Nonlinear Error-Correction Models for Interest Rates in the Netherlands

Dick van Dijk*
Tinbergen Institute, Erasmus University Rotterdam
and
Philip Hans Franses
Rotterdam Institute for Business Economic Studies and
Econometric Institute, Erasmus University Rotterdam

January 13, 1997

Abstract

In this chapter we investigate empirical specification of smooth transition error correction models (STECMs). These models can be used to describe linear long-run relationships between nonstationary variables where adjustment towards equilibrium is nonlinear and can depend on exogenous variables. The various steps involved in specifying an appropriate model are discussed for a monthly bivariate interest rate series for The Netherlands. Using simulations we first establish that standard (linearity-based) cointegration tests can be used to examine joint long-run properties. Second, we apply various tests for nonlinearity to decide on an appropriate function for the adjustment of disequilibrium errors. When we estimate an STECM, we find indications that nonlinearity is due to only two observations. We investigate the relevance of these data points by applying robust tests for linearity and by considering less aggregated, i.e. weekly, data. We conclude with some suggestions for practitioners.

Keywords: Nonlinear Error-correction, Outliers, Aggregation, Interest Rates

JEL classification: C12, C22

*We would like to thank Heather Anderson, Herman Bierens and Robert Kunst for useful comments and suggestions. Correspondence to Dick van Dijk, Tinbergen Institute, Burg. Oudlaan 50, NL-3062 PA Rotterdam, The Netherlands, `e-mail: dvandijk@tir.few.eur.nl`

1 Introduction

Many economic variables, while nonstationary individually, are linked by long-run equilibrium relationships. The concept of cointegration, introduced by Granger (1981) and Engle and Granger (1987), together with the corresponding error-correction models, allows these two characteristics to be modelled simultaneously. In the 'standard' error-correction model adjustment towards the long-run equilibrium is linear, i.e., it is always present and of the same strength under all circumstances. There are however economic situations for which the validity of this assumption might be questioned. Recently, there have been several attempts to construct econometric models which allow for nonlinear adjustment, see Anderson (1995), Dwyer *et al.* (1996), Hansen and Kim (1996), Kunst (1992,1995), and Swanson (1996). It appears that relevant forms of nonlinear error-correction often concern some sort of asymmetry, i.e., distinction is made between adjustment of positive or negative and of large or small deviations from equilibrium. Both types of asymmetry arise in a rather natural way when applying cointegration techniques to modelling prices of so-called equivalent assets in financial markets, see Yadav *et al.* (1994) and Anderson (1995) for an elaborate discussion. Equivalent assets in a certain sense represent the same value, examples include stock and futures, and bonds of different maturity. Since these assets are traded in the same market, or in markets which are linked by arbitrage-related forces, their prices should be such that investors are indifferent between holding either one of the equivalent assets. If prices deviate from equilibrium, arbitrage opportunities are created which will result in the prices being driven back together again. However, market frictions can give rise to asymmetric adjustment of such deviations. For example, due to short-selling restrictions, the response to negative deviations will be different from the response to positive deviations from equilibrium. Alternatively, transaction costs prevent adjustment of equilibrium errors as long as the benefits from adjustment, which equal the price difference, are smaller than those costs. Additionally, short-selling restrictions and transaction costs are not the same for all market participants. Because of this heterogeneity among traders, it might be expected that the aggregate force on the prices to return to equilibrium might be gradually changing, see Anderson (1995).

The purpose of this chapter is twofold. First, we document that both types of asymmetric adjustment discussed above can be modelled by means of Smooth Transition Error-

Correction Models (STECMs). Second, and more important, we aim to review the practical issues involved in the empirical specification of STECMs and to provide useful guidelines for practitioners. These practical issues concern (i) cointegration, i.e., how can one establish whether there is a linear long-run relation while there is nonlinear adjustment? (ii) nonlinearity, i.e., which form of nonlinear adjustment is appropriate?, (iii) outliers, i.e., can we prevent that our results are due to only a few influential data points?, and (iv) aggregation, i.e., what is the effect of aggregation on finding nonlinear adjustment? We address these issues using an example of a monthly bivariate interest rate series in The Netherlands, of which we also have weekly observed data.

The outline of this chapter is as follows. Section 2 introduces the general idea of smooth transition error-correction by discussing a simple model, which subsequently will be used in simulation experiments. This section also contains an outline of an empirical specification procedure for STECMs. Section 3 focuses on the first step in this specification procedure by presenting some Monte Carlo evidence on the performance of standard linearity-based tests for cointegration, when applied in the presence of nonlinear error-correction. Section 4 reviews the results of Lagrange Multiplier (LM) tests for nonlinear error-correction when applied to the interest rate series, and it also presents estimates of an STECM for these series. Section 5 deals with the issues of outliers and sampling frequency. For our sample series we find that these issues have a substantial impact on finding an adequate empirical model. Finally, Section 6 contains some recommendations for practitioners.

2 Smooth transition error-correction

The concept of smooth transition error-correction can conveniently be introduced by considering the following system for a bivariate time series $\{(y_t, x_t)', t = 1, \ldots, T\}$,

$$y_t + \beta x_t = z_t, \quad z_t = (\rho_1 + \rho_2 F(z_{t-d}))z_{t-1} + \varepsilon_t, \tag{1}$$

$$y_t + \alpha x_t = w_t, \quad w_t = w_{t-1} + \eta_t, \tag{2}$$

where the so-called transition function $F(z_{t-d})$ is continuous and bounded between 0 and 1, $d \in \{1, 2, \ldots\}$, $\alpha \neq \beta$, and

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim \text{i.i.d}(0, \Sigma), \quad \Sigma = \begin{pmatrix} 1 & \theta\sigma \\ \theta\sigma & \sigma^2 \end{pmatrix}. \tag{3}$$

The standard linear set-up, which is used by, *inter alia*, Banerjee *et al.* (1986) and Engle and Granger (1987), is obtained by taking $F(z_{t-d})$ equal to zero and imposing the restriction $|\rho_1| < 1$. The series y_t and x_t then are cointegrated with cointegrating vector $(1, \beta)'$. Put differently, the series y_t and x_t are linked by the (long-run) relationship $y_t = -\beta x_t$, and z_t represents the deviation from this 'equilibrium'. In the general system (1)-(3), z_t is assumed to follow a smooth transition autoregressive (STAR) model, see Granger and Teräsvirta (1993) and Teräsvirta (1994) for elaborate discussions of this class of nonlinear time series models. For y_t and x_t still to be cointegrated in this case, z_t has to be stationary. This implies that, depending on the specific form of the function $F(z_{t-d})$, certain restrictions have to be put on ρ_1 and ρ_2. For example, $|\rho_1| < 1$ and $|\rho_1 + \rho_2| < 1$ are sufficient, but not necessary, conditions for z_t to be stationary for all possible choices of $F(z_{t-d})$.

It is useful to rewrite the system (1)-(2) in error-correction format as,

$$\Delta y_t = \frac{\alpha(\rho_1 + \rho_2 F(z_{t-d}) - 1)}{\alpha - \beta} z_{t-1} + \xi_{1t}, \qquad (4)$$

$$\Delta x_t = \frac{-(\rho_1 + \rho_2 F(z_{t-d}) - 1)}{\alpha - \beta} z_{t-1} + \xi_{2t}, \qquad (5)$$

where $\xi_{it}, i = 1, 2$, are linear combinations of ε_t and η_t. From (4)-(5), the term smooth transition error-correction is obvious. For example, in the equation for Δy_t, the strength of error-correction changes smoothly from $\alpha(\rho_1 - 1)/(\alpha - \beta)$ to $\alpha(\rho_1 + \rho_2 - 1)/(\alpha - \beta)$ as $F(z_{t-d})$ changes from 0 to 1.

The function $F(z_{t-d})$ can be used to obtain many different kinds of nonlinear error-correction behaviour. As argued in the introduction, in empirical applications involving financial variables, one might be especially interested in modelling asymmetric adjustment. Asymmetric effects of positive and negative deviations can be obtained by setting $F(z_{t-d})$ equal to the logistic function,

$$F(z_{t-d}) \equiv F(z_{t-d}; \gamma, c) = (1 + \exp\{-\gamma(z_{t-d} - c)\})^{-1}, \quad \gamma > 0 \, . \qquad (6)$$

In the resulting logistic STECM (LSTECM), the strength of attraction of z_t to zero changes monotonically from ρ_1 to $\rho_1 + \rho_2$ for increasing values of z_{t-d}. The logistic model therefore allows for different effects of positive and negative deviations (relative to the threshold c) from the equilibrium. The parameter γ determines the speed of the transition; the higher γ, the faster the change from ρ_1 to $\rho_1 + \rho_2$. If $\gamma \to 0$, the LSTECM becomes linear, while if $\gamma \to \infty$, the logistic function approaches an Heaviside function, taking the value 0 for $z_{t-d} \leq c$ and 1 for $z_{t-d} > c$.

