Event Clustering and Abnormal Returns: Reassessing the Informational Value of Bets

Massimiliano Castellani
Pierpaolo Pattitoni
Roberto Patuelli

Quaderni - Working Paper DSE N° 817
Event Clustering and Abnormal Returns: 
Reassessing the Informational Value of Bets

Massimiliano Castellani, * Pierpaolo Pattitoni, † Roberto Patuelli ‡

Abstract

We analyse the links between soccer match results, bets and stock returns of all listed European soccer teams. Using an event study approach, we measure abnormal returns following wins, ties and losses. Wins are associated with positive abnormal returns, and ties and losses with negative abnormal returns. Additionally, we analyse the role of bets in shaping market reactions to unexpected results, which we find to be non-significant. We propose an alternative econometric approach, using seemingly unrelated regression models, to take into account the problem of overlapping events. While our results concerning match results are confirmed, abnormal returns following unexpected results are found to be statistically significant and to magnify the positive (negative) effects of wins (losses).

Keywords: Information and Market Efficiency; Event Studies; Soccer; Bets; Event Clustering; Seemingly Unrelated Regression Equation (SUR)

JEL Classification: G14; L83; C30

This version: January 27, 2012

* Department of Economics, University of Bologna, Bologna, Italy; The Rimini Centre for Economic Analysis (RCEA), Rimini, Italy. E-mail: m.castellani@unibo.it.
† Department of Management, University of Bologna, Italy; The Rimini Centre for Economic Analysis (RCEA), Rimini, Italy. E-mail: pierpaolo.pattitoni@unibo.it.
‡ Department of Economics, University of Bologna, Bologna, Italy; The Rimini Centre for Economic Analysis (RCEA), Rimini, Italy. E-mail: roberto.patuelli@unibo.it.
1 – Introduction

In the literature of the last twenty years, several studies have suggested an association between sport events and individual mood: fans usually experience a strong positive reaction when their team performs well and a strong negative reaction when their team performs poorly. More specifically, sport wins are generally associated with improvements in mood and self-esteem (Hirt et al. 1992). Instead, sport losses are followed by an increase in heart attacks, crimes and suicides (White 1989; Trovato 1998; Wann 2001; Carroll et al. 2002; Berthier and Boulay 2003; Chi and Kloner 2003).

Other studies have found a relationship between investor mood and stock prices. Several contributions in behavioral finance literature investigate the effect of mood and investor sentiment on asset prices. Remarkable examples are Barberis et al. (1998), Daniel, Hirshleifer and Subrahmanyam (1998), and Baker and Wurgler (2006).¹

A number of further authors were able to join these two streams of literature and show that soccer matches may influence stock returns. Ashton et al. (2003) find a strong association between the performance of the England soccer team and the general stock market index returns. Using international soccer results as a mood variable, Edmans et al. (2007) find an economically and statistically significant market reaction to losses, but not to wins. Brini and Palomba (2011) find a relationships between the stock returns of Italian soccer teams, their sport performances in terms of match outcomes, and the importance of the matches.

Two recent studies by Bernile and Lyandres (2011) and by Palomino et al. (2009) consider the information content of pre-match bets in studying the relationship between soccer match results and soccer team market returns.² Using bets as a proxy for investor expectations, Bernile and Lyandres (2011) analyse the returns of a sample of listed European soccer teams around important matches. They find that investors are overly optimistic about their teams’ prospects ex ante, and disappointed ex post. This result leads to negative post-event abnormal returns. Using a sample of soccer teams listed on the London Stock Exchange, Palomino et al. (2009) find a strong market reaction to match results and a lack of reaction to betting odds. They conclude that this lack of reaction may be due to the lack of informational content of betting information.

Our paper fits in the above stream of literature. In particular, we analyse the relationship between soccer match results, bets and stock returns of all listed European soccer teams. Since sport events affect mood, expected and unexpected soccer match results provide an unique way of studying the

¹ A review of this stream of the literature can be found in Hirshleifer (2001) and Shiller (2000).
² The use of pre-match bets as an indicator of match uncertainty is quite common in the applied literature (e.g., see Czarnitzki and Stadtmann 2002).
link between investor mood and stock market prices. Our analysis aims at answering several questions.

i. First, splitting the matches between wins, ties and losses, we wonder if these events lead to a different market reaction.

ii. Second, we test whether the intensity of the result, as measured by goal difference, has an effect on listed soccer team returns.

iii. Third, we ask whether the location where a match is played (at home or away) influences the market reaction.

iv. Fourth, we test if any type of competition or date (year and month) effect influences the market reaction following wins, ties and losses.

v. Finally, using betting data, we divide all match results between expected and unexpected, and test whether unexpected events are associated with a larger effect on listed soccer team returns.

Our main findings show positive abnormal returns following wins and negative abnormal returns following both ties and losses. This result strongly persists after controlling for the location where the match is played, the type and the date of competition. Furthermore, using the information of pre-match betting odds, we show that abnormal returns are magnified by unexpected results.

Our study contributes to the existing literature in several respects:

i. The data at our disposal allow us to corroborate known results with new empirical evidence and offer new insights on the effect of moods on soccer team stock prices.

ii. The existing literature focuses on national and international matches in single countries, or on international matches in several countries. Instead, our dataset includes all national and international match results of all 23 listed European soccer teams, in the three year period 2007–09. These teams play in 9 national championships and in all UEFA³ international competitions, and represent the whole universe of the listed soccer teams.⁴ Thus, our results contribute to widen the previous empirical evidence on the effect of soccer match results on market returns.

iii. The role of betting is analysed only by Bernile and Lyandres (2011) and Palomino et al. (2009). Both studies consider the information contents of the bets market as a proxy for in-

---

³ UEFA is the Union des Associations Européennes de Football.
⁴ Although these 23 soccer teams represent the entire universe of the listed soccer teams, they are clearly just a small part of all the European soccer teams.
vestor expectations and provide controversial results. Our dataset includes all pre-match betting odds of all 23 listed European soccer teams and allows us to deepen the empirical evidences on their information content.

iv. To double check our findings, we use an alternative econometric methodology with respect to the traditional event study approach used in most of the previous studies. Standard event studies generally deal with non-overlapping events, for which it is reasonable to assume that abnormal returns on individual securities are uncorrelated. However, soccer matches are played in the same days, thus the non-overlapping hypothesis no longer holds. To accommodate this problem of event clustering, we apply a Seemingly Unrelated Regression (SUR) model, which allows the covariances between the abnormal returns to be non-null. Our results shows that the problem of overlapping events is not negligible.

