporting events across the world. In 2012, the global market for online gambling was USD 21 billion, with Europe accounting for 43% of the world market.[3] In 2012, betting on sports events was legal in many European countries; however, it was illegal in many states in the United States. A sample list of countries where sports betting was legalized along with their 2012 online betting market share is shown in **Exhibit 1**. Even countries where sports' betting was banned, such as India, accounted for 0.5% of the market share in 2012, was ranked 40th out of 246 countries according to the website www.top100bookmakers.com.[4]

Many of the betting companies were also the official sponsors of a few sports clubs. In August 2012, Manchester United, one of the most popular football clubs in the world, signed a 3-year contract with Bwin, making it the club's official online betting and gaming partner.[5] Many football clubs in England were listed on the London Stock Exchange (LSE). There is evidence that the stock market reacts strongly to the outcomes of a match and generates abnormal returns and trading volumes.[6] According to behavioral finance researchers, sports betting firms benefit from the emotional bias of the bettors. As in many investment choices, home bias exists among sports bettors; that is, bettors have a tendency to bet in favor of their favorite team. The bias could also mean favorite teams, underdogs, and so on. The primary objective of this case is to analyze past data from the English Premier League (EPL) to develop winning strategies for in-play betting to verify Peter's belief.

Sports' betting has many different characteristics compared to other forms of gambling. For example, the odds are generally set in favor of the gambling house so that the house wins in the majority of instances; that is, gambling

[1] Source: https://home.bwin.com/page.aspx?view=aboutus (last accessed on January20, 2013).
[2] Source: FIFA, (March 26, 2011), Sports betting under the microscope, http://www.fifa.com/aboutfifa/organisation/footballgovernance/news/newsid=1406983/index.html (last accessed on January 20, 2013).
[3] Source: European Gaming & Betting Association Newsletter, Issue 10, (January 2012), http://www.egba.eu/pdf/EGBA-Newsletter-10.pdf (last accessed on January 20, 2013).
[4] Source: http://www.top100bookmakers.com/countries/in.php (last accessed on January 20, 2013).
[5] Source: BBC, (August 16, 2012), Manchester United in sponsorship deal with Bwin, http://www.bbc.co.uk/news/business-19284809 (last accessed on January 20, 2013).
[6] Source: Palomino F., Renneboog L., and Zhang C., (2009), Information salience, investor sentiment and stock returns: The case of British soccer betting, *Journal of Corporate Finance,* Vol. 15, 368–387.

V Sandeep and H Satyabala prepared this case for classroom discussion under the guidance of Professor U. Dinesh Kumar. This case is not intended to serve as an endorsement or a source of primary data, or to show the effective/inefficient handling of decision or business processes.

houses set odds such that the house has a much higher edge (known as house edge) over players and always wins in the long run. The probability of a gambler winning in a game of roulette (with 37 slots) is 2.7%; the house has a winning probability of approximately 97.3% in each game. The corresponding odds in favor of the gambler are 1:36; that is, if he/she plays roulette 37 times, he/she is likely to win once. Fixing the odds in sports is difficult since the outcome in sports events is unknown and varied. The focus of this case is on one of the popular betting types known as in-play betting. In-play betting is one of the fastest growing activities in sports betting. Richard Glynn, CEO at Ladbrokes (a British gaming company), stated in 2011 that in-play betting was a key battleground for betting operators.[7] Peter's objective was to determine optimal strategies that he could use for betting while the football match was in progress, based on the information available until half-time in a football match. Peter used data collected from four sessions between 2006 and 2011 of the EPL to test the optimal half-time betting strategies.

SPORTS BETTING

Sports betting started in the 17thcentury with horse racing in England.[8] From horse racing, it moved to many other sports, and being the most popular sport, football attracted a significant proportion of sports betting. Many authors have claimed that the rules of the sport were designed with betting interests in mind (Forrest and Simmons 2003).[9] There are more than 60 different betting possibilities in a single football match. The bets are not just about events in a match. In 2012, bookmakers (via Ladbrokes) offered a wager that Lionel Messi would play in the EPL at some point.[10] This bet was offered despite the fact that Messi's contract with the football club Barcelona carried a USD 312 million buyout clause after his contract was renewed in December 2012. A few popular betting types in football matches are listed below.

