Gone in 60 seconds: the absorption of news in a high-frequency betting market

Babatunde Buraimo, *University of Central Lancashire*
David Peel
Rob Simmons

Available at: https://works.bepress.com/babatunde_buraimo/17/
Gone in 60 Seconds: The Absorption of News in a High-Frequency Betting Market

Babatunde Buraimo¹, David Peel² and Rob Simmons²*

¹University of Central Lancashire, ²Lancaster University

March 2008

Abstract: This paper tests for efficiency in a betting market that offers high-frequency data, the Betfair betting exchange for wagering on outcomes of English Premier League soccer matches. We find clear evidence of rapid adjustment of prices to large disturbances (news). Full adjustment takes place within a one minute interval after the news. This suggests that this particular wagering market is not just efficient at pre-match prices but is also efficient in the face of events within games.

Keywords: betting, efficiency, news

JEL classification: G13, L13, L83

* Corresponding author: Department of Economics, The Management School, Lancaster University, Lancaster LA1 4YX, UK, Phone 0044 1524 594234, fax 0044 1524 594244, Email: r.simmons@lancaster.ac.uk.
Introduction

As is now well recognized wagering markets have a very convenient structure for testing hypotheses on the formation of expectations or market efficiency. As Thaler and Ziemba (1988) pointed out in an early contribution on wagering markets each asset or bet has a well-defined termination point at which its value becomes certain. This structure radically simplifies the pricing problem since, for example, the property obviates the need to measure future fundamentals. The purpose of this paper is to examine market efficiency by employing high frequency data in a particular betting market. There have been numerous studies of the properties of the probabilities derived from fixed odds bets (see Forrest, 2008 for a survey). But relatively few of these studies involve high frequency data. In an innovative paper, Gil and Levitt (2008) studied futures contracts traded in real time as football games were played in the 2002 World Cup. To test the response of these markets to a goal being scored, they analyzed the pattern of prices in the 30 minute window bracketing each score (15 minutes before and 15 minutes after the goal). They found that the prices in soccer betting markets moved sharply when a goal was scored, but that prices 10-15 minutes after the goal is scored were statistically significantly higher than immediately after the goal. This is indicative of market inefficiency.

However their analysis has some questionable features. First, trading occurred somewhat infrequently so that data were highly aggregated. Second their trading data were recorded in Greenwich Mean Time, whereas the timing of goals is based on the minutes of play that have elapsed in the game. This led to discrepancies of one to two minutes. Given that the null hypothesis of efficiency embodies essentially instantaneous reaction this introduces potentially important measurement errors. Third, and most importantly, they treated goals as homogenous news items so they analyzed the pattern of prices in the 30

---

1 For studies of North American sports wagering markets ‘in the running’ see Hartzmark and Solomon (2007) and Hakes, Sauer and Waller (2005).
minute window bracketing each score (15 minutes before and 15 minutes after the goal) employing regression analysis. Unfortunately goals scored are not homogenous from the perspective of news. A goal in the first minute is not the same as a goal in the last minute, ceteris paribus. Similarly a goal scored when 3-0 up is not the same as a goal scored when 1-1. Similar issues relate to other news items, such as sending offs, cautions and corner kicks. The news in a particular event has to be measured relative to the probabilities prevailing at the time of the news impact. Its degree of surprise is the issue and we set out an appropriate model below. Given this we follow the standard test procedures of efficiency. We set out the various tests below. These include response to new information and also response to the size of new information.

**Some Analysis**

In an efficient market the revisions to the rational expectations of an event with a fixed terminal horizon should be unpredictable and due to new information (Samuelson, 1965). This must also be the case for probabilities. However the process will be heteroskedastic due to the bounded nature of the probabilities. Following Makropoulou and Markellos (2007) we will assume that the true probability follows a continuous process. Given that the probability can only take values in the range \([0, 1]\), assuming a geometric Brownian motion over the duration of the match would be inappropriate, since this would imply that true probabilities can take values in the range \([0, \infty)\). Suppose instead that the true probability \(p\) of the outcome evolves according to the following stochastic equation:

\[
dp = p(1-p)u
\]  

(1)
where $p$ is the probability and $u$ is news.$^2$

This equation is heteroskedastic and has boundaries at zero and unity. The level of the probability is seen to exhibit a unit root. An important feature of (1) is that the impact of news on the probability is not independent of the probability at the time of the news event. Intuitively if $p$ is close to zero a given magnitude of good news, positive $u$, will have less impact on the change in the probability if we are very close to unity as opposed to say 0.5. This is the aspect that Gil and Levitt neglected in their analysis.

