

Exploring the Efficient Market Hypothesis through Time; Evidence from the Total Goals Football Betting Market



Anastasios Oikonomidis, Johnnie Johnson
University of Southampton

Objective

Analyze football data to test whether market prices are driven towards efficiency in cases where *obvious* inefficiencies have existed

The Efficient Market Hypothesis (EMH)

- In an efficient market prices fully reflect any relevant information

Fama (1970)

- Anomalies tend to disappear in the long-term

Fama (1997)

The EMH through Time

Often Inefficiencies apparent in several time periods disappear later:

- Movement of prices towards weak form efficiency (Toth and Kertesz 2005 - NYSE, Jefferis and Smith 2005 - Asian Markets)
- Returns on past winners are high only for a limited period (De Bondt and Thaler 1985, 1987, Chan, Jegadeesh and Lakonishok 1996)

Investors become aware of value strategies and adjust their portfolios, driving market prices towards efficiency; Chan and Lakonishok (2004)

The EMH in Betting Markets

In wagering markets each asset / bet has a well-defined termination point at which its value becomes certain

Thaler and Ziemba (1988)

The EMH through Time in Football Betting

- Pope and Peel (1988) Cain Law and Peel (2000) conclude that betting on favourites would be profitable in seasons 1981/1982 and 1991/1992 respectively. Opposite evidence is presented by Cain, Law and Peel (2003) for the season 1992/1993
- Forrest, Goddard and Simmons (2005) test a statistical model to predict match outcomes in period 1998-2003 and find that it produces superior results to market probabilities in the early years only

Total Goals

- Among the most popular markets in today's football fixed-odds betting
- Over: Total goals >2 | Under: Total Goals < 3
- Fixed-odds; a winning bet yields the same profit independently from how far the outcome is from 2.5 goals

Fixed-Odds

In an efficient fixed-odds betting market, betting on each alternative (j) of event (i) has fixed negative expected profit determined by the level of market's trading costs (c).

$$E(\textit{Profit}_{ij}) = -\frac{c}{c+1}$$

The trading cost (over-round) of betting on event (i) can be calculated through the returns (X_{ij}) offered for the (j) alternatives of event (i)

$$c = \sum_{j=1}^n \frac{1}{X_{ij}} - 1$$

Data

- The dataset consists of about 33,000 games
- Eighteen European football betting markets are investigated in order to test whether possible inefficiencies are homogeneously expressed
- The sample covers seven years (2002-2008) so that the persistence of possible inefficiencies can be tested
- Gamebookers [a significant internet operator (customer base of 140,000 punters)] is the company whose odds are analyzed

Over / Under Bias

Conditional logistic regression is employed, Mcfadden (1979). The probability that the alternative (j) of event (i) is successful:

$$P(Y_{ij} = 1) = \frac{e^{Z_{ij}}}{\sum_{j=1}^n e^{Z_{ij}}}$$

$$Z_{ij} = b_0 \ln p_{ij} + b_1 d_{unij}$$

	Coef.	Std.Err.	z	P>z	[95% Confidence]	
success						
lnp	1.152811	0.048054	23.99	0	1.058628	1.246994
under	0.0794471	0.011848	6.71	0	0.056226	0.102668

The event specific dummy variable 'under' derives a significantly positive coefficient. The market underestimates the probability of fewer than 3 goals to be scored in a football game

Over / Under Bias by League

success	Coef.	Std.Err.	z	P>z	[95% Confidence]	
Inp	1.067	0.073	14.65	0.000	0.924	1.210
Greece*under	0.222	0.057	3.87	0.000	0.109	0.334
France1*under	0.196	0.049	4.02	0.000	0.100	0.291
Portugal*under	0.185	0.057	3.26	0.001	0.074	0.295
Belgium*under	0.139	0.047	2.99	0.003	0.048	0.230
Italy2*under	0.127	0.048	2.63	0.009	0.032	0.221
Spain2*under	0.104	0.048	2.18	0.030	0.010	0.198
France2*under	0.075	0.052	1.44	0.151	-0.027	0.178
Spain1*under	0.070	0.043	1.63	0.104	-0.014	0.155
Germany2*under	0.070	0.053	1.32	0.187	-0.034	0.174
Scotland2*nder	0.069	0.086	0.8	0.422	-0.100	0.239
Scotland1*under	0.063	0.057	1.11	0.268	-0.048	0.175
England2*under	0.061	0.035	1.73	0.083	-0.008	0.131
England1*under	0.052	0.043	1.22	0.223	-0.032	0.135
Germany1*under	0.019	0.049	0.39	0.699	-0.077	0.115
Italy1*under	0.017	0.045	0.38	0.704	-0.071	0.105
Netherlands*under	0.011	0.052	0.22	0.828	-0.091	0.114
Turkey*under	-0.003	0.055	-0.06	0.951	-0.110	0.104

Focusing on Six Leagues

The event specific dummy 'under' is split into two binary variables to distinguish the leagues in which the inefficiency is more pronounced

success	Coef.	Std.Err.	z	P>z	[95% Confidence]	
lnp	1.100329	0.051288	21.45	0	0.999807	1.200852
under1	0.151041	0.020873	7.24	0	0.110131	0.191951
under2	0.0461467	0.014278	3.23	0.001	0.018162	0.074131

