Olympic news and attitudes towards the Olympics: a compositional time-series analysis of how sentiment is affected by events

This item was submitted to Loughborough University’s Institutional Repository by the/an author.


Additional Information:

- This is an Accepted Manuscript of an article published in the Journal of Applied Statistics on 19 Dec 2013, available online: http://www.tandfonline.com/10.1080/02664763.2013.868417

Metadata Record: https://dspace.lboro.ac.uk/2134/19908

Version: Accepted

Publisher: © Taylor & Francis

Rights: This work is made available according to the conditions of the Creative Commons Attribution-Non Commercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Olympic news and attitudes towards the Olympics: A compositional time-series analysis of how sentiment is affected by events

Peter Dawsona*, Paul Downwardb and Terence C. Mills c

a School of Economics, University of East Anglia, Norwich Research Park, Norwich, Norfolk, UK
b School of Sport, Exercise and Health Sciences, Loughborough University, Leicestershire, UK
c School of Business and Economics, Loughborough University, Leicestershire, UK

* Corresponding Author. Email: Peter.Dawson@uea.ac.uk
Abstract

Sentiment affects the evolving economic valuation of companies through the stock market. It is unclear how ‘news’ affects the sentiment towards major public investments like the Olympics. In this paper we consider, from the context of the pre-event stage of the 30th Olympiad, the relationship between attitudes towards the Olympics and Olympic-related news; specifically the bad news associated with an increase in the cost of provision, and the good news associated with Team GB’s medal success in 2008. Using a unique data set and an event-study approach that involves compositional time-series analysis, it is found that ‘good’ news affects sentiments much more than ‘bad’ but that the distribution of such sentiment varies widely. For example, a much more pronounced effect of good news is identified for females than males, but ‘bad’ news has less of an impact on the young and older age groups. Caution should be exercised in basing policy on such sentiments.

Keywords: Attitudes; news; Olympic Games; compositional time-series.

JEL codes: C25, D12, L83
1. Introduction

Hosting the Olympics, and the Summer Olympics in particular, involve large scale public investment. Despite this there is competition to host the Olympics. In the case of the London 2012 Olympics the rationale was set out in Game Plan and ‘despite the difficulty in quantifying the impact, there appears to be a positive impact for the nation as a whole’ [15, p.70].

Most of the economic literature suggests negligible and even negative impacts through crowding out [4, 14, 16]. For example, focusing on the 1984 Los Angeles and 1996 Atlanta Summer Olympics Baade and Matheson [5] found that the (ex-post) impact on unemployment in both regions was entirely transitory and concluded that long-run changes to steady-state equilibrium were only possible when new infrastructure and facilities were appropriate for the present and future economy. Madden [31] and Giesecke and Madden [20] concluded that initial ex-ante assessments of the impact of the 2000 Sydney Olympics over-estimated the financial gain because the assumptions relating to induced tourism and the responsiveness of the labour market were overly optimistic.

It is partly because of such difficulties that the claim that major sporting events such as the Olympics can provide important intangible benefits has emerged. For example, hosting such events, or success in them, appears to influence optimism in stock markets [2, 17, 28, 30, 42] and subjective well-being [29, 40].
In this study, therefore, it is hypothesised that the same sort of impact of news stories could take place during the preparation stage, and consequently the impact of the two most significant Olympic-related stories that occurred during the run-up to London 2012 are examined. These are news of the substantial increase in cost (March 2007) and the success of the British team during the Beijing 2008 Olympics (August 2008). In respect of these issues, initial estimates by the London Organising Committee placed the total cost of hosting the event at £2.4 billion. However, in March 2007 it was announced that the total cost had almost quadrupled to just under £9.3 billion.

In the 2008 Beijing Olympics, the British team won 47 medals, including 19 gold, finishing fourth in the medal table, which was a position that was widely unexpected. Confirmation of better than expected performance is provided by the UK Sporting Preferences Survey 2008 [41], who carried out a pre-Beijing survey asking interviewees where in the Olympic medal table they expected the British team to finish. Only 16 per cent of respondents expected the team to finish in the top five. Further verification is provided by a number of econometric models of the determination of medal counts that generally, and in some cases significantly, under-predicted Britain’s medal total [18].

This paper examines the changes in the stated monthly support, as measured as a scale variable, for hosting the 2012 games over time. The impacts of these news events are revealed using compositional time-series techniques that, unlike traditional time-series methods, account for the restrictions placed on the data by the use of a scale in eliciting opinion: in other words, that increases in the proportion of one category of response must be matched by reductions elsewhere on the scale.
2. Data and Methodology

This study uses data from the Taking Part Survey (TPS), commissioned by the Department for Culture, Media and Sport (DCMS). The TPS is a continuous (repeated cross-section, but not a panel) national survey of England which was first undertaken in 2005. In this study, data from the first four waves (2005-6, 2006-7, 2007-08 and 2008-09) are employed, generating over 35,000 usable observations.