The remainder of this paper is organized as follows: Section 2 describes the dataset; Section 3 presents our hypotheses and modeling framework; Section 4 shows the results of our event study analysis; in Section 5, we retest some of our hypotheses using the SUR methodology; finally, Section 6 draws the conclusions.

2 – Variable description and descriptive statistics

We base our empirical analysis on a unique dataset we collected from the archives of SNAI (Società Nazionale Agenzie Ippiche), the leading agency for sports betting in Italy. Our dataset includes all the results and the pre-match betting odds of 2,157 matches played by all 23 listed soccer teams in Europe in the period 2007–09. Market data on these soccer teams and market index returns are collected from Datastream.

In what follows, we propose a short description of the variables we use in our econometric analysis.

Abnormal returns

$Ar_i$ is the percentage abnormal return of the soccer team $i$ at the event date, defined as $Ar_i = r_i - r_m$, where $r_i$ is the logarithmic return of the soccer team $i$ and $r_m$ is the logarithmic market return (represented by the

---

3 The 23 listed European soccer teams are Celtic, Borussia Dortmund, Tottenham Hotspur, Afc Ajax, Lazio, Brondby If B, Arhus Elite, Sport Lisboa E Benfica, Silkeborg, Futebol Clube Do Porto, Watford, Galatasaray, Aalborg Boldspilklub, Millwall Hldg, Sport, Aik Football, Parken Sport & Entertainment, Trabzonspor Sportif Yatir, Fenerbahce Sportif Hizmet, Juventus, Besiktas, As Roma, Olympique Lyonnais.
STOXX® Europe 600 index\(^6\). We consider a market-adjusted return to take into account the systematic component of stock returns.

**Match result**

\(\text{Win}_i, \text{Tie}_i, \text{Loss}_i\) are three dummy variables indicating if the soccer team \(i\) won, tied, or lost the match.\(^7\) Of 2,157 matches considered, 1,088 are wins (50%), 519 are ties (24%) and 550 are losses (26%). The prevalence of wins on losses and ties shows that listed teams generally belong to the upper tier of their respective national leagues.

**Delta score**

\(\Delta_i\) is the goal difference of the match, defined as the goals scored by the home team minus the goals scored by the away team. This variable assumes positive, negative, or null values when the soccer team \(i\) wins, loses, or ties the match respectively. \(\Delta_i\) exhibits a symmetric distribution with a mean (median) value of 0.54 (1) and a minimum (maximum) of \(-8\) (8). Extreme positive or negative values of this variable (less than \(-5\) or more than 5) are very rare (0.74% of the observations).

**Site of competition**

\(\text{Away}_i, \text{Home}_i\) are two dummy variables indicating if the soccer team \(i\) played the match away or at home. 1085 (50%) matches are played at home and 1078 (50%) matches are played away.

**Type of competition**

\(\text{Ncha}_i, \text{Clea}_i, \text{Uefa}_i, \text{Icup}_i, \text{Ncup}_i\) are a set of dummy variables indicating the type of competition played by the soccer team \(i\). 1663 (76.88%) of competitions are national championship matches, 144 (6.66%) are champions league matches, 128 (5.92%) are UEFA league matches, 12 (0.55%) other international cup matches, and 216 (9.99%) are national cup matches.

**Year of competition**

\(\text{Y07}_i, \text{Y08}_i, \text{Y09}_i\) are a set of year dummies. Of the total number of matches, 1,062 were played in 2008 (49% of the total), 566 in 2007 (26%), and 529 in 2009.

---

\(^6\) Using country-specific market indices might seem more appropriate if we consider that an investor is more likely to invest in her country. This hypothesis, called home bias, is often justified by the tendency of people to invest in what they believe to know better. To take this point into account, we recalculate all abnormal returns using country-specific indices as a robustness check. However, the results are not qualitatively different from the ones reported in the paper.

\(^7\) All the dummies are equal to one in correspondence of the first trading day following the match. If the match is played during the week, the first available trading day is the day following the match. If the match is played during the weekend, the first available trading day is on Monday.
(25%). The average number of matches by year is 719.

Month of competition
Jani, Febi, Mari, Apri, Mayi, Juni, Juli, Augi,
Sepi, Oci, Novi, Deci
are a set of month dummies. The number of matches is almost uniformly distributed across months (about 8% per month). The only exceptions are the summer months when less matches are played.

Expected and unexpected results
Exp, Unei
are two dummy variables indicating if the match results was expected or unexpected based on the information of the last pre-match betting odds. The match result with the smaller probability of occurrence is considered an unexpected event. 952 out of 2,157 match results are unexpected (44% of the total): 142 are unexpected wins (13% of the total number of wins), 519 are unexpected ties (100% of the total number of ties), and 291 are unexpected losses (53% of the total number of losses). Therefore, while all ties are unexpected,8 about 90 per cent of wins are expected, while 47% of losses are expected.

3 – Hypotheses and modeling framework

We use a standard event study methodology (Campbell et al. 1997) to examine the behavior of teams’ stock returns around matches, i.e., the events. The response variable used in all regression models is Ar.9

We start from a basic model analyzing the market reaction following wins, ties and losses. The model equation is

\[ \text{AR}_i = \alpha_1 \text{Win}_i + \alpha_2 \text{Tie}_i + \alpha_3 \text{Loss}_i + \epsilon_i. \]

8 Since a tie is the most probable result when teams are equally skilled, the absence of unexpected ties may be indicate a disparity of skills between listed and unlisted soccer teams. However, this notion is merely speculative, thus future research might consider investigating this issue.