Win-Draw-Win	Predicting the match result (win for home team, draw, or win for the away team)
Half-time/Full-time	Predict the team will be leading at half-time, then again at full-time; draw can also be used as an option
Win both halves	Predict the team that will win both halves of the match
Correct score	Predict the full-time correct score for home and away teams
First goal scorer	Predict the player who will score the first goal in the match
Half-time result	Predict half-time result, win for home team, draw, or win for away team
Total goals	Predict total goals in a match
Team to score first goal	Predict the team that will score the first goal (includes no goals as one of the choices)
Half with most goals	Predict which half will have the most goals (with tie as an option)
To win either half	Predict which team will win either the first or the second half
In-play betting	Betting on match outcome or specific events when the match is in progress (in most instances, the bet may be placed any time during the match)

Pool betting (parimutuel betting) is another popular betting system in which bettors attempt to predict the outcomes of a set of matches that are played each week. The winners—usually those who predict at least 9 out of 10 match results correctly—share the prize money; if there are no winners, the betting is rolled over to the next week.

---

[7]Source: Gambling Data, (September 2011), In-play Betting Report, http://media.sportingindex.com/ZmsMedia/Sporting%20Solutions/Documents/GamblingData_inplayReport.pdf (last accessed on January 20, 2013).
[8]Source: Munting, R., *An economic and social history of gambling in Britain and the USA*. Manchester, UK/ New York: Manchester University Press, 1996.
[9]Source: Forrest, D. and Simmons, R., Sport and gambling, *Oxford Review of Economic Policy*, 2003, Vol. 19, 598–611.
[10] Source: Bailey, R., 5 Teams who could buy out Leo Messi's €250m release clause, *Bleacher Report*, December 19, 2012, http://bleacherreport.com/articles/1451018-5-teams-who-could-buy-out-leo-messis-eur250m-release-clause/page/4 (last accessed on January 20, 2013).

ODDS IN SPORTS BETTING

Many formats are used to convey the odds in football betting. The most popular formats are fractional odds, decimal odds (European odds), and American odds. Fractional odds, popular among bookmakers in the United Kingdom, are expressed in fractions; for example, A/B indicates that for every wager amount B, the bettor will gain amount A if he/she wins the bet. Fractional odds of A/B are also the total (A+B) that will be paid to the bettor if he/she wins. For example, if the fractional odds quoted by a bookmaker for Manchester United to win against Arsenal are 1/5, then the bettor will win one dollar for every 5-dollar bet on Manchester United to win. A decimal odd represents the potential return on a bet, including the betting amount. A decimal odd of 1.20 for Barcelona to win an *El Clasico* match against Real Madrid implies that if a bettor bets on Barcelona to win, he/she would gain 1.2 times the betting amount if the wager were successful. Decimal odds are relatively easier to understand than fractional odds and are popular among bookmakers in continental Europe, which is why they are commonly referred to as European odds. American odds, also known as money line odds, use positive and negative values for favorites and underdogs, respectively. For example, if the New York Yankees are the favourite to win a match and a bookmaker offers odds of –120, the bettor has to risk USD 120 to make a profit of USD 100. If the Texas Rangers are the underdog in a match and a bookmaker offers odds of +120, a bettor who bets USD 100 on the Texas Rangers will gain USD 120 if the Texas Rangers win. A comparison of all three odds with same rate of returns is shown in **Exhibit 2**. Most bookmakers provide odds in all three formats and a bettor can choose the format with which he/she is most comfortable.

IN-PLAY BETTING

In-play betting allows bettors to place their bets while the match is in progress. In most instances, the bettors can place bets any time during the match and can bet on the final score, match result, next goal scorer, and so on. Unlike other bets, in-play betting odds are updated in real-time by the betting companies. In-play betting also provides an opportunity to the bettors to mitigate the risk on their other bets. For example, consider a bettor who has placed a bet on a particular team to win before the start of the play. During the match, if the other team leads, he/she can mitigate the risk of losing money by placing another in-play bet. The success of in-play betting can be attributed to the benefit of informed choice that a bettor can utilize before placing the bet. William-Hill, an online betting firm, claimed that in-play betting grew by 95% in the first half of 2011 compared to the preceding year, accounting for more than 63% of all sports bets excluding horse racing.[11] Ladbrokes claimed that during non-horse racing, in-play betting accounted for 49% of all bets in 2011, an increase of 7.7% from 2010.[12] A screenshot of the in-play betting odds for two of the matches obtained from the website of an online betting firm is shown in **Exhibit 3**.

