

NOTICE: this is the author's version of a work that was accepted for publication in *Economic Modelling*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Economic Modelling*, 19.1 (2002), DOI: 10.1016/S0264-9993(00)00061-4

Testing for non-stationarity and cointegration allowing for the possibility of a structural break: an application to EuroSterling interest rates

Chris Brooks* and Alistair G. Rew

July 2000

Abstract

In this paper we examine the order of integration of EuroSterling interest rates by employing techniques that can allow for a structural break under the null and/or alternative hypothesis of the unit root tests. In light of these results we investigate the cointegrating relationship implied by the single, linear expectations hypothesis of the term structure of interest rates employing two techniques, one of which allows for the possibility of a break in the mean of the cointegrating relationship. The aim of the paper is to investigate whether or not the interest rate series can be viewed as $I(1)$ processes and furthermore to consider whether there has been a structural break in the series. We also determine whether, if we allow for a break in the cointegration analysis, the results are consistent with those obtained when a break is not allowed for. The main results reported in this paper support the conjecture that the 'short' Euro-Currency rates are characterised as $I(1)$ series that exhibit a structural break on or near Black Wednesday, 16 September 1992 whereas the 'long' rates are $I(1)$ series that do not support the presence of a structural break. The evidence from the cointegration analysis suggests that tests of the expectations hypothesis based on data sets that include the ERM crisis period, or a period that includes a structural break, might be problematic if the structural break is not explicitly taken into account in the testing framework.

Keywords: Eurocurrency Interest rates, rational expectations, unit roots, structural breaks, augmented Dickey Fuller test, Perron test, cointegration tests.

1. Introduction

A large and ever increasing body of econometric analysis has sought to test the validity of the rational expectations hypothesis of the term structure of interest rates, henceforth the expectations hypothesis. The expectations hypothesis defines long rates as a weighted-average of current and expected future short rates and consequently it follows that a rise in the long rates relative to the short rates can be justified on the basis that the short-term rates are expected to increase in the future.

The results outlined in many of the articles investigating the expectations hypothesis (see Shiller (1990) for a survey of some of the literature) have been far from conclusive with respect to either rejecting or not rejecting it. Evidence provided by Campbell and Shiller (1987), Fama (1984), Mankiw and Miron (1986), employing US data, and MacDonald and Speight (1988), Hurn, Moody and Muscatelli (1995) and Cuthbertson (1993), employing UK data, all support the main conjecture of the expectation hypothesis. The apparent success of the expectations hypothesis is however brought into question by evidence documented in Shiller (1979), for the UK, and Shiller, Campbell and Schoenholtz (1983) and Campbell and Shiller (1984, 1991), for the US. In all cases the authors report that the spread predicts the opposite of the actual change in the long rate. Consequently, an increase in the spread was followed by a decrease in the long rate in the next period whilst the expectations

* Address for correspondence: Chris Brooks, ISMA Centre, Department of Economics, PO Box 242, Whiteknights Park, University of Reading RG6 6BA, UK. Phone: 0118 987 5123 Fax: 0118 931 4741; e-mail: C.Brooks@ismacentre.reading.ac.uk . We would like to thank Simon Burke, Olan Henry and an anonymous referee for useful discussions and comments on an earlier draft of this paper respectively. The usual disclaimer applies.

hypothesis would have predicted that the long rate will actually increase in the next period. Hence, extant studies of the validity of the expectations hypothesis provide conflicting conclusions.

The problems that we are faced with in attempting to reconcile these apparent contradictory findings are, firstly, why do tests of the expectations hypothesis seem to provide different results from paper to paper and, secondly, is this evidence of a fundamental problem with the expectations hypothesis or merely with the existing testing-framework? We feel that there is evidence that supports the expectations hypothesis in many of the results that have been reported and, in particular, if we consider those for the US then when the sample period 1979 to 1982 is not included the results, on balance, are more supportive of the expectations hypothesis than not. This is of interest since Hamilton (1988) and Sola and Driffill (1994) argue that the apparent failure of the expectations hypothesis over this period may have been caused by a change in the monetary policy operations of the Federal Reserve during this period¹. In particular, the failure seems to have occurred as a consequence of the fact that the monetary policy change induced a change in the level and volatility of interest rates, i.e. a structural break in the series. If, as appears to be the case, the series did indeed experience a structural break then the reported failures of the expectations hypothesis could have arisen from the fact that the break has not been incorporated within the testing framework. If we consider the expectations hypothesis, the linearised version assumes that there is not a break in the data generating process, henceforth the DGP, of the short rate. If there is indeed a break in the short rate, then the linearisation will be violated and therefore effects on long rates will be found to be inconsistent with forecasts of future short rates derived from a series that has had a stable DGP imposed on it. To this end, Hamilton (1988) developed a model of stochastic regime changes that is able to accommodate such a change in the DGP of the short rates and is able to explain the effect on long rates which is inconsistent with a stable DGP for forecasts of future short rates.

It is important that we investigate the possibility of a structural break in the DGP of interest rate series since if there is evidence that supports its existence, then any analysis of the expectations hypothesis is likely to prove problematic. A structural break, represented by a shift in the mean or/and trend of the series, is likely to cause the effect on the long rate, as prescribed by the expectations hypothesis, to be inconsistent with forecasts of future short rates that have been derived from a non-breaking and therefore stable DGP, hence explaining the conflicting results that have previously been reported. It may be possible to reconcile the expectations hypothesis with the data in question if we can incorporate any structural breaks within the model itself. Further, the majority of the testing frameworks developed for the expectations hypothesis require that the rates are $I(1)$ and thus it is important that we correctly classify the series prior to such analysis. The existence of a structural break in a series is likely to cause a bias in the traditional unit root testing framework and consequently we may not reject the null of a unit root even if the series was known to be $I(0)$ with a structural break. If we incorrectly classify the series as $I(1)$ rather than $I(0)$ with a structural break, then the testing framework developed for the expectations hypothesis will be inappropriate. Similarly, it is feasible that the presence of a structural break in the series has in

¹ In October 1979 the Federal Reserve changed their operating procedure from targeting interest rates to targeting non-borrowed reserves. The outcome was a significant change in the average level of the rates, from 7.6% to 15.3%, and their variance, from 5.7% to 10.5% (see Lewis, 1991).

fact biased the testing procedures such that evidence documented in support of the expectations hypothesis is in fact spurious and simply a bi-product of a mis-specified testing framework.

We enter the debate for two reasons. Firstly we consider and apply methodologies that allow for structural breaks in the unit root and cointegration testing frameworks. In particular, we identify the most appropriate characterisation of the EuroSterling interest rate series analysed and in light of these results, we assess whether allowing for a structural break in the cointegration testing framework has any impact on the evidence found for the expectations hypothesis. Secondly, the effect of a change in monetary policy on the behaviour of interest rates has been addressed for the US over the period 1979 to 1982 but there has not, to our knowledge, been an assessment of the expectations hypothesis over the ERM crisis of 1992, which resulted in a significant change in monetary policy in the UK and we wish to remedy this omission.

The rest of the paper is ordered as follows. Section 2 details the linearised-expectations hypothesis of the term structure of interest rates, henceforth the EH. Section 3 outlines the testing procedures employed to determine the order of integration of the individual rates and to test for cointegration between pairs of the rates. Section 4 reports our empirical results, which are then discussed at more length in Section 5, before drawing our conclusions in Section 6.

2. The expectations hypothesis of the term structure

The following notation, to simplify the overview of the theory closely, replicates that employed by Campbell and Shiller (1991). The single, linear-expectations theory of the term structure used to represent the EH defines a relationship between an n-period interest rate or yield, denoted $R_t^{(n)}$, and an m-period interest rate, denoted $R_t^{(m)}$, where $n > m$. Hence $R_t^{(n)}$ is the interest rate or yield on a longer term period relative to a shorter term interest rate or yield, $R_t^{(m)}$. More precisely, the EH states that the expected return from investing in an n-period rate will equal the expected return from investing in m-period rates up to n-m periods in the future plus a constant risk-premium, c , which can be expressed as:

$$R_t^{(n)} = \frac{1}{q} \sum_{i=0}^{q-1} E_t R_{t+mi}^{(m)} + c \quad (1)$$

where $q = n / m$. Consequently, the longer-term interest rate, $R_t^{(n)}$ can be expressed as a weighted-average of current and expected shorter-term interest rates, $R_t^{(m)}$, plus a constant risk premium, c . If we consider (1), it can be seen that by subtracting $R_t^{(m)}$ from both sides of the relationship we have:

$$R_t^{(n)} - R_t^{(m)} = \frac{1}{q} \sum_{i=1}^{q-1} \sum_{j=1}^{j=i} E_t \left[\Delta^{(m)} R_{t+jm}^{(m)} \right] + c \quad (2)$$