9 We use the event parameter approach (also known as regression-based approach). The event parameter approach produces results that are numerically equivalent to the standard event study method à la Brown and Warner (1980, 1985). All our inference is based on heteroskedasticity consistent standard errors and is double checked by constructing bootstrap confidence intervals. Bootstrap confidence intervals are constructed using the procedure described in Greene (2007) by carrying out (with replacement) 5000 samples from the full sample of observations.
In this model, $\alpha_1$, $\alpha_2$ and $\alpha_3$ measure the average abnormal returns following wins, ties and losses respectively. Coherently with previous findings in the literature and in financial economic theory we expect

- $H_1$: $\alpha_1 > 0$, $\alpha_2 \geq 0$, $\alpha_3 < 0$, and $|\alpha_1| < |\alpha_3|$.

The first hypothesis predicts positive abnormal returns following wins and negative abnormal returns following losses, but no abnormal returns following ties. Furthermore, given the well-known asymmetric reaction associated with positive and negative events, we expect the abnormal returns following losses to be greater in absolute terms than those following wins. All the predictions of $H_1$ are motivated by the effect of moods on stock prices.

As a further step in the analysis, we test whether the intensity of the result, as measured by goal difference, has a (possibly nonlinear) effect on abnormal returns. The second model is

$$AR_i = \alpha + \delta_1 Delta_i + \delta_2 Delta_i^2 + \varepsilon_i.$$  \hspace{1cm} (2)

In this model, $dAR_i/dDelta_i = \delta_1 + 2\delta_2 Delta_i$ is the marginal effect of the goal difference on the abnormal returns. We expect

- $H_2$: $\delta_1 > 0$, $\delta_2 < 0$, and $|\delta_1| > |\delta_2|$.

If the market reaction is driven by mood, we expect abnormal returns to increase with the goal difference.\(^{10}\) However, we expect the effect of the goal difference to be nonlinear and to reduce for large values of this variable, i.e., the positive mood associated with a win gradually decreases as $Delta_i$ increases.

A third concern that may arise is whether the location where a match is played (at home or away) influences abnormal returns. To answer this question, we estimate the following model:

$$AR_i = \alpha_1 Win_i \times Away_i + \alpha_2 Tie_i \times Away_i + \alpha_3 Loss_i \times Away_i +$$
$$\alpha_4 Win_i \times Home_i + \alpha_5 Tie_i \times Home_i + \alpha_6 Loss_i \times Home_i + \varepsilon_i.$$ \hspace{1cm} (3)

\(^{10}\) Beck and Meyer (Beck and Meyer 2012) use goal differences as the outcome of a soccer match. The motivation for their decision is that marginal performance variations might be reflected by variations in observed goal differences.
In this model, $\alpha_1$, $\alpha_2$ and $\alpha_3$ measure the average abnormal returns following wins, ties and losses when the soccer team $i$ plays the match away; $\alpha_4$, $\alpha_5$ and $\alpha_6$ measure the average abnormal returns following wins, ties and losses when the soccer team $i$ plays the match at home. We expect

- $H_3$: $\alpha_4 > \alpha_2 > 0$, and $\alpha_6 < \alpha_3 < 0$.

We expect a win away to produce a larger positive abnormal return than a win at home or a tie away. Conversely, we expect a loss at home to cause larger negative abnormal returns than a loss or a tie away. Indeed, wins away or at home and ties away are generally perceived as a good news, while losses at home or away and ties at home are commonly considered as a bad news. Playing at home is expected to provide an advantage, because the team can benefit from the support of its fans.

We also check if any type or date (year and month) of competition influences the abnormal returns following wins, ties and losses. To answer this question, we estimate four regression models (one for each effect). We setup the equation of each of the four models, not reported here to preserve space, by interacting the match result dummies with the type of competition, year, and month dummies. Joint Wald tests on the difference of the coefficients associated with one effect should evidence any influence on abnormal returns. In what follows, we state and briefly comment the hypotheses associated with these regression models.

- $H_4$: The type of competition played by the soccer team $i$ ($Ncha_i$, $Clea_i$, $Uefa_i$, $Icup_i$, $Ncup_i$), has an effect on the magnitude of abnormal returns following wins, ties and losses.

Since not all type of competitions are equally important, we expect a significant Wald test associated with this model (e.g., champion league matches are generally considered more important than UEFA league matches and can elicit larger market reactions).

- $H_5$: The year of competition in which the match takes place ($Y07_i$, $Y08_i$, $Y09_i$) has an effect on the magnitude of abnormal returns following wins, ties and losses.

Since the attention of the media around soccer and the interest of the investors on the listed soccer teams may vary over time, we could expect matches played in different years to be characterized by systematic differences in terms of abnormal returns.
• \( H_6 \): The month in which the match takes place (\( \text{Jan}, \text{Feb}, \text{Mar}, \text{Apr}, \text{May}, \text{Jun}, \text{Jul}, \text{Aug}, \text{Sep}, \text{Oct}, \text{Nov}, \text{Dec} \)), has an effect on the magnitude of abnormal returns following wins, ties and losses.

The period of the year in which the match takes place (beginning or end of the season) may have an effect on fan mood and, thus, on the abnormal returns (e.g., the final match of UEFA Champions League, which is one of the most important event for the soccer fans, is played in May or June).

The last hypothesis that we test concerns the market reaction following expected and unexpected match results. The equation of the estimated model is

\[
AR_i = \alpha_1 \text{Win}_i \times \text{Exp}_i + \alpha_2 \text{Loss}_i \times \text{Exp}_i + \\
\alpha_3 \text{Win}_i \times \text{Une}_i + \alpha_4 \text{Tie}_i + \alpha_5 \text{Loss}_i \times \text{Une}_i + \epsilon_i. \tag{4}
\]

In this model, \( \alpha_1 \) and \( \alpha_2 \) measure the average abnormal returns following expected wins and losses; \( \alpha_3, \alpha_4 \) and \( \alpha_5 \) measure the average abnormal returns following unexpected wins, ties and losses. It is worth noting that, as already pointed out, all the ties in our dataset are unexpected, thus it is not possible to identify the average abnormal return following expected ties. Our hypotheses can be stated as

• \( H_7: \alpha_3 > \alpha_1 > 0, \alpha_5 < \alpha_2 < 0, \) and \( \alpha_4 = 0. \)

We expect larger abnormal returns following unexpected events, i.e. unexpected wins and losses should determine larger abnormal returns than expected wins and losses respectively.