The TPS provides a unique opportunity to investigate attitudes and behaviour during the preparation stage of the Olympics, as these waves of the survey have a question specifically asking respondents about their feelings towards the UK hosting the 2012 Summer Olympic Games in London. Five options are presented in a measurement scale: strongly supportive ($x_{5t}$), slightly supportive ($x_{4t}$), neither supportive or against ($x_{3t}$), slightly against ($x_{2t}$), strongly against ($x_{1t}$).

Importantly, the survey is administered in such a way that there are enough observations to consider month-by-month variations in public attitudes, rather than simple wave-specific (yearly) effects. This provides a rich data set that allows a more precise investigation of the relationship between Olympic-related news and public opinion. A preliminary view of the attitudes and motivation of respondents on a month-by-month basis (from January 2006 to March 2009) is presented in Figure 1.

**FIGURE 1 [about here]**
Panel (a) shows the proportion of respondents against the UK hosting the 2012 Olympics and Panel (b) the proportion of respondents who are supportive. Visual inspection suggests the presence of longer term trends though there appears to be a spike in the data associated with cost overruns for those against hosting the games in Panel (a). Likewise in Panel (b) against a declining trend, a spike in support for hosting the Games might be seen to follow from August 2008. To unpick these effects more forensically, a time-series analysis is undertaken, taking account of the compositional nature of the data as implied by the measurement scale.

2.1 Compositional Data

A compositional data set is one in which the \( n \) observations on \( d + 1 \) variables, denoted \( \mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{in})' \), \( i = 1, 2, \ldots, d + 1 \), are such that \( x_i > 0 \) and \( x_1 + x_2 + \ldots + x_{d+1} = 1 \), where \( \mathbf{1} \) is the unit vector, here of dimension \( d + 1 \). The matrix \( \mathbf{X}^{(d)} = \begin{bmatrix} x_1 & x_2 & \cdots & x_d \end{bmatrix} \) then lies in the \( d \)-part simplex, \( S^d \), with \( x_{d+1} = 1 - \sum_{i=1}^{d} x_i \) being the vector of ‘fill-up’ values: \( \mathbf{X} = \begin{bmatrix} \mathbf{X}^{(d)} & x_{d+1} \end{bmatrix} \) is thus the complete data matrix.

Because of the difficulties of analysing \( \mathbf{X} \) within the simplex sample space, Aitchison [1] proposed mapping \( \mathbf{X}^{(d)} \) from \( S^d \) to the \( d \)-dimensional real space \( \mathbb{R}^d \) and then examining the statistical properties of the transformed data within \( \mathbb{R}^d \). Several transformations have been proposed for doing this, the most popular being the additive-logratio transformation defined as
\[ Y = \begin{bmatrix} y_1 & y_2 & \ldots & y_d \end{bmatrix} = a_d(x^{(d)}) = \begin{bmatrix} \log \left( \frac{x_1}{x_{d+1}} \right) & \log \left( \frac{x_2}{x_{d+1}} \right) & \ldots & \log \left( \frac{x_d}{x_{d+1}} \right) \end{bmatrix} \]

The inverse transformation, known as the additive-logistic, is

\[ X^{(d)} = a_d^{-1}(Y) = \begin{bmatrix} \frac{\exp(y_1)}{1 + \sum_{i=1}^d \exp(y_i)} & \frac{\exp(y_2)}{1 + \sum_{i=1}^d \exp(y_i)} & \ldots & \frac{\exp(y_d)}{1 + \sum_{i=1}^d \exp(y_i)} \end{bmatrix} \]

\[ x_{d+1} = \frac{1}{1 + \sum_{i=1}^d \exp(y_i)} \]

Thus, an analysis which is difficult in the ‘awkward’ sample space \( S^d \) may be transformed using \( a_d \) to \( \mathbb{R}^d \), in which more tractable statistical analyses may be performed before using \( a_d^{-1} \) to return to the original variables.

3. Results

Consider the case where the \( x_i \) are time series of proportions, and are now written

\[ x_i = (x_{i1}, x_{i2}, \ldots, x_{iT})' \]

so that there are \( T \) observations available. Accordingly, denote \( X_i \) to be the \( t \)th row of \( X \) and \( Y_i \) to be the \( t \)th row of the additive-logratio transformed data matrix \( Y \). Brunsdon and Smith [9] consider modelling \( Y_i \) as a vector ARMA process but other frameworks are available in which covariates and trend functions may be introduced: see Mills [32, 33, 34] for a variety of examples. It should be emphasised that the modelling is invariant to the choice of fill-up value, so that any of the \( x_i \) can be chosen.
The survey data available thus constitute a compositional time series of dimension $d + 1 = 5$ of proportions from January 2006 to April 2009, a total of 40 observations, as defined earlier. They are plotted as Figure 2. Attention focuses on the two ‘events’ of the March 2007 announcement of the increase in the cost of staging the Olympics and the success of Team GB at the Beijing Olympics of August 2008. These events are modelled as dummies, taking zeros everywhere except that $d_{07}$, takes the value 1 in March 2007 and $d_{08}$, takes the value 1 in August 2008.