Examination of (2) generates some interesting restrictions. If the interest rates that are being analysed, say $R_t^{(n)}$ and $R_t^{(m)}$, are $I(1)^2$ series then by definition $\Delta R_t^{(n)}$ and $\Delta R_t^{(m)}$ will be stationary series. Further, since c is a constant then it is by definition a stationary series. Consequently, if the EH is to hold, given that c and $\Delta R_t^{(m)}$ are $I(0)$ implying that the right-hand side of (2) is stationary, then $R_t^{(n)} - R_t^{(m)}$ must by definition be stationary, otherwise we will have an inconsistency in the order of integration between the right and left-hand sides of the relationship. $R_t^{(n)} - R_t^{(m)}$ is commonly known as the spread between the n -period and m -period rates, denoted $S_t^{(n,m)}$, which in turn gives an indication of the slope of the term structure. Consequently, it follows that if the EH is to hold then the spread will be found to be stationary and therefore $R_t^{(n)}$ and $R_t^{(m)}$ will cointegrate with a cointegrating vector $(1, -1)$ for $[R_t^{(n)}, R_t^{(m)}]$. Therefore, the integrated process driving each of the two rates is common to both and hence it can be said that the rates have a common stochastic trend. As a result, since the EH predicts that each interest rate series will cointegrate with the one period interest rate, it must be true that the stochastic process driving all the rates is the same as that driving the one-period rate, i.e. any combination of rates formed to create a spread should be found to cointegrate with a cointegrating vector $(1, -1)^3$. Consequently, if the rates do cointegrate, we should find that e_t in (3) is stationary:

$$R_t^{(n)} = \alpha + \beta R_t^{(m)} + e_t \quad (3)$$

Further, the EH predicts that $\alpha = 0$ and $\beta = 1$, but since we estimate (3) via OLS, we cannot statistically test the joint restriction on the coefficient values. However, since our main interest is in determining whether we can detect cointegration between the rates, the inability to test the statistical significance of the coefficients in no way hinders our analysis. Further, Cuthbertson and Nitzsche (1994), p. 7, state that “*there is not a great deal of difference between the point estimates from the OLS cointegrating regression and the fully modified estimator*”. Therefore, we can expect that our estimates will be close in magnitude to those that would be generated by a more suitable estimation procedure such as the Phillips and Hansen (1990) fully modified estimators⁴.

The analysis that we undertake involves determining the order of integration of the interest rate series and if applicable, we test for cointegration between pairs of rates. To this end, we now provide a brief review of the main statistical procedures used to test for a unit root against an alternative hypothesis of a stationary series, a unit root against an alternative hypothesis of a stationary series that exhibits a one time structural break, and a unit root process that exhibits a structural break against an alternative hypothesis of a stationary series that

² There is a general acceptance that interest rates, Treasury Bill yields, etc are well described as $I(1)$ processes and this can be seen in Campbell and Shiller (1988), Stock and Watson (1988), Hurn, Moody and Muscatelli (1995) and Mills (1991).

³ This would be expected since “*the spread vector associated with any two yields is just a linear combination of two spread vectors defined using the one period yield*”, Hall, Anderson and Granger (1992), p. 118, and given that two spreads vectors, defined using the one period rate, are stationary then the linear combination of the two stationary spreads must itself be stationary.

⁴ One procedure that is commonly employed is termed the Adjusted-Hall procedure, see Hall (1986), which involves estimating the Phillips and Hansen (1990) fully modified estimators for the two regressions given by $R_t^{(n)} = \alpha + \beta R_t^{(m)} + e_t$ and then $R_t^{(m)} = \alpha + \beta R_t^{(n)} + e_t$ for which Hall (1986) notes that the two regressions should provide a band for the point estimates of the cointegrating parameter β . However, since our main interest is in determining whether the application of different testing procedure results in different conclusions with respect to whether the series are individually $I(1)$ processes and whether the rates cointegrate, we do not place too much emphasis on the estimates of the estimated cointegrating vector.

exhibits a structural break⁵. Additionally, we outline the two cointegration procedures that we employ in our analysis.

3. The testing procedures

(i) The unit root testing procedures

The first test that we consider is the standard augmented Dickey-Fuller test of a unit root (see Dickey and Fuller (1979, 1981)). This test is based on the t-statistic for $\alpha = 1$ in the regression given by:

$$y_t = \mu + \beta t + \alpha y_{t-1} + \sum_{j=1}^p c_j \Delta y_{t-j} + e_t \quad (4)$$

Typically, rejection of the null hypothesis would be taken as strong evidence of trend stationarity, whilst non-rejection would infer that the series is non-stationary. However, a major problem with this conclusion is the well-documented fact (see Perron (1989)) that the power of the test in the presence of a structural break is low and thus the test statistic may suggest that the series in question is non-stationary, when in fact it is not⁶. In a similar fashion, a structural break occurring early in the sample under the null hypothesis has been shown by Leybourne, Mills and Newbold (1998) to generate a ‘spurious rejection’ of the null hypothesis. Therefore, the rejection (non-rejection) of the null hypothesis should not in fact be regarded as irrefutable evidence supporting the alternative of trend stationarity (null hypothesis). Consequently, the importance of allowing for a structural break in the DGP can be seen whether we reject the null hypothesis or not. To this end Banerjee et al (1992) and Zivot and Andrews (1992) developed tests that allow for the possibility of a break under the alternative hypothesis. The advantage of this methodology is that it overcomes the power problem encountered by the traditional ADF test in that, by explicitly allowing for a break under the alternative hypothesis, the test should be able to differentiate a stationary process with a structural break from a unit root process.

Following Banerjee et al (1992), we can consider three econometric techniques that allow for us to assess the order of integration of a series allowing for a structural break under the alternative hypothesis – the recursive, rolling and sequential test statistics⁷. The recursive test statistics explicitly consider the traditional ADF equation as prescribed in (4). In particular, we estimate (4) using ordinary least squares (OLS) and compute the relevant test statistics over the sub-samples $t = 1, \dots, k$, for $k = k_0, \dots, T$, where k_0 is termed the start-up value, given by $k_0 = T\delta_0$, where δ_0 is a ‘trimming parameter’, and T is the full sample size. Banerjee et al (1992) propose four statistics given as: (i) the full sample Dickey-Fuller test t-test statistic, testing $\alpha = 1$, denoted \hat{t}_{DF} ; (ii) the maximal Dickey-Fuller t-test statistic, denoted $\hat{t}_{DF}^{\max} \equiv \max_{k_0 \leq k \leq T} \hat{t}_{DF}(k/T)$; (iii) the minimal Dickey-Fuller t-test statistic, denoted $\hat{t}_{DF}^{\min} \equiv \min_{k_0 \leq k \leq T} \hat{t}_{DF}(k/T)$; (iv) the difference between \hat{t}_{DF}^{\max} and \hat{t}_{DF}^{\min} , denoted $\hat{t}_{DF}^{\text{diff}} = \hat{t}_{DF}^{\max} - \hat{t}_{DF}^{\min}$. The rolling test statistics are computed by estimating (4) over sub-samples that are a constant fraction (δ_0) of the total data set. Consequently, we consider the sub-samples given by $t = k - [T\delta_0] + 1, \dots, k$

⁵ The procedures that we consider explicitly assume that the date of the structural break, known as the break-date, is unknown *a priori* (see Christiano (1992)).

⁶ Perron (1989), p. 1370, reports that “one could hardly reject the unit root hypothesis even if the series is that of a trend (albeit with a break) with i.i.d disturbances”.

⁷ The notation employed is based on that employed by Banerjee et al (1992). For a more formal derivation and explanation, we direct the reader to the original article.

where k is the ‘trimmed’ start value of the sub-samples. The test statistics of interest are: (i) the maximal Dickey-Fuller t-test statistic, denoted $\bar{t}_{DF}^{\max} \equiv \max_{k_0 \leq k \leq T} \bar{t}_{DF}(k/T, \delta_0)$; (ii) the minimum Dickey-Fuller t-test statistic, denoted $\bar{t}_{DF}^{\min} \equiv \min_{k_0 \leq k \leq T} \bar{t}_{DF}(k/T, \delta_0)$; (iii) the difference between \bar{t}_{DF}^{\max} and \bar{t}_{DF}^{\min} , denoted $\bar{t}_{DF}^{\text{diff}} = \bar{t}_{DF}^{\max} - \bar{t}_{DF}^{\min}$.

The sequential test statistics are based on full sample regressions of an ADF-type equation, given in (5), that allows for a break under the alternative hypothesis:

$$y_t = \mu_0 + \mu_1 \tau_{1t}(k) + \mu_2 t + \alpha y_{t-1} + \beta(L) \Delta y_{t-1} + \varepsilon_t \quad (5)$$

for $t = 1, \dots, T$, where $\beta(L)$ is a p^{th} order lag polynomial, $k = T_B / T$. The possibility of a shift is explicitly taken into account by the $\tau_{1t}(k)$ regressor, for which we consider two possibilities: a shift in the trend, for which $\tau_{1t}(k) = (t - k)$ for $t > k$, zero otherwise; or a shift in the mean, for which $\tau_{1t}(k) = 1$ if $t > k$, zero otherwise. The test statistics of interest are computed by sequentially estimating (5) across the range $\Lambda = [.001, .999]$, allowing for the two different specifications of $\tau_{1t}(k)$ and where the order of $\beta(L)$ is allowed to change for each sample. In particular, the test statistics that are considered are given by: (i) the Quandt LR statistic, denoted $QLR(p)$ where p is the number of lags of Δy_t in (2)⁸; (ii) the maximum of the sequential F-statistics testing $\mu_1 = 0$ in (5), denoted $\tilde{F}_T^{\max} \equiv \max_{k_0 \leq k \leq T-k_0} \tilde{F}_T(k/T)$; (iii) the sequential Dickey-Fuller t-test statistic for the value of k that generates \tilde{F}_T^{\max} , denoted $\tilde{t}_{DF}(\tilde{\delta})$; (iv) the minimal Dickey-Fuller t-test statistic, denoted $\tilde{t}_{DF}^{\min*} = \min_{k_0 \leq k \leq T-k_0} \tilde{t}_{DF}(k/T)$; and finally (v) the absolute extreme t statistic computed for testing $\mu_1 = 0$, for the mean break model, under the restriction that $\mu_2 = 0$ and $\alpha = 1$, denoted $\left| \text{ext}_{\delta \in (\delta_0, 1-\delta_0)} \tilde{t}_{t1}(\delta) \right|^9$.