4 – Event study results

In this section, we present the results of our event study analysis. The following subsections are devoted to the test of the hypotheses stated in the previous section.

4.1 – Market reaction to match results, goal difference, and site of competition

The first hypothesis we test is \( H_1 \). The results of the estimation of Model (1) are presented in Table 1, which shows statistically significant positive market reactions following wins (0.792%) and negative ones following losses (−1.142%). Furthermore, the average abnormal return following
losses is larger in absolute terms than the average abnormal return following wins. All these results are in line with expectations (and with previous findings in the literature). However, we reject the hypothesis of no market reaction associated with ties: ties seem to have a negative effect in terms of abnormal returns, even if their effect is not as large as the one associated with losses (−0.371%). Thus, our results shed new light on the informational content of ties: investors seem to consider ties as negative rather than neutral events. This result could be explained considering that ties reduce the probability of winning a championship or a cup, and that all soccer teams in our sample are upper tier teams in their respective leagues.

Table 1. Match result and abnormal returns.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>p-value</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss</td>
<td>−1.142</td>
<td>0.169</td>
<td>0.000</td>
<td>***</td>
</tr>
<tr>
<td>Tie</td>
<td>−0.371</td>
<td>0.175</td>
<td>0.034</td>
<td>**</td>
</tr>
<tr>
<td>Win</td>
<td>0.792</td>
<td>0.138</td>
<td>0.000</td>
<td>***</td>
</tr>
</tbody>
</table>

***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

Once we have verified that abnormal returns are affected by match results, we test the second hypothesis ($H_2$) by estimating Equation (2). Indeed, the results (see Table 2) seem to support our hypothesis: abnormal returns increase with the goal difference. Furthermore, as expected, the effect of the goal difference is nonlinear and decreases for larger goal differences.

Table 2. Delta score and abnormal returns.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>p-value</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>−0.088</td>
<td>0.105</td>
<td>0.400</td>
<td></td>
</tr>
<tr>
<td>Delta</td>
<td>0.457</td>
<td>0.056</td>
<td>0.000</td>
<td>***</td>
</tr>
<tr>
<td>Squared Delta</td>
<td>−0.042</td>
<td>0.016</td>
<td>0.008</td>
<td>***</td>
</tr>
</tbody>
</table>

***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

The results of this second model are consistent with those of the first model: match results affect abnormal returns. However, this effect is amplified (in a gradually decreasing way) by the intensity of the result. This nonlinear pattern (depicted in Figure 1) can be interpreted as follows: the positive coefficient on the linear term (0.457%) identifies win- or loss-related moods (e.g., winning is preferable to tying or losing); the negative coefficient on the quadratic term (−0.042%) may imply satiety (a greater overall goal difference at the end of a league rarely brings to significant advantages) or a disparity in the strength of the two teams, which thus reduces the importance of the win itself.
Now we test the third hypothesis ($H_3$). The results are presented in Table 3. First, we note that the estimated coefficients of Model (3) are in line with those presented in Table 1: a positive market reaction follows wins (1.021% at home and 0.476% away), a negative market reaction follows losses (−1.317% at home and −1.045% away), and a negative market reaction is suggested for ties (a non-significant −0.300% at home and a slightly significant −0.437% away). Second, we test whether the location where a match is played (at home or away) influences abnormal returns. A joint Wald test for the site of competition, based on a re-parameterization of the model not reported for brevity, shows that there is no statistically significant difference between the abnormal returns of matches played at home or away (Wald test = 1.514, p-value = 0.209). Thus, the presence of a site of competition effect is not supported by the data. This evidence is not consistent with the stylized fact that playing at home is an important advantage for a soccer team and leads to a rejection of $H_3$.

4.2 – Competition type and date effect

In this section, we test our hypotheses regarding the competition type and date. The first hypothesis ($H_4$) concerns the influence of the type of competition on abnormal returns. We estimate a model distinguishing among different competition types (i.e., national championship, champions league, UEFA league, other international cup, and national cup matches), and we find a strongly statistical difference in abnormal returns associated with different types of competition (Wald test = 4.0529, p-value < 0.000). Instead of reporting the results in a table (there would be a total of 15 coeffi-
cients), we prefer to plot the regression coefficients in a figure (see Figure 2). The abnormal returns registered for the national championship completely agree with those reported in Model (1): wins elicit positive abnormal returns; ties and losses are associated with negative abnormal returns. The same finding applies to Champions League and national cup matches, even if with differences in the magnitude of abnormal returns. The only anomalies we observe concern UEFA cup wins (negative abnormal returns) and international cup losses and ties (positive abnormal returns). However, this anomaly is only apparent. In fact, the coefficients associated with these events are not significant (thus no abnormal returns are detected). This statistical insignificance may be justified by the small number of observations for these competition types. However, all in all the empirical evidence seems to support our hypothesis ($H_4$) that not all type of competitions are equally important.