**Figures 2 and 3 [about here]**

The proportions were additive-logratio transformed using $x_{5i}$ as the fill-up value, thus defining $y_{it} = \log(x_{it}/x_{5i})$, $i = 1, \ldots, 4$. These ratios are shown in Figure 3 and were then modelled as a dummy-augmented vector AR(1) process, with this order being selected by a variety of criteria, the fitted model being estimated as
\[
\begin{bmatrix}
  y_{1t} \\
y_{2t} \\
y_{3t} \\
y_{4t}
\end{bmatrix} = \begin{bmatrix}
  0.054 & 0.272 & 0.600 & -0.056 \\
  0.202 & -0.005 & 1.128 & -0.398 \\
  0.019 & 0.138 & 0.818 & -0.104 \\
-0.142 & 0.183 & 0.462 & 0.078
\end{bmatrix} \begin{bmatrix}
  y_{1,t-1} \\
y_{2,t-1} \\
y_{3,t-1} \\
y_{4,t-1}
\end{bmatrix}
\]

\[
\begin{bmatrix}
  0.303 & -0.285 & -1.022 \\
  0.283 & -0.266 & -1.119 \\
-0.063 & -0.381 & 0.167 \\
  0.037 & -0.280 & -0.164
\end{bmatrix} \begin{bmatrix}
  d_{07,t} \\
d_{08,t} \\
  1
\end{bmatrix} + \begin{bmatrix}
  u_{1t} \\
u_{2t} \\
u_{3t} \\
u_{4t}
\end{bmatrix}
\]

(3)

Standard errors are in parentheses and coefficients significant at the 10% level are indicated by an asterisk. Few coefficients are individually significant and the model is clearly over-parameterized, making interpretation problematic. The following restricted model was therefore developed.
The restrictions are either coefficient omissions or cross-equation coefficient equalities: in total a set of 19 restrictions are imposed, with an accompanying likelihood ratio statistic of 13.16, so the set is certainly acceptable. The nine coefficients that are freely estimated are all significant at less than the 1% level. The system is being driven by \( y_3 = \log(x_3/x_5) \), while both events appear to have had significant effects on the system, with \( y_1 \) and \( y_2 \) being positively affected by the March 2007 event, and \( y_3 \) and \( y_4 \) being unaffected, while all ratios are negatively affected by the August 2008 event, movements that can clearly be seen in Figure 2. The system is more interpretable if log-proportions are considered. The model (4) can thus be written as

\[
\begin{align*}
\begin{bmatrix}
y_{1t} \\
y_{2t} \\
y_{3t} \\
y_{4t}
\end{bmatrix} &= \begin{bmatrix}
- & - & 0.956_{(0.029)} \\
- & - & 1.095_{(0.084)} \\
- & - & 0.956 \\
- & - & 0.601_{(0.044)}
\end{bmatrix} \begin{bmatrix}
y_{1,t-1} \\
y_{2,t-1} \\
y_{3,t-1} \\
y_{4,t-1}
\end{bmatrix} \\
&+ \begin{bmatrix}
0.377_{(0.142)} & -0.290_{(0.087)} & -1.418_{(0.029)} \\
0.377 & -0.290 & -1.300_{(0.045)} \\
- & -0.397_{(0.097)} & - \\
- & -0.290 & -0.201_{(0.022)}
\end{bmatrix} \begin{bmatrix}
d_{07,t} \\
d_{08,t} \\
1
\end{bmatrix} + \begin{bmatrix}
u_{1t} \\
u_{2t} \\
u_{3t} \\
u_{4t}
\end{bmatrix}
\end{align*}
\] (4)

\[
\begin{align*}
\log x_{1t} &= 0.96 \log x_{3,t-1} + \Delta \log x_{5t} + 0.04 \log x_{5,t-1} + 0.38 d_{07,t} - 0.29 d_{08,t} - 1.42 \\
\log x_{2t} &= 1.09 \log x_{3,t-1} + \Delta \log x_{5t} - 0.09 \log x_{5,t-1} + 0.38 d_{07,t} - 0.29 d_{08,t} - 1.30
\end{align*}
\] (5)
\[ \log x_{3f} = 0.96 \log x_{3,f-1} + \Delta \log x_{5f} + 0.04 \log x_{5,f-1} - 0.40 d_{08, f} \]

\[ \log x_{4f} = 0.60 \log x_{3,f-1} + \Delta \log x_{5f} - 0.40 \log x_{5,f-1} - 0.29 d_{08, f} - 0.20 \]

