Vector Autoregressive Processes
with Nonlinear Time Trends
in Cointegrating Relations*

Antti Ripatti
Bank of Finland
P.O. Box 160
FIN-00101 Helsinki
FINLAND
antti.ripatti@bof.fi

Pentti Saikkonen
Department of Statistics
P.O. Box 54
FIN-00014 University of Helsinki
FINLAND
pentti.saikkonen@helsinki.fi

Running head:

Nonlinear Time Trends in Cointegrating Relations

or if it is too long:

Nonlinear Time Trends

Proofs to:

Antti Ripatti
Bank of Finland
P.O. Box 160
FIN-00101 Helsinki
FINLAND
phone: +358-9-1832527
fax: +358-9-1832560
e-mail: antti.ripatti@bof.fi

2

Abstract

We extend the conventional cointegrated VAR model to allow for general nonlinear deterministic trends. These nonlinear trends can be used to model gradual structural changes in the intercept term of the cointegrating relations. A general asymptotic theory of estimation and statistical inference is reviewed and a diagnostic test for testing the correct specification of an employed nonlinear trend is developed. The methods are applied to Finnish interest rate data. A smooth level shift of the logistic form between the own-yield of broad money and the short-term money market rate is found appropriate for these data. The level shift is motivated by the deregulation of issuing certificates of deposit and its inclusion in the model solves the puzzle of 'missing cointegration vector' found in a previous study.

Keywords: cointegrated VAR model, gradual structural change, nonlinear deterministic trend

1 Introduction

A fairly common finding in empirical analysis of cointegrated time series is that the number of cointegrating relations supported by the data turns out to be smaller than expected on the basis of economic theory. The low power of cointegration tests is often blamed for this but the reason may also be that the data have been affected by structural changes not taken into account in the employed (standard) model. Changes in institutions, policy and technology are typical examples one can think of here. If the changes occur as sudden structural breaks at known points of time they may be modeled by conventional dummy variables. However, as recently discussed by Leybourne, Newbold and Vougas (1998) in the context of unit root tests, it may sometimes be more reasonable to consider structural changes which are gradual and smooth (see also Lin and Teräsvirta (1994) and the references therein). An example is a structural change due to adopting a new technology which usually takes time to completely realize and no clear date of change can be pointed out. On the other hand, the realized change may be gradual even if a particular date could be related to it. For instance, regulations can change at a particular date but economic agents may start adjusting their behavior in advance if they know of the forthcoming mandatory change.

Choosing an incorrect number of cointegrating vectors is a rather extreme misspecification which can result from ignoring structural changes. Even if this misspecification is avoided serious distortions can occur in other inference procedures on cointegrating vectors. In this paper we shall therefore consider an extension of a standard cointegrated vector autoregressive (VAR) model in which smooth or continuous deterministic changes are allowed in the constant term of the cointegrating relations. In a typical example, also relevant for the empirical part of the paper, the constant term changes smoothly from a level to another. In the same way as in Lin and Teräsvirta (1994) and Leybourne et al. (1998) we use a logistic function of time to model this. However, in our general model almost any conceivable smooth function of time can be used to allow for additive structural changes in the cointegrating relations and, although not emphasized in this paper, sudden structural breaks are also possible provided the date(s) of break(s) is (are) a priori known and not estimated from current data.

Recently Saikkonen (2000a, b) has developed a general asymptotic theory of estimation and statistical inference applicable to the model considered in this paper. Since that work is purely theoretical our purpose is here to discuss related empirical aspects and provide a motivation for the considered model. A further contribution of this paper is that we develop diagnostic tests which can be used to check whether a deterministically changing constant term is really needed in the cointegrating relations of a standard VAR model and, more generally, whether a diagnosed structural change can be adequately described by a chosen function of time. These tests, which are based on the Lagrange multiplier (LM) principle, are similar to the tests developed by Lin and Teräsvirta (1994) and Eitrheim and Teräsvirta (1996) for stationary models. On the other hand, since they are also similar to the variable addition test Park (1990) proposed for testing the null of cointegration they have power when the cointegrating rank is chosen too large. This and the lack of a statistical test for cointegration in the presence of a continuous structural change means that our tests are mainly designed for cases where strong prior information about the cointegrating rank is available.

In the empirical part of the paper we analyze a data set of four Finnish interest rates. The same data were previously used by Luukkonen, Ripatti and Saikkonen (1999) who found two cointegrating vectors instead of the expected three. Here we show that this unexpected finding can be attributed to a known structural change which appears as a smooth level shift in one of the three cointegrating relations.