ENGLISH PREMIER LEAGUE DATA

The English Premier League has been one of the most popular football tournaments in terms of viewership and economic value. The EPL has been contested by 20 teams and 380 matches are played every year between August and May. Minute-by-minute data on various parameters was collected for 1,520 EPL matches played between 2006 and 2011 up to half-time. The data was publically available on the Internet.[13] The performance of the home team, which was measured using percentages of wins, draws, and losses, is presented in **Exhibit 4.** The performance of the home team for different half-time goal differences (Number of goals scored by home team at half-time – Number of goals scored by the away team at half-time in a match) is shown in **Exhibit 5.**The in-play betting problem can be modeled as a classification problem in which each match outcome can be classified into three classes: win, draw, and loss for the home team. In this instance, we have used two classification techniques, namely, multinomial logistic regression and chi-square automatic interaction detection (CHAID). The covariates that are used in the multinomial regression and CHAID models are listed in **Exhibit 6**. The multinomial logistic regression model developed using the covariates listed in **Exhibit 6** is shown in **Exhibit 7**. Loss to the home team (coded as 0) is used as the reference category. The multinomial regression model in **Exhibit 7** can be used to take appropriate decisions during in-play betting. The classification table obtained from the multinomial regression is shown in **Exhibit 8**. The

[11] Source: Gambling data, (September 2011), In-play Betting Report, http://media.sportingindex.com/ZmsMedia/Sporting%20Solutions/Documents/GamblingData_inplayReport.pdf (last accessed on January20, 2013).
[12] *Ibid.*
[13] Data was collected from two websites: http://www.statto.com/football/stats/england/premier-league/2006-2007/custom-table and http://www.pedwards.co.uk/ (last accessed on January20, 2013).

classification tree obtained using the CHAID model is shown in **Exhibit 9**. **Exhibit 10** provides a list of 20 EPL matches played over two weekends in 2012, with the values of covariates and the outcome of each match.

STUDY QUESTIONS

Instructions: Use **Exhibit 7** to answer Questions 1–4.

1. Write the equation that Peter can use for predicting the probability of win for the home team (coded as 2) using statistically significant variables (use $\alpha = 0.05$).
2. What is the influence on the match output of red cards conceded by the home and the away team? Discuss the possible reasons for the empirical evidence from the model.
3. Is it relevant to use the points scored by a team in the previous season for predicting the outcome of a match?
4. What is the probability that the home team will win the match for the values shown in the following table?

HTGD	FGS	RED_H	RED_A	POINTS_H	POINTS_A	TOTAL_H_P	TOTAL_A_P
2	1	0	0	15	18	40	30

5. If the first goal is scored by the away team, is it advisable to bet in favor of the away team? Answer by controlling for all the other variables in the regression model.
6. What conclusions can you derive from the classification table shown in **Exhibit 8**? Is it advisable to bet on draws (based on the model developed)?
7. Using the CHAID decision tree shown in **Exhibit 9**, frame rules that may be used for betting.
8. **Exhibit 10** lists 20 matches played over two weekends in 2012 along with the values of the covariates. Use multinomial logistic regression to predict the match outcome in all 20 cases listed in **Exhibit 10**.
9. Apply the CHAID decision tree on the 20 matches listed in **Exhibit 10** and compare the results with your answers obtained using multinomial logistic regression.
10. If Peter were to choose one match from the list of 20 matches for betting, which match should he choose? Discuss the reasons for your suggestion.

Exhibit 1: Sample list of countries where sports betting is legal along with their online betting market share in 2012[14]

S. No.	Name of the country	Size of online betting market (%)
1	Australia	2.5
2	Belgium	1.4
3	Czech Republic	1.7
4	Denmark	0.3
5	France	0.9
6	Germany	3.4
7	Italy	5.0
8	South Africa	0.3
9	Spain	1.6
10	United Kingdom	18

Exhibit 2: Comparison of fractional, decimal, and American odds with same returns

Team	Fractional Odds	Decimal Odds	American Odds
Barcelona	1/5	1.20	–500
Manchester United	1/2	1.50	–200
New York Giants	1/1	2.00	(±)100

[14] Source: http://www.top100bookmakers.com/countries/us.php (last accessed on January 20, 2013).

Exhibit 3: Screenshot of sample in-play betting odds from an online betting firm

Exhibit 4: Home team performance

Home team performance	Matches	Percentage
WIN	724	47.63158
DRAW	404	26.57895
LOSS	392	25.78947
Total	1520	100

Source: Values calculated based on the data collected by the authors.