Through a likelihood ratio test, the hypothesis that no significant difference between the 6 leagues and the rest in terms of over-under bias exists is rejected, LR chi2=17.39

Betting on 'Under'

<i>Greece</i>		Bets	Profit
Under		1295	0.008
Over		1295	-0.200
<i>France</i>			
Under		2214	-0.004
Over		2214	-0.195
<i>Italy II</i>			
Under		1953	-0.025
Over		1953	-0.156
<i>Spain II</i>			
Under		2301	-0.043
Over		2301	-0.153
<i>Portugal</i>			
Under		1335	-0.007
Over		1335	-0.181
<i>Belgium</i>			
Under		1989	-0.024
Over		1989	-0.141

- Betting on 'under' yields significantly less loss than betting on 'over' in the total sample
- In Greece and France, betting on 'under' breaks even
- In the rest of the countries betting on 'under' yields losses

<i>All</i>		Bets	Profit
	Under	11087	-0.018
	Over	11087	-0.169

The Over / Under Bias through Time

Splitting the sample into two sub-samples according to the date of the game shows that 'under' used to be profitable in the earlier years but the situation changed later

	1st Half	2nd Half
Under	0.01	-0.04
Over	-0.21	-0.13

	1st Half	2nd Half
France	0.04	-0.06
Greece	0.03	-0.02

An interaction term between the days elapsed since the first game in the sample and the event specific dummy 'under' indicates the decreasing trend in the bias, LR chi2 = 10.67

success	Coef.	Std.Err.	z	P>z	[95% Confidence]	
Inp	1.073024	0.085886	12.49	0	0.90469	1.241358
Under	0.2358712	0.031361	7.52	0	0.174405	0.297338
Under*days^2	-4.88E-08	1.49E-08	-3.27	0.001	-7.80E-08	-1.95E-08

Frequency of 'Under' through Time

Is there a decreasing frequency in the occurrence of 'under'?

Year	2002	2003	2004	2005	2006	2007	2008
f(success)	0.555	0.565	0.557	0.595	0.599	0.56	0.59

An interaction term between elapsed days and 'under' does not add forecasting power to the model if the odds are excluded, indicating no trend in the frequency of the event

success	Coef.	Std.Err.	z	P>z	[95% Confidence]	
Under	0.2613672	0.037636	6.94	0	0.187603	0.335132
Under*days	0.0000376	3.02E-05	1.24	0.214	-2.2E-05	9.68E-05

The betting opportunity offered by betting on under did not vanish because under happens more rarely than it used to

The Odds for 'Under' through Time

Is there a decreasing trend in the level of odds on 'under'?

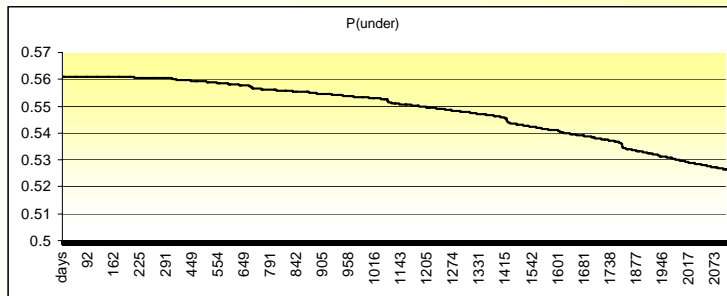
Year	2002	2003	2004	2005	2006	2007	2008
M(odds)	1.79	1.79	1.74	1.73	1.65	1.68	1.66

Regressing the odds probability for under on a quadratic combination of the elapsed days documents the decrease in the level of odds through time

P(under)	Coef.	Std.Err.	z	P>z	[95% Confidence]	
days	0.0000377	3.83E-06	9.85	0	3E-05	0.0000452
days2	-7.87E-09	1.72E-09	-4.58	0	-1.1E-08	-4.5E-09
_cons	0.5610379	0.001858	301.93	0	0.5574	0.5646802

Moving towards Efficiency

During the sample the odds for 'under' decreased significantly, moderating the degree of the bias and vanishing any opportunity for profit



Year	M(success)	M(odds)
2002	0.555	0.52514
2003	0.565	0.52514
2004	0.557	0.54023
2005	0.595	0.543353
2006	0.599	0.569697
2007	0.56	0.559524
2008	0.59	0.572289

This decrease is not expected to continue in the same form as most of the bias is already removed

Inp	1.155871	0.195616	5.91	0	0.772471	1.53927
under	0.035939	0.056175	0.64	0.522	-0.07416	0.146041

Conclusion

- The probability of the 'under' event is underestimated by the market in this sample, leading to an over-under bias
- The degree of the bias is not homogeneously expressed through time. It was more pronounced in earlier years, leaving space for profit
- The market moved towards efficiency, because market makers noticed the bias and corrected their odds, even though no change in the frequency of the event was documented

These findings support conclusions derived from the literature of stock markets which suggest that inefficiency is a temporary state in markets and therefore should not be considered as regularity