The rest of the paper is organized as follows. The general model and some particular cases are discussed in Section 2. The main points of Gaussian maximum likelihood (ML) estimation and asymptotic inference are summarized in Section 3. The diagnostic tests of the paper are developed in Section 4 and the empirical example is presented in Section 5. Section 6 concludes.

2   Model

Let y_t, $t = 1,...,T$, be an s-dimensional time series generated by a VAR process of order p. Using the error correction form of the process the series is modeled as

$$\Delta y_t = d_t + \Pi y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t, \quad t = 1,...,T, \tag{2.1}$$

where Δ is the usual difference operator, Π $(s{\times}s)$ and Γ_j $(s{\times}s)$ are unknown parameters and d_t $(s{\times}1)$ is a deterministic sequence to be discussed below. Furthermore, the initial values $y_{-p+1},...,y_0$ are observable and ε_t is Gaussian white noise, that is, $\varepsilon_t \sim \text{NID}(0,\Omega)$ with Ω positive definite. This last assumption is only made to facilitate the discussion of likelihood based methods and could be relaxed. It is also assumed that the matrix Π is of rank r $(0 < r < s)$ so that we can write

$$\Pi = \alpha\beta\text{'} \tag{2.2}$$

where α and β are $s{\times}r$ matrices of full column rank. Since we are interested in time series whose stochastic components are integrated of order one and cointegrated we also assume that the parameters of the model satisfy the conditions of Johansen's (1995, p. 49) version of Granger's representation theorem. Thus, it is assumed that the roots of the characteristic equation $\det\left[\left(I_n - \sum_{j=1}^{p-1}\Gamma_j z^j\right)(1-z) - \alpha\beta\text{'}z\right] = 0$ are equal to one or lie outside the unit circle and the matrix $\alpha_\perp\text{'}\left(I_n - \sum_{j=1}^{p-1}\Gamma_j\right)\beta_\perp$ has full rank $s{-}r$. Here, $\beta_\perp$ for example denotes an $s{\times}(s{-}r)$ matrix of full column rank and such that $\beta'\beta_\perp = 0$. The above assumptions imply that, with a suitable specification of initial values, both Δy_t and $\beta' y_t$ are stationary around deterministic trends (see Johansen (1995, p. 49)).

In particular cases of the above model, which can now be called standard, the deterministic sequence d_t may contain an intercept term, seasonal dummies and perhaps also a linear time trend. In this paper we are mainly interested in nonlinear deterministic trends. The idea is to allow for the possibility that a standard model may fail because the mean of the error correction term $\beta' y_t$ changes in a nonlinear fashion. For simplicity we only consider the case where the

5

deterministic term d_t can be absorbed into the cointegrating relations. This means that the sequence d_t is of the form

$$d_t = -\alpha g_t(\mu) \tag{2.3}$$

where $g_t(\mu)$ is a generally nonlinear deterministic function of the time index t and a parameter vector μ. Thus, inserting (2.2) and (2.3) into (2.1) yields

$$\Delta y_t = \alpha(\beta' y_{t-1} - g_t(\mu)) + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_j, \quad t = 1, ..., T. \tag{2.4}$$

A convenient way to specify the sequence $g_t(\mu)$ is to follow the approach used in nonparametric regression and assume that $g_t(\mu) = g(t/T;\mu)$ where $g(\cdot;\mu)$ is a suitable function defined on the interval *[0,1]*. Two examples which illustrate how the function $g(\cdot;\mu)$ may be specified are given by