The above tests require the user to select trimming parameters, δ_0 , which restrict the actual tests to sub-sets of the full sample, where the trimming is applied to the beginning and end of the sample period. We employ the trimming parameters that are specified in Banerjee et al (1992), where for the recursive tests and the QLR statistics $\delta_0 = 0.25$; for the rolling statistics, $\delta_0 = 1/3$; and for the sequential statistics $\delta_0 = 0.15$. Further, the above tests require the selection of the lag truncation parameter, i.e. how many lags of Δy_t we allow for. In much empirical work, the truncation lag is simply chosen and fixed and is not considered as an integral part of the model specification (see Banerjee et al (1992)). However, Ng and Perron (1995), Perron and Vogelsang (1992) and Hall (1994) provide evidence that employing data-dependent methods to select the lag will result in test statistics based on the regressions having stable size and higher power¹⁰. To this end we employ two different methodologies to select the truncation lag – the t-test and F-test methods following Perron (1997), where the maximum number of lags is set at 15.

⁸ See Quandt (1960) and Engle and McFadden (1994), p. 2807, for an explanation of this testing procedure. This test statistic can be computed by considering the Wald version of the Chow (1960) test for each of the possible break dates and then selecting the maximum value from the array of subsequent test statistics, for which the resulting value is denoted the QLR test statistic. Stock (1994, p.2807) notes that the Wald, LM and LR are equivalent when implemented asymptotically.

⁹ Consequently, this final test statistic is only of interest if we cannot reject the hypothesis that the series in question has a unit root and that the series does not exhibit a time trend. If this were not the case, then we would not be able to accept the necessary restriction on (2) that allows us to compute the test statistic.

In light of the potential size issue associated with the DF/ADF procedures when a structural break occurs under the null hypothesis we also consider the procedure proposed by Perron (1989) and extended by Perron (1997) that allows for a break under both the null and alternative hypotheses. The version of the test that we consider and outline below allows only for a break in the mean of the series since the evidence from the analysis executed under the Banerjee et al (1992) procedure (see results in Section 4) suggests that there is no evidence of a break in the slope of the DGP. The test equation of interest is that given by Perron (1997), which can be expressed as:

$$y_t = \mu + \theta DU_t(\lambda) + \beta t + \delta D(T_b)_t + \alpha y_{t-1} + \sum_{j=1}^p c_j \Delta y_{t-j} + e_t \quad (6)$$

Where $DU_t(\lambda) = 1$ if $t > T\lambda$, and zero otherwise and $D(T_b)_t = 1$ for $t = T_b + 1$, and zero otherwise. The latter dummy variable is the one that accounts for a structural break at time T_b . The first test statistic of interest is derived from testing the hypothesis that $\alpha = 1$, where α is estimated sequentially. In particular, we consider the test statistic given by $t_\alpha^* = t_{\hat{\alpha}}(T_b^*, p)$, where T_b^* is such that $t_\alpha = \text{Min}_{T_b \in (k+1, T)} t_{\hat{\alpha}}(T_b, p)$ ¹¹. The second test statistic of Perron (1997) chooses the break date associated with the minimal t-statistic on the change-in-the-intercept variable, denoted $t_{\hat{\theta}}^{12}$. The t-statistic for this unit root hypothesis test is denoted $t_{\alpha, \theta}^*$ and is more precisely defined as $t_{\alpha, \theta}^* = t_{\hat{\alpha}}(T_b^*, p)$ where T_b^* is such that $t_{\hat{\theta}}(T_b^*) = \text{Min}_{T_b \in (k+1, T)} t_{\hat{\theta}}(T_b, p)$. Alternatively, the break date is selected by choosing the maximum of the absolute value of the t-statistic on DU_t for which the associated unit root test statistic is denoted $t_{\alpha, |\theta|}^*$ and is more precisely defined as $t_{\alpha, |\theta|}^* = t_{\hat{\alpha}}(T_b^*, p)$, where T_b^* is such that $t_{\hat{\theta}}(T_b^*) = \text{Max}_{T_b \in (k+1, T)} |t_{\hat{\theta}}(T_b, p)|$ ¹³. In each case, we sequentially estimate (6) across the sample range $\Lambda = [.001, .999]$, again allowing p to change for each tentative break-date, where p is selected using the procedures outlined previously. The critical values that we must consider are those given in Table 1, (a), (b) and (c), Perron (1997), p. 362.

We argue that this specification of the Perron (1997) test equation is the most relevant for several reasons. Firstly, the types of events that could have a major impact on financial prices, such as the EuroSterling interest rates examined here, are likely to have an immediate impact (or a very short reaction period of 1 to 2 days) on the path of the series. For example, a new piece of information regarding the ‘health’ of the economy is likely to be interpreted by the markets very rapidly and as such is likely to be reflected in the process almost instantaneously. The reaction of financial markets is usually dramatic, represented by a shift in the level of the series and to this end we feel that the test equation should allow for the possibility of an instantaneous change in

¹⁰ In addition, techniques based on information criteria, such as Akaike’s (1974) procedure, tend to result in very parsimonious models and this could lead to unstable size and low power in the resulting tests based on the regression (see Ng and Perron (1995))

¹¹ It can be seen that the above test is analogous to the sequential test statistic proposed by Banerjee et al (1992), $\tilde{t}_{DF}^{\min*} = \min_{k_0 \leq k \leq T-k_0} \tilde{t}_{DF}(k/T)$, and to the test of Zivot and Andrews (1992), but is executed on a different model that allows for a structural break under the null.

¹² Under this methodology, we are imposing the mild *a priori* restriction of a one sided change, i.e. allowing the date of the change to be unknown *a priori* but restricting the analysis to the cases of a ‘crash’.

¹³ Under this methodology, we do not impose the *a priori* that the change in the intercept is one-sided, and we are now allowing for either a decrease or increase in the DGP of the rates.

the level of the underlying series¹⁴. Further, we feel that in highly liquid markets dominated by major, well informed agents, such as those considered in this study, it is likely that the reaction will be sudden rather than gradual. If we were considering economic data which, due to the complex interactions between all the agents in the economy, could take a longer period of time to react to new information, then the argument above is less robust and possibly a rather extreme assumption. As a result, an alternative specification of the test equation that allows for a gradual adjustment rather than an instantaneous one is likely to be more appropriate. To this end, it has been suggested (by an anonymous referee on an earlier version of this paper) that a Kalman-filter model or an approach consistent with the smooth transition models as discussed in Teräsvirta and Anderson (1992), Granger and Teräsvirta (1993) and generalised in Teräsvirta (1994) could be employed. In the empirical work that we undertake we feel that the assumption of an instantaneous jump is more appropriate and further, the Kalman filter/smooth transition models have not been incorporated within the parametric extensions of the Dickey-Fuller testing framework, which is the focus of this paper.

(ii) *Testing for cointegration between the rates*

In our discussion of the EH, we noted that we need to be able to test for cointegration between the different rates. To this end, we employ the standard Engle and Granger (1987) procedure, estimating the cointegrating relationship given by:

$$y_{1t} = \mu + \alpha y_{2t} + e_t \quad (7)$$

We then construct the estimated residuals and estimate an ADF type regression for which the regression equation is given by:

$$\Delta \hat{e}_t = \zeta \hat{e}_{t-1} + \upsilon_1 \Delta \hat{e}_{t-1} + \upsilon_2 \Delta \hat{e}_{t-2} + \dots + \upsilon_p \Delta \hat{e}_{t-p} + \varepsilon_t \quad (8)$$

Where \hat{e}_t are the residuals generated by estimating (7) and k is selected by using the $p(t\text{-sig})$ procedure with a maximum value for p of 15. We now consider a t -test for $\zeta = 0$ which is consistent with testing $\alpha = 1$ in (8), for which the critical values are those reported in Engle and Granger (1987).

Similarly, we consider the Gregory and Hansen (1996) testing procedure, allowing for a break in the mean¹⁵, where the cointegrating relationship that is estimated, for each tentative break date is given by:

$$y_{1t} = \mu_1 + \mu_2 \phi_{1\tau} + \alpha y_{2t} + e_t \quad (9)$$

Where $t = 1, \dots, n$ and $\phi = \begin{cases} 0 & \text{if } t \leq [\lambda T] \\ 1 & \text{if } t > [\lambda T] \end{cases}$. We construct the estimated residuals \hat{e}_t and then estimate the

regression given by:

¹⁴ Additionally, as reported in our results for the other approaches that allow for the possibility of a break in the slope of the series, mainly Banerjee et al (1992) and Zivot and Andrews (1992), we do not find any evidence to support this definition and therefore we do not consider it in this approach.