The discrete representation of (1) we employ in our empirical analysis, $Y$, is given by

$$Y_i = \frac{p_t - p_{t-1}}{p_{t-1}(1 - p_{t-1})} = u_i$$

(2)

This transformation will be random under the null of efficiency.

One important advantage of tests of efficiency in our framework is that we can more accurately measure the sign and magnitude of news under the null of efficiency. As pointed out by Gil and Levitt, tests of how quickly new information is incorporated quickly and fully into markets involve analysis of special cases in which new information becomes available to market participants at an identifiable point in time such as announcements about macroeconomic variables in the context of prices on stock markets (Pearce and Roley 1985, Kliesen and Schmid (2006). However, these tests have limitations. The survey data underlying the expectations is often dated for example and

$^2$ Note if the probability of an outcome is $p = \frac{e^{bx}}{e^{bx} + e^{cz}}$ where $b$, $c$ are constants, $x$, $z$ variables. If we assume $dx = random$ error and $dz = random$ error we obtain (1).
its precise mapping into market expectations unclear. In other areas some agents may have inside information. (Jarrell and Poulsen (1989)).

Measuring the magnitude of the news item is another important feature of our analysis. In their classic study De Bondt and Thaler (1985) suggested that agents systematically overreacted to unexpected and dramatic news events, resulting in substantial weak-form inefficiencies in the stock market (see also Barberis et al (1998) De Bondt and Thaler (1990) and Bremer and Sweeney (1991)). Sturm (2003) suggests that investors respond differently to negative price shocks than to positive price shocks. In particular, large price decreases generally drive positive post-event abnormal returns, while large price increases do not drive positive or negative abnormal returns.

Data and Hypotheses

Our data come from Betfair.com and comprise betting odds on outcomes (home win, away win and draw) of 296 football matches in the 2006/07 and 2007/08 English Premier League seasons. Betting odds are available for each second of a match, in real time, so long as the market is running, i.e. not suspended. Betfair is not an orthodox bookmaker; rather it is a market maker, allowing gamblers to take positions, either side of the market, that accept a bet or lay off a bet. We focus here on match outcomes and not secondary betting activity on specific events, such as identity of goal scorers or dismissed players. Whereas high street bookmakers offer an over-round on fixed-odds bets on football matches of around 12 per cent (Forrest, 2008), Betfair has a much smaller over-round of just 1.82 per cent over our sample of 296 games. Bettors pay a small commission fee to successful punters, charged at 0.5 per cent of gains. Bets are not

---

3 Also the survey data measure of news is only a measure of the outcome relative to its expectation. It does not measure the news that can occur due to future revisions to expectations of future outcomes due to the announcement. For instance current inflation above expectation in Japan would be good news in Japan in the 80’s but bad news in US.
subject to tax. As Table 1 shows the volume of betting activity on this market exchange is considerable, although it varies widely across games.

Table 1. Descriptive statistics of betting volumes for 296 English Premier League games 2006-08

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Before game</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home probability</td>
<td>691,165</td>
<td>894,705</td>
<td>22,680</td>
<td>5,336,772</td>
</tr>
<tr>
<td>Away probability</td>
<td>446,545</td>
<td>775,249</td>
<td>10,117</td>
<td>5,643,094</td>
</tr>
<tr>
<td>Draw</td>
<td>86,065</td>
<td>98,423</td>
<td>7,812</td>
<td>615,127</td>
</tr>
<tr>
<td>Total</td>
<td>1,223,774</td>
<td>1,062,825</td>
<td>66,629</td>
<td>6,163,497</td>
</tr>
<tr>
<td><strong>During game</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home probability</td>
<td>840,941</td>
<td>1,012,245</td>
<td>7,611</td>
<td>5,033,787</td>
</tr>
<tr>
<td>Away probability</td>
<td>588,178</td>
<td>956,165</td>
<td>1,635</td>
<td>6,028,895</td>
</tr>
<tr>
<td>Draw</td>
<td>327,467</td>
<td>467,017</td>
<td>2,494</td>
<td>2,780,773</td>
</tr>
<tr>
<td>Total</td>
<td>1,756,586</td>
<td>1,582,990</td>
<td>26,275</td>
<td>8,552,785</td>
</tr>
</tbody>
</table>