The second type of asymmetry, which distinguishes between small and large equilibrium errors, is obtained when $F(z_{t-d})$ is taken to be the exponential function,

$$F(z_{t-d};\gamma,c) = 1 - \exp\{-\gamma(z_{t-d}-c)^2\}, \quad \gamma > 0\ , \tag{7}$$

which results in gradually changing strength of adjustment for larger (both positive and negative) deviations from equilibrium. In the resulting STECM, the strength of mean reversion changes from $\rho_1 + \rho_2$ to ρ_1 and back again with increasing z_{t-d}, and this change is symmetric around c. A possible drawback of this choice for the transition function is that both if $\gamma \to 0$ or $\gamma \to \infty$, the model becomes linear. This can be avoided by using the 'quadratic logistic' function

$$F(z_{t-d};\gamma,c_1,c_2) = (1+\exp\{-\gamma(z_{t-d}-c_1)(z_{t-d}-c_2)\})^{-1}, \quad \gamma > 0\ , \tag{8}$$

as proposed by Jansen and Teräsvirta (1995). In this case, if $\gamma \to 0$, the model becomes linear, while if $\gamma \to \infty$, the function $F(\cdot)$ is equal to 1 for $z_{t-d} < c_1$ and $z_{t-d} > c_2$, and equal to 0 inbetween. Hence, the model nests the three regime threshold error-correction model of Balke and Fomby (1997) as a limiting case. Note that for finite γ, the minimum value taken by the function (8), which is attained for $z_{t-d} = (c_1 + c_2)/2$, is not equal to zero. This has to be kept in mind when interpreting estimates from models with this particular transition function.

It is fairly straightforward to extend the specification strategy for STAR models of Teräsvirta (1994) to the error-correction case considered here. Empirical specification of an STECM then involves the following steps: (i) testing for cointegration and estimating the cointegrating relationship, (ii) testing for nonlinearity of the adjustment process and investigating the type of nonlinearity, and (iii) estimating and evaluating the STECM. Each of these steps is addressed in turn in the following sections.

3 Testing for cointegration

In this section we address the first step involved in specifying an STECM, i.e., testing for cointegration and estimating the cointegrating relationship. Escribano and Mira (1996) show that the cointegrating vector(s) can still be estimated superconsistently in the presence of neglected nonlinearity in the adjustment process, see also Corradi *et al.* (1995). In this section we evaluate these theoretical results using Monte Carlo experiments. Additionally, we also examine the finite sample properties of linearity-based tests for cointegration. These simulations complement and extend the Monte Carlo results in Pippenger and Goering (1993) and Balke and Fomby (1997). Both studies only consider the case of threshold error-correction and, furthermore, Pippenger and Goering (1993) only consider situations in which the cointegrating relationship can be assumed to be known, something which may not always be possible in practice.

We consider those cointegration tests that are most popular among practitioners: the residual-based test suggested by Engle and Granger (1987) and the likelihood ratio test introduced by Johansen (1988). For simplicity, we discuss the tests only for (bivariate) cases where no deterministic regressors are included in the model. The residual-based test for cointegration is performed via the two-step procedure of Engle and Granger (1987). That is, we first estimate the cointegrating regression

$$y_t = -\beta x_t + u_t \ , \tag{9}$$

by ordinary least squares (OLS) and, second, test for the presence of a unit root in the regression residuals $\hat{u}_t = y_t + \hat{\beta} x_t$. The latter is done by using the Augmented Dickey-Fuller (ADF) test of Dickey and Fuller (1979), which is the familiar t-ratio of ρ in the auxiliary AR(p) regression

$$\Delta \hat{u}_t = \rho \hat{u}_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta \hat{u}_{t-i} + \eta_t \ . \tag{10}$$

The ADF statistic requires the choice of an appropriate value for p in (10). The number of lagged differences included should be such that the residuals $\hat{\eta}_t$ obtained from (10) resemble white noise. In our Monte Carlo experiments, we follow Gregory (1994) by initially setting p fairly large (equal to 6) and then reducing this number until the last lag included is significant at the 5% level, using normal critical values.

The likelihood ratio tests developed by Johansen (1988) are derived from the vector ECM (VECM), for the 2×1 vector time series $\{X_t = (y_t, x_t)', t = 1 - p, \ldots, T\}$,

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{p-1} ,\ {}_i \Delta X_{t-i} + \varepsilon_t \ , t = 1, \ldots T \ , \tag{11}$$

where $\varepsilon_t \sim NID(0, \Sigma)$. If y_t and x_t are cointegrated, the matrix Π has rank 1 and can be decomposed as $\Pi = \alpha\beta'$ for 2×1 vectors α and β. Johansen (1988) advocates to test for cointegration by testing the rank r of Π. This can be done by applying likelihood ratio (LR) tests to test the significance of the squared partial canonical correlations between ΔX_t and X_{t-1}, denoted $\hat{\lambda}_1 and \hat{\lambda}_2$, which can be obtained by solving a generalized eigenvalue problem. Ordering them such that $\hat{\lambda}_1 > \hat{\lambda}_2$, the trace statistics can be used to test $H_0 : r = r_0$ against the alternative hypothesis $H_1 : r \geq r_0 + 1$ for $r_0 = 0, 1$, and is given by

$$LR_{trace} = -T \sum_{i=r_0+1}^{2} \ln(1 - \hat{\lambda}_i) . \tag{12}$$

The asymptotic distribution of the trace statistic is non-standard and depends on the number of zero canonical correlations, see Johansen (1988,1991). If the trace test points towards cointegration between y_t and x_t, an estimate of the cointegrating vector β is given by the eigenvector corresponding to the largest canonical correlation $\hat{\lambda}_1$. In our Monte Carlo experiments, the VAR-order p in (11) is determined by minimizing the Schwarz-criterion BIC.

3.1 Monte Carlo experiments

The tests for cointegration discussed above assume that the adjustment process driving the variables towards the equilibrium is linear. In this section we investigate the empirical rejection frequencies of the tests, as well as the corresponding estimates of the cointegrating vector, when the series of interest are characterized by smooth transition error-correction. For convenience, we denote the rejection frequency of the cointegration tests as 'power'.

Monte Carlo design

The generalized bivariate system (1)-(3) is used as DGP for the artificial time series y_t and x_t. Both types of asymmetric error-correction which have been dicussed before are investigated, by using (6) and (8) as transition functions with z_{t-1} as transition variable, i.e., $d = 1$. In the Monte Carlo experiments, we investigate the effects of the autoregressive parameters in the STAR model for z_t, ρ_1 and ρ_2, on the power of the tests for cointegration and the estimates of the cointegrating parameter β.

In case (6) is used as transition function (which will be referred to as case I), the threshold c is fixed at zero, such that adjustment is different for positive and negative equilibrium errors, while the parameter γ, which determines the speed of the transition, is set equal to .5 and 5. Finally, ρ_1 and ρ_2 are chosen such that ρ_1 and $\rho_1+\rho_2$, which are the effective autoregressive parameters for $F(\cdot)=0$ and $F(\cdot)=1$, respectively, vary between 0.2 and 1.

In case (8) is taken to be transition function (case II), the thresholds c_1 and c_2 are varied between 0 and 8, with $c_1=-c_2\equiv c$, while γ is set equal to .1 and 1. For $\gamma=1$, the transition is already almost instantaneous at the thresholds, so it is not very useful to consider larger values for this parameter. In all experiments ρ_1 is fixed at 1, while ρ_2 is varied between 0 and -.8. Hence, adjustment is stronger for larger deviations from equilibrium. Finally, it should be remarked that the function $F(\cdot)$ is scaled by applying the transformation $F^*(\cdot)=(F(\cdot)-F(0))/(1-F(0))$. The function $F^*(\cdot)$ attains a minimum value of 0 at $z_{t-1}=0$ and approaches 1 for large negative and positive values of z_{t-1}. Hence, no error-correction is present at $z_{t-1}=0$ (and for $\gamma=1$, almost no error-correction occurs for $-c<z_{t-1}<c$).

The remaining parameters in the model are fixed at the following values for both case I and II: $\beta=-1$, $\alpha=-2$, $\sigma^2=1$, $\theta=0$. For each experiment we generate 2500 series of $T=100$ or 250 observations. The starting values for both y_t and x_t are set equal to zero and the first 100 observations are discarded. All calculations are performed using GAUSS.

Strictly speaking, power comparisons of the various tests are possible only when size-adjusted critical values are used. The power calculations presented in this section are made using asymptotic critical values, since this corresponds to empirical practice. The asymptotic critical values for the ADF test are taken from Phillips and Ouiliaris (1990), Tables IIa, while Table 0 in Osterwald-Lenum (1992) gives critical values for the Johansen trace test.

Results of Monte Carlo experiments

The results for case I are set out in Tables 1 and 2. Table 1 shows the rejection frequencies of the ADF and trace tests. Note that the cells corresponding to $\rho_1=1.0$ and $\rho_1+\rho_2=1.0$ denote the size of the tests, and the cells corresponding to $\rho_1=\rho_1+\rho_2$ denote the power of the tests in case of linear error-correction. The remaining cells contain estimates of the power in case of asymmetric error-correction.