¹⁵ The specification of the cointegration test equation allows for a break in the mean. Given our previous discussion concerning the specification of the Perron (1997) test equation, allowing for an instantaneous break in the mean of the series, we feel that this is the relevant specification for the cointegrating regression. Further, allowing this specification of the cointegrating relationship ensures a consistent approach for the unit root and cointegration tests.

$$\Delta \hat{e}_{t\lambda} = \zeta \hat{e}_{t-1\lambda} + \nu_1 \Delta \hat{e}_{t-1\lambda} + \nu_2 \Delta \hat{e}_{t-2\lambda} + \dots + \nu_l \Delta \hat{e}_{t-p\lambda} \quad (10)$$

We then execute an ADF test on the above regression, testing the hypothesis of $\zeta = 0$ using a standard t-statistic, defined as $ADF(\lambda) = tstat(\hat{e}_{t-1\lambda})$, where λ indicates the date of the tentative break date. We run the procedure across all the possible break dates, i.e. for values of $\lambda = [0, 1]$, and select the smallest, most negative value from the associated array of $ADF(\lambda)$ test statistics, denoted $ADF^* = \inf_{\lambda \in T} ADF(\lambda)$. The critical values are those reported in Gregory and Hansen (1996), pp. 109, Table 1.

4. Empirical Results

The data that we use in this study are EuroSterling rates for overnight, one-week and one-, three-, six- and twelve-month maturities. We employ daily observations for the period 1 January 1981 to 1 September 1997, a total of 4348 observations. All series were obtained from Datastream International. The period considered is of particular interest because it contains some significant monetary policy changes that may have generated a structural break in the DGP. Driffill, Zacharias and Sola (1997) suggest that the introduction of the Medium Term Financial Strategy, announced in the March 1980 Budget, and the removal of foreign exchange controls may well have had an impact on the interest rate series and, in particular, may have generated a structural break. The sample period investigated also covers a change in the UK's monetary policy during the 1980's from strictly targeting monetary growth to the consideration of "a broad range of monetary indicators". In addition, Tzavalis (1995) argues that it is possible that UK interest rates could also have been affected by changes in the monetary policy of other countries. Consequently, the beginning of the data set is likely to embody the effects of the change in the Federal Reserve's operating procedures and may therefore exhibit a structural break. In addition, German Re-unification had a significant impact on the German monetary policy set by the Bundesbank and this could also be expected to have fed through and affected the UK rates. In general, monetary policy changes in other countries, such as the South East Asian 'Tiger Economies' or those in Latin America, could also have impacted the characteristics of the UK data. Intuitively, this could be expected since global financial markets and economies are becoming increasingly integrated such that a change in policy that impacts on one market is rapidly transmitted around the globe. Finally, one of the major changes in UK monetary policy, certainly in the last 10 years and perhaps over the entire data set, occurred when the UK was forced to withdraw from the ERM on 16 September 1992, which has become known as 'Black Wednesday'. The resulting significant change in monetary policy may well have caused a structural break in the DGP to occur.

By considering series that are of daily rather than monthly or quarterly frequency, we should be able to more accurately pin-point possible causes for any structural break identified in the testing procedures since the methodology will pin point a particular day. If we can find an economic rationale for the existence of the break on a particular day then this will help explain and justify the statistical evidence identifying it. Previous studies that have looked at structural breaks have often used annual data and have justified the break on the basis of a major economic event in that year. Due to the frequency of the data employed, it usually has not been possible to identify a particular day, month or quarter. Our data frequency allows the possibility of pinpointing a

particular event and therefore removes the spuriousness and subjectiveness of tying in a date and an event on the basis of only the information that a break occurred in a particular year.

We examine the data sets for the individual interest rates using the methodologies outlined previously. The first set of tests that we considered were those suggested by Dickey and Fuller. In all cases, we were unable to reject the null hypothesis of a unit root¹⁶ and thus the specification for each of the interest rates considered in the Banerjee et al (1992) tests were those of a unit root and a non-zero drift parameter. Although the models proposed by Banerjee et al (1992) can include additional regressors, we selected not to allow for any¹⁷.

The test results for the recursive and rolling tests based on (5) did not reject the null hypothesis of a unit root for any of the series that were examined. Table 1 summarises the results where $\delta_0 = 0.25$ for the recursive and rolling tests. It can be seen that none of the recursive statistics reject the null hypothesis of a unit root at the 10% significance level. The same conclusion can be drawn from the rolling statistics apart from those for 1-month and 3-month series for which the \bar{t}_{DF}^{diff} rejects the null hypothesis of a unit root. However, Banerjee et al (1992) report that these test statistics are unreliable in Monte Carlo analysis stating that “*the relatively low power of these tests found in Section 3 suggests that these non-rejections might be uninformative*” [p. 279], and consequently it is necessary to consider in more detail the sequential test statistics.

The first sequential test that we consider is that of QLR, which tests whether the coefficients in the regression of y_t on Δy_{t-i} , t , 1 and y_{t-1} , where $i = 1$ to p , are constant¹⁸. For each sample considered, the lag length was reset to p^* and the optimal lag was found via the t and F -test approaches. Consequently, we did not impose the restriction that each sample must have the same lag structure but allowed it to be an integral part of the procedure and therefore data dependent. As a result, our test statistics should have stable size and higher power. Table 2 reports the test statistics, where p was selected using the t -test approach, although the results are also consistent for the F -test approach, and Table 3 reports the appropriate critical values¹⁹. The critical values are seen to be consistent with those reported by Banerjee et al (1992), Table 2, p. 278, in that as p increases, the critical values increase; as the significance level decreases, the critical values increase; and the critical values increase with the increase in the sample size. The results indicate that the null hypothesis of constant coefficients in the regression equations can be rejected at the 2.5% level for each of the rates. This may be caused by a stationary model with a break in the trend or by an $I(1)$ model with a shift in the mean and/or a break in the trend, and it is to this issue that we now turn²⁰.

¹⁶ These results are not shown due to space constraints but are available on request.

¹⁷ This is consistent with their empirical application of the tests.

¹⁸ This test involves estimating the regression over the full sample and then over the sub-samples 1 to i , where $i = T\delta_0, T\delta_0+1, \dots, (T-T\delta_0)$, and $i+1$ to T . The residual sum of squares for each of the three regression samples (RSS_T, RSS_1 and RSS_2 respectively) is then used to calculate the Chow (1960) statistic which is given by $F_T(r/T) = (RSS_T - (RSS_1 + RSS_2)) / ((RSS_1 + RSS_2) / (T - (2 \times nregs)))$, where $nregs$ is the number of regressors in the full sample regression. We can then compute the QLR statistic from the array of test statistics generated from the $F_T(r/T)$ test by selecting the highest value, i.e. $QLR = \max_{r=T\delta_0, \dots, T-T\delta_0} F_T(r/T)$.

¹⁹ It is necessary to generate the critical values for the QLR statistics as Banerjee et al (1992) note that the test statistics have a limiting distribution that is dependent on p . The remaining tests do not have a limiting distribution that is dependent on p and therefore we do not generate the relevant critical values.

²⁰ The analysis was repeated using the F -test approach to determined the lag length and the only result that changed was that for the 3 month rate where 15 lags rather than 12 were selected. The QLR test was reconsidered in light of this, however, the conclusion that the coefficients appeared to not be constant remained unchanged.

To consider whether or not a stationary model with a breaking trend causes the rejection of constant coefficients, we consider the $\tilde{t}_{DF}(\tilde{\delta})$ and \tilde{F}_T^{\max} test statistics, for which the critical values that we employ are those in Zivot and Andrews (1992) and Banerjee et al (1992). We considered both a break in the trend and a shift in the mean, for which the results for the former indicated that we could not reject the null hypothesis of a unit root for any of the series. The test results are reported in Table 4 and the corresponding critical values are Zivot and Andrews (1992), p. 256, Table 3A and Banerjee et al (1992), p. 278, Table 2. The two methods for selecting the break date identify the same day for the three longer maturities, however they result in different days for the three short rates. However, the conclusion that there is not a break in the trend is not affected, i.e. the unit-root/no-break null cannot be rejected in favour of the stationary/trend-shift alternative.

The above results support the conjecture that there is not a break in the trend of the DGP of any of the rates. We must now consider the possibility that there is a break in the mean of the DGP and to achieve this we will consider the mean-shift statistics. The results for the tests are summarised in Table 5 and the critical values are those reported in Zivot and Andrews (1992), p. 256, Table 2A and Banerjee et al (1992), p. 278, Table 2.

If we consider the two methods of estimating the break-date, the \tilde{F}_T^{\max} and $\tilde{t}_{DF}^{\min*}$ statistics, then it can be seen that the two procedures lead to very similar results. For the short rates, the same break-date is chosen by the two methods. The only inconsistent results arose for the 3 and 6-month rates, which will be shown not to exhibit evidence of a structural break, and therefore the inconsistency is uninformative. If we now consider the test statistics for a break in the mean, then it can be seen that in the case of the three short rates the null hypothesis can be rejected at the 10% significance level. In the case of the three longer rates, we are unable to reject the null hypothesis, indicating that the series appear to be I(1) processes.