Table 2. Descriptive statistics of betting probabilities for 296 English Premier League games 2006-08

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Before game</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home probability</td>
<td>0.450</td>
<td>0.195</td>
<td>0.066</td>
<td>0.902</td>
</tr>
<tr>
<td>Away probability</td>
<td>0.289</td>
<td>0.172</td>
<td>0.022</td>
<td>0.757</td>
</tr>
<tr>
<td>Draw</td>
<td>0.261</td>
<td>0.048</td>
<td>0.076</td>
<td>0.327</td>
</tr>
<tr>
<td><strong>During game</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home probability</td>
<td>0.445</td>
<td>0.284</td>
<td>0.001</td>
<td>0.997</td>
</tr>
<tr>
<td>Away probability</td>
<td>0.280</td>
<td>0.256</td>
<td>0.001</td>
<td>0.995</td>
</tr>
<tr>
<td>Draw</td>
<td>0.275</td>
<td>0.155</td>
<td>0.002</td>
<td>0.991</td>
</tr>
</tbody>
</table>

An important operational feature of the Betfair market is suspension of trading period at various points within games. It is common for news to be associated with suspensions. But in some instances, suspensions occur with no obvious emergence. Let us now define ‘news’. News is based on the size of the disturbance term from the regression:

\[ Y_t = \alpha_0 + \alpha_1 Y_{t-1} + u_t \]  

(3)

where \( Y_t = \frac{p_t - p_{t-1}}{p_{t-1}(1 - p_{t-1})} \) and \( p_t \) is the home win probability.
News is taken to be the smallest 0.025% (i.e. negative values), and the largest 0.025% values (positive values) of $\mu$. Our analysis is based on the period immediately after news emerges (post-news period).\(^4\)

Given the above definition of news, there are 1,445 newsbreaks in the dataset. This represents an average of 4.88 news items per match for all 296 games. There are 1,635 trading suspensions, of which 454 were because of breaking news in the previous second.

To make our discussion of news more concrete, consider a particular game, Arsenal versus Aston Villa in August 2006. Arsenal was the team favoured to win. The horizontal axis shows real time with the game commencing at 15.00 hours GMT. The vertical lines denote the half-time interval. The game ended as a 1-1 draw and there were no cautions or dismissals. We see three items of news. The first, just after kick-off cannot be accounted for by a specific event as far as we can tell. The second, at 53 minutes of play, registers an away goal by Aston Villa. As would be expected, the away win probability rises while the home win probability falls. The third piece of news is an equalizing goal by Arsenal at 84 minutes play. Now the home probability rises but not by as much as the away probability falls. The draw probability shows a large increase that persists until the end of the game, reflecting the limited amount of time left in the game for either team to score a winning goal. Tactically, as the home team, Arsenal would have been more likely than Aston Villa to make efforts to score again.

\(^4\) Our analysis rests on the assumption that both $\gamma$ and the primitive betting probability contain unit roots in levels and are stationary in differences. We performed Perron-Phillips unit root tests which verified that this in indeed the case.
Our analysis enables us to test for weak form and strong form efficiency. For weak form efficiency we examine the time series properties of $y_i$ in equation 2 above, following large positive or negative movements in the residual $u$. Specifically, we test for $a_i = 0$ as the null of market efficiency i.e. the probabilities of match outcomes should not be serially correlated following news. We perform this test for all items of news across games and for several distinct time periods after the news breaks, from 2 minutes down to 2 seconds. For short time periods, we expect persistence in our constructed probability variable. For longer periods, we expect persistence to be absent as news is fully incorporated into betting probabilities. Note that this approach does not involve use of dubious econometric methods to perform a regression of match outcomes against bookmaker probabilities. Methods such as linear probability modeling and probit regression are not well-suited to test for market efficiency in sports betting markets, for a number of reasons as Forrest (2008) makes very clear. Nor do we need to measure and
model the fundamentals of team performance characteristics that determine match results, as attempted by Goddard (2005). All we need to do is examine the properties of Betfair probabilities in response to news.

To test for strong-form efficiency we consider a variety of trading rules based on trading at different points in the match, e.g. betting favorite at beginning of match, after large positive or negative news. We ask whether the returns from within-game rules, applied following news, lead to superior or inferior gains compared with a benchmark pre-match position.