- insert Table 1 -

From Table 1, it is seen that the power of both tests is almost not affected by the nonlinearity of the error-correction process. The only exception to this general observation is the case where either ρ_1 or $\rho_1 + \rho_2$ is equal to 1, i.e., in case there is no correction at all of either negative or positive errors, respectively.

Table 2 shows the means and standard deviations of the bias in the estimates of the cointegrating parameter β, obtained from the static regression (9) (OLS) and the Johansen procedure (VECM). Only the results for $\gamma = 0.5$ are shown, the results for $\gamma = 5.0$ are very similar. It should be noted that all entries are only based on those replications for which the respective test procedures detect cointegration at the 5% significance level.

- insert Table 2 -

It is seen that on average, the cointegrating parameter is over- and underestimated by the static regression and maximum likelihood procedure, respectively. This however can simply be a consequence of the choice of the DGP. Some conclusions which emerge from Table 2 are that the mean of the bias from the static regression is larger, but the variance is smaller. Both the mean and the variance of the bias decrease as the strength of attraction of the equilibrium error becomes stronger, i.e., for increasing values of ρ_1 and $\rho_1 + \rho_2$.

The results for case II are shown in Tables 3 and 4.

- insert Table 3 -

It appears that overall the simple ADF test is more powerful than the trace statistic, although the difference in power is not very large. Increasing values of c imply that the strength of error-correction increases more slowly as z_{t-1} gets larger (in absolute value). It is seen that the power of the tests decreases accordingly. More negative values of ρ_2 imply that the strength of attraction of z_t to zero becomes larger for given values of z_{t-1}. It might be expected that in this case the power of the tests increases, which is confirmed by Table 3. Finally, increasing γ, while keeping c and ρ_2 fixed, has two opposing effects. On the one hand, for large values of γ the strength of error-correction is virtually zero as long as $z_{t-1} \in (-c, c)$, while for small γ, error-correction becomes active as soon as there is a deviation from equilibrium. This effect might be expected to decrease the power of the cointegration tests as γ increases. On the other hand, for larger values of γ, the transition to the maximum strength of attraction is much quicker, which might be expected to increase the power of the tests. The simulation results seem to suggest that the second effect dominates, since for a very large majority of combinations of c and ρ_2 the power of the tests is higher for $\gamma = 1.0$.

Table 4 displays the means and standard deviations of the bias in the estimates of the cointegrating parameter β for case II. From this table, roughly the same conclusions can be drawn as for case I. The main difference is that for $T = 100$ and $\gamma = 0.1$, the mean of the bias from the static regression now is smaller (in absolute value) than the mean bias from the VECM.

- insert Table 4 -

In general, we observe that the bias in estimating the cointegrating rank and the cointegrating vector is not larger for asymmetric and nonlinear adjustment when compared to linear adjustment. These findings serve to substantiate some of the theoretical results in Escribano and Mira (1996) and Corradi *et al.* (1995).

3.2 Data analysis

In this subsection we examine the cointegration properties of our bivariate sample series. It is generally accepted that interest rates can be characterized as nonstationary processes or, to be more precise, processes which are integrated of order 1 ($I(1)$). Hall *et al.* (1992) argue that many theories of the term structure of interest rates imply that n interest rates of different maturity are cointegrated with cointegrating rank $n-1$, with the differences between the interest rates, or spreads, being the stationary linear combinations. If interest rates are such that the spread deviates from its equilibrium value, arbitrage opportunities are created, and these will drive the interest rates back towards equilibrium. Anderson (1995) argues that, due to market imperfections such as transaction costs, asymmetric error-correction of large and small deviations may play an important role.

To investigate the empirical usefulness of STECMs, we consider a monthly bivariate interest rate series for the Netherlands, consisting of one- and twelve-month interbank rates. We denote these as $R_{1,t}$ and $R_{12,t}$, respectively. The sample runs from January 1981 until December 1995, giving 180 monthly observations in total. The series are graphed in Figure 1.

- insert Figure 1 -

Table 5 shows the results from applying univariate ADF tests to the interest rates and the spread S_t, which is defined as $R_{12,t} - R_{1,t}$. The tests clearly indicate that the series are individually $I(1)$, while the spread seems to be stationary at the 5% significance level.

- insert Table 5 -

In order to check whether it is appropriate to use the spread as cointegrating relationship, i.e., to impose the cointegrating vector to be equal to (1,-1), a regression from $R_{1,t}$ on $R_{12,t}$ (including a constant as well) renders an estimate of 0.999, which seems close enough to unity.

The Johansen trace test is also computed, including one lagged difference of the series in the VECM (a VAR order of 2 for the levels is indicated by the Schwarz criterion), and no constants. The trace test for testing $r = 0$ and $r = 1$ are equal to 15.30 and 1.88. When compared with the appropriate 5% critical values, we conclude that the tests point towards cointegration. The estimate of the normalized cointegrating vector is (1, -1.029). Since the estimated standard error of the second element equals 0.018, we cannot reject the hypothesis that it is equal to unity. Hence, in the remainder of the analysis we assume that the spread amounts to a stationary linear combination of our two interest rates. Furthermore, the estimates of the parameters in the linear VECM (not shown here) reveal that the error-correction variable S_{t-1} is not significant in the equation for the twelve-month interest rate. For this reason, we focus on the equation for the short-term rate $R_{1,t}$ by conditioning on $R_{12,t}$.

The fitted linear conditional error-correction model (CECM) for $R_{1,t}$ is

$$\Delta R_{1,t} = \underset{(0.02)}{-0.02} + \underset{(0.04)}{0.13}\, S_{t-1} + \underset{(0.04)}{0.93}\, \Delta R_{12,t} - \underset{(0.07)}{0.16}\, \Delta R_{1,t-1} + \underset{(0.08)}{0.09}\, \Delta R_{12,t-1} \tag{13}$$

$\hat{\sigma}_\varepsilon = 0.236$, DW = 2.00, SK = 0.00, EK = 4.23, JB = 132.92(0.00), ARCH(1) = 16.53(0.00), ARCH(4) = 19.03(0.00), LB(8) = 7.84(0.45), LB(12) = 19.42(0.08), BIC = −2.767, where standard errors are given in parentheses below the parameter estimates, $\hat{\sigma}_\varepsilon$ is the residual standard deviation, DW is the Durbin-Watson statistic, SK is skewness, EK excess kurtosis, JB the Jarque-Bera test of normality of the residuals, ARCH is the LM test of no autoregressive conditional heteroscedasticity, LB is the Ljung-Box test of no autocorrelation and BIC is the Schwarz criterion. The figures in parentheses following the test statistics are p-values.

This linear model seems quite satisfactory, with reasonable values for all coefficients. Due to the large kurtosis, normality of the residuals is strongly rejected. Closer inspection of the residuals reveals that this may be entirely caused by only three observations in the beginning of the sample for which the residuals are very large (in absolute value). These aberrant observations may also cause the ARCH tests to reject homoscedasticity. On the other hand, it may also be that these significant test values are cauased by neglected nonlinearity. In the next section, we focus on a nonlinear extension of (13).

4 Testing for smooth transition error-correction

Once the presence of an equilibrium relationship has been established, the next question is whether possible nonlinearity in the adjustment process can be detected. Alternatively, if, perhaps contrary to one's prior expectations, cointegration may not have been found, the application of linearity tests may provide some insight in the causes for this finding. The Lagrange Multiplier (LM) type tests developed by Luukkonen *et al.* (1988) for general smooth transition nonlinearity can easily be adapted to test for smooth transition error-correction, see also Swanson (1996). The objective of testing for nonlinearity is threefold. First, we want to obtain an impression of whether the error-correction process is indeed nonlinear. Second, we need to determine the appropriate transition variable, i.e., obtain an estimate of the lag d. Third, we want to obtain an idea of the most appropriate form of the nonlinearity in the error-correction, i.e., we want to select between the forms of nonlinearity implied by (6) on the one hand and (7) or (8) on the other. In our empirical example, we confine our analysis to selecting between the logistic function (6) and the quadratic logistic function (8).

Consider a general CECM for y_t,

$$\Delta y_t = \pi_1' w_t + F(z_{t-d}; \gamma, c)\pi_2' w_t + \eta_t \ , \qquad (14)$$

where $w_t = (1, \tilde{w}_t)', \tilde{w}_t = (z_{t-1}, \Delta y_{t-1}, \ldots, \Delta y_{t-p+1}, \Delta x_t, \ldots, \Delta x_{t-p+1})'$, $z_t = y_t + \beta x_t$, $\pi_i = (\pi_{i0}, \pi_{i1}, \ldots, \pi_{im})'$, for $i = 1, 2$, $m = 2p - 1$. The noise process $\{\eta_t\}$ is assumed to be normally distributed with mean zero and variance σ_η^2. Compared to (4), constants and lagged first differences of y_t and x_t have been added to allow for more general dynamic structures.