We now need to address whether the rejection of the null hypothesis by the short rates has in fact been generated by the mis-specification of the null in the sense that they are actually characterised by a unit root process with a structural break. To this end we must employ the Perron (1997) approach. With respect to the long rates, we would also like to investigate whether or not there is a structural break in the series and this can easily be achieved by imposing the unit root restriction on the model and then testing for the significance of the shift-parameter, as per the Banerjee et al (1992) procedure. This approach is not applicable to the short rates because it relies on the fact that the null for the long-rates was not rejected and thus we can impose the unit root restriction on the DGP for the long-rates. For the short rates, since the null was rejected, it is inconsistent to impose the unit root hypothesis and then test the significance of the break parameter, and to this end the Perron (1997) approach is followed.

If we consider the long-rates then we need to impose the restriction that $\alpha = 1$ and $\mu = 0$ in (4) and then estimate the equation:

$$\Delta y_t = \mu_1 \tau_{it}(k) + \mu_2 t + \beta(L)\Delta y_{t-1} + \varepsilon_t \quad (11)$$

we then compute the $\tilde{t}_{\tau_1}(\cdot)$ statistic, which is simply the sequential t-test on the break parameter, i.e. a test of the null hypothesis that $\mu_1 = 0$ and therefore the break parameter is statistically insignificant. This is then compared to the critical values tabulated in Banerjee et al (1992), Table 2A, p. 278. Table 6 summarises the results. It can be seen that we cannot reject the null hypothesis and we can therefore conclude that the evidence seems to support the conjecture that the long rates are I(1) series that do not exhibit a structural break.

The results for the Perron (1997) approach applied to the three short-rates are reported in Table 7. Under the Banerjee et al (1992) approach, we did not incorporate a structural break under the null hypothesis and the outcome was that we rejected the null. When we incorporate a structural break under the null, as per the Perron approach, it can be seen that for none of the short rates can we reject the null hypothesis. In addition, if we consider the break parameter then we can see that in the short and 7-day rates the break term is significant. There is a slight discrepancy in the choice of break-dates for the 7-day rate, however the 14 September 1992 is very close to Black Wednesday and therefore can be justified on the same basis as if the 16 September 1992 had been chosen. In addition, Perron (1997, p. 371) states that “the tests (when T_b is chosen to minimise $t_{\hat{\theta}}$) have higher power” and therefore when faced with the contradictory evidence it would seem more appropriate to consider the results from the more powerful tests, which are the $t_{\hat{\theta}}$ results. The 1-month rate also fails to reject the null hypothesis of a unit root and a structural break; however, the break parameter does not appear to be significant. The results for the day prior to Black Wednesday indicate that the break parameter is significant and this seems to confirm that there was indeed a structural break that occurred on Black Wednesday²¹. Similarly, the results for the 7-day rates for the Tuesday indicate that the break parameter was most significant on Black Wednesday again supporting the conjecture that the rates did indeed exhibit a structural break on Black Wednesday.

Finally, the short rate indicates that the break date was in fact later than Black Wednesday; however, there was also a significant break parameter the day prior to Black Wednesday, which we feel is again evidence that there was a structural break on Black Wednesday. We also feel that the results indicating the 21 October 1992 as the break-date still fully support the conjecture that the UK being forced out of the ERM generated a structural break in the interest rate series. Furthermore on the 20 October 1992 John Major, the Prime Minister of the UK at the time, issued a statement indicating that he was ‘going for growth’, which can be interpreted as a confirmation of the expected change in the monetary policy generated by the ERM crisis, which in turn generated the break in the DGP of the rates, i.e. confirmation of a shift in government economic policy. Consequently, identifying a break at this date can be justified since there has obviously been a change in the economic policy and hence the

²¹ This is an analogy to the Perron (1997) finding that in some of the series that were examined the break-date appeared to be 1928, the year prior to the start of Great Depression. However on closer inspection it could be seen that the dummy variable that allows for the structural break and takes the value of unity in the next period was improving the fit of the model and therefore the economic interpretation of the results, i.e. there was a structural break caused by the Great Depression, could still be made. We feel that the same approach is applicable with our data, i.e. though the break parameter is not itself significant for Black Wednesday this is because the break has been picked up in the regression estimated for the day before.

monetary policy of the government, however it is very unlikely that such a re-direction of policy could have occurred if the UK had remained within the ERM²².

The finding that some of the rates support a unit root/break in mean characterisation whilst some support a unit root/no break characterisation could well generate problems when testing for cointegration among pairs of the rates. In particular, if the cointegrating relationship exhibits a structural break and this is not accounted for then the power and size characteristics of the Engle and Granger (1987) tests are likely to be affected. To this end we now consider the results for the cointegration analysis among pairs of the rates when we employ both the Engle and Granger (1987) and Gregory and Hansen (1996) procedures.

The cointegration results, reported in Table 8, show that in all cases, the Engle and Granger (1987) procedure rejects the null hypothesis of no cointegration, in favour of the alternative hypothesis of cointegration. In particular, it can be seen that the results for pairs of maturities involving short maturities, for example, those involving the short-rate, 1 month and 3 month rates, suggest that the evidence against the null hypothesis is very significant, as indicated by very large absolute test statistics. However, as the maturities increase the evidence against the null hypothesis, though still supporting the rejection of the null hypothesis, decreases. In particular, by the time that the spread involves the 6 month and 1 year rates, the test statistic is a mere -3.33 compared to a value of -15.18 when the spread contains the short rate and 7 day rate. However, the presence of a structural break in the cointegrating relationship, if ignored as in the Engle and Granger (1987) testing framework, may generate problems with the estimated test statistics. To this end, if we consider the results for the Gregory and Hansen (1996) procedure, the evidence is consistent with that based on the Engle and Granger (1987) procedure. In particular, the spreads involving the short maturities still provide the most significant evidence against the null hypothesis and as the maturities of the rates analysed in the spreads increases the evidence against the null hypothesis decreases.

There exists one very significant difference in the results obtained by the two procedures when the spreads being analysed contain the longer maturity rates. In particular, when the spreads contain the 6-month and 1-year rates, the evidence based on the Gregory and Hansen (1996) procedure appears to indicate that the null hypothesis of no cointegration cannot be rejected (except when combined with the short rate). Consequently, the evidence based on this procedure implies that cointegration, as predicted by the EH, between pairs of rates that involve either the 6-month and/or the 1-year rate is not supported once we allow for the possibility of a structural break in the cointegrating relationship. As a result, it can be seen that there is a significant conflict in the evidence provided by the two testing procedures and this difference can be seen to be due to the fact that one of the procedures explicitly allows for the possibility of a structural break, whilst the other does not.

If we consider Table 9, which reports the estimated Gregory and Hansen (1996) regression equations, it can be seen that the spreads for which the two approaches report conflicting results exhibit the most significant

²² The entire analysis was reconsidered with the initial maximum lag length being allowed to take the values 5, 10, 20, 25 and 50 lags. The results from the analysis confirmed that the results reported above are robust to the chosen maximum lag length. In almost all cases and for all of the rates, the conclusions that could be drawn from the results remained unchanged and the same break date was identified.

structural breaks. In particular, the breaks in these spreads typically have t-values in excess of 20, and in the case of the 6-month/1-year spread as high as approximately 31. The remaining spreads for which the two procedures support the EH typically exhibit t-values of 13 or less. Consequently, it appears that when a significant break exists (t-value in excess of approximately 20), the Engle and Granger (1987) procedure is in fact biased towards rejection of the null hypothesis whereas when the possibility of a break is explicitly incorporated in the testing framework, the evidence does not support rejection.

Finally, considering the selected break dates then it can be seen that there is no pattern or consistency in these dates for the different spreads. As a result, we can state that the breaks in the cointegrating relationships are not generated by a single regime change since if this was the case, we would expect to find a similar break date for each of the spreads. Consequently, the evidence suggests that the events generating the breaks appear to be idiosyncratic and unlike the breaks in the individual series, cannot be associated with the ERM crisis. We note, however, that if the events were not idiosyncratic, the breaks would have been consistent across the rates and it would therefore be unlikely that a structural break would have been generated in the cointegrating relationship, since the breaks in each of the rates would have effectively cancelled one another out and hence a 'large' break is unlikely to have been detected.

5. Discussion of Results

We have found that there appears to have been a structural break in the short rates on or close to the 16 September 1992 and we have stated that it can therefore be linked directly to the ERM crisis. However, we have seen that a structural break is commonly associated with a change in monetary policy and we therefore need to be able to justify our results in the light of this.

One may ask why the ERM crisis would have a direct bearing on the DGP of the rates. In answering this question, one needs to acknowledge the fact that the monetary policy that prevailed prior to the ERM crisis was aimed at keeping UK exchange rates within the bounds set by the agreement. In the period leading up to the crisis, Germany was increasing her interest rates, in an attempt to stem strong inflationary pressure caused by the Reunification of East and West Germany. The traditional and preferred economic environment that the Bundesbank had was one of low inflation. In an endeavour to achieve this, they were employing monetary policy in an attempt to reduce the inflationary pressure that was building in the domestic markets. The consequence of this was that the Deutsche mark was strengthening against most other currencies including the British Pound and it was anticipated that German rates would have to increase further to stave off additional inflationary pressure. At the same time the UK was entering into a severe recessionary period and her desired monetary policy was low interest rates in an attempt to stimulate demand. Consequently, there existed a conflict of objectives in that on one side, the UK needed to maintain its interest rate parity with Germany so that it could remain within the permitted currency bounds under the ERM agreement, but on the other hand, the UK's domestic economy required lower interest rates. The Bank of England was aware of this conflict and in order to minimise the increases that were necessary to ensure that the UK remained in the permitted currency bands, they were having to play an ever-increasing role in supporting the currency. In addition, the financial markets were aware of the

UK's conflict and that the Bank was having to increase its operations within the currency markets and speculators entered into the market, taking massive short Sterling positions, to profiteer from this situation anticipating that a devaluation would eventually occur. Internal pressures were too strong to allow interest rates to increase any further to prevent the large speculative flows and the Bank's reserves were being severely undermined. As a consequence of these factors, the UK was forced to resign from the ERM on 16 September 1992.