$$g(x;\mu) = v + \left[1 + exp\{-\gamma(x-\tau)\}\right]^{-1} \delta \tag{2.5}$$

and

$$g(x;\mu) = v + \left[1 - exp\{-\gamma(x-\tau)^2\}\right]\delta. \tag{2.6}$$

Here $\mu = [v' \ \delta' \ \gamma \ \tau]'$ with v and δ unknown $r \times 1$ parameter vectors while γ and τ are scalar parameters with $\gamma > 0$ and $0 < \tau < 1$. These functions have been used to model a smooth or continuous structural change in the coefficients of a dynamic regression model. A recent paper in this area is Lin and Teräsvirta (1994) which also contains references to earlier work. In (2.5) the smooth change is modeled by a logistic function while in (2.6) the density function of the normal distribution is essentially used. To see the idea of (2.5), suppose that the components of δ are positive. Then the parameter vector v is the constant term of the cointegrating relations which applies at the beginning of the sample while $v + \delta$ is a new constant term which applies after a smooth increasing change. The parameter τ determines the average location of the change while γ is a slope parameter which indicates how rapid the change is. The smaller the value of the parameter γ is the longer it takes for the constant term to reach its new level. When the value of γ is "large" we are close to the case where a single structural break occurs. In fact, this case is obtained in the limit by letting $\gamma \to \infty$ because then (2.5) approaches $v + I(x \geq \tau)\delta$ where $I(\cdot)$ is an indicator function taking the value one if the indicated condition is true and zero elsewhere. When the value of the parameter τ is known this amounts to using a conventional step dummy. In the case of (2.6) the change in the constant term is nonmonotonic and symmetric about τ. This specification can be used, for instance, to describe the situation where the mean of the cointegration relations first decreases smoothly from v to δ and then increases smoothly back to v. Again, γ is a slope parameter which determines how rapid these changes are.

Obvious extensions of the above specifications are obtained by defining the function $g(x;\mu)$ as a linear combination of functions of the type given on the right

6

hand sides of (2.5) and (2.6). Another extension, proposed by Lin and Teräsvirta (1994), is given by

$$g(x;\mu) = v + \left[1 + exp\left\{ -\gamma\left(x^k + \tau_1 x^{k-1} + ... + \tau_{k-1} x + \tau_k \right) \right\} \right]^{-1} \delta \qquad (2.7)$$

where $\tau_1, ..., \tau_k$ $(k >1)$ are scalar parameters and the rest of the notation is as in (2.5). When $k=1$ (2.7) reduces to (2.5). In addition to this choice Lin and Teräsvirta (1994) consider the value $k=3$. With these extensions, one can, for instance, allow for more than a single smooth transition in the constant term.

As the above discussion shows, our approach includes a variety of interesting possibilities to model structural changes in cointegrating relations. In some applications continuous changes may be more natural than sudden breaks which have recently received considerable attention. Whether the change is continuous or not the above examples make clear that in these cases it is quite natural to proceed as in nonparametric regression and use the scaled time index t/T in the trend model instead of t. A similar scaling has also been used by other authors to model trends or varying parameters. For instance, Phillips and Hansen (1990) use a similarly scaled time index in cointegrated systems although only in models which are linear in parameters. In cointegrating regressions the same idea is also used by Park and Hahn (1999) who develop nonparametric methods to analyze deterministic changes in cointegrating vectors. Andrews and McDermott (1995), who scale the time index in a slightly different way, consider general nonlinear deterministic trends in a parametric framework but do not allow integrated processes. Dahlhaus (1996a, b) makes extensive use of this approach in his locally stationary models and provides an insightful discussion of its motivation.

The above discussion also implies that we are mainly thinking of deterministic trends which are not "large" like those implied by the linear specification $d_t = \mu_1 + \mu_2 t$. Assuming a linear trend in our formulation means that $d_t = \mu_1 + \mu_2(t/T)$ which is "small" in the sense that it remains bounded as the sample size T tends to infinity. It is worth noting, however, that in the case of a linear trend the inference procedures to be discussed in the next section are valid even if the time index is not scaled by the sample size. The scaling is only required in nonlinear functions of time. On the other hand, the test procedures to be developed in Section 3 are affected by the presence of a linear trend in the cointegrating relations so that, for simplicity, this case will be excluded there.

Since $g_t(\mu) = g(t/T; \mu)$ is assumed the process y_t also depends on the sample size so that, strictly speaking, a notation like y_{tT} should be used. Fortunately, however, this dependence is very simple and therefore does not cause any theoretical complications. Indeed, a version of Granger's representation theorem given by Saikkonen (2000a) shows that the dependence of y_t on T is only due to the deterministic sequence $g(t/T; \mu)$. The stochastic part of y_t is independent of T and identical to its counterpart in Johansen's (1995, p. 49) version of Granger's representation theorem. The assumptions imposed on the function $g(x; \mu)$ also imply that the processes Δy_t and $\beta y_{t-1} - g_t(\mu)$, which are stationary in standard cases, can here be considered as nearly stationary or asymptotically stationary (see Saikkonen, 2000a).