The null hypothesis of linear error-correction in (14) with (6) or (8) can be formulated as $H_0 : \gamma = 0$. It is immediately seen that under the null hypothesis the model is not identified and, hence, the usual asymptotic theory cannot be applied to derive LM tests, see Davies (1977,1987) for a general discussion of such identification problems. Luukkonen *et al.* (1988) suggest to solve this by replacing the transition function $F(z_{t-d}; \gamma, c)$ in (14) by a suitable approximation around $\gamma = 0$. In the reparameterized model, the identification problem is no longer present and linearity can easily be tested.

A general test against smooth transition error-correction emerges when $F(z_{t-d})$ is replaced by a third-order Taylor approximation. Rearranging terms yields the reparameterized model,

$$\Delta y_t = \phi' w_t + \phi_1' \tilde{w}_t z_{t-d} + \phi_2' \tilde{w}_t z_{t-d}^2 + \phi_3' \tilde{w}_t z_{t-d}^3 + \eta_t \ , \tag{15}$$

It should be noted that when $d > p$, $\tilde{w}_t$ should be replaced by w_t because z_{t-d} is not present as an (implicit) regressor in w_t. The original null hypothesis of linearity, $H_0 : \gamma = 0$, is easily shown to be equivalent to the hypothesis that all coefficients of the auxiliary regressors $\tilde{w}_t z_{t-d}^j, j = 1, 2, 3$ are zero, i.e., $H_0' : \phi_1 = \phi_2 = \phi_3 = 0$. The LM-type test for this null hypothesis can be carried out in a few steps:

1. Estimate the parameters of the model under the null hypothesis by regressing Δy_t on w_t, with z_t replaced by $\hat{z}_t = y_t + \hat{\beta} x_t$, where $\hat{\beta}$ is obtained from preliminary cointegration analysis. The value of p, necessary for the construction of w_t can be taken from the linear model. Compute the sum of squared residuals $SSR_0 = \sum \hat{\eta}_t^2$, where $\hat{\eta}_t = y_t - \hat{\pi}_1' w_t$.

2. Estimate the parameters ϕ and $\phi_j, j = 1, 2, 3$ from the auxiliary regression

$$\hat{\eta}_t = \phi' w_t + \phi_1' \tilde{w}_t z_{t-d} + \phi_2' \tilde{w}_t z_{t-d}^2 + \phi_3' \tilde{w}_t z_{t-d}^3 + \nu_t \ , \tag{16}$$

   and compute the sum of squared residuals $SSR_1 = \sum \hat{\nu}_t^2$.

3. The LM-type test statistic can now be computed as

$$\mathrm{LM}_0 = T(SSR_0 - SSR_1)/SSR_0 \ . \tag{17}$$

The test statistic has an asymptotic χ^2 distribution with $3m$ degrees of freedom, where it is assumed that prior estimation of β does not affect the asymptotic distribution. In small samples it usually is recommended to use an F version of the test, i.e.,

$$\mathrm{LM}_0 = \frac{SSR_0 - SSR_1)/(3m)}{SSR_1/(T-4m)} \ , \tag{18}$$

which is approximately F distributed with $3m$ and $T-4m$ degrees of freedom under the null hypothesis of linearity.

To decide upon the most appropriate lag of z_t to use as transition variable, the test should be carried out for a number of different values of d, say $d = 1, \ldots, D$. If linearity is rejected for several values of d, the one with the smallest p-value is selected as the transition variable. This rule is motivated by the notion that the test might be expected to have maximum power if the true transition variable is used, see Granger and Teräsvirta (1993).

Deciding between the transition functions (6) and (8) can be done by a short sequence of tests nested within H_0. This testing sequence is motivated by the observation that if a logistic alternative is appropriate, the second order derivative in the Taylor expansion is zero. Hence, when $\phi_2 = 0$, the model can only be a logistic model. The null hypotheses to be tested are as follows

$$\begin{aligned} &H_{03} : \phi_3 = 0 \ , \\ &H_{02} : \phi_2 = 0 | \phi_3 = 0 \ , \\ &H_{01} : \phi_1 = 0 | \phi_3 = \phi_2 = 0 \ . \end{aligned} \tag{19}$$

Granger and Teräsvirta (1993) suggest to carry out all three tests, independent of rejection or acceptance of the first or second test, and use the outcomes to select the appropriate transition function. The decision rule is to select the quadratic logistic function (8) only if the p-value corresponding to H_{02} is the smallest, and select the logistic function (6) in all other cases. There is however no guarantee that this sequence will give the right answer. For practical purposes it therefore seems useful to estimate models with both transition functions and to base a decision between the two on other criteria.

- insert Table 6 -

We compute the LM-type test statistics for the various null hypotheses for the one-month Dutch interest rate in the estimated CECM (13). We set d equal to 1 through 6. The first panel of Table 6 shows the p-values of the standard LM-type tests. From the results for H_0, it is seen that linearity is rejected for both $d = 1$ and $d = 2$. Based upon the p-values, we select $d = 1$ as the appropriate transition variable. Unfortunately, the p-values of the test sequence for testing H_{03}, H_{02} and H_{01}, are not very conclusive with respect to the appropriate transition function. The p-values are equal to $0.075, 0.093$ and 0.004. Hence, if we would adopt the decision rule of Granger and Teräsvirta (1993), a logistic model seems most appropriate. When we estimate STECMS with (6) and (8), we find however that the logistic function (6) does not render sensible results. Therefore, in the sequel we only present models which assume (8) as the transition function.

We estimate the parameters of our STECM by non-linear least squares (NLS). We follow the suggestions of Teräsvirta (1994) and standardize the exponent of $F(S_{t-1})$ by dividing it by the variance of the transition variable, $\sigma^2_{S_{t-1}} = 0.229$, such that γ is a scale-free parameter. The estimation results are

$$
\begin{aligned}
\Delta R_{1,t} \;=\; & \underset{(0.03)}{-0.03} + \underset{(0.07)}{0.12}\, S_{t-1} + \underset{(0.04)}{0.90}\, \Delta R_{12,t} - \underset{(0.08)}{0.11}\, \Delta R_{1,t-1} + \underset{(0.09)}{0.11}\, \Delta R_{12,t-1} + \\
& (\underset{(0.12)}{0.32} + \underset{(0.24)}{0.42}\, S_{t-1} + \underset{(0.19)}{0.15}\, \Delta R_{12,t} + \underset{(0.19)}{0.25}\, \Delta R_{1,t-1} - \underset{(0.34)}{0.73}\, \Delta R_{12,t-1}) \qquad (20) \\
& \times (1 + \exp[-\underset{(5.67)}{3.74}\, (S_{t-1} + \underset{(0.08)}{0.40})(S_{t-1} - \underset{(0.12)}{1.25})/\sigma^2_{S_{t-1}}])^{-1} + \varepsilon_t
\end{aligned}
$$

$\hat{\sigma}_\varepsilon = 0.225$, DW $= 1.89$, SK $= -0.14$, EK $= 4.88$, JB $= 177.25(0.00)$, ARCH(1) $= 0.01(0.92)$, ARCH(4) $= 3.48(0.48)$, BIC $= -2.672$.

The large standard error of the estimate of γ is due to the fact that a wide range of values of this parameter renders about the same transition function. Accurate estimation of γ then requires a large number of observations close to c_1 and c_2, see Teräsvirta (1994) for a discussion. The estimate of γ is such that transition from $F(S_{t-1}) = 0$ to $F(S_{t-1}) = 1$ is almost instantaneous at the thresholds -0.40 and 1.25. The estimates of the coefficients of the error-correction term S_{t-1} are such that adjustment is stronger if the series is in the upper or lower regime, i.e., if the spread lagged one period is larger (in absolute value). Note that for the lower regime ($S_{t-1} < -0.40$) this is counteracted considerably by the change in the intercept. In fact, S_{t-1} needs to be smaller than -0.81, approximately, for the first effect to dominate.

Also notice that the ARCH test statistics have become insignificant, i.e., the previous evidence of ARCH in the linear model may have been due to neglected nonlinearity.

- insert Figure 2 -

Figure 2 shows some graphs which serve to illustrate the estimated smooth transition model. From the residual plot in the lower left panel it appears that the model still fails to capture some of the large interest rate movements in the beginning of the sample. The upper right panel shows how the transition function evolves over time. It is seen that the nonlinearity mainly serves to explain the behavior in 1993-1994, when the shape of the term structure was inverted, i.e., the short term rate exceeds the long term rate. Apart from this period, a few observations in the beginning of the 1980s are picked up by the nonlinear function, when the spread was more than 1.25%. From the graph in the lower right panel, it is seen that there are in fact only two observations in the regime $S_{t-1} > 1.25$. In the next section we examine whether these two observations might be regarded as outliers, or whether the monthly sampling frequency does not lead to sufficient observations in the different regimes, and hence that aggregation has resulted in 'less nonlinearity'.

