Long Memory and Structural Breaks in the Spanish Stock Market Index

G.M. Caporale[1] and L.A. Gil-Alana[*,2]

[1]*Centre for Empirical Finance, Brunel University, London, UK*

[2]*Department 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:

$$y_t \;=\; \beta' z_t \;+\; x_t, \qquad t \;=\; 1, 2, ..., \tag{1}$$

*Address correspondence to this author at the Department of Quantitative Methods, Faculty of Economics, University of Navarra, Pamplona, Spain; E-mail: alana@unav.es

where y_t is the time series we observe at t = 1, 2, …T; β 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:

$$(1 \;-\; L)^d\, x_t \;=\; u_t, \qquad t \;=\; 1, 2, ... \tag{2}$$

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

$$H_o: \quad d \;=\; d_o, \tag{3}$$

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

$$\hat{\beta} \;=\; \left(\sum_{t=1}^{T} w_t\, w_t'\right)^{-1} \sum_{t=1}^{T} w_t\,(1 \;-\; L)^{d_o}\, y_t;$$

$$w_t = (1 \;-\; L)^{d_o} z_t; \;\; \hat{u}_t = (1-L)^{d_o}\, y_t - \hat{\beta}' w_t,$$

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

$$\hat{r} \;=\; \left(\frac{T}{\hat{A}}\right)^{1/2} \frac{\hat{a}}{\hat{\sigma}^2}, \tag{4}$$

$$\hat{a} \;=\; \frac{-2\pi}{T}\sum_{j=1}^{T-1} \psi(\lambda_j)\, g(\lambda_j;\hat{\tau})^{-1}\, I(\lambda_j);$$

$$\hat{\sigma}^2 \;=\; \frac{2\pi}{T}\sum_{j=1}^{T-1} g(\lambda_j;\hat{\tau})^{-1}\, I(\lambda_j);$$

$$\hat{A} = \frac{2}{T}$$

$$\left(\sum_{j=1}^{T-1}\psi(\lambda_j)^2 \;-\; \sum_{j=1}^{T-1}\psi(\lambda_j)\,\hat{\varepsilon}(\lambda_j)' \;\times \left(\sum_{j=1}^{T-1}\hat{\varepsilon}(\lambda_j)\,\hat{\varepsilon}(\lambda_j)'\right)^{-1} \times \sum_{j=1}^{T-1}\hat{\varepsilon}(\lambda_j)\,\psi(\lambda_j)\right)$$

$$\psi(\lambda_j) = \log\left|2\sin\frac{\lambda_j}{2}\right|;$$

$$\hat{\varepsilon}(\lambda_j) = \frac{\partial}{\partial\tau}\log g(\lambda_j;\hat{\tau}); \quad \lambda_j = \frac{2\pi j}{T}.$$

$I(\lambda_j)$ is the periodogram of $\hat{u}_t$, and g above is a function coming from the spectral density of u_t: $f(\lambda;\tau) = \frac{\sigma^2}{2\pi}g(\lambda;\tau), \quad -\pi < \lambda \le \pi$, evaluated at $\hat{\tau}$ = arg $\min_{\tau \in T} \sigma^2(\tau)$. 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, $g \equiv 1$, and if u_t is an AR process of the form $\phi(L)u_t = \varepsilon_t$, $g = |\phi(e^{i\lambda})|^{-2}$, with $\sigma^2 = V(\varepsilon_t)$, so that the AR coefficients are a function of τ.

Based on the null hypothesis H_o (3), Robinson [14] showed that under certain regularity conditions,

$$\hat{r} \rightarrow_d N(0,1) \quad as \quad T \rightarrow \infty, \tag{5}$$

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_o against the alternative: H_a: $d > d_o$ ($d < d_o$) if $\hat{r} > z_\alpha$ ($\hat{r} < -z_\alpha$), where the probability that a standard normal variate exceeds z_α is α. 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 five 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.,

$$y_t = \beta + x_t, \qquad t = 1, 2, \ldots \tag{6}$$

$$(1 - L)^d x_t = u_t, \quad t = 1, 2, \ldots, \tag{7}$$

testing H_o (3) for values d_o = 0, (0.25), 2, and different types of disturbances. In particular, we assume white noise and also Bloomfield [23] disturbances. The latter is a non-parametric approach for modelling u_t which produces autocorrelations decaying exponentially as in the autoregressive (AR) processes.