3 ML estimation and statistical inference

The parameters in (2.4) with a chosen specification of the sequence $g_t(\mu)$ or, equivalently, the function $g(x;\mu)$ can be estimated by ML. Since the relevant asymptotic estimation theory is developed in Saikkonen (2000a, b) we shall here only briefly summarize the main points. First note that Saikkonen (2000a, b) also proves results when some nuisance parameters are not identified but, unless otherwise stated, identifiability will be assumed here. Thus, it is assumed that the cointegrating vectors can be written as $\beta' = [I_r - A(\phi)]$ where $A(\phi)$ is a continuously differentiable function of the underlying identifiable parameter vector ϕ. The function $g(x;\mu)$ is assumed to be continuously differentiable with respect to the latter argument. A number of technical regularity conditions are also imposed on the function $g(x;\mu)$. These conditions are satisfied by the examples given in (2.5) and (2.6) if the values of the parameters γ and τ are restricted as $0 < c_1 \leq \gamma \leq c_2 < \infty$ and $0 < c_3 \leq \tau \leq c_4 < 1$. Many other choices of the function $g(x;\mu)$ are also allowed. A general sufficient condition is that $g(x;\mu)$ is continuously differentiable as a function of (x,μ). Since the function $g(x;\mu)$ may be discontinuous with respect to its first argument conventional dummy variables are also included. The most important case which is excluded is that of structural breaks with unknown brake dates or dummy variables with dates of jump depending on unknown parameters. Finally, note that the results of Saikkonen (2000a, b) also allow for the possibility that the short-run parameters α and Γ are smooth functions of an underlying structural parameter vector but we shall not discuss this extension here.

Denote $\theta = [\theta_1' \, \theta_2']'$ where $\theta_1 = [\phi' \, \mu']'$ and $\theta_2 = \text{vec}[\alpha \, \Gamma]$. Here vec signifies the usual columnwise vectorization operator. Thus, conditioning on the initial values $y_{-p+1}, \dots, y_0$ we can write the log-likelihood function of the data as

$$L_T(\theta,\Omega) = -\frac{T}{2}\log\det(\Omega) - \frac{1}{2}\text{tr}\left\{\Omega^{-1}\sum_{t=1}^{T}\varepsilon_t(\theta)\varepsilon_t(\theta)'\right\} \tag{3.1}$$

where

$$\varepsilon_t(\theta) = \Delta y_t - \alpha(y_{1,t-1} - A(\phi)y_{2,t-1} - g_t(\mu)) - \Gamma q_t$$

with $\Gamma = [\Gamma_1 \dots \Gamma_{p-1}]$, $q_t = [\Delta y'_{t-1} \dots \Delta y'_{t-p+1}]'$ and $y_t = [y'_{1t} \, y'_{2t}]'$ partitioned in an obvious way. ML estimators of θ and Ω, denoted by $\hat{\theta} = [\hat{\theta}_1' \hat{\theta}_2']'$ and $\hat{\Omega}$, are obtained by maximizing the function $L_T(\theta,\Omega)$. This maximization problem is of course highly nonlinear. Saikkonen (2000a) shows that, under suitable regularity conditions, the ML estimators $\hat{\theta}$ and $\hat{\Omega}$ exist with probability approaching one and are consistent. The limiting distribution of $\hat{\theta}$ is derived in Saikkonen (2000b). The estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ are asymptotically independent and also asymptotically independent of the estimator $\hat{\Omega}$. The limiting distribution of $\hat{\theta}_1$ is mixed normal and that of $\hat{\theta}_2$ is normal. To be able to describe these results more precisely, define

$$G_{1t}(\theta) = -\begin{bmatrix} (\partial \operatorname{vec} A(\phi)'/\partial \phi)(y_{2,t-1} \otimes \alpha') \\ (\partial g_t(\mu)'/\partial \mu) \otimes \alpha' \end{bmatrix}$$

and

$$G_{2t}(\theta_1) = z_t(\theta_1) \otimes I_s$$

where the symbol $\otimes$ signifies Kronecker's product and $z_t(\theta_1) = [q'_t \ u_{t-1}(\theta_1)']'$ with $u_{t-1}(\theta_1) = y_{1,t-1} - A(\phi)y_{2,t-1} - g_t(\mu)$. Note that here, for example, $\partial g_t(\mu)'/\partial \mu = (\partial g_t(\mu)/\partial \mu')'$ as in Lütkepohl (1996, p. 173). From Saikkonen (2000b) we can now conclude that, under regularity conditions,