5 Nonlinearity, outliers and sampling frequency

In this section we investigate whether our findings in the previous section based on monthly data may be caused by only a few observations by applying tests for nonlinearity which are robust to additive outliers. We also address the importance of sampling frequency or aggregation level of the series. For this purpose, we investigate model (20) for weekly data.

Testing for nonlinearity in the presence of outliers

Using theoretical derivations and extensive Monte Carlo simulations, Van Dijk *et al.* (1996) show that evidence for nonlinearity based on the above LM-type tests can be due to only a few additive outliers. For practical purposes it is important to investigate this possibility in order to prevent the empirical specification process to be governed by only a few data points. As an example, the c_2 parameter in (8) appears to be quantified on the basis of only two observations in our monthly data set.

We apply the robust LM-type tests for nonlinearity, as they are proposed in Van Dijk *et al.* (1996), to our monthly data set, and we report the p-values of the test statistics in the second panel of Table 6. The robust test involves the same steps as the standard test outlined in the previous section. The difference is that the linear model under the null hypothesis is estimated using a robust method, which downplays the effect of additive outliers. The auxiliary regression (16) is estimated using both weighted residuals and weighted regressors, where the weights indicate the relative importance of the observations in the robust estimation procedure. The asymptotic distributions of the various tests are still χ^2 and F. The results in Table 6 show that evidence for nonlinearity seems to vanish, i.e., the null hypothesis of overall linearity is now rejected at about the 12% level or more. Additionally, the test results for H_{03}, H_{02}, and H_{01} less clearly point towards a specific choice of a nonlinear adjustment function.

Sampling frequency

So far, we have considered monthly data to fit our STECMs for the bivariate interest rate series. Although nonlinear error-correction can be motivated by arbitrage arguments, it is unclear at what speed such arbitrage would take place. When arbitrage would take, say, three weeks to become effective, and we sample our data only monthly, one can expect nonlinear adjustment to be reflected only in a single observation. Would one, however, consider weekly data, one may obtain three data points which are informative for nonlinear modeling.

- insert Table 7 -

To evaluate our empirical STECM in (20) in the light of sampling frequency, we collect weekly observed data for the same bivariate interest rate series. Similar to the monthly data we calculate standard and robust LM-type tests for the various hypotheses on nonlinear error-correction, and we report the results in Table 7. From the first panel of this table, which contains the standard tests, we can conclude that there is substantial evidence for nonlinearity in these weekly data. For the robust tests, shown in the second panel, we observe that the p-values are generally smaller than the comparable ones in Table 6, although the overall evidence for nonlinearity is still weak. Only when d equals 6 we can reject H_0, H_{03} and H_{02} quite convincingly, and when $d = 1$ we can reject H_{02} at the 5% level.

In order to compare the effect of sampling frequency, we decide to estimate the same model as in (20) for the weekly data. The estimation results are

$$\begin{aligned}
\Delta R_{1,t} \;=\; & \underset{(0.01)}{-0.01} + \underset{(0.02)}{0.05}\, S_{t-1} + \underset{(0.03)}{0.81}\, \Delta R_{12,t} - \underset{(0.04)}{0.04}\, \Delta R_{1,t-1} + \underset{(0.04)}{0.07}\, \Delta R_{12,t-1} + \\
& (\underset{(0.02)}{0.12} + \underset{(0.03)}{0.10}\, S_{t-1} + \underset{(0.08)}{0.41}\, \Delta R_{12,t} + \underset{(0.10)}{0.21}\, \Delta R_{1,t-1} - \underset{(0.12)}{0.21}\, \Delta R_{12,t-1}) \qquad (21) \\
& \times (1 + \exp[-\underset{(9.02)}{7.38}\, (S_{t-1} + \underset{(0.04)}{0.42})(S_{t-1} - \underset{(0.04)}{1.03}) / \sigma^2_{S_{t-1}}])^{-1} + \varepsilon_t
\end{aligned}$$

$\hat{\sigma}_\varepsilon = 0.131$, DW $= 1.97$, SK $= 0.32$, EK $= 4.80$, JB $= 762.59(0.00)$, ARCH(1) $= 63.45(0.00)$, ARCH(4) $= 89.10(0.00)$, BIC $= -4.063$.

Compared with the estimated model for the monthly data, two things are most noteworthy. First, the coefficients for the error-correction (as well as the intercepts) are smaller, which intuitively makes sense, and, second, the estimate for the threshold c_2 has become smaller, as well as the corresponding standard error.

- insert Figure 3 -

In Figure 3 we present similar graphs as in Figure 2. The most relevant difference between these two Figures appears in the lower panel on the right, containing the function $F(\cdot)$ versus the transition variable S_{t-1}. As opposed to the model for the monthly data, there are now several observations in the upper regime, and, hence, we can have more confidence in the precision of the estimate of c_2 in the transition function (8). In other words, it pays off to consider less aggregate data for this bivariate interest rate series.

6 Concluding remarks

In this paper we have analyzed the empirical specification of a smooth transition error-correction model for a bivariate Dutch interest rate series, where we used monthly and weekly observed data. Using simulation experiments we substantiated the conjecture that standard linearity-based cointegration tests can be used to test for the presence of cointegration and to estimate the corresponding cointegrating vector. From our empirical results we must conclude that tests for nonlinearity should be used with caution when one aims to specify the nonlinear adjustment function in the STECM. First of all, our (unreported, tentative) estimation results show that key parameters like transition lag and type of transition function may not always be indicated by formal test results. We therefore recommend the practitioner to estimate various models and to base model selection also on the empirical sensibility of the estimated transition function. Secondly, additive outliers can spuriously suggest nonlinearity, and may lead to the specification of complicated nonlinear functions for only one or two data points. We recommend the use of a robust test for smooth transition nonlinearity in order to prevent one from putting too much effort in fitting a small number of observations. In fact, it may be that a robust test suggests linearity or another form of nonlinearity. When robust tests give such deviating results, one may consider other sampling frequencies, if such data are available. In fact, the third conclusion from our empirical results is that less aggregated data can lead to more precise estimates of nonlinear adjustment functions. In practice, the optimal level of sampling can be based on the available data at hand. Whether any theoretical arguments for some optimal level of aggregation for nonlinear modeling exist is left for further research.

References

Anderson, H.M. (1995) Transaction costs and nonlinear adjustment towards equilibrium in the US Treasury Bill market, working paper, University of Texas, Austin.

Balke, N.S. and T.B. Fomby (1997) Threshold cointegration, *International Economic Review*, forthcoming.

Banerjee, A., J.J. Dolado, D.F. Hendry and G.W. Smith (1986) Exploring equilibrium relationships in econometrics through static models: some Monte Carlo evidence, *Oxford Bulletin of Economics and Statistics*, **48**, 253–277.

Corradi, V., N.R. Swanson and H. White (1995) Testing for stationarity-ergodicity and for comovements between nonlinear discrete time Markov processes, working paper, University of Pennsylvania.

Davies, R.B. (1977) Hypothesis testing when a nuisance parameter is present only under the alternative, *Biometrika*, **64**, 247–254.

Davies, R.B. (1987) Hypothesis testing when a nuisance parameter is present only under the alternative, *Biometrika*, **74**, 33–43.

Dickey, D.A. and W.A. Fuller (1979) Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, **74**, 427–431.

Dwyer, G.P., P. Locke and W. Yu (1996) Index arbitrage and nonlinear dynamics between the S&P 500 futures and cash, *Review of Financial Studies*, **9**, 301–332.

Engle, R.F. and C.W.J. Granger (1987) Co-integration and error-correction: representation, estimation and testing, *Econometrica*, **55**, 251–276.

Escribano, A. and Mira, S. (1996) Nonlinear cointegration and nonlinear error-correction models, working paper, Universidad Carlos III de Madrid.

Granger, C.W.J. (1981) Some properties of time series data and their use in econometric model specification, *Journal of Econometrics*, **16**, 121–120.

Granger, C.W.J. and T. Teräsvirta (1993) *Modelling Nonlinear Economic Relationships*, Oxford University Press, Oxford.

Gregory, A.W. (1994) Testing for cointegration in linear quadratic models, *Journal of Business and Economic Statistics*, **12**, 347–360.

Hall, A.D., H.M. Anderson and C.W.J. Granger (1992) A cointegration analysis of Treasury Bill yields, *Review of Economics and Statistics*, **74**, 116–126.

Hansen, G. and J.-R. Kim (1996) Nonlinear cointegration analysis of German unemployment, mimeo, Institute of Statistics and Econometrics, Christian Albrechts University at Kiel.

Jansen, E.S. and T. Teräsvirta (1995) Testing parameter constancy and super exogeneity in econometric equations, working paper, Stockholm School of Economics.

Johansen, S. (1988) Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control*, **12**, 231–254.

Johansen, S. (1991) Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica*, **59**, 1551–1580.

Kunst, R.M. (1992) Dynamic patterns in interest rates: threshold cointegration with ARCH, working paper, Institute for Advanced Studies, Vienna.