Table 3. Site of competition and abnormal returns.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>p-value</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss Away</td>
<td>-1.045</td>
<td>0.211</td>
<td>0.000</td>
<td>***</td>
</tr>
<tr>
<td>Tie Away</td>
<td>-0.437</td>
<td>0.248</td>
<td>0.078</td>
<td>*</td>
</tr>
<tr>
<td>Win Away</td>
<td>0.476</td>
<td>0.215</td>
<td>0.027</td>
<td>**</td>
</tr>
<tr>
<td>Loss Home</td>
<td>-1.317</td>
<td>0.281</td>
<td>0.000</td>
<td>***</td>
</tr>
<tr>
<td>Tie Home</td>
<td>-0.300</td>
<td>0.245</td>
<td>0.221</td>
<td></td>
</tr>
<tr>
<td>Win Home</td>
<td>1.021</td>
<td>0.179</td>
<td>0.000</td>
<td>***</td>
</tr>
</tbody>
</table>

***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

Figure 2. Competition type and abnormal returns

The fifth hypothesis ($H_5$) focuses on the year of competition (2007, 2008 and 2009). We hypothesized potential differences in the abnormal returns associated with these three years. Figure 3 seems to support this hypothesis. However, a joint Wald test on the year effect rejects the hypothesis of
changing abnormal returns over years on a statistical ground (Wald test = 1.647, p-value = 0.130). Thus, $H_5$ is not supported by the data.

Figure 3. Year of competition and abnormal returns

In addition, we test for the presence of a month effect ($H_6$). Figure 4 show that the abnormal returns in different months mimic those of Model (1) (with few exceptions). Furthermore, the figure seems to suggest that the month in which the match takes place influences the magnitude of the abnormal returns. A joint Wald test on the month effect supports this hypothesis (Wald test = 2.060, p-value = 0.000). The strongest reaction is registered at the end of season in July (despite the appearance the abnormal returns in June are not statistically significant), supporting our hypothesis that the period of the year in which the match takes place affects the importance of the match itself. However, it is worth noting that only 14 matches were played in July.\textsuperscript{11} Thus, caution is needed in interpreting this finding.

Figure 4. Month of competition and abnormal returns.

\textsuperscript{11}Most championships have already ended by July.
4.3 – **Informational value of bets**

In this section, we focus on the eighth hypothesis ($H_7$), which is related to the informational value of bets. To test this hypothesis, we estimate Model (3) and present the results in table 4. The abnormal returns associate with wins and losses follow the usual pattern, both following expected and unexpected events. Furthermore, unexpected ties (the only available) register negative abnormal returns. All these results are in line with those of Model (1). Focusing more in depth on $H_7$, a joint Wald test does not reject the null of no effect of unexpected events (Wald test = 0.2255, p-value = 0.7981). Thus, according to these results, we reject $H_7$, which predicts a stronger market reaction following unexpected events. Since this result is counterintuitive, it needs to be analysed further.

**Table 4. Informational value of bets.**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>p-value</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp Loss</td>
<td>–1.028</td>
<td>0.266</td>
<td>0.000</td>
<td>***</td>
</tr>
<tr>
<td>Exp Win</td>
<td>0.808</td>
<td>0.139</td>
<td>0.000</td>
<td>***</td>
</tr>
<tr>
<td>Une Loss</td>
<td>–1.244</td>
<td>0.250</td>
<td>0.000</td>
<td>***</td>
</tr>
<tr>
<td>Tie</td>
<td>–0.371</td>
<td>0.188</td>
<td>0.048</td>
<td>**</td>
</tr>
<tr>
<td>Une Win</td>
<td>0.685</td>
<td>0.359</td>
<td>0.057</td>
<td>*</td>
</tr>
</tbody>
</table>

***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

5 – **The problem of overlapping events: a SUR approach**

A problem with the standard event study methodology is that it generally deals with non-overlapping events. In case of non-overlapping events, it is reasonable to assume abnormal returns on individual securities to be uncorrelated between teams (one of the main hypothesis in event study testing procedures). However, matches are played in the same days, thus the non-overlapping hypothesis does not hold. Ignoring event date clustering would result in biased standard errors, which in turn may lead to misleading results in hypothesis testing.

In our dataset, the average number of events per day is around 12. Thus, the problem of overlapping events is a concern. To accommodate this potential problem, we apply an SUR model (Zellner 1962) to retest $H_1$ and $H_7$. The SUR model takes into account the cross sectional correlation due to clustering of event days, allowing the covariances between the abnormal returns to be non-null (Schipper and Thompson 1985). We focus only on $H_1$ and $H_7$ because, by construction, the SUR
model does not allow to retest the remaining hypotheses. To do so, the dataset should include observations for each possible combination of variables (e.g., wins and losses for each team and for each month).

Formally, the SUR model can be seen as a generalization of a linear regression model consisting of several regression equations. For each team $i$, let $r_i$ denote a $T \times 1$ vector of daily returns, with generic element $r_{it}$ ($t = 1, \ldots, T$), $X_i$ a $T \times k$ matrix of $k$ covariates, $\varepsilon_i$ a $T \times 1$ vector of error terms and $\beta_i$ a vector of $k$ parameters. Then, the SUR model consists of $N$ equations (one for each soccer team) of the form

$$r_i = X_i \beta_i + \varepsilon_i, \ i = 1, \ldots, N. \tag{5}$$

Two main assumptions on the errors of a SUR model are

$$E(\varepsilon_i | X_1, \ldots, X_N) = 0, \ i = 1, \ldots, N; \text{ and}$$

$$\text{Cov}(\varepsilon_i, \varepsilon_j | X_1, \ldots, X_N) = \begin{cases} \sigma_i^2 I_T & \text{if } i = j \\ \sigma_{ij} I_T & \text{if } i \neq j \end{cases},$$

where $I_T$ is the identity matrix.

It is important to note that the system in Equation (4) does not provide a structural relationship among the different equations (i.e., the dependent variable of one equation never appears among the covariates of other equations). However, it is clear from the covariance equation above that the $N$ equations are linked together through the error covariances $\sigma_{ij}$. This assumption is crucial because it allows the correlation across the abnormal returns of different teams to assume any values and implies unobserved variables with correlated values between equations.