The test statistic reported in Table **1** (and also in Table **2**) is the one-sided statistic given by $\hat{r}$ in (4). Thus, for a given d_o, significantly positive values of $\hat{r}$ are consistent with orders of integration higher than d_o, whereas significantly negative ones imply orders of integration smaller than the one under the null. Table **1** (i) reports the results based on the whole sample, while Tables **1** (**ii**) and (**iii**) correspond respectively to the first and the second subsample (4.01.94 - 31.12.97 and 2.01.98 - 29.11.01). It can be seen that the only non-rejection value of d corresponds to d = 1, i.e. to the unit

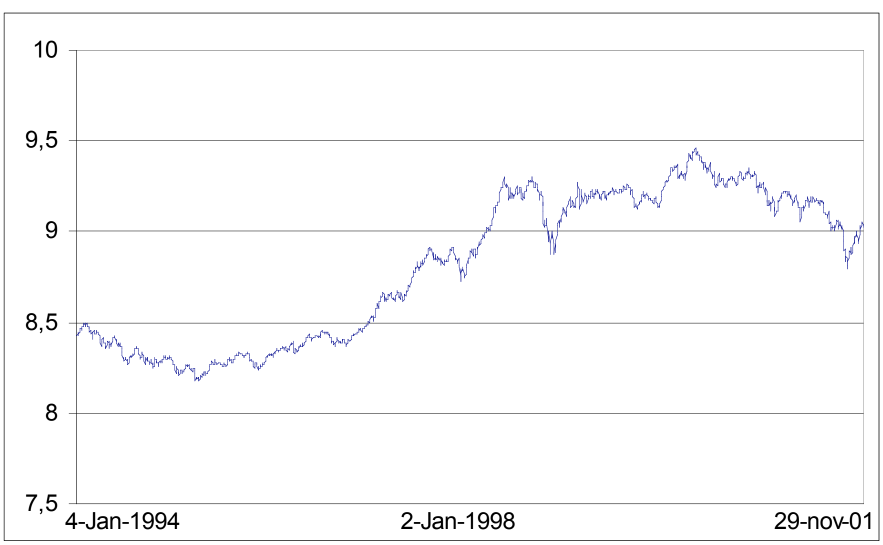

Fig. (1). Log of the Spanish stock market prices.

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_o when H_o (3) cannot be rejected at the 95% significance level, and those which produce the lowest $|\hat{r}|$ across d_o. We find that all confidence intervals include the unit root, and the lowest statistics occur when d_o is equal to 1 or smaller. The results are similar in both subsamples, though the values of d_o 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_o 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:

$$y_t = \alpha_1 + x_t; \qquad (1-L)^{d_1} x_t = u_t, \quad t = 1, \dots, T_b \tag{8}$$

$$y_t = \alpha_2 + x_t; \qquad (1-L)^{d_2} x_t = u_t, \quad t = T_b + 1, \dots, T, \tag{9}$$

where the α's are the coefficients corresponding to the intercepts; d_1 and d_2 can be real values, and represent the orders of integration for each subsample; u_t is I(0) and T_b is the time of the break that is assumed to be unknown. Note that the model above can also be written as:

$$(1 - L)^{d_1} y_t = \alpha_1 \tilde{1}_t(d_1) + u_t, \quad t = 1, \dots, T_b, \tag{10}$$

$$(1 - L)^{d_2} y_t = \alpha_2 \tilde{1}_t(d_2) + u_t, \quad t = T_b + 1, \dots, T, \tag{11}$$

where $\tilde{1}_t(d_i) = (1 - L)^{d_i} 1$, i = 1, 2.