Kunst, R.M. (1995) Determining long-run equilibrium structures in bivariate threshold autoregressions: a multiple decision approach, working paper, Institute for Advanced Studies, Vienna.

Luukkonen, R., P. Saikkonen and T. Teräsvirta (1988) Testing linearity against smooth transition autoregressive models, *Biometrika*, **75**, 491–499.

Phillips, P.C.B. and S. Ouliaris (1990) Asymptotic properties of residual based tests for cointegration, *Econometrica*, **58**, 165–193.

Pippenger, M.K. and G.E. Goering (1993) A note on the empirical power of unit root tests under threshold processes, *Oxford Bulletin of Economics and Statistics*, **55**, 473–481.

Swanson, N.R. (1996) Finite sample properties of a simple LM test for neglected nonlinearity in error-correcting regression equations, working paper, University of Pennsylvania.

Teräsvirta, T. (1994) Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association*, **89**, 208–218.

van Dijk, D., P.H. Franses and A. Lucas (1996) Testing for smooth transition nonlinearity in the presence of additive outliers, Econometric Institute Report 9622/A, Erasmus University Rotterdam.

Yadav, P.K., P.F. Pope and K. Paudyal (1994) Threshold autoregressive modeling in finance: the price difference of equivalent assets, *Mathematical Finance*, **4**, 205–221.

Table 1: Size and power of cointegration tests, case I[1]

T	γ	ρ_1	Test	$\rho_1+\rho_2$ 1.0	0.8	0.6	0.4
100	0.5	1.0	ADF	6.6	19.2	39.8	62.8
			LR_{trace}	4.4	11.1	27.6	52.6
		0.8	ADF	17.4	80.3	93.4	95.9
			LR_{trace}	10.8	60.5	88.4	94.0
		0.6	ADF	39.8	93.8	96.2	97.2
			LR_{trace}	28.5	89.4	94.8	94.7
		0.4	ADF	62.7	95.6	96.9	98.0
			LR_{trace}	51.2	94.1	94.6	94.6
	5.0	1.0	ADF	5.7	10.2	10.3	10.6
			LR_{trace}	3.2	6.3	7.0	7.2
		0.8	ADF	10.1	79.4	90.0	89.8
			LR_{trace}	5.4	60.0	81.8	87.7
		0.6	ADF	10.2	90.0	96.3	97.0
			LR_{trace}	5.8	83.5	93.4	93.6
		0.4	ADF	10.0	90.5	97.0	98.0
			LR_{trace}	6.0	88.2	93.7	93.8
250	0.5	1.0	ADF	4.6	27.7	55.2	78.2
			LR_{trace}	4.6	21.1	48.4	74.0
		0.8	ADF	29.7	99.8	100.0	100.0
			LR_{trace}	22.2	94.4	94.4	94.6
		0.6	ADF	56.5	100.0	100.0	100.0
			LR_{trace}	49.2	94.4	94.6	94.5
		0.4	ADF	77.8	100.0	100.0	100.0
			LR_{trace}	74.1	94.4	94.5	94.5
	5.0	1.0	ADF	4.6	10.3	10.0	9.9
			LR_{trace}	4.6	7.2	7.5	7.0
		0.8	ADF	11.4	99.8	99.9	99.8
			LR_{trace}	7.7	94.4	94.5	94.5
		0.6	ADF	11.2	100.0	100.0	100.0
			LR_{trace}	7.8	94.4	94.6	94.6
		0.4	ADF	10.8	99.8	100.0	100.0
			LR_{trace}	7.7	94.5	94.5	94.5

[1] Rejection frequencies at 5% significance level using asymptotic critical values for series generated by (1)-(3) and (6). The table is based on 2500 replications. Critical values are taken from Phillips and Ouiliaris (1990), table IIa, for the ADF test and from Osterwald-Lenum (1992), table 0, for the LR_{trace} tests.

Table 2: Mean and standard deviation of $\hat{\beta} - \beta$, case I[1]

				$\rho_1 + \rho_2$			
T	γ	ρ_1		1.0	0.8	0.6	0.4
100	0.5	1.0	OLS	0.505(0.607)	0.135(0.380)	0.073(0.163)	0.050(0.118)
			VECM	0.553(1.593)	−0.021(0.805)	−0.012(1.179)	0.393(14.49)
		0.8	OLS	0.126(0.302)	0.049(0.084)	0.043(0.071)	0.037(0.062)
			VECM	0.023(0.754)	−0.078(0.332)	−0.026(0.185)	−0.035(0.868)
		0.6	OLS	0.070(0.177)	0.043(0.069)	0.036(0.056)	0.032(0.050)
			VECM	−0.060(0.434)	−0.213(8.906)	−0.010(0.090)	−0.019(0.358)
		0.4	OLS	0.048(0.120)	0.037(0.062)	0.031(0.050)	0.029(0.044)
			VECM	−0.060(0.289)	0.006(0.506)	−0.012(0.046)	−0.014(0.052)
250	0.5	1.0	OLS	0.519(0.475)	0.094(0.240)	0.058(0.133)	0.043(0.104)
			VECM	−0.294(6.776)	−0.065(0.754)	−0.028(0.653)	−0.417(17.51)
		0.8	OLS	0.099(0.228)	0.035(0.049)	0.026(0.038)	0.022(0.034)
			VECM	−0.028(0.380)	−0.013(0.074)	−0.017(0.040)	−0.020(0.030)
		0.6	OLS	0.055(0.135)	0.027(0.039)	0.021(0.029)	0.018(0.025)
			VECM	−0.072(1.386)	−0.018(0.035)	−0.021(0.026)	−0.022(0.018)
		0.4	OLS	0.040(0.086)	0.022(0.034)	0.018(0.025)	0.016(0.022)
			VECM	−0.029(0.329)	−0.020(0.031)	−0.022(0.017)	−0.022(0.014)

[1] Mean(standard deviation) of $(\hat{\beta} - \beta)$ for series generated by (1)-(3) and (6). OLS and VECM refer to the estimates obtained from the cointegrating regression (9) and the vector error-correction model (11) respectively. The entries for the respective estimators are based on those replications for which the ADF and LR_{trace} statistic reject the null of no cointegration.

Table 3: Size and power of cointegration tests, case II[1]

T	γ	ρ_2	Test	c = 0	2	4	6	8
100	0.1	−0.2	ADF	27.8	25.2	17.5	11.6	9.4
			LR_{trace}	13.9	13.0	9.6	6.2	5.7
		−0.4	ADF	59.2	52.3	31.0	14.0	10.2
			LR_{trace}	32.7	28.2	16.2	8.3	6.1
		−0.6	ADF	79.7	73.9	46.2	16.6	10.3
			LR_{trace}	57.3	49.4	24.3	9.7	6.6
		−0.8	ADF	88.0	84.4	60.1	19.3	10.9
			LR_{trace}	78.8	70.2	34.8	11.1	6.8
	1.0	−0.2	ADF	74.2	59.2	17.1	11.1	9.1
			LR_{trace}	52.1	33.3	10.1	8.1	6.4
		−0.4	ADF	95.2	89.4	27.7	13.6	10.7
			LR_{trace}	94.3	83.8	16.5	9.5	7.4
		−0.6	ADF	96.8	92.9	43.1	17.6	12.3
			LR_{trace}	94.7	94.3	26.6	11.3	7.9
		−0.8	ADF	97.5	95.1	63.5	21.8	14.6
			LR_{trace}	94.6	94.8	44.1	14.3	9.0
250	0.1	−0.2	ADF	93.3	91.3	78.2	32.6	13.3
			LR_{trace}	81.8	76.4	51.4	18.2	8.6
		−0.4	ADF	98.6	98.1	93.6	58.5	16.6
			LR_{trace}	94.2	94.2	88.8	33.4	10.5
		−0.6	ADF	99.6	99.5	97.1	74.4	20.0
			LR_{trace}	94.3	94.4	94.1	49.4	12.4
		−0.8	ADF	99.9	99.7	98.8	83.5	23.5
			LR_{trace}	94.4	94.3	94.3	65.2	14.2
	1.0	−0.2	ADF	99.8	98.9	78.3	23.4	12.6
			LR_{trace}	94.4	94.3	56.0	15.0	8.4
		−0.4	ADF	100.0	99.9	88.8	40.0	16.4
			LR_{trace}	94.5	94.4	89.8	25.4	12.1
		−0.6	ADF	100.0	100.0	95.7	65.1	27.1
			LR_{trace}	94.6	94.4	94.4	48.8	17.4
		−0.8	ADF	100.0	100.0	98.4	82.7	42.8
			LR_{trace}	94.5	94.5	94.5	73.8	28.0

[1] Rejection frequencies at 5% significance level using asymptotic critical values for series generated by (1)-(3) and (8). The table is based on 2500 replications. Critical values are taken from Phillips and Ouiliaris (1990), table IIa, for the ADF test and from Osterwald-Lenum (1992), table 0, for the LR_{trace} tests.