The SUR model can be estimated by means of the feasible generalized least squares (FGLS) method. As $T$ diverges, the FGLS estimator is consistent, asymptotically efficient and asymptotically normal.\footnote{The model can also be estimated equation-by-equation using ordinary least squares (OLS). Even if the OLS estimator is consistent, it is not efficient because it ignores the cross-correlation among errors.}

After presenting the main methodological feature of the SUR model, we now focus on retesting $H_1$. To do so, we estimate an SUR model in which each equation (one for each team) can be formalized as follows:
\[ r_{it} = \alpha_i + \beta_i^W \text{Win}_{it} + \beta_i^T \text{Tie}_{it} + \beta_i^L \text{Loss}_{it} + \text{con}_i' \gamma_i + \epsilon_{it}. \] (6)

In Equation (6), \( \text{Win}_{it} \), \( \text{Tie}_{it} \) and \( \text{Loss}_{it} \) are dummy variables indicating if the soccer team \( i \) won, tied, or lost the match played at time \( t \). \( \beta_i^W \), \( \beta_i^T \) and \( \beta_i^L \) measure the market reaction following wins, ties and losses, respectively. Furthermore, \( \gamma_i \) is the vector of coefficients associated with the control variables included in \( \text{con}_i \).\(^{13}\) In particular, we estimate three models, which differ for the control variables included in \( \text{con}_i \). In the first case, we do not include any controls (i.e., we impose \( \gamma_i = 0 \)). In this model, \( \alpha_i \) measures the average return that occurs when no matches are played, and the abnormal returns are defined as a deviation from \( \alpha_i \) when matches are instead played. In the second case, to take into account the systematic component of stock returns, we add the variable \( \text{Mkt}_t \) among the controls.\(^{14}\) \( \text{Mkt}_t \) is the STOXX® Europe 600 index’s logarithmic daily return, used as a proxy for the market return. Finally, in the third case, we consider a possible year effect by adding year dummies among the controls.

After estimating our SUR model, we test the joint significance of all the coefficients \( \beta_i^W \), \( \beta_i^T \) and \( \beta_i^L \) (one for each team) by means of three tests: Wald test (F-distributed), Wald test and Likelihood ratio test (both Chi-squared-distributed).\(^{15}\)

Since for each team we estimate a different \( \beta_i^W \), \( \beta_i^T \) and \( \beta_i^L \), as a summary measure we report the average value of all coefficients:

\[ Ar^W = \frac{1}{23} \sum_{i=1}^{23} \beta_i^W, \quad Ar^T = \frac{1}{23} \sum_{i=1}^{23} \beta_i^T, \quad Ar^L = \frac{1}{23} \sum_{i=1}^{23} \beta_i^L. \] (7)

where \( Ar^W \), \( Ar^T \) and \( Ar^L \) are the average abnormal returns following wins, ties and losses, respectively.

Focusing on \( H_7 \), we estimate an augmented SUR model in which each equation is written as

\(^{13}\) It is worth noting that the control variables are the same for each soccer team, thus we do not need the team-specific index \( i \) to characterize the vector \( \text{con}_i \).

\(^{14}\) More sophisticated models can be used to assess abnormal returns (e.g., the Fama-French three factor model, or Carhart’s four-factor model). However, previous literature on short-term event studies (Campbell et al. 1997) points out that results are virtually unaffected by the choice of the model. This is especially true in our case, because the stocks considered belong to the same sector and are quite homogeneous.

\(^{15}\) While the latter two, which are asymptotically equivalent, are especially indicated for large samples, the former has better finite sample properties.
\[
\begin{align*}
    r_{it} = & \alpha_i + \beta_i^W \text{Win}_{it} + \beta_i^T \text{Tie}_{it} + \beta_i^L \text{Loss}_{it} + \\
    & \delta_i^W (\text{Une}_{it} \times \text{Win}_{it}) + \delta_i^T (\text{Une}_{it} \times \text{Loss}_{it}) + \text{con}_i, \gamma_t + \varepsilon_{it}.
\end{align*}
\] (8)

In Equation (8), \( \text{Win}_{it} \), \( \text{Tie}_{it} \) and \( \text{Loss}_{it} \) indicate, as in Equation (6), if the soccer team \( i \) won, tied, or lost the match played at time \( t \); \( \text{Une}_{it} \) is a dummy variable indicating if the match result was unexpected based on the information of the pre-match betting odds. Based on this model specification, \( \beta_i^W \) and \( \beta_i^L \) measure the market reaction following expected wins and losses, while \( \delta_i^W \) and \( \delta_i^L \) measure the difference in abnormal returns when the events are unexpected. Thus, \( \beta_i^W \delta_i^W \) and \( \beta_i^L \delta_i^L \) indicate the total abnormal returns associated with unexpected wins and losses. \( \beta_i^T \) represents the abnormal return associated with unexpected ties. Once more, \( \gamma_t \) is the vector of coefficients associated with the control variables included in \( \text{con}_i \).

Similarly to what was done for \( H_1 \), we test the joint significance of all the coefficients \( \beta_i^W \), \( \beta_i^T \), \( \beta_i^L \), \( \delta_i^W \) and \( \delta_i^L \) by means of the three aforementioned tests, and report the average value of all coefficients as a summary statistic:

\[
\begin{align*}
    \text{Ar}^w = & \frac{1}{23} \sum_{i=1}^{23} \beta_i^w, \text{Ar}^T = \frac{1}{23} \sum_{i=1}^{23} \beta_i^T, \text{Ar}^L = \frac{1}{23} \sum_{i=1}^{23} \beta_i^L, \\
    \Delta \text{UneAr}^w = & \frac{1}{23} \sum_{i=1}^{23} \delta_i^w, \Delta \text{UneAr}^T = \frac{1}{23} \sum_{i=1}^{23} \delta_i^T,
\end{align*}
\] (9)

where \( \text{Ar}^w \), \( \text{Ar}^T \) and \( \text{Ar}^L \) are the average abnormal returns following expected wins, unexpected ties, and expected losses, respectively, while \( \Delta \text{UneAr}^w \) and \( \Delta \text{UneAr}^L \) measure the average differential abnormal returns following unexpected wins and losses.