On doing so, the UK no longer had the obligation of maintaining their currency bounds against the Deutschmark and hence the main aim of monetary policy could be redirected towards domestic concerns, i.e. the deepening recession. The result was a fast and dramatic decline in interest rates as an attempt to kick-start the economy and a significant change in monetary policy. Consequently, a structural break in the short-rates, which are the main vehicle for monetary policy used by the Bank, can be justified, thus strengthening the econometric evidence that we have reported.

We conjectured that the presence of a structural break in the individual rates may in fact have generated a structural break in the cointegrating relationship. The results confirm that the cointegrating relationships involving the longer rates do in fact exhibit significant structural breaks and if these are not incorporated within the testing framework, as is the case in the Engle and Granger (1987) approach, then the inferences based on these results may in fact be spurious. To this end, we then considered the Gregory and Hansen (1996) procedure which explicitly allows for the possibility of a structural break and the results confirmed the conjecture that the Engle and Granger (1987) results maybe spurious in the presence of a significant structural break. In the cases where the t-value on the break coefficient is found to be large, the inferences with respect to whether the rates cointegrate or not is shown to be sensitive to whether the break is allowed for within the testing procedure. It appears that once the possibility of a break is allowed for, the evidence in support of the EH is not as significant , especially when the rates in the spread involve the long rates. The identified dates and therefore the associated events that are linked to the structural break in the tests of the spreads are found to be idiosyncratic and unrelated to the ERM crisis.

6. Conclusions

The analysis that we have executed was aimed at determining whether or not EuroSterling interest rates are correctly classified as unit root processes or whether in fact they are best classified as unit root processes that exhibit a structural break or a stationary process around a broken trend. Additionally, in light of the problems associated with the Engle and Granger (1987) procedure in the presence of structural breaks, we also considered the cointegrating evidence associated with testing the expectations hypothesis of the term structure of interest rates when we explicitly allow for the possibility of a break in the cointegrating relationship, as per the Gregory and Hansen (1996) procedure.

Taken as a whole, the results appear to support the conjecture that the short rates contain a unit root and a structural break and the long rates only a unit root, where the break-date has been identified as the 16 September

1992, 'Black Wednesday', in two of the three cases. This suggests that the structural break was generated by the ERM crisis. More specifically the break appears to have been generated by the change in monetary policy that occurred as a result of leaving the ERM. With respect to the cointegration evidence, we have showed that when the possibility of a break is allowed for then the results are in conflict to those assuming no break has occurred. Though the breaks are not associated with the period surrounding the ERM crisis the results indicate the importance of allowing for the possibility of a break within the testing framework. It is interesting to note that when a 'significant' break occurs in the cointegrating relationship, and it is not incorporated within the testing framework, the evidence for the rates and sample considered in our analysis supports the expectations hypothesis. However, if we incorporate the break, the results do not support cointegration and therefore do not support the expectations hypothesis. When the break is relatively small, we can conclude that the results are not sensitive to whether a break is allowed for within the testing framework. Our results therefore seem to be consistent with the theoretical impact that a structural break would have on the testing procedures. In particular, as the magnitude of the break increases, the importance of allowing for it within the testing framework increases and the impact on the results/inferences if it is not allowed for becomes problematic.

These results are interesting for a number of reasons, especially when considered within the context of the expectations hypothesis of the term structure of interest rates. First, the existence of a structural break in the short rates and not the long rates suggests that the linear relationship as proposed by the expectations hypothesis may not be valid. This can be extended to the general case that the existence of a structural break in the DGP governing the interest rate series over any sample is likely, if not taken into account, to result in a rejection of the expectations hypothesis. Second, the results reported in this article suggest that when a data sample of interest rates is selected that incorporates the ERM crisis, it will be necessary to allow for the structural break in the testing of the expectations hypothesis. Finally, we have to consider the effect of structural breaks on cointegration tests. The evidence that we report suggests that it is imperative that the possibility of a structural break in the cointegrating relationship is allowed for when testing for cointegration. Further, our results indicate that the breaks are likely to have an impact on the results/inferences of the cointegration tests associated with idiosyncratic events, as these are more likely to generate a large break relative to a regime shift that results in all the rates being affected. Consequently, it is important that an approach such as the Gregory and Hansen (1996) procedure that endogenously selects the break date is employed.

If the above results and discussion are considered within the light of the existing literature, then we should be wary of placing too much emphasis on the results that have been reported when the procedures employed have not explicitly allowed for the possibility of a structural break. We speculate that if the structural breaks are explicitly accounted for then the evidence, in some cases, maybe in conflict with that reported using the traditional techniques. Finally, the development of a set of tests that allow for a more flexible and general class of structural breaks and the extension to a multi-break testing framework, whether in the spirit of Perron (1997)/Gregory and Hansen (1996) or Hall and Sola (1993)/Hall, Psaradakis and Sola (1997) should prove to be an empirically useful and relevant extension to the existing literature. We leave this to future research.

Table 1. Unit Root Tests Based on Recursive and Rolling Statistics

Rate	Recursive Statistics				Rolling Statistics		
	t_{DF}	\hat{t}_{DF}^{\max}	\hat{t}_{DF}^{\min}	$\hat{t}_{DF}^{\text{diff}}$	\bar{t}_{DF}^{\max}	\bar{t}_{DF}^{\min}	$\bar{t}_{DF}^{\text{diff}}$
SR	-2.43941	-1.32873	-3.28799	1.95926	-7.9E-05	-3.87949	3.879411
7D	-1.95183	-1.32791	-3.18918	1.86127	-0.00028	-3.85816	3.85788
1M	-1.81841	-1.06947	-2.89679	1.82732	-0.00151	-4.54646	4.54495
3M	-1.80400	-1.02052	-2.75277	1.73225	-0.00017	-4.59078	4.59061
6M	-1.86405	-1.00456	-2.85105	1.84649	-0.00088	-3.63758	3.6367
1Y	-1.96502	-0.74098	-2.88007	2.13909	-0.00018	-3.54327	3.54309

Note: The critical values for t_{DF} , which is estimated over the full sample, are those reported in Fuller (1976, pp. 373). At the 10% level the critical value is -3.13. The critical values for the remaining tests are those from Table 1 in Banerjee et al (1992), pp. 277. The 10% critical values are: $\hat{t}_{DF}^{\max} = -1.66$; $\hat{t}_{DF}^{\min} = -3.88$; $\hat{t}_{DF}^{\text{diff}} = 3.21$; $\bar{t}_{DF}^{\max} = -1.25$; $\bar{t}_{DF}^{\min} = -4.55$; $\bar{t}_{DF}^{\text{diff}} = 4.31$.

Table 2. QLR Statistics Testing for Changing Coefficients in Equation 2

Series	p	QLR(p) Statistic
SR	15	204.2392
7D	15	363.7741
1M	15	271.4564
3M	12	87.88327
6M	10	61.0585
1Y	10	100.1522

Note: The above values have been estimated by calculating the QLR(p) statistic for each tentative break date in the trimmed sample. The maximum value is then selected and is tabulated above. p is the number of lags included in (2). The critical values are those tabulated in Table 3 and are based on Monte Carlo simulations.

Table 3. Critical Values for QLR Tests

p	Critical QLR(p) 2.5%	Critical QLR(p) 5%	Critical QLR(p) 10%
15	49.69472	46.46325	43.51393
12	46.75683	43.6674	40.59882
10	44.56217	41.6353	38.49935

Note: The critical values were estimated as directed by Banerjee et al (1992, Table 2). They are estimated using data generated by $\Delta y_t = \varepsilon_t$, where ε_t is iid $N(0,1)$. The values are based on 10,000 Monte Carlo replications for $T = 5000$, where the first 652 values were removed to avoid start-up value bias. Consequently these values are sample specific for the sample size that we are considering, i.e. 4348. The actual test values were computed by calculating the Wald form of the Chow (1960) test statistic for the regression of y_t on Δy_{t-i} , t , 1 and y_{t-1} , where $i = 1$ to p . This test involves estimating the regression over the full sample and then over the sub-samples 1 to i , where $i = T\delta_0, T\delta_0+1, \dots, (T-T\delta_0)$, and $i+1$ to T . The residual sum of squares for each of the three regression samples (RSS_T, RSS_1 and RSS_2 respectively) is then used to calculate the Chow (1960) statistic which is given by $F_T(t/T) = (RSS_T - (RSS_1 + RSS_2)) / ((RSS_1 + RSS_2) / (T - (2 \times \text{nregs})))$, where nregs is the number of regressors in the full sample regression. We can then compute the QLR statistic from the array of test statistics generated from the $F_T(t/T)$ test by selecting the highest value, i.e. $QLR = \max_{t=T\delta_0, \dots, T-T\delta_0} F_T(t/T)$.