$$\hat{M}_{1\cdot2}^{1/2}(\hat{\theta}_1 - \theta_1) \xrightarrow{d} N(0,I) \tag{3.2}$$

where

$$\hat{M}_{1\cdot2} = \sum_{t=1}^{T} \hat{G}_{1t}\hat{\Omega}^{-1}\hat{G}'_{1t} - \sum_{t=1}^{T} \hat{G}_{1t}\hat{\Omega}^{-1}\hat{G}'_{2t}\left(\sum_{t=1}^{T} \hat{G}_{2t}\hat{\Omega}^{-1}\hat{G}'_{2t}\right)^{-1}\sum_{t=1}^{T} \hat{G}_{2t}\hat{\Omega}^{-1}\hat{G}'_{1t}$$

with $\hat{G}_{1t} = G_{1t}(\hat{\theta})$ and $\hat{G}_{2t} = G_{2t}(\hat{\theta}_1)$. For $\hat{\theta}_2$ we have

$$\hat{M}_{2\cdot1}^{1/2}(\hat{\theta}_2 - \theta_2) \xrightarrow{d} N(0,I) \tag{3.3}$$

where $\hat{M}_{2\cdot1}$ is defined in the same way as $\hat{M}_{1\cdot2}$ except that the roles of the subscripts 1 and 2 are interchanged. Note that (3.2) would also hold if $\hat{M}_{1\cdot2}$ were replaced by the first matrix in its defining equation and similarly for (3.3). Since $G_{1t}(\theta)$ contains integrated processes and $G_{2t}(\theta_1)$ asymptotically stationary processes this fact together with (3.2) and (3.3) explains why the limiting distribution of $\hat{\theta}_1$ is mixed normal and that of $\hat{\theta}_2$ is normal.

Approximate standard errors can be obtained for the components of $\hat{\theta}_1$ and $\hat{\theta}_2$ by taking square roots of the diagonal elements of the matrices $\hat{M}_{1\cdot2}$ and $\hat{M}_{2\cdot1}$, respectively. More generally, it is shown in Saikkonen (2000b) that Wald tests with asymptotic chi-square distributions under the null hypothesis can be constructed in the usual way to test (possibly nonlinear) hypotheses on the parameters θ_1 and θ_2 and similar results are also obtained for corresponding likelihood ratio (LR) and LM tests. In these Wald tests the matrices $\hat{M}_{1\cdot2}$ and $\hat{M}_{2\cdot1}$ provide natural estimators of the needed information matrices while their constrained counterparts can be used in corresponding LM tests. It should be noted, however, that these tests assume that the parameters of the model are identified. This particularly means that if (2.5) or (2.6) is specified it is not possible to test the hypothesis $\delta = 0$ because it implies that the parameters γ and τ are not identified. Test procedures for this kind of hypotheses are considered in the next section.

4 Test procedures

An interesting question in the model introduced in Section 2 is whether any nonlinear time trend is actually needed or whether a specified nonlinear time trend, like (2.5), (2.6) or (2.7), can really adequately describe a diagnosed structural change. Tests based on the LM principle appear convenient because they only require estimating parameters of the null model which in the present context is a considerably simpler task than estimating parameters of the unrestricted model.

For ease of exposition, we shall first discuss how to obtain a test for the null hypothesis which states that the sequence $g_t(\mu)$ in (2.4) reduces to a constant. It is obvious that parametric tests cannot be obtained without suitable assumptions of the sequence $g_t(\mu)$ or the function $g(x;\mu)$. Our assumptions are fairly general and apply in a number of cases including those discused in section 2. Partition the parameter vector μ into three subvectors as $\mu = [\nu' \, \lambda' \, \gamma']'$ and assume that for some known value of γ, denoted by γ_0, we have $g(x;\nu,\lambda,\gamma_0) = \nu$ for all ν and λ. Thus, under the null hypothesis $\gamma = \gamma_0$ we have a standard model with a constant term in the cointegrating relations. This formulation applies to the examples in (2.5) and (2.6) although in the former the notation has to be redefined. In both cases the parameter γ is scalar with $\gamma_0 = 0$ while $\lambda = [\delta' \, \tau]'$.