Table 4: Mean and standard deviation of $\hat{\beta} - \beta$, case II[1]

				c				
T	γ	ρ_2		0	2	4	6	8
100	0.1	−0.2	OLS	0.067(0.147)	0.072(0.156)	0.066(0.187)	0.086(0.281)	0.110(0.382)
			VECM	−0.161(0.430)	−0.158(0.446)	−0.110(0.413)	−0.125(0.655)	−0.037(0.602)
		−0.4	OLS	0.056(0.106)	0.057(0.112)	0.067(0.143)	0.085(0.223)	0.087(0.351)
			VECM	−0.133(0.394)	−0.120(0.365)	−0.110(0.508)	−0.172(0.463)	−0.120(0.581)
		−0.6	OLS	0.052(0.090)	0.053(0.094)	0.059(0.120)	0.074(0.201)	0.085(0.337)
			VECM	−0.102(0.333)	−0.101(0.345)	−0.152(0.616)	−0.163(0.456)	−0.125(0.605)
		−0.8	OLS	0.050(0.082)	0.051(0.086)	0.058(0.108)	0.066(0.183)	0.087(0.317)
			VECM	−0.053(0.271)	−0.075(0.452)	−0.116(0.475)	−0.116(0.476)	−0.112(0.503)
	1.0	−0.2	OLS	0.051(0.089)	0.056(0.105)	0.066(0.319)	0.050(0.118)	0.122(0.396)
			VECM	−0.094(0.330)	−0.127(0.463)	−0.156(0.537)	−0.105(0.523)	−0.009(0.618)
		−0.4	OLS	0.040(0.063)	0.049(0.078)	0.073(0.272)	0.037(0.062)	0.098(0.354)
			VECM	−0.012(0.193)	−0.046(0.469)	−0.140(0.417)	−0.119(0.456)	−0.050(0.543)
		−0.6	OLS	0.033(0.051)	0.043(0.068)	0.075(0.243)	0.032(0.050)	0.104(0.320)
			VECM	−0.012(0.051)	−0.001(0.260)	−0.102(0.531)	−0.096(0.447)	0.292(3.903)
		−0.8	OLS	0.029(0.044)	0.040(0.061)	0.078(0.207)	0.029(0.044)	0.114(0.293)
			VECM	−0.014(0.036)	−0.006(0.065)	−0.111(0.616)	−0.050(0.673)	−0.069(0.504)
250	0.1	−0.2	OLS	0.057(0.079)	0.058(0.082)	0.063(0.096)	0.071(0.146)	0.092(0.224)
			VECM	−0.063(1.220)	−0.037(0.761)	−0.100(0.479)	−0.167(0.699)	−0.131(0.488)
		−0.4	OLS	0.043(0.061)	0.046(0.064)	0.054(0.076)	0.063(0.111)	0.086(0.199)
			VECM	−0.006(0.060)	−0.003(0.067)	−0.025(0.204)	−0.134(0.410)	−0.104(0.526)
		−0.6	OLS	0.036(0.051)	0.038(0.054)	0.047(0.067)	0.064(0.100)	0.084(0.184)
			VECM	−0.011(0.047)	−0.010(0.051)	0.003(0.154)	−0.107(0.401)	−0.097(0.438)
		−0.8	OLS	0.032(0.045)	0.034(0.047)	0.043(0.061)	0.063(0.093)	0.078(0.175)
			VECM	−0.012(0.158)	−0.014(0.040)	−0.006(0.062)	−0.086(0.386)	−0.094(0.495)
	1.0	−0.2	OLS	0.037(0.052)	0.044(0.061)	0.065(0.099)	0.077(0.177)	0.109(0.245)
			VECM	−0.010(0.057)	−0.007(0.088)	−0.090(0.780)	−0.071(0.458)	−0.102(0.488)
		−0.4	OLS	0.024(0.033)	0.031(0.043)	0.057(0.086)	0.074(0.142)	0.099(0.212)
			VECM	−0.026(0.292)	−0.016(0.042)	−0.020(0.219)	−0.091(0.814)	−0.101(0.449)
		−0.6	OLS	0.019(0.026)	0.026(0.035)	0.053(0.076)	0.075(0.118)	0.087(0.183)
			VECM	−0.021(0.018)	−0.022(0.160)	−0.010(0.523)	−0.102(0.376)	−0.061(1.076)
		−0.8	OLS	0.017(0.022)	0.023(0.031)	0.051(0.068)	0.078(0.109)	0.091(0.149)
			VECM	−0.022(0.014)	−0.020(0.022)	−0.001(0.060)	−0.064(0.710)	−0.097(0.401)

[1]Mean(standard deviation) of $(\hat{\beta} - \beta)$ for series generated by (1)-(3) and (8). OLS and VECM refer to the estimates obtained from the cointegrating regression (9) and the vector error-correction model (11) respectively. The entries for the respective estimators are based on those replications for which the ADF and LR_{trace} statistic reject the null of no cointegration.

Table 5: ADF statistics for interest rates[1]

	$R_{1,t}$	$R_{12,t}$	S_t	5%crit. value
Level	-2.15	-2.06	-3.00	-2.88
First Difference	-5.65	-8.35	–	-1.94

[1] ADF tests applied to monthly interest rates and spread. Test statistics for levels are $\hat{\tau}_\mu$ while those for the first differences are $\hat{\tau}$. Number of lagged differences in each regression were chosen such that the last lag included is significant at 5% level, using normal critical values.

Table 6: Standard and outlier robust LM-type tests for smooth transition error-correction in a CECM for monthly data on the one-month interest rate[1]

		d					
Test	Null	1	2	3	4	5	6
Standard	H_0	0.002	0.039	0.968	0.880	0.485	0.721
	H_{03}	0.075	0.988	0.828	0.327	0.313	0.669
	H_{02}	0.093	0.406	0.829	0.995	0.543	0.214
	H_{01}	0.004	0.001	0.804	0.757	0.478	0.956
Robust	H_0	0.151	0.124	0.936	0.523	0.700	0.889
	H_{03}	0.390	0.976	0.555	0.188	0.452	0.578
	H_{02}	0.724	0.545	0.903	0.599	0.452	0.576
	H_{01}	0.027	0.006	0.839	0.738	0.820	0.957

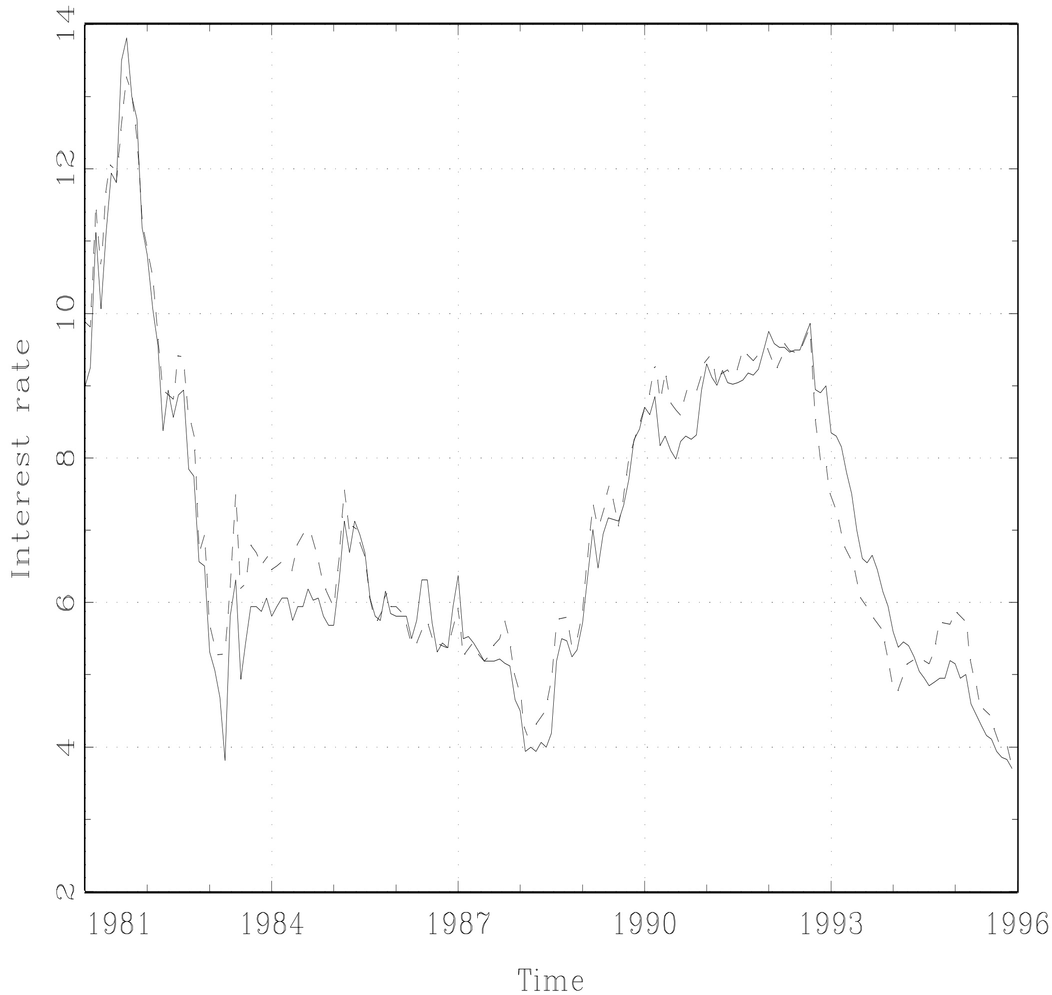
[1] p-values for LM-type tests for smooth transition error-correction in one month Dutch interest rate. The upper panel gives p-values for standard tests, the lower panel for LM-type tests which are robust to additive outliers. The null hypotheses are given in the text.