This method is based on the least squares principle. First we choose a grid for the values of the fractionally differencing parameters d_1 and d_2, for example, d_{io} = 0, 0.01, 0.02, …, 2, i = 1, 2. Then, for a given partition $\{T_b\}$ and given d_1, d_2-values, (d_{1o}, d_{2o}), we estimate the α's and the β's by minimising the sum of squared residuals,

$$\min \sum_{t=1}^{T_b} \left[(1-L)^{d_{1o}} y_t - \alpha_1 \tilde{1}_t(d_{1o}) \right]^2 + \sum_{t=T_b+1}^{T} \left[(1-L)^{d_{2o}} y_t - \alpha_2 \tilde{1}_t(d_{2o}) \right]^2 .$$

$$w.r.t. \{\alpha_1, \alpha_2\}$$

Let $\hat{\alpha}(T_b;\, d_{1o}^{(1)}, d_{2o}^{(1)})$ denote the resulting estimates for partition $\{T_b\}$ and initial values $d_{1o}^{(1)}$ and $d_{2o}^{(1)}$. Substituting these estimated values into the objective function, we obtain RSS(T_b; $d_{1o}^{(1)}$, $d_{2o}^{(1)}$), and minimising this expression over all values of d_{1o} and d_{2o} in the grid we get: $RSS(T_b) = \arg\min_{\{i,j\}} \; RSS(T_b; d_{1o}^{(i)}, d_{2o}^{(j)})$. Then, the esti-

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

u_t/d_o	0.00	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	Conf. Int.	d*
i) Whole Sample (4.01.94 - 29.11.01)											
White noise	195.94	183.26	113.58	26.44	**-0.64**	-9.42	-13.49	-15.85	-17.41	[0.97 - 1.02]	0.99
Bloomfield 1	125.71	110.07	63.07	15.39	**-1.05**	-6.92	-9.66	-11.11	-12.10	[0.94 - 1.01]	0.97
Bloomfield 2	97.81	86.00	40.85	12.16	**-1.36**	-8.23	-10.75	-11.75	-12.09	[0.98 - 1.00]	0.99
ii) First Subsample (4.01.94 - 31.12.97)											
White noise	120.54	114.10	79.03	20.40	**-0.67**	-6.94	-9.79	-11.43	-12.50	[0.95 - 1.02]	0.99
Bloomfield 1	75.06	65.55	42.08	12.11	**-0.71**	-5.17	-7.21	-8.25	- 9.07	[0.93 - 1.03]	0.98
Bloomfield 2	55.03	44.69	35.52	10.43	**-0.11**	-3.50	- 5.78	-6.68	-7.01	[0.94 - 1.07]	1.00
iii) Second Subsample (2.01.98 - 29.11.01)											
White noise	101.78	80.50	45.51	14.83	**-0.51**	-6.61	- 9.53	-11.24	-12.35	[0.95 - 1.03]	0.97
Bloomfield 1	59.67	42.56	22.89	8.00	**-0.84**	-4.77	-6.76	- 7.81	- 8.72	[0.91 - 1.04]	0.96
Bloomfield 2	42.72	28.53	14.12	5.63	**-1.11**	-3.56	- 5.58	- 6.39	-6.899	[0.91 - 1.05]	0.92

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.

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

u_t/d_o	0.00	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00	Conf. Int.	d^*
i) with a Shift Dummy											
White noise	162.06	0.25	0.50	0.75	1.00	1.25	1.50	1.75	-17.40	[0.95 - 1.01]	d^*
Bloomfield 1	98.02	136.09	86.61	25.70	**-0.64**	-9.42	-13.49	-15.84	-12.09	[0.93 - 1.01]	0.96
Bloomfield 2	65.62	75.52	46.52	14.62	**-1.04**	-6.92	-9.67	-11.10	-9.72	[0.91 - 1.00]	0.96
ii) with a Slope Dummy											
White noise	193.11	179.28	112.36	27.36	**-0.62**	-9.42	-13.49	-15.85	-17.40	[0.94 - 1.01]	0.95
Bloomfield 1	125.47	107.32	62.03	16.41	**-1.02**	-6.92	-9.66	-11.10	- 12.30	[0.92 – 1.00]	0.94
Bloomfield 2	86.11	83.52	39.95	9.99	**-1.32**	-4.15	-7.98	-9.05	-9.68	[0.90 - 1.00]	0.91

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.