Focusing on the impact of the two events, it can be seen that \( x_1 \) and \( x_2 \) were positively affected by the March 2007 event, \( x_3 \) and \( x_4 \) were unaffected and, by the nature of the composition, \( x_5 \) was negatively affected. In other words, those proportions already hostile to the Olympics increased in response to the announcement that the games would be more expensive, those proportions that were indifferent or mildly supportive remained the same, while the proportion strongly in favour declined somewhat. Again these shifts can be seen in Figure 2.

The response to the August 2008 event was that the first four proportions all declined in size, in particular those who were indifferent, while the proportion most supportive increased in size, shifts that can also be seen in Figure 2. The performance of the British team in the Beijing Olympics thus had a pronounced positive influence on support for the London Olympics. The first row of Table 1 provides estimates of the changes in these proportions (measured in percentage points) obtained via simulation of the model. Thus the March 2007 event increased \( x_1 \) and \( x_2 \) by 2.3 and 2.4 percentage points respectively, with \( x_5 \) being correspondingly reduced by 4.7 points. The August 2008 event led to a decline of 0.5, 0.6, 4.8 and 1.5 percentage points for the first four categories, with \( x_5 \) correspondingly increasing by 7.4 points. Table 1 also presents analogous results at a disaggregated level by analysing the geographic groups 'excluding London (Ex
Ldn’), ‘north’, ‘south’ and ‘midlands and east (M&E)’, the age groups 16–29, 30–44, 45–59 and 60+, and males and females. For brevity, details of the models are omitted.

**TABLE 1 [about here]**

The results show some considerable variation in impact. Broadly speaking, the good news of Olympic success produces relatively large increases in strong support for the Olympics, as the above literature review might suggest. However, this is not uniform. It is most prevalent in the younger age-groups and females and more so away from London. This might be because the Olympics reaches an audience that is usual for sport, that of young males, but it also attracts female support because, unlike most professional sports, it is not male dominated. That London and the south seem to respond less to success could be indicative of the concerns of residents of the host areas noted earlier. Interestingly, the impact of the cost overruns being announced in March 2007 are accompanied by increases in support for the Games in the north and for females, which might suggest that, for the reasons noted above, the benefits of hosting the games exceed the costs for these groups. In general, however, for males and those of middle age there is a clear shift away from support for the Games in the context of bad news.

4. Conclusion

Although differences in sentiment play a part in the evolving economic valuation of companies through the stock market, it is less well known how ‘news’ affects sentiment towards major public investments like the Olympics. In this paper the
relationship between attitudes towards the Olympics, the bad news associated with an increase in the cost of provision and the good news associated with Team GB’s medal success in 2008 are analysed. Using a unique data set and a novel event-study approach that involves compositional time-series analysis enables us to identify that ‘good’ news affects sentiment much more than ‘bad’ news, but that the distribution of such sentiment varies widely. We identify that the good news of Olympic success produces large increases in ‘strong support’ towards the Olympics particularly for younger age groups, females and non-London residents. For males and those of middle-age, there is a general reduction in sport for the Games in the context of bad news.
References


[41] UK Sport, *UK Sporting Preferences Survey 2008*.


Figure 1: Feelings Towards the Olympics

Panel (a): Slightly or Strongly Against

Panel (b): Slightly or Strongly Supportive
Figure 2   Proportions having various levels of support for the London Olympics (overall)
Figure 3  Additive log-ratios of proportions
Table 1  Change in support in response to the two events (percentage points)

<table>
<thead>
<tr>
<th>March 2007</th>
<th>August 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>$x_2$</td>
</tr>
<tr>
<td>All</td>
<td>+2.3</td>
</tr>
<tr>
<td>Ex Ldn</td>
<td>+2.2</td>
</tr>
<tr>
<td>North</td>
<td>0</td>
</tr>
<tr>
<td>South</td>
<td>0</td>
</tr>
<tr>
<td>M&amp;E</td>
<td>+1.7</td>
</tr>
<tr>
<td>16-29</td>
<td>0</td>
</tr>
<tr>
<td>30-44</td>
<td>+2.2</td>
</tr>
<tr>
<td>45-59</td>
<td>+2.8</td>
</tr>
<tr>
<td>60+</td>
<td>0</td>
</tr>
<tr>
<td>Male</td>
<td>+5.0</td>
</tr>
<tr>
<td>Female</td>
<td>0</td>
</tr>
</tbody>
</table>