Table 4. The Sequential $\tilde{t}_{DF}(\tilde{\delta})$ for a Break in the Trend

Rate	T_b	\tilde{F}_T^{\max}	$\tilde{t}_{DF}(\tilde{\delta})$	T_b	$\tilde{t}_{DF}^{\min*}$
SR	10 July 1989	2.96996	-2.98516	19 July 1989	-2.98528
7D	2 Jan 1989	2.2136	-2.43689	25 May 1989	-2.44387
1M	1 Mar 1995	3.40397	-1.70474	9 Jan 1989	-2.31916
3M	25 Nov 1988	2.03735	-2.27985	25 Nov 1988	-2.27985
6M	25 Nov 1988	1.7825	-2.2817	25 Nov 1988	-2.2817
1Y	25 Nov 1988	1.68844	-2.34669	25 Nov 1988	-2.34669

Note: The critical values employed are those reported in Zivot and Andrews (1992, Table 3A) and Banerjee et al (1992, Table 2). The critical values at the 10% level are: $\tilde{F}_T^{\max} = 13.20$, $\tilde{t}_{DF}(\tilde{\delta}) = -4.12$ and $\tilde{t}_{DF}^{\min*} = -4.11$. The test statistics reported are based on the regression of y_t on Δy_{t-i} , t , 1, y_{t-1} , and $\tau_{1i}(k)$ for which $\tau_{1i}(k) = (t-k)$ for $t > k$, otherwise zero DU where $i = 1$ to p . The results reported are based on a maximum lag length of 15 however the conclusions remained unchanged when we considered 5, 10, 20, 25 and 50 lags. These results are available on request.

Table 5. The Sequential $\tilde{t}_{DF}(\tilde{\delta})$ for a Break in the Mean

Rate	T_b	\tilde{F}_T^{\max}	$\tilde{t}_{DF}(\tilde{\delta})$	T_b	$\tilde{t}_{DF}^{\min*}$
SR	23 Oct 1992	17.6612	-4.7916	23 Oct 1992	-4.7916
7D	16 Sept 1992	34.39187	-5.65089	16 Sept 1992	-5.65089
1M	16 Sept 1992	22.74302	-4.78102	16 Sept 1992	-4.78102
3M	25 Aug 1992	14.24301	-4.01411	15 Sept 1992	-4.01758
6M	25 Aug 1992	14.59592	-4.08902	16 Sept 1992	-4.09953
1Y	16 Sept 1992	18.88103	-4.5527	16 Sept 1992	-4.5527

Note: The critical values employed are those reported in Zivot and Andrews (1992, Table 2A) and Banerjee et al (1992, Table 2). The critical values at the 10% level are : $\tilde{F}_T^{\max} = 16.78$, $\tilde{t}_{DF}(\tilde{\delta}) = -4.77$ and $\tilde{t}_{DF}^{\min*} = -4.58$. The test statistics reported are based on the regression of y_t on Δy_{t-i} , t , 1 , y_{t-1} , and $\tau_{1t}(k)$ for which $\tau_{1t}(k) = 1$ for $t > k$, otherwise zero where $i = 1$ to p . The results reported are based on a maximum lag length of 15 however the conclusions remained unchanged when we considered 5, 10, 20, 25 and 50 lags. These results are available on request.

Table 6. Testing for a Structural Break in the Long-Rates

Rate	T_b	$\text{ext}_{\delta} \tilde{t}_{\tau_1}(\delta)$
3M	17 Feb 1994	1.43131
6M	6 Dec 1993	1.47129
1Y	9 Dec 1993	1.54067

Note: The critical values employed are those reported in Banerjee et al (1992, Table 2), which takes a value of 2.66 at the 10% level . The test statistic is derived from the sequential t-statistic on the shift parameter in the regression of Δy_t on Δy_{t-i} , t , 1 , y_{t-1} , and $\tau_{1t}(k)$ for which $\tau_{1t}(k) = 1$ for $t > k$, otherwise zero where $i = 1$ to p . The results reported are based on a maximum lag length of 15 however the conclusions remained unchanged when we considered 5, 10, 20, 25 and 50 lags. These results are available on request.

Table 7. The Perron Testing Methodology

Series	T_b	μ	θ	δ	β	α
SR	21 Oct 1992	5.420e-002	-1.813e-002	0.220	1.062e-006	0.977
		(4.65527)	(-4.28705)	(4.17526)	(1.03052)	(-4.85933)
7D	14 Sept 1992	0.0347	-0.0123	1.0148	0.00000080	0.9851
		(5.60142)	(-5.53414)	(36.77211)	(1.45523)	(-5.9028)
	16 Sept 1992	3.546e-002	-1.408e-002	-0.176	1.365e-006	0.985
		(5.00389)	(-5.53570)	(-4.60127)	(2.17946)	(-5.34266)
1M	16 Sept 1992	0.0141	-5.2302e-003	-5.9811e-003	4.5101e-007	0.9939
		(4.47738)	(4.73669)	(-0.42419)	(1.68832)	(-4.7512)

Note: t-statistics are in parentheses. In the case of α the test statistic tests the null hypothesis that $\alpha = 1$, in all other cases the t-test if testing a null of zero. p is determined by the t-test approach, however the results are robust to the F-test approach. In each case the first row of coefficient estimates and associated t-statistics are those generated when we employ the first procedure to estimate the break date, i.e. $t_{\alpha} = \text{Min}_{T_b \in (k+1, T)} t_{\hat{\alpha}}(T_b, k)$, for which the critical values are those reported in Perron (1997). If there is an additional row then this corresponds to the estimated equation when we have employed the second and third breakpoint date selection criteria, i.e. $t_{\hat{\delta}}(T_b^*) = \text{Min}_{T_b \in (k+1, T)} t_{\hat{\delta}}(T_b, k)$ and $t_{\alpha|\theta}^* = t_{\hat{\alpha}}(T_b^*, k)$, which were found to generate a different breakpoint date.

Table 8. The Cointegration Results

	Engle and Granger (1987) ADF	Gregory and Hansen (1996) ADF*	Break Date
SR			
7D	-12.51	-12.49	6 July 1984
1M	-10.21	-10.23	1 June 1989
3M	-8.05	-8.20	30 July 1990
6M	-6.72	-6.78	30 Aug 1984
1Y	-5.45	-5.65	27 Apr 1994
7D			
1M	-8.54	-8.63	15 May 1990
3M	-6.81	-6.93	6 Sept 1990
6M	-5.60	-4.52	26 Apr 1994
1Y	-4.60	-4.25	21 Mar 1991
1M			
3M	-6.06	-6.85	21 Jan 1994
6M	-4.98	-4.40	19 Jan 1994
1Y	-4.79	-4.34	30 Mar 1994
3M			
6M	-4.52	-4.01	1 Mar 1994
1Y	-3.53	-4.15	11 July 1994
6M			
1Y	-3.38	-3.93	15 Dec 1993

The titles in bold (first column – i.e., SR, 7D, 1M, 3M and 6M) are each used in separate pair-wise comparisons with the maturities listed beneath them. The values in the column denoted Engle and Granger (1987) are the estimated test statistics for the spreads when the cointegrating relationship is defined as in (7). The values in the column Gregory and Hansen (1996), ADF* are the estimated test statistics for $\zeta = 0$ in (10) such that $ADF(\lambda) = tstat(\hat{c}_{t-1\lambda})$ when the cointegrating relationship is given by (9). The break date is the date associated with ADF*. The critical values, at the 5% significance level, are -3.34 for the Engle and Granger (1987) procedure and -4.61 for the Gregory and Hansen (1996) procedure. The results that are in bold show the spreads for which the results for the two procedures are different and are therefore sensitive to the inclusion of the possibility of a structural break in the testing procedure.

Table 9. The Estimated Gregory and Hansen (1996) Cointegrating Regression Equations

	Breakdate	μ_1	μ_2	α	t-value(μ_2)
SR					
7D	6 July 1984	0.0704 (0.054)	-0.0108 (0.0018)	0.9745 (0.0021)	-6.0297
1M	1 June 1989	0.1824 (0.0075)	-0.0276 (0.0021)	0.9318 (0.0031)	-13.0029
3M	30 July 1990	0.3639 (0.0104)	-0.0644 (0.0049)	0.8595 (0.0042)	-12.8941
6M	30 Aug 1984	0.3509 (0.0105)	-0.0378 (0.0034)	0.8673 (0.0042)	-10.9770
1Y	27 Apr 1994	0.3885 (0.0146)	0.0323 (0.0053)	0.8388 (0.0062)	6.1045
7D					
1M	15 May 1990	0.1097 (0.0046)	-0.0174 (0.0013)	0.9595 (0.0019)	-13.5566
3M	6 Sept 1990	0.2453 (0.0076)	-0.040 (0.0041)	0.9058 (0.0031)	-8.7570
6M	26 Apr 1994	0.1598 (0.0103)	0.0339 (0.0017)	0.9345 (0.0043)	19.2354
1Y	21 Mar 1991	0.6279 (0.0140)	-0.0855 (0.0040)	0.7478 (0.0057)	-21.5779
1M					
3M	21 Jan 1994	0.0217 (0.0042)	0.0190 (0.0014)	0.9913 (0.0018)	13.4276
6M	19 Jan 1994	0.0460 (0.0075)	0.0440 (0.0025)	0.9791 (0.0031)	17.3294
1Y	30 Mar 1994	0.1312 (0.0101)	0.0833 (0.0035)	0.9423 (0.0042)	23.8908
3M					
6M	1 Mar 1994	0.0026 (0.0037)	0.0316 (0.0012)	0.9969 (0.0015)	25.7235
1Y	11 July 1994	0.1214 (0.0068)	0.0562 (0.0024)	0.9469 (0.0029)	23.6781
6M					
1Y	15 Dec 1993	0.0458 (0.0044)	0.0424 (0.0014)	0.9799 (0.0018)	30.6961

The table reports the estimated cointegrating regressions for the spreads when the Gregory and Hansen (1996) procedure is employed. We report the coefficient estimates and associated standard errors (in parenthesis) for each of the spreads. The final column denoted t-value(μ_2) is the t-test statistic on the structural break coefficient, for which the bold values indicate the values associated with the spreads that do not reject the null hypothesis of no cointegration.

References

Akaike, H. (1974). "A New Look at the Statistical Model Identification", *IEEE Transactions on Automotive Control*, Vol. AC-19, pp. 716-723.

Banerjee, A., Lumsdaine, R. L., and Stock, J. H. (1992). "Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypotheses: Theory and International Evidence", *Journal of Business and Economic Statistics*, Vol. 10, pp. 271-287

Brown, R. L., Durbin, J., and Evans, J. M. (1975). "Techniques for Testing the Constancy of Regression Relationships Over Time", with comments, *Journal of the Royal Statistical Society, Ser. B*, Vol. 37, pp. 149-192.

Campbell, J. Y., Shiller, J., and Schoenholtz, K. L. (1983). "Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates", *Brookings Papers on Economic Activity*, Vol. 1, pp. 173-217.

— — (1991). "Yield Spreads and Interest Rate Movements: A Bird's Eye View", *Review of Economic Studies*, Vol. 58, pp. 495-514.

— — (1987). "Cointegration and Tests of Present Value Models", *Journal of Political Economy*, Vol. 95, pp. 1062-1088.

— — (1984). "A Simple Account of the Behaviour of Long-Term Interest Rates", *American Economic Review*, Vol. 74, pp. 44-48.

Chow, G. C. (1960). "Tests of equality between sets of coefficients in two linear regressions", *Econometrica*, 28, 591-605.

Christiano, L. J. (1992). "Searching for a Break in GNP", *Journal of Business and Economic Statistics*, Vol. 10, pp. 237-250.

Cuthbertson, K. D. (1993). "The Expectations Hypothesis of the Term Structure: The UK Interbank Market", Discussion Paper No. 92-04, University of Newcastle Upon Tyne.

— and Nitzsche (1994). "Are German Money Market Rates Well Behaved?", Mimeo.

Dickey, D. A., (1976). *Introduction to Statistical Time Series*, Wiley, New York,

— and Fuller, W. A. (1979). "Distribution of the Estimators for Autoregressive Time Series with a unit Root", *Journal of the American Statistical Association*, Vol. 74, pp. 427-431.

— — (1981). "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, Vol. 49, pp. 1057-1072.

Driffil, J., Psaradakis, Z., and Sola, M. (1997) A Reconciliation of Some Paradoxical Empirical Results on the Expectations Model of the Term Structure *Oxford Bulletin of Economics and Statistics* Vol. 59, 29-42.

Engle, R. F. and Granger, C. W. J. (1987). "Cointegration and error correction: Representation, estimation and testing". *Econometrica*, 55, pp. 251-276.

— and McFadden, D. L. (1994). *Handbook of Econometrics*, Volume IV.

Fama, E. F. (1984). "The Information in the Term Structure", *Journal of Financial Economics*, Vol. 13, pp. 509-528.

— (1990). "Term Structure Forecasts of Interest Rates, Inflation and Real Returns", *Journal of Monetary Economics*, Vol. 25, pp. 59-76.

Fuller, W.A. (1976) *Introduction to Statistical Time Series* Wiley, N.Y.

Granger, C. W. J. and Teräsvirta, T. (1993). *Modeling nonlinear relationships*. Oxford University Press, Oxford.

Gregory, A. W. and Hansen, B. E. (1996). "Residual-based test for cointegration in models with regime shifts". *Journal of Econometrics*, 70, pp. 99-126.

Hall, A. D. (1994). "Testing for a Unit Root in Time Series with Pretest Data Based Model Selection", *Journal of Business and Economic Statistics*, Vol. 12, pp. 461-470.

— Anderson, H. M., and Granger, C. W. J. (1992). "A Cointegration Analysis of Treasury Bill Yields", *The Review of Economics and Statistics*, Vol. 74, pp. 117-126.

Hall, S. G. (1986). "An application of the Granger and Engle two-step estimation procedure to UK aggregate wage data". *Oxford Bulletin of Economics and Statistics*, Vol. 48, pp. 229-239.

— Psaradakis, Z. and Sola, M. (1997). "Cointegration and changes in regime: the Japanese consumption function". *Journal of applied Econometrics*, Vol. 12 (2), pp. 151-168.

— and Sola, M. (1993). "Testing for collapsing bubbles: An endogenous switching ADF test". Discussion Paper, no. 15-93 (London Business School, London).

Hamilton, J. D. (1988). "Rational Expectations Econometric Analysis of Changes in Regime", *Journal of Economic Dynamics and Control*, Vol. 12, pp. 385-423.

Hurn, A. S., Moody, T., and Muscatelli, V. A. (1995). "The Term Structure of Interest Rates in the London Interbank Market", *Oxford Economic Papers*, Vol. 47, pp. 418-436.

Lewis, K. (1991). "Was there a Peso Problem in the US Term Structure of Interest Rates", *International Economic Review*, Vol. 32, pp. 159-173

Leybourne, S. J., Mills, T. C. and Newbold, P. (1998). "Spurious rejections by Dickey-Fuller tests in the presence of a break under the null". *Journal of Econometrics*, Vol. 87, pp. 191-203.

MacDonald, R., and Speight, A. E. H. (1988). "The Term Structure of Interest Rates in the UK", *Bulletin of Economic Research*, Vol. 40, pp. 287-299.

Mankiw, N. G. (1986). "The Term Structure of Interest Rates Revisited", *Brookings Papers on Economic Activity*, Vol. 1, pp. 61-96.

— and Miron, J. A. (1986). "The Changing Behaviour of the Term Structure of Interest Rates", *Quarterly Journal of Economics*, Vol. 101, pp. 211-228.

Mills, T. C. (1991). "The Term Structure of UK Interest Rates: Test of the Expectations Hypothesis", *Applied Economics*, Vol. 23, pp. 599-606.

Ng, S., and Perron, P. (1995). "Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag", *Journal of the American Statistical Association*, Vol. 90, pp. 268-281.

Perron, P. (1989). "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis", *Econometrica*, Vol. 57, pp. 1361-1401.

— (1997). "Further Evidence on Breaking Trend Functions in Macroeconomic Variables", *Journal of Econometrics*, Vol. 80, pp. 355-385.

— and Vogelsang, T. J. (1992). "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis: Corrections and Extensions of Some Asymptotic Results", unpublished manuscript, Princeton University, Department of Economics.

- Phillips, P. C. B. and Hansen, L. E. (1990). "Statistical inference in instrumental variable regression with I(1) processes". *Review of Economic Studies*, 57, pp. 99-125.
- Quandt, R. E. (1960). "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes", *Journal of the American Statistical Association*, Vol. 55, pp. 324-330.
- Said, S. E., and Dickey, D. A. (1984). "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order", *Biometrika*, Vol. 71, pp. 599-608.
- Shiller, R. J. (1979). "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure", *Journal of Political Economy*, Vol. 87, pp. 1190-1219.
- (1990). "The Term Structure of Interest Rates", in Friedman, B. M., and Hahn, F. H. (eds). *Handbook of Monetary Economics*, North-Holland, Amsterdam.
- Sola, M., and Driffill, J. (1994). "Testing the Term Structure of Interest Rates Using a Stationary Vector Autoregression with Regime Switching", *Journal of Economic Dynamics and Control*, Vol. 18, pp. 601-628.
- Stock, J. H. (1994). "Unit Roots, Structural Breaks and Trends", in *Handbook of Econometrics*, Volume IV, 2739-2841, eds. Engle, R. F. and McFadden, D. L.
- and Watson, W. (1988). "Testing for common trends". *Journal of the American Statistical Association*, 83, pp. 1097-1107.
- Teräsvirta, T. (1994). "Specification, estimation and evaluation of smooth transition autoregressive models". *Journal of the American Statistical Association*, 89, pp. 208-218.
- and Anderson, H. M. (1992). "Characterizing non-linearities in business cycles using smooth transition autoregressive models". *Journal of Applied Econometrics*, 7, pp. 119-136.
- Tzavalis, E. (1995). "Common Shifts in Real Ex Ante Interest Rates Across Countries", mimeo, Department of Economics, University of Exeter.
- Zivot, E., and Andrews, D. W. K. (1992). "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis", *Journal of Business and Economic Statistics*, Vol. 10, pp. 251-270.