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

	Break-Date	d_1	d_2	α_1	α_2	ρ_1	ρ_2
White noise.	12. Jan. 1998	1.02	0.95	8.4252 (616.24)	9.1257 (578.44)	---	---
AR (1)	12. Jan. 1998	0.00	1.02	9.23132 (996.46)	7.89020 (488.77)	0.9992	-0.0730
AR (1)	12. Jan. 1998	1.00	1.02	8.62895 (631.42)	7.89020 (488.77)	0.0423	-0.0730

mated break date, $\hat{T}_k$, is such that: $\hat{T}_k = \arg\min_{i=1,\dots,m} RSS(T_i)$, where the minimisation is done over all partitions $T_1, T_2, \dots, T_m$, such that $T_i - T_{i-1} \geq |\varepsilon T|$. Then, the regression parameter estimates are the associated least-squares estimates of the estimated k-partition, i.e., $\hat{\alpha}_i = \hat{\alpha}_i(\{\hat{T}_k\})$, and their corresponding differencing parameters, $\hat{d}_i = \hat{d}_i(\{\hat{T}_k\})$, 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 (1.02), while d_2 (the order of integration after the break) is 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 subsample 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 model with AR disturbances there are two competing forces trying to capture time dependence between the observations: the fractional differencing parameter and the AR coefficient. Theoretically, the former describes the long-run behaviour while the latter concerns the short-run dynamics. However, in practice, this distinction is not clear.

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 structural breaks. The results show that the inclusion of a dummy variable for the break slightly reduces the order of integration 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% significance 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 integgrated semiparametric techniques. Int J Th Appl Finance 2002; 25: 647- 657.

[2] DePenya FJ, Gil-Alana LA. Do Spanish stock market prices follow a random walk? Eur Rev Econ Fin 2004; 3: 3-13.

[3] Klemes V. The Hurst phenomenon. A Puzzle? Water Resour Res 1994; 10: 675-688.

[4] Hidalgo FJ, Robinson PM. Testing for structural change in a long memory environment. J Econometrics 1996; 70: 159-174.

[5] Lobato IN, Savin NW. Real and spurious long memory properties of stock market data, J Bus Econ Stat 1997; 16: 261-283.

[6] Kunsch HR. Discrimination between monotonic trends and long range dependence. J Appl Prob 1986; 23: 1025-1030.

[7] Teverovcsky V, Taqqu MS. Testing for long range dependence in the presence of shifting means or a slowly declining trend, using a variance-type estimator. J Time Ser Anal 1997; 18: 279-304.

[8] Diebold FX, Inoue A. Long memory and regime switching. J Econometrics 2001; 105: 131-159.

[9] Granger CWJ, Hyung N. Occasional structural breaks and long memory with an application to the S&P 500 absolute stock return. J Emp Fin 2004; 11: 299-421.

[10] Gil-Alana LA. A fractionally integrated model with a mean shift for the US and the UK real oil prices. Econ Model 2001; 18: 643-658.

[11] Beran J, Terrin N. Testing for a change of the long memory parameter. Biometrika 1996; 83: 627-683.

[12] Bos CS, Franses PH, Ooms M. Long memory and level shifts. Reanalyzing inflation rates. Emp Econ 1999; 4: 427-449.

[13] Bos CS, Franses PH, Ooms M. Inflation, forecast intervals and long memory regression models. Int J Forecasting 2002; 18: 243-431.

[14] Robinson PM. Efficient tests of nonstationary hypotheses. J Am Stat Assoc 1994; 9: 1420-1437.

[15] Gil-Alana LA. Fractional integration and structural breaks at an unknown point in time. J Time Ser Anal 2008, forthcoming.

[16] Bhargava A. On the theory of testing unit roots in observed time series. Rev Econ Stat 1986; 53: 369-384.

[17] Schmidt P, Phillips PCB. LM tests for a unit root in the presence of deterministic trends. Oxford B Econ Stat 1992; 54: 257-287.

[18] Gil-Alana LA, Robinson PM. Testing of unit roots and other non-stationary hypotheses in macroeconomic time series. J Econometrics 1997; 80; 241-268.

[19] Gil-Alana LA. Mean reversion in the real exchange rates. Econ Lett 2000; 69: 285-288.

[20] Gil-Alana LA, Robinson PM. Testing of seasonal fractional integration in the UK and Japanese consumption and income series. J Appl Econometrics 2001; 16: 95-114.

[21] Gil-Alana LA. Fractional integration with monthly data. Econ Model 1999; 16: 613-629.

[22] Gil-Alana LA. Testing stochastic cycles in macroeconomic time series. J Time Ser Anal 2001; 22: 411-430.

[23] Bloomfield P. An exponential model in the spectrum of a scalar time series. Biometrika 1973; 60: 217-226.

Received: January 17, 2008 Revised: February 12, 2008 Accepted: February 15, 2008