In the above set-up the previously mentioned identification problem appears in the parameter λ which is not identified under the null hypothesis. There has recently been a great interest in this problem and significant progress has been made by Andrews (1993), Andrews and Ploberger (1994) and Hansen (1996). However, instead of following the approach of these authors we shall proceed as in Lin and Teräsvirta (1994) and obtain a simple LM type test by replacing the function $g(x;\mu)$ by a Taylor approximation. In other words, we assume that the function $g(x;\mu)$ is continuously differentiable with respect to μ (or at least γ) and consider the approximation

$$g(x;\nu,\lambda,\gamma) \approx \nu + \frac{\partial g(x;\nu,\lambda,\gamma_0)}{\partial \gamma'}(\gamma - \gamma_0). \tag{4.1}$$

Since the parameter λ appears on the right hand side we still have an identification problem which cannot be solved unless suitable assumptions are made for the partial derivatives therein. We assume that the vector $\partial g(x;\nu,\lambda,\gamma_0)/\partial\gamma$ can be written as a polynomial of x so that for some integers $1 \leq n_1 < ... < n_q$

$$\frac{\partial g(x;\nu,\lambda,\gamma_0)}{\partial \gamma} = \underline{\psi}_0 + \underline{\psi}_1 x^{n_1} + ... + \underline{\psi}_q x^{n_q} \tag{4.2}$$

where the parameters $\underline{\psi}_0,...,\underline{\psi}_q$ $(r \times 1)$ are functions of ν and λ. This assumption is satisfied in (2.5) and (2.6). In (2.5) $q = 1$, $n_1 = 1$ and $\partial g(x;\mu)/\partial\gamma$ evaluated at $\gamma = 0$ is $0.25(x - \tau)\delta$, which is of the form (4.2). In (2.6) we have $q = 2$, $n_1 = 1$ and $n_2 = 2$. Now $\partial g(x;\mu)/\partial\gamma$ evaluated at $\gamma = 0$ is $\delta(x - \tau)^2$ which is a special case of (4.2). The same conclusion clearly applies to (2.7) and extensions of (2.5) and

10

(2.6) where the function $g(x;\mu)$ is a linear combination of functions of the type (2.5) and (2.6).

Now, using (4.1) and (4.2) leads us to replace the original model (2.4) by the approximation

$$\Delta y_t = \alpha(\,\beta'\,y_{t-1} - v - \Psi w_t\,) + \Gamma q_t + e_t, \quad t = 1,\dots,T, \tag{4.3}$$

where $\Psi = [\psi_1 \dots \psi_q]$, $w_t = [(t/T)^{n_1} \dots (t/T)^{n_q}]'$ and e_t is an error term which equals the true error term ε_t when the linearity hypothesis holds or when $\Psi = 0$. Note that here ψ_i stands for $\underline{\psi}_i(\gamma - \gamma_0)$ ($i = 1, \dots, $ q) and, for convenience, $\underline{\psi}_0 = 0$ has been assumed so that $v + \underline{\psi}_0(\gamma - \gamma_0) = v$ holds. It is a simple matter to test the null hypothesis $\Psi = 0$ in (4.3). One possibility is to apply the reduced rank regression method directly to (4.3) and test the null hypothesis $\Psi = 0$ by using the LR test based on the assumption that the error term e_t is Gaussian white noise (see Johansen (1995)). An alternative possibility is to use the corresponding LM test which only requires the application of reduced rank regression to the null model or to (4.3) with the constraint $\Psi = 0$. This is the test considered in this paper. However, we shall not describe it explicitly here because it can be obtained as a special case of the more general test to be developed shortly.

Before generalizing the above test we note that a test obtained for the null hypothesis $\Psi = 0$ in (4.3) can obviously be seen as a version of the variable addition test Park (1990) proposed for testing the null of cointegration. This means that a rejection of the null hypothesis $\Psi = 0$ may also be an indication that the cointegrating rank has been erroneously specified too large. To look this from another angle, suppose economic theory has suggested the cointegrating rank $r = r_0$ and one starts with the standard model, that is, (4.3) with $\Psi = 0$. Suppose further that the data clearly supports the choice $r = r_0-1$ against $r = r_0$ (e.g. the LR test of the null hypothesis $r = r_0-2$ is clearly rejected but the null hypothesis $r = r_0-1$ cannot be rejected at any reasonable significance level). This might also happen when $r = r_0$ is the correct specification but the data are generated by (2.4) with $g_t(\mu)$ given by (2.5) or (2.6). In this situation testing the null hypothesis $\Psi = 0$ in (4.3) with $r = r_0$ is also likely to lead to a rejection. Thus, if there is no prior information about the cointegrating rank it is difficult to say whether one should adopt (2.4) with $r = r_0-1$ and $g_t(\mu) = v$ or alternatively $r = r_0$ and $g_t(\mu)$ modeled by a nonlinear function of time. To be able to discriminate between these two cases a statistical test for cointegration in the general model (2.4) would be needed but, to the best of our knowledge, there is no such test available at the moment. This problem can be avoided, however, when the specification of the cointegrating rank is based on economic theory which, we believe, is fairly typical in practise. Another point worth emphasizing here is that if a standard model is found inadequate and a nonlinear trend is added to the cointegrating relations a meaningful interpretation of the specified nonlinear trend and reasons leading to the failure of the standard model would be reasonable to have. Otherwise, including a nonlinear time trend in the cointegrating relations seems rather pointless.