Results

Table 3. Results of regressions of $y_t = \alpha_0 + \alpha_1 y_{t-1} + u_t$ for different post-news time periods

<table>
<thead>
<tr>
<th>Post-news period: seconds</th>
<th>Coefficient</th>
<th>Absolute $t$ statistic</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td>-0.0034</td>
<td>1.48</td>
<td>85,270</td>
</tr>
<tr>
<td>60</td>
<td>-0.0033</td>
<td>1.51</td>
<td>45,228</td>
</tr>
<tr>
<td>30</td>
<td>-0.0034</td>
<td>1.56</td>
<td>24,717</td>
</tr>
<tr>
<td>10</td>
<td>-0.0036</td>
<td>1.76</td>
<td>8,972</td>
</tr>
<tr>
<td>5</td>
<td>-0.0039</td>
<td>2.00</td>
<td>4,699</td>
</tr>
<tr>
<td>2</td>
<td>-0.0046</td>
<td>2.32</td>
<td>1,940</td>
</tr>
</tbody>
</table>

There are six regression results and these reflect the size of the post-news period: 2 minutes; 1 minute; 30 seconds; 10 seconds; 5 seconds and 2 seconds. For the shorter periods, it should be noted that many observations are dropped as once the market resumes after a suspension, the post-news period may have elapsed i.e. if the market was suspended for 9 seconds, then this particular newsbreak will not feature in the regression results in which the post-news periods are 2 and 5 seconds. Our reported $t$ statistics are derived from standard errors that are clustered by match and robust to heteroskedasticity. We see that for 2 and 5 second intervals, the market is not efficient as previous prices do serve to inform future prices. But between 5 and 10 seconds, the
coefficient on the lagged dependent variable becomes insignificant at the conventional level. By 60 seconds, past prices do not usefully predict future prices and the news has been fully absorbed by the betting market. This represents very fast adjustment to news; this is a market with quick response of participants to large disturbances. An interesting supplementary question is whether this rapid adjustment of prices to news reflects sophistication by agents and/or a well-designed market system. We suspect that the Betfair market contains a higher incidence of knowledgeable and sophisticated players compared to the traditional bookmaker fixed-odds betting sector. Hence, efficiency in the Betfair market could co-exist with (short-term) inefficiency in the before-the-match fixed-odds market.

Turning to strong form efficiency, we can identify returns from betting on strategies to back home favourite, home underdog, away favourite and away underdog across 294 games in our sample. Prior to the match, the percentage returns form these strategies, net of 5% commission on games, are 4.87%, 16.95%, -11.7% and -31.6%. The volume of bets at a unit stake are 208, 86, 86 and 208 respectively. Hence, in this particular sample, backing home teams at pre-match odds generates positive returns while backing away teams generates losses. Moreover, returns are higher for backing home underdogs relative to home favourites and markedly inferior for away underdogs compared to away favourites. Now consider the returns to following a given strategy throughout the match immediately following big news in favour of the chosen strategy. How would the returns compare? As an example, given the volume of news across the 294 matches, the total number of unit bets on backing home underdogs after large news is 315. The return from this is 142.14, a loss of 17%. In fact, all four strategies of placing extra money in the running result in losses. In the case of backing home teams, positive profits are transformed into losses (48% and 55% for favourites and underdogs respectively) while
for away team bets, losses are exacerbated (17% and 75% for favourites and underdogs respectively). Of course, these results could be an artefact of our particular sample of matches but it appears that backing the favoured and/or home team after large news clearly generates inferior returns, and not superior returns. Strong form efficiency is therefore not violated.

**Conclusion**

Betting exchanges are increasingly popular vehicles for sports betting. They also offer researchers an excellent opportunity to test for market efficiency using hitherto unavailable high-frequency data. By identifying news in an appropriate manner, and with a suitable test, we find that news is incorporated into prices very quickly indeed, within 10 seconds. Consistent with this finding, it does not appear possible to generate a simple trading rule from betting in the running that delivers superior returns to taking a pre-match position. We suggest that the Betfair exchange for football betting exhibits clear market efficiency properties.
References


Hakes, Jahn, Sauer, Raymond and Waller, Kerry (2005), The Progress of the Betting on a Baseball Game, Clemson University.