5.1 – Reassessment of market reaction

We first retest \( H_1 \). To do so, we estimate three SUR models, which differ according to the control variables included (indicated at the bottom of Table 5). While the first model includes no control variables, the second and the third model introduce among the covariates the market returns and the year dummies. It is worth noting that the results of the three models agree in the sign and magnitude of the average coefficients reported.
Table 5. Average market reaction to match result.

<table>
<thead>
<tr>
<th>Abnormal Return (AR percent)</th>
<th>SUR 1</th>
<th>SUR 2</th>
<th>SUR 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Ar^W$</td>
<td>0.589</td>
<td>0.627</td>
<td>0.631</td>
</tr>
<tr>
<td>Wald test (F-test)</td>
<td>3.914</td>
<td>&lt;0.001</td>
<td>4.230</td>
</tr>
<tr>
<td>Wald test (Chi-squared-test)</td>
<td>90.030</td>
<td>&lt;0.001</td>
<td>97.286</td>
</tr>
<tr>
<td>Likelihood ratio test (Chi-squared-test)</td>
<td>90.030</td>
<td>&lt;0.001</td>
<td>97.286</td>
</tr>
<tr>
<td>$Ar^T$</td>
<td>-0.687</td>
<td>-0.626</td>
<td>-0.631</td>
</tr>
<tr>
<td>Wald test (F-test)</td>
<td>2.410</td>
<td>&lt;0.001</td>
<td>2.262</td>
</tr>
<tr>
<td>Wald test (Chi-squared-test)</td>
<td>55.430</td>
<td>&lt;0.001</td>
<td>52.034</td>
</tr>
<tr>
<td>Likelihood ratio test (Chi-squared-test)</td>
<td>55.430</td>
<td>&lt;0.001</td>
<td>52.034</td>
</tr>
<tr>
<td>$Ar^L$</td>
<td>-1.476</td>
<td>-1.451</td>
<td>-1.443</td>
</tr>
<tr>
<td>Wald test (F-test)</td>
<td>6.126</td>
<td>&lt;0.001</td>
<td>6.165</td>
</tr>
<tr>
<td>Wald test (Chi-squared-test)</td>
<td>140.900</td>
<td>&lt;0.001</td>
<td>141.800</td>
</tr>
<tr>
<td>Likelihood ratio test (Chi-squared-test)</td>
<td>140.900</td>
<td>&lt;0.001</td>
<td>141.800</td>
</tr>
<tr>
<td>con, Mkt'</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year dummies</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 5 presents the values of $Ar^W$, $Ar^T$ and $Ar^L$ as defined above. Consistently with Section 4.1, our results show a statistically significant positive market reaction following wins, and a statistically significant negative market reaction following ties and losses. The average abnormal return following losses is larger in absolute terms than the average abnormal return following wins. In addition, ties have a negative effect in terms of abnormal returns, even if their effect is not as large as the one associated with losses.

We now retest $H_7$. Table 6 reports the estimation of our three SUR models, which agree in sign and magnitude of the average estimated coefficients: $Ar^W$, $Ar^T$, $Ar^L$, $\Delta UneAr^W$ and $\Delta UneAr^L$.

Our results show a statistically significant positive (negative) market reaction following expected wins (losses). Once again, the absolute average abnormal return following losses is larger than the one following wins. Furthermore, ties seem to have a negative effect in terms of abnormal returns. These results agree with the findings of Section 4.3. However, when looking at $\Delta UneAr^W$ and $\Delta UneAr^L$, we observe that unexpected events generate a larger market reaction than expected event. In particular, in all three models, $\Delta UneAr^W$ is positive and $\Delta UneAr^L$ is negative. Thus, unexpected events are associated with larger abnormal returns on average. This result is consistent with $H_7$ and contrasts with our findings in Section 4.3, suggesting that not accounting for simultaneous events may lead to unreliable inference.
### Table 6. Average market reaction to expected and unexpected results.

<table>
<thead>
<tr>
<th>Abnormal Return (AR percent)</th>
<th>SUR 4</th>
<th>SUR 5</th>
<th>SUR 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Ar^W )</td>
<td>0.509</td>
<td>0.550</td>
<td>0.553</td>
</tr>
<tr>
<td>Wald test (F-test)</td>
<td>3.621 0.000</td>
<td>3.9169 0.000</td>
<td>3.941 0.000</td>
</tr>
<tr>
<td>Wald test (Chi-squared-test)</td>
<td>83.272 0.000</td>
<td>90.088 0.000</td>
<td>90.642 0.000</td>
</tr>
<tr>
<td>Likelihood ratio test (Chi-squared-test)</td>
<td>80.352 0.000</td>
<td>86.724 0.000</td>
<td>87.435 0.000</td>
</tr>
<tr>
<td>( Ar^T )</td>
<td>–0.702</td>
<td>–0.637</td>
<td>–0.641</td>
</tr>
<tr>
<td>Wald test (F-test)</td>
<td>2.491 0.000</td>
<td>2.323 0.000</td>
<td>2.301 0.000</td>
</tr>
<tr>
<td>Wald test (Chi-squared-test)</td>
<td>57.288 0.000</td>
<td>53.421 0.000</td>
<td>52.929 0.000</td>
</tr>
<tr>
<td>Likelihood ratio test (Chi-squared-test)</td>
<td>56.305 0.000</td>
<td>52.657 0.000</td>
<td>52.388 0.000</td>
</tr>
<tr>
<td>( Ar^L )</td>
<td>–1.158</td>
<td>–1.169</td>
<td>–1.166</td>
</tr>
<tr>
<td>Wald test (F-test)</td>
<td>3.012 0.000</td>
<td>2.924 0.000</td>
<td>2.934 0.000</td>
</tr>
<tr>
<td>Wald test (Chi-squared-test)</td>
<td>69.272 0.000</td>
<td>67.243 0.000</td>
<td>67.478 0.000</td>
</tr>
<tr>
<td>Likelihood ratio test (Chi-squared-test)</td>
<td>67.177 0.000</td>
<td>65.091 0.001</td>
<td>65.493 0.000</td>
</tr>
<tr>
<td>( \Delta UneAr^W )</td>
<td>0.493</td>
<td>0.492</td>
<td>0.494</td>
</tr>
<tr>
<td>Wald test (F-test)</td>
<td>1.572 0.040</td>
<td>1.726 0.017</td>
<td>1.704 0.019</td>
</tr>
<tr>
<td>Wald test (Chi-squared)</td>
<td>36.153 0.040</td>
<td>39.686 0.017</td>
<td>39.183 0.018</td>
</tr>
<tr>
<td>Likelihood ratio test (Chi-squared)</td>
<td>35.090 0.051</td>
<td>38.505 0.022</td>
<td>38.152 0.025</td>
</tr>
<tr>
<td>( \Delta UneAr^T )</td>
<td>–0.639</td>
<td>–0.596</td>
<td>–0.591</td>
</tr>
<tr>
<td>Wald test (F test)</td>
<td>2.348 0.000</td>
<td>2.133 0.000</td>
<td>2.137 0.000</td>
</tr>
<tr>
<td>Wald test (Chi-squared)</td>
<td>53.998 0.000</td>
<td>49.050 0.000</td>
<td>49.141 0.000</td>
</tr>
<tr>
<td>Likelihood ratio test (Chi-squared)</td>
<td>52.455 0.000</td>
<td>47.765 0.002</td>
<td>47.979 0.002</td>
</tr>
<tr>
<td>con,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Mkt_t )</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year dummies</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