Now consider the case where (2.4) with a chosen functional form of $g_t(\mu)$ has been adopted and the question is whether an additional or alternative nonlinear term should be used in the model. In this context we proceed in the same way as

in Eitrheim and Teräsvirta (1996) and assume that the correct specification of the nonlinear trend can be written as

$$g_t^*(\mu, \upsilon) = g_t(\mu) + h_t(\upsilon) \tag{4.4}$$

where $h_t(\upsilon)$ depends on a parameter vector υ and $h_t(\upsilon) = 0$ under the null hypothesis. Moreover, $h_t(\upsilon)$ is supposed to be such that $g_t^*(\mu, \upsilon)$ satisfies the regularity conditions required from $g_t(\mu)$ for the theoretical results of Section 2 to apply. In particular, we assume that $h_t(\upsilon) = h(t/T; \upsilon)$ for a suitable function $h(x; \upsilon)$ similar to $g(x; \mu)$. The sequence $g_t(\mu)$ is supposed to contain an additive constant term although this would not be necessary. Now, the obvious idea is to linearize the function $h(x; \upsilon)$ in the same way as was done for the function $g(x; \mu)$ in (4.1) and assume that the related vector of partial derivatives can be written as a polynomial similar to (4.2). Thus, instead of (4.3) we here consider its generalization

$$\Delta y_t = \alpha(\beta' y_{t-l} - g_t(\mu) - \Psi w_t) + \Gamma q_t + e_t, \quad t = 1, ..., T, \tag{4.5}$$

where the error term e_t equals the true error term ε_t when the null hypothesis $h_t(\upsilon) = 0$, or equivalently, $\Psi = 0$ holds. We shall derive a LM test for this latter form of the null hypothesis by assuming that $e_t \sim N(0, \Omega)$ holds in (4.5). The test requires a related constrained ML estimation which can be carried out by specializing the discussion of Section 2 to the present context. Thus, we again assume that $\beta' = [I_r -A(\phi)]$ and that $g_t(\mu)$ satisfies the regularity conditions discussed in Section 2. In what follows, we use the symbol "^" to indicate constrained ML estimators obtained from (4.5) with the constraint $\Psi = 0$. We need the score of vec Ψ based on (4.5) with the assumption $e_t \sim NID(0, \Omega)$. It is straightforward to see that this score evaluated at constrained ML estimators is $\sum_{t=1}^{T} \hat{G}_{3t} \hat{\Omega}^{-1} \hat{\varepsilon}_t$ where $\hat{G}_{3t} = w_t \otimes \hat{\alpha}'$ and $\hat{\varepsilon}_t$ the obvious constrained residual. Specializing the general LM test of Saikkonen (2000b) to the present context leads to the test statistic

$$S = \left(\sum_{t=1}^{T} \hat{G}_{3t} \hat{\Omega}^{-1} \hat{\varepsilon}_t \right)' \hat{M}_{3 \cdot 12}^{-1} \left(\sum_{t=1}^{T} \hat{G}_{3t} \hat{\Omega}^{-1} \hat{\varepsilon}_t \right) \tag{4.6}$$

where $\hat{M}_{3 \cdot 12}$ is defined in the same way as $\hat{M}_{1 \cdot 2}$ in (4.2) except that $\hat{G}_{1t}$ is replaced by $\hat{G}_{3t}$ and $\hat{G}_{2t}$ is replaced by $[\hat{G}'_{1t} \ \hat{G}'_{2t}]'$. The limiting null distribution of test statistic S can be derived from the general results of Saikkonen (2000b) by observing that under the null hypothesis we really have $e_t \sim NID(0, \Omega)$. Thus, we can conclude that, under the null hypothesis and appropriate regularity conditions, test statistic S has a standard chi-square limiting distribution, that is,