Table 7: Standard and outlier robust LM-type tests for smooth transition error-correction in a CECM for weekly data on the one-month interest rate[1]

		d					
Test	Null	1	2	3	4	5	6
Standard	H_0	0.000	0.000	0.001	0.001	0.000	0.000
	H_3	0.000	0.000	0.000	0.000	0.517	0.867
	H_2	0.000	0.000	0.001	0.006	0.018	0.006
	H_1	0.003	0.005	0.001	0.001	0.000	0.000
Robust	H_0	0.107	0.181	0.236	0.402	0.422	0.008
	H_3	0.520	0.689	0.625	0.134	0.594	0.027
	H_2	0.048	0.054	0.160	0.114	0.150	0.017
	H_1	0.240	0.324	0.212	0.402	0.607	0.454

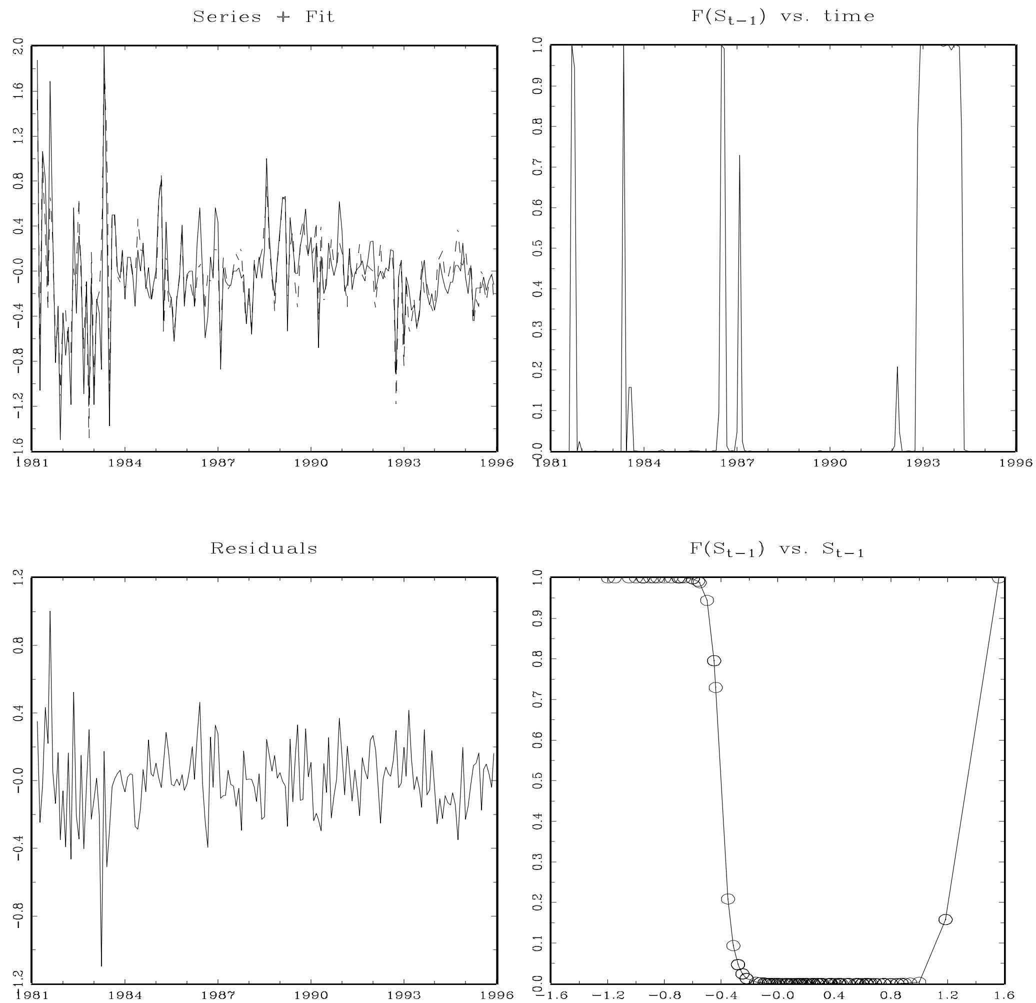
[1] p-values for LM-type tests for smooth transition error-correction in weekly observations on the one month Dutch interest rate. The upper panel gives p-values for standard tests, the lower panel for LM-type tests which are robust to additive outliers. The null hypotheses are given in the text.

Figure 1: Monthly Dutch short- and long-term interest rates



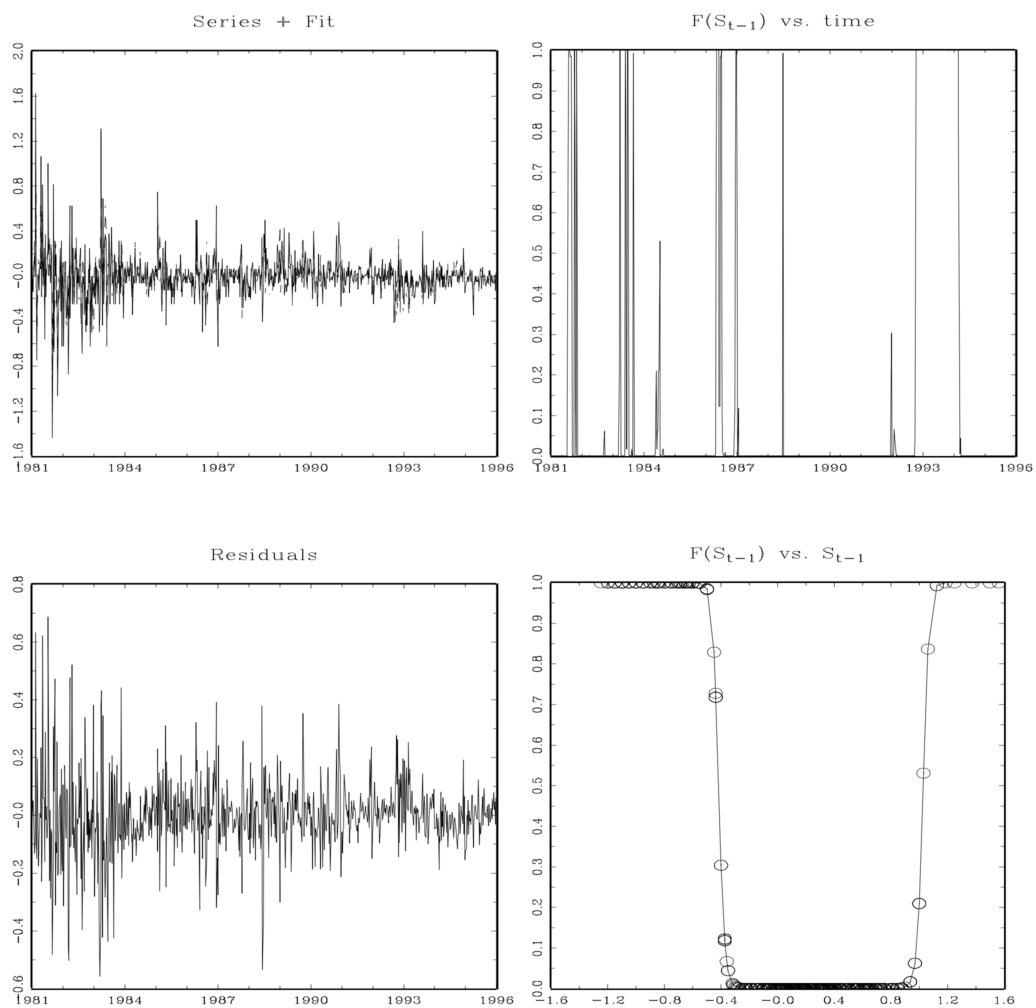
Note: Monthly Dutch short- and long-term interest rates, Jan 1981 - Dec 1995, — short term rate $R_{1,t}$, - - - long-term rate $R_{12,t}$.

Figure 2: Quadratic logistic STECM



Note: Graphical representation of STECM estimated for monthly Dutch short- and long-term interest rates, Jan 1981 - Dec 1995. The parameters of this model are given in (20).

Figure 3: Quadratic logistic model - weekly observations



Note: Graphical representation of STECM estimated for weekly Dutch short- and long-term interest rates, Jan 1981 - Dec 1995. The parameters of this model are given in (21).