#### 5.2 – The distribution of individual abnormal returns

Since in the previous section we find that the SUR model leads to different findings with respect to the standard event study approach for H7, it is worth inspecting the entire distribution of individual abnormal returns. In order to analyse the potential heterogeneity of the individual coefficients\(^{16}\) \( Ar^W, Ar^L, Ar^W + \Delta UneAr^W, Ar^T \) and \( Ar^L + \Delta UneAr^L \), in Figure 5 we report five boxplots. Boxplots show the distribution of abnormal returns conditionally to expected and unexpected match results. The extremes of the box correspond to the first and third quartiles (\( Q_1 \) and \( Q_3 \)). The vertical height of the box represents the inter-quartile difference (\( \delta = Q_3 - Q_1 \)), a measure of variability. The horizontal bold line inside each box represents the median, while the black point is the average. Finally, the two external lines correspond to \( Q_1 - 1.5\delta \) and \( Q_3 - 1.5\delta \), respectively. Observations beyond these lines indicate the presence of extreme values. While the analysis of the previous section was

---

\(^{16}\) We use the coefficients estimated for the third SUR model (SUR 6, which controls for the market and year effects), but our results are virtually unaffected by this choice.
limited to providing the average abnormal return related to match results (a measure of central tendency), the boxplots are more informative.

![Figure 5. Distribution of individual abnormal returns among soccer teams.](image)

The plots of the first column of Figure 5 present the abnormal returns in case of expected and unexpected wins respectively. Coherently with the result in the previous section, the average (and median) abnormal return in the two plots is positive with a larger abnormal return in case of unexpected wins. Furthermore, the two plots show that the main part of the distribution of individual teams’ abnormal returns is characterized by positive abnormal returns. The second plot of the second row of Figure 5 presents the abnormal returns in case of (unexpected) ties. In this case, the average (and median) abnormal return is negative, but the distribution appears to be asymmetric: the median is close to the upper extreme of the box. Finally, the plots of the third column of Figure 5 present the abnormal returns in case of expected and unexpected losses. In this case, the negative market reaction following this event is evident: the average (and median) abnormal return is strongly negative in both plots (with a larger abnormal return in case of unexpected losses) and all the abnormal returns appear to be concentrated in the negative quadrant of the plot around the median.\(^\text{17}\) Thus, the SUR model results appear to be largely confirmed when inspecting the distribution of the team-specific coefficients.

\(^{17}\)The only exception is a positive (non-significant) abnormal return associated to Futebol Clube Do Porto.
Conclusions

Several studies suggest that match results influence investor mood that in turn affects sport team stock returns. Wins are generally associated with a good mood and losses with a bad mood. These emotional reactions affect investor behavior on the market, which reacts positively to wins and negatively to losses.

In this work, we analyse the relationship between soccer match results, bets and stock returns of all listed European soccer teams. Our dataset includes all national and international match results and pre-match betting odds of all 23 listed European soccer teams, in the three year period 2007–09. To measure the aggregate market reaction to match results, we use an event study approach. Our main results can be summarized as follows.

i. While wins are followed by positive abnormal returns, ties and losses are followed by negative abnormal returns. Furthermore, abnormal returns following losses are larger in magnitude with respect to those following wins. Positive and negative abnormal returns are nonlinearly related to the intensity of a match result, as measured by goal difference.

ii. Abnormal returns are not affected by the location where the match is played (home vs. away) nor by the year of the competition. However, both the type and the month of competition influence abnormal returns.

iii. When we use a SUR model to accommodate the problem of event clustering, positive and negative abnormal returns are magnified by unexpected match results. From this point of view, our study makes an empirical contribution to the existing literature.

The objective of this paper was not to identify a profitable trading strategy. However, albeit indirectly, our results suggest some trading rules that may be applied by investors. For example, investors may choose to rebalance their personal portfolios by buying stocks of the teams that are expected to win, and by selling those of the teams that are expected to lose. Future extensions of this paper could attempt to measure the expected profit of this investment strategy. Furthermore, since positive and negative abnormal returns are magnified by goal difference, future research could explore the relationship between the pre-match betting odds on the goal number and the actual goal difference.18

---

18 In this regard, Goddard and Asimakopoulos (2004) and Goddard (2005) propose a model to test betting strategies with a positive expected return.
Acknowledgements

We are indebted to Enrico Maria Cervellati for his comments on an earlier version of this paper. We would also like to thank Lorenzo Zirulia for his advice. The usual disclaimers apply.

References


Berthier F, Boulay F (2003) Lower Myocardial Infarction Mortality in French Men the day France Won the 1998 World Cup of Football. Heart 89 (5):555-556. doi:10.1136/heart.89.5.555