Exhibit 5: Home team performance based on half-time goal difference (HTGD)

HTGD	4	3	2	1	0	–1	–2	–3	–4	–5	Total
WIN	4	37	120	293	231	36	3	0	0	0	724
DRAW	0	1	6	62	258	73	3	0	1	0	404
LOSS	0	0	5	23	142	158	54	8	1	1	392
Total	4	38	131	378	631	267	60	8	2	1	1520

Source: Values calculated from data collected by the authors.

Exhibit 6: List of covariates and their description

Covariate Name	Description	Data Type
HTGD	Half-time goal difference (Number of goals scored by home team at half-time – Number of goals scored by the away team at half-time)	Numerical
FGS	First goal scored	Categorical 1: Home Team 0: Away team 2: No goals scored
RED_H	Red cards conceded by the home team	Numerical
RED_A	Red cards conceded by the away team	Numerical
POINTS_H	Points earned by the home team in the league until that match	Numerical
POINTS_A	Points earned by the away team in the league until that match	Numerical
TOTAL_H_P	Total points earned by the home team in the previous season	Numerical
TOTAL_A_P	Total points earned by the away team in the previous season	Numerical
MATCH_O	Match outcome	Categorical 2: Home team won 1: Draw 0: Away team won

Exhibit 7: Multinomial logistic regression output from SPSS (2 for WIN, 1 for DRAW, and 0 for LOSS)

Parameter Estimates									
								95% Confidence Interval for Exp(B)	
Match_O[a]		*B*	Standard Error	Wald	*df*	Sig.	Exp(B)	Lower Bound	Upper Bound
1	Intercept	3.535	0.462	58.524	1	0.000			
	RED_H	0.301	0.572	0.276	1	0.599	1.351	0.440	4.146
	RED_A	0.463	0.540	0.736	1	0.391	1.589	0.552	4.575
	POINTS_H	0.024	0.009	7.318	1	0.007	1.024	1.007	1.042
	POINTS_A	–0.018	0.008	5.055	1	0.025	0.982	0.967	0.998
	HTGD	0.511	0.120	18.108	1	0.000	1.667	1.318	2.110
	TOTAL_H_P	0.000	0.004	0.003	1	0.960	1.000	0.993	1.007
	TOTAL_A_P	–.010	0.004	7.169	1	0.007	0.990	0.982	0.997
	[FGS=0]	–3.521	0.410	73.892	1	0.000	0.030	0.013	0.066
	[FGS=1]	–2.819	0.426	43.776	1	0.000	0.060	0.026	0.137
	[FGS=2]	0[b]	.	.	0	.	.	.	.
2	Intercept	3.313	0.470	49.620	1	0.000			
	RED_H	–0.811	0.743	1.189	1	0.275	0.445	0.104	1.908
	RED_A	0.983	0.547	3.237	1	0.072	2.673	0.916	7.802
	POINTS_H	0.035	0.009	14.590	1	0.000	1.036	1.017	1.055
	POINTS_A	–0.035	0.009	16.160	1	0.000	0.966	0.950	0.982
	HTGD	1.618	0.143	127.225	1	0.000	5.045	3.808	6.683
	TOTAL_H_P	0.010	0.004	7.227	1	0.007	1.010	1.003	1.018
	TOTAL_A_P	–0.015	0.004	13.788	1	0.000	0.985	0.978	0.993
	[FGS=0]	–3.320	0.413	64.555	1	0.000	0.036	0.016	0.081
	[FGS=1]	–2.473	0.430	33.080	1	0.000	0.084	0.036	0.196
	[FGS=2]	0[b]	.	.	0	.	.	.	.

[a] The reference category is 0.
[b] This parameter is set to zero because it is redundant.

Exhibit 8: Classification table output for the multinomial logistic regression

Observed	Predicted			
	0	**1**	**2**	**Correct Percentage**
0	313	10	69	79.8
1	120	103	181	25.5
2	75	78	571	78.9
Overall Percentage	33.4	12.6	54.0	64.9

Source: Output from SPSS software based on the data collected by the authors.