$$S \xrightarrow{d} \chi_{qr}^2. \tag{4.7}$$

Of course, large values of the test statistic are critical for the null hypothesis. We shall not give formal results about the properties of our test under the alternative

hypothesis but note that its power should be reasonable when the functions $g(x;v,\lambda,\gamma)$ and $h(x;v)$ are not orthogonal to the polynomial defined by the components of w_t. This is clearly the case in the examples discussed in section 2. It is worth noting, however, that from a significant value of test statistic S it is not possible to deduce the exact form of the function $g(x;v,\lambda,\gamma)$ or $h(x;v)$ one should entertain. This is of course clear because choosing $w_t = t/T$ was found reasonable in the case of (2.5) but this choice is also used when the need of a linear trend is tested. The situation remains the same even if a linear trend can be ruled out a priori. For instance, if (2.6) is suspected it is reasonable to choose $w_t = [(t/T) (t/T)^2]'$ so that the resulting test should also have power against (2.5). Graphical methods may be useful when the specification of the function $g(x;v,\lambda,\gamma)$ or $h(x;v)$ is considered. This point will be exemplified in the next section.

The above discussion also makes clear that our test with $w_t = t/T$ breaks down if the null model is augmented by a linear trend. This is simply because then the additional regressor $w_t = t/T$ used in the test already appears in the model. Of course this difficulty could be circumvented by using higher powers $(t/T)^2$, $(t/T)^3$, ... but we will not pursue this matter explicitly in this paper (cf. Lin and Teräsvirta (1994)).

The expression of test statistic S reveals that it can be computed as a LM test statistic for the need of the regressor vector $\hat{G}'_{3t}$ in the auxiliary regression of $\hat{\varepsilon}_t$ on $\hat{G}'_{1t}$, $\hat{G}'_{2t}$ and $\hat{G}'_{3t}$ with the error term treated as normal with covariance matrix $\hat{\Omega}$. Although the derivation of test statistic S may appear somewhat ad hoc it can in some cases, like those given by (2.5) and (2.6), be motivated by the LM principle. To demonstrate this, consider (2.6) in the case where the null hypothesis implies that $g_t(\mu) = v$. This null hypothesis can be formulated by restricting $\gamma = 0$ in (2.6). The score of γ evaluated at $\gamma = 0$ is $\delta'\alpha'\Omega^{-1}\sum_{t=1}^{T}((t/T)-\tau)^2\varepsilon_t$. If the values of the parameters δ and τ were known a proper LM test could be based on this score. Then we could use test statistic S with $\hat{G}_{3t} = ((t/T)-\tau)^2\delta'\hat{\alpha}'$. Of course this test statistic is infeasible because it depends on the values of the nuisance parameters δ and τ which are unknown in practice. A standard approach in a case like this is to take the supremum of the test statistic over the possible values of the unknown nuisance parameters (see Davies (1977, 1987), Andrews (1993) and Andrews and Ploberger (1994)). However, using the above auxiliary regression interpretation of test statistic S and well-known properties of least squares theory it can be seen that our feasible version of test statistic S is obtained precisely in this way. Thus, in this special case our test can be motivated as a "supLM test". The same result is also obtained in the case of (2.5) and therefore we may call our test statistic a LM type or a score type test statistic.

An alternative motivation of test statistic S is obtained by observing that its derivation is based on the idea of replacing the function $g(x;v,\lambda,\gamma)$ or $h(x;v)$ by a polynomial approximation. This makes the time trend linear in parameters and thereby solves the involved identification problem. This interpretation is to some extent nonparametric. Indeed, such an approximation can be made arbitrarily accurate by taking the degree of the approximating polynomial large enough. However, when one has prior information about the possible form of the function $g(x;v,\lambda,\gamma)$ or $h(x;v)$ it is worthwhile to make use of this information and choose the polynomial approximation accordingly. For instance, in (2.5) and (2.6) first

and second order polynomials seem quite reasonable and parsimonious choices even if they could not be motivated by the LM principle.

We close this section by discussing test statistic S in the special case where the function $g(x;\mu)$ is of the form

$$g(x;\mu) = v + \varphi f(x,\mu_2). \tag{4.8}$$

Here the parameters v and φ are unrestricted and φ may be a matrