Long Memory and Structural Breaks in the Spanish Stock Market Index

G.M. Caporale and L.A. Gil-Alana

1Centre for Empirical Finance, Brunel University, London, UK
2Department of Quantitative Methods, Faculty of Economics, University of Navarra, Pamplona, Spain

Abstract: It is well known that the Spanish stock market index (IBEX 35) exhibits unit roots. However, the implications of possible structural breaks in this series have not been deeply investigated. In this paper, we show that, when including a break at the beginning of 1998, the order of integration of the series becomes slightly smaller, strengthening the evidence of mean-reverting behaviour. When the break date is supposed to be unknown, it is found to be January 1998, with both subsamples still being characterised by a high degree of persistence.

INTRODUCTION

Two recent papers of DePenya and Gil-Alana [1, 2] show that Spanish stock market index (IBEX 35) can be modelled as a unit root or I(1) process. These authors use both parametric and semiparametric methods to estimate and test the fractional differencing parameter, and conclude that, although fractional degrees of integration (slightly smaller than one) may be plausible in some cases, the unit root null hypothesis cannot be rejected, implying that mean reversion does not occur. In this paper, we examine whether these conclusions are affected by the presence of structural breaks. The interaction of long memory with structural change has been analysed in a number of papers, including applied hydrology [3], econometrics ([4, 5]), and mathematical statistics ([6, 7]). More recently, Diebold and Inoue [8] provide both theoretical and Monte Carlo evidence that structural breaks-based models and long-memory processes are not easily distinguished. Granger and Hyung [9] also analysed theoretically the links between the two types of models, and Gil-Alana [10] showed that the order of integration of some series is reduced by the inclusion of dummy variables for the breaks. Other recent articles of fractional integration with structural change are those of Beran and Terrin [11] and Bos, Franses and Ooms [12, 13].

The outline of this paper is as follows. In the following section we briefly describe a procedure due to Robinson [14], which is suitable to test I(d) statistical models including structural breaks. Then, this procedure is applied to the Spanish stock market index, and also a recently developed procedure (see Gil-Alana, [15]) is applied to test for fractional integration in the presence of a structural break at an unknown point in time. The final section concludes.

THE TESTING PROCEDURE

Following the contributions of Bhargava [16], Schmidt and Phillips [17] and others on the parameterisation of unit-root models, we consider the following specification:

\[ y_t = \beta^* z_t + x_t, \quad t = 1, 2, \ldots, \]  

where \( y_t \) is the time series we observe at \( t = 1, 2, \ldots T \); \( \beta \) is a (kx1) vector of unknown parameters; \( z_t \) is a (kx1) vector of deterministic regressors that may include, for example, dummy variables to incorporate structural breaks, and \( x_t \) is given by:

\[ (1 - L)^d x_t = u_t, \quad t = 1, 2, \ldots \]  

with I(0) \( u_t \). In general, we want to test the null hypothesis:

\[ H_o: \quad d = d_o, \]  

in (1) and (2) for any real value \( d_o \). Based on (3), the least-squares estimate of \( \beta \) and residuals are:

\[ \hat{\beta} = \left( \sum_{j=1}^{T} w_j w_j' \right)^{-1} \sum_{j=1}^{T} w_j (1 - L)^{d_o} y_j; \]

\[ w_j = (1 - L)^{d_o} z_j; \quad \hat{u}_t = (1 - L)^{d_o} y_t - \hat{\beta}^* w_t, \]

and the test statistic proposed by Robinson [14], which is based on the Lagrange Multiplier (LM) principle, is then given by

\[ \hat{r} = \left( \frac{T}{\hat{\sigma}^2} \right)^{1/2} \frac{\hat{\alpha}}{\hat{\sigma}^2}, \]

\[ \hat{r} = \frac{-2\pi}{T} \sum_{j=1}^{T-1} \psi(\lambda_j) g(\lambda_j; \hat{\gamma})^{-1} I(\lambda_j); \]

\[ \hat{\sigma}^2 = \frac{2\pi}{T} \sum_{j=1}^{T-1} g(\lambda_j; \hat{\gamma})^{-1} I(\lambda_j); \]

\[ \hat{\nu} = \frac{2}{T} \left( \sum_{j=1}^{T-1} \psi(\lambda_j)^2 - \sum_{j=1}^{T-1} \psi(\lambda_j) \hat{\nu}(\lambda_j) \hat{\psi}(\lambda_j) \right) \times \left( \sum_{j=1}^{T-1} \hat{\nu}(\lambda_j) \hat{\psi}(\lambda_j) \right)^{-1} \times \left( \sum_{j=1}^{T-1} \hat{\nu}(\lambda_j) \hat{\psi}(\lambda_j) \right)^{-1} \]

where \( \psi(\lambda_j) \) is the characteristic function of the model
\[ \psi(\lambda_j) = \log \left| 2 \sin \frac{\lambda_j}{2} \right|; \]
\[ \hat{v}(\lambda_j) = \frac{d}{d\tau} \log g(\lambda_j; \hat{\tau}); \quad \lambda_j = \frac{2\pi j}{T}. \]

\( I(\lambda_j) \) is the periodogram of \( \hat{u}_t \), and \( g \) above is a function coming from the spectral density of \( u_t \):
\[ f(\lambda; \tau) = \frac{\sigma^2}{2\pi} g(\lambda; \tau), \quad -\pi < \lambda \leq \pi, \]
 evaluated at \( \hat{\tau} = \arg \min_{\tau} \sigma^2(\tau). \) Note that these tests are purely parametric and, therefore, they require specific modelling assumptions about the short-memory specification of \( u_t \). Thus, if \( u_t \) is white noise, \( g = 1 \), and if \( u_t \) is an AR process of the form \( \phi(L)u_t = \epsilon_t, \) \( g = |\phi(\epsilon_j)|^2 \), with \( \sigma^2 = \text{V}(\epsilon_t) \), so that the AR coefficients are a function of \( \tau. \)

Based on the null hypothesis \( H_0 \) (3), Robinson [14] showed that under certain regularity conditions,
\[ \hat{\tau} \rightarrow_d N(0,1) \quad \text{as} \quad T \rightarrow \infty, \quad (5) \]
and also the Pitman efficiency theory against local departures from the null applies. Thus, an approximate one-sided 100\( \alpha \)% level test of (3) will reject \( H_0 \) against the alternative: \( H_1: \hat{\tau} > z_\alpha \) (\( \hat{\tau} < -z_\alpha \)), where the probability that a standard normal variate exceeds \( z_\alpha \) is \( \alpha. \) This version of the tests was used in empirical applications in Gil-Alana and Robinson [18] and in Gil-Alana [19] and, other applied studies of the tests based on seasonal (quarterly and monthly) and cyclical models are Gil-Alana and Robinson [20] and Gil-Alana [21, 22] respectively.

**EMPIRICAL RESULTS**

The time series analysed in this section is the Spanish stock market index (IBEX 35), daily, for the time period 4 January 1994 to 26 November 2001, obtained from the Spanish Stock Exchange Interconnection System (SIBE). The IBEX-35 is a value-weighted index that includes the thirty most traded stocks of the Spanish stock market. Every semester, the effective trading volumes of all stocks are recorded in order to adjust the stocks and their weights and compute the index in the following semester. We use this series in our analysis in order to be able to make a direct comparison with the results obtained in [1] and [2], where the same data were used and the authors found strong evidence of unit roots, but they did not allow for possible breaks. Moreover, the IBEX is a relatively homogeneous market, directly comparable to any other European financial market.

Fig. (1) displays the log-transformed series. Visual inspection reveals a clear change in the mean occurring around the beginning of 1998. Therefore, we perform the analysis for the pre-and post-1998 subperiods as well as for the whole sample period.

Denoting the log-transformed series \( y_t \), we employ throughout model (1) and (2), with \( z_t = 1 \), i.e.,
\[ y_t = \beta + x_t, \quad t = 1, 2, \ldots \quad (6) \]
\[ (1 - L)^d x_t = u_t, \quad t = 1, 2, \ldots \quad (7) \]
and also the Pitman efficiency theory against local departures from the null applies. Thus, an approximate one-sided 100\( \alpha \)% level test of (3) will reject \( H_0 \) against the alternative: \( H_1: \hat{\tau} > z_\alpha \) (\( \hat{\tau} < -z_\alpha \)), where the probability that a standard normal variate exceeds \( z_\alpha \) is \( \alpha. \) This version of the tests was used in empirical applications in Gil-Alana and Robinson [18] and in Gil-Alana [19] and, other applied studies of the tests based on seasonal (quarterly and monthly) and cyclical models are Gil-Alana and Robinson [20] and Gil-Alana [21, 22] respectively.

![Fig. (1). Log of the Spanish stock market prices.](image-url)
root case, and this is so regardless of the sample used. The last two columns in the table report, respectively, the confidence intervals for the values of $d_0$ when $H_0$ (3) cannot be rejected at the 95% significance level, and those which produce the lowest $|\hat{r}|$ across $d_0$. We find that all confidence intervals include the unit root, and the lowest statistics occur when $d_0$ is equal to 1 or smaller. The results are similar in both subsamples, though the values of $d_0$ are slightly smaller in the second one.

Table 2 reports the results of the same statistic as in Table 1 but including a structural break. We set $z_t = (1, S_t)$, first, with $S_t = I(t \geq T_b)$ and then, with $S_t = (t - T_b) I(t \geq T_b)$, $T_b = 2.01.1998$. In other words, we introduce a shift and a slope dummy variable respectively for the break in the regression model (1). Similarly to Table 1, the unit root null cannot be rejected, though the confidence intervals are now smaller, and the values of $d_0$ which produce the lowest statistics are much smaller, especially in the case of the slope dummy. Therefore, it appears that the inclusion of a dummy variable for the break reduces the order of integration of the series, restoring mean-reverting behaviour.

Next, we use an alternative approach to test for fractional integration in the presence of a structural break. Here, unlike in the previous method where it was set a priori, the break date is assumed to be unknown, and is endogenously determined by the model. This procedure was developed by Gil-Alana [15], and it allows for possibly changing intercepts and fractional differencing parameters. The employed model is the following:

\[
y_t = \alpha_1 + \beta_1 \cdot 1_{t \leq T_b} + \epsilon_t, \quad \beta_1 \leq 1
\]

\[
y_t = \alpha_2 + \beta_2 \cdot 1_{t > T_b} + \epsilon_t, \quad \beta_2 > 1
\]

\[
y_t = \alpha_3 + \beta_3 \cdot 1_{t \in \{T_1, T_2\}} + \epsilon_t
\]

where $\alpha_1$, $\beta_1$, and $\beta_2$, $\beta_3$ are the intercepts and fractional differencing parameters, respectively.

The model above can also be written as:

\[
(1 - L)^d y_t = \alpha_1 \cdot (1 - L)^d \cdot 1_{t \leq T_b} + \epsilon_t, \quad t = 1, ..., T_b
\]

\[
(1 - L)^d y_t = \alpha_2 \cdot (1 - L)^d \cdot 1_{t > T_b} + \epsilon_t, \quad t = T_b + 1, ..., T
\]

Table 1. Testing $H_0$ (3) in (1) and (2) with $\hat{r}$ given by (4) in the Log of the Spanish Stock Market Index

<table>
<thead>
<tr>
<th>$\alpha_0/d_0$</th>
<th>0.00</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>1.00</th>
<th>1.25</th>
<th>1.50</th>
<th>1.75</th>
<th>2.00</th>
<th>Conf. Int.</th>
<th>$d^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>White noise</td>
<td>195.94</td>
<td>183.26</td>
<td>113.58</td>
<td>26.44</td>
<td>-0.64</td>
<td>-9.42</td>
<td>-13.49</td>
<td>-15.85</td>
<td>-17.41</td>
<td>[0.97 - 1.02]</td>
<td>0.99</td>
</tr>
<tr>
<td>Bloomfield 1</td>
<td>125.71</td>
<td>110.07</td>
<td>63.07</td>
<td>15.39</td>
<td>-1.05</td>
<td>-6.92</td>
<td>-9.66</td>
<td>-11.11</td>
<td>-12.10</td>
<td>[0.94 - 1.01]</td>
<td>0.97</td>
</tr>
<tr>
<td>Bloomfield 2</td>
<td>97.81</td>
<td>86.00</td>
<td>40.85</td>
<td>12.16</td>
<td>-1.36</td>
<td>-8.23</td>
<td>-10.75</td>
<td>-11.75</td>
<td>-12.09</td>
<td>[0.98 - 1.00]</td>
<td>0.99</td>
</tr>
<tr>
<td>ii) First Subsample (4.01.94 - 31.12.97)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White noise</td>
<td>120.54</td>
<td>114.10</td>
<td>79.03</td>
<td>20.40</td>
<td>-0.67</td>
<td>-6.94</td>
<td>-9.79</td>
<td>-11.43</td>
<td>-12.50</td>
<td>[0.95 - 1.02]</td>
<td>0.99</td>
</tr>
<tr>
<td>Bloomfield 1</td>
<td>75.06</td>
<td>65.55</td>
<td>42.08</td>
<td>12.11</td>
<td>-0.71</td>
<td>-5.17</td>
<td>-7.21</td>
<td>-8.25</td>
<td>-9.07</td>
<td>[0.93 - 1.03]</td>
<td>0.98</td>
</tr>
<tr>
<td>Bloomfield 2</td>
<td>55.03</td>
<td>44.69</td>
<td>35.52</td>
<td>10.43</td>
<td>-0.31</td>
<td>-3.50</td>
<td>-5.78</td>
<td>-6.68</td>
<td>-7.01</td>
<td>[0.94 - 1.07]</td>
<td>1.00</td>
</tr>
<tr>
<td>iii) Second Subsample (2.01.98 - 29.11.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White noise</td>
<td>101.78</td>
<td>80.50</td>
<td>45.51</td>
<td>14.83</td>
<td>-0.51</td>
<td>-6.61</td>
<td>-9.53</td>
<td>-11.24</td>
<td>-12.35</td>
<td>[0.95 - 1.03]</td>
<td>0.97</td>
</tr>
<tr>
<td>Bloomfield 1</td>
<td>59.67</td>
<td>42.56</td>
<td>22.89</td>
<td>8.00</td>
<td>-0.84</td>
<td>-4.77</td>
<td>-6.76</td>
<td>-7.81</td>
<td>-8.72</td>
<td>[0.91 - 1.04]</td>
<td>0.96</td>
</tr>
<tr>
<td>Bloomfield 2</td>
<td>42.72</td>
<td>28.53</td>
<td>14.12</td>
<td>5.63</td>
<td>-1.11</td>
<td>-3.56</td>
<td>-5.58</td>
<td>-6.39</td>
<td>-6.899</td>
<td>[0.91 - 1.05]</td>
<td>0.92</td>
</tr>
</tbody>
</table>

In bold: The non-rejection values of the null hypothesis at the 95% significance level. $d^*$ is the value of $d$ producing the lowest statistic across $d$. The Open Operational Research Journal, 2008, Volume 2
with a high degree of dependence. Note that in a fractional
determining the appropriate order of integration in series
important problem in econometrics, which is the difficulty of
coefficient extremely close to 1 (0.9992). This illustrates an
slight, the I(0) model for the first subsample having an AR
properties, though the differences in specification are only

The estimated break date, $\hat{T}_1$, is such that:
$\hat{T}_1 = \arg\min_{i=1, \ldots, n} RSS(T_i)$, where the minimisation is
done over all partitions $T_1, T_2, \ldots, T_m$, such that $T_i - T_{i-1} \geq |\varepsilon|$. Then, the regression parameter estimates are the associated
least-squares estimates of the estimated k-partition, i.e.,
$\hat{\alpha}_i = \hat{\alpha}_i((\hat{T}_i))$, and their corresponding differencing parameters, $\hat{d}_i = \hat{d}_i((\hat{T}_i))$, for $i = 1$ and 2. Several Monte Carlo
experiments conducted in [15] show that this procedure performs relatively well even for small sample sizes.

The results for the two cases of white noise and AR(1)
disturbances are displayed in Table 3. The estimated break
date is January 12th, 1998 for the two types of disturbances.
Starting with the white noise case, one can see that the order
of integration for the first subsample is slightly above 1
(below unity (0.95), implying mean reversion.

When allowing for an AR(1) structure in the error term,
the model corresponding to the lowest RSS is the one with $d_1$
= 0 and $d_2$ = 1.02, implying short memory for the first sub-
and long memory (no mean reversion) after the
break. However, a rival model, with a slightly higher RSS, is
the one with $d_1$ = 1 and the same $d_2$ as before (1.02). Note
that these two models have completely different statistical
properties, though the differences in specification are only
slight, the I(0) model for the first subsample having an AR
coefficient extremely close to 1 (0.9992). This illustrates an
important problem in econometrics, which is the difficulty of
determining the appropriate order of integration in series
with a high degree of dependence. Note that in a fractional

| Table 2. Testing $H_0$ (3) in (1) and (2) with $\hat{r}$ given by (4) in the Log of the Spanish Stock Market Index |

<table>
<thead>
<tr>
<th>Break-Date</th>
<th>$d_1$</th>
<th>$d_2$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\rho_1$</th>
<th>$\rho_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>White noise</td>
<td>12. Jan. 1998</td>
<td>0.00</td>
<td>0.95</td>
<td>8.4252 (616.24)</td>
<td>9.1257 (578.44)</td>
<td>---</td>
</tr>
<tr>
<td>AR (1)</td>
<td>12. Jan. 1998</td>
<td>0.00</td>
<td>1.02</td>
<td>9.23132 (996.46)</td>
<td>7.89020 (488.77)</td>
<td>0.9992</td>
</tr>
<tr>
<td>AR (1)</td>
<td>12. Jan. 1998</td>
<td>1.00</td>
<td>1.02</td>
<td>8.62895 (631.42)</td>
<td>7.89020 (488.77)</td>
<td>0.0423</td>
</tr>
</tbody>
</table>

| Table 3. Estimates in a Fractional Model with a Single Structural Break |

<table>
<thead>
<tr>
<th>Break-Date</th>
<th>$d_1$</th>
<th>$d_2$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\rho_1$</th>
<th>$\rho_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>White noise</td>
<td>12. Jan. 1998</td>
<td>0.25</td>
<td>0.50</td>
<td>0.75</td>
<td>1.00</td>
<td>1.25</td>
</tr>
<tr>
<td>Bloomfield 1</td>
<td>136.09</td>
<td>86.61</td>
<td>25.70</td>
<td>-0.64</td>
<td>-9.42</td>
<td>-13.49</td>
</tr>
<tr>
<td>Bloomfield 2</td>
<td>75.52</td>
<td>46.52</td>
<td>14.62</td>
<td>-1.04</td>
<td>-6.92</td>
<td>-9.67</td>
</tr>
</tbody>
</table>

In bold: The non-rejection values of the null hypothesis at the 95% significance level.
$d^*$ is the value of $d$ producing the lowest statistic across $d$.

CONCLUSIONS

In this paper we have examined the stochastic behaviour
of the Spanish stock market index (IBEX 35) using fractionally
integrated techniques, also allowing for possible struc-
tural breaks. The results show that the inclusion of a dummy
variable for the break slightly reduces the order of integra-
tion of the series, and thus the mean-reversion property of
prices is reinforced. Admittedly, even when allowing for a
break the unit root null still cannot be rejected at the 5% sig-
ificance level. However, in all cases the test statistic is very
close to the boundary of the confidence interval (see the last
two columns in Table 2), and therefore the evidence of mean
reversion is strengthened compared to the case without a
break. When the break date is assumed to be unknown, the
results support the mean-reversion hypothesis for the second
subsample if the underlying disturbances are white noise.
However, when allowing for weak autocorrelated terms, the
I(0) and the I(1) hypotheses are difficult to tell apart in
the first subsample.

REFERENCES

[1] DePenya FJ, Gil-Alana LA. Mean reversion in the Spanish stock
market prices using fractionally integrated semiparametric tech-
Long Memory and Structural Breaks in the Spanish Stock Market Index