Exhibit 9: Chi-Squared Automatic Interaction Detection (CHAID) output from SPSS

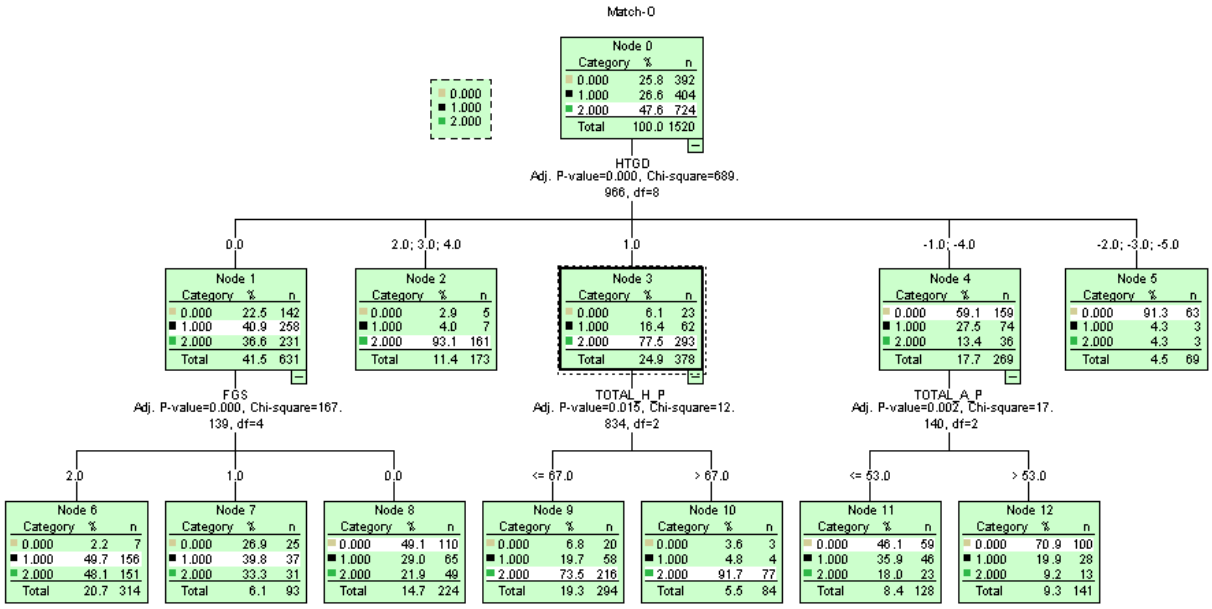

Exhibit 10: Twenty EPL matches played over two weekends in 2012 with the values of covariates and the match outcome

Match Number	Match b/w	HTGD	FGS	RED_H	RED_A	POINTS_H	POINTS_A	TOTAL_H_P	TOTAL_A_P	MATCH_O
1	Swansea City vs. Everton	–2	0	0	0	7	7	47	56	0
2	Chelsea vs. Stoke City	0	2	0	0	10	4	64	45	2
3	Southampton vs. Aston Villa	–1	0	0	0	0	4	0	38	2
4	West Bromwich Albion vs. Reading	0	2	0	0	7	1	47	0	2
5	West Ham United vs. Sunderland	–1	0	0	0	7	3	0	45	1
6	Wigan Athletic vs. Fulham	–1	0	0	0	4	6	43	52	0
7	Liverpool vs. Manchester United	0	2	1	0	2	9	52	89	0
8	Newcastle United vs. Norwich City	1	1	0	0	5	3	65	47	2
9	Manchester City vs. Arsenal	1	1	0	0	8	8	89	70	1
10	Tottenham Hotspur vs. Queens Park Rangers	–1	0	0	0	5	2	69	37	2
11	Arsenal vs. Chelsea	0	0	0	0	9	13	70	64	0
12	Everton vs. Southampton	2	0	0	0	10	3	56	0	2
13	Fulham vs. Manchester City	0	1	0	0	9	9	52	89	0
14	Norwich City vs. Liverpool	–2	0	0	0	3	2	47	52	0
15	Reading vs. Newcastle United	0	2	0	0	1	8	0	65	1
16	Stoke City vs. Swansea City	2	1	0	0	4	7	45	47	2
17	Sunderland vs. Wigan Athletic	0	2	0	0	4	4	45	43	2
18	Manchester United vs. Tottenham Hotspur	–2	0	0	0	12	8	89	69	0
19	Aston Villa vs. West Bromwich Albion	0	2	0	0	4	10	38	47	1
20	Queens Park Rangers vs. West Ham United	–2	0	0	0	2	8	37	0	0

Source: http://soccernet.espn.go.com/results//date/20121001/league/eng.1/barclays-premier-league?cc=4716 (last accessed on January20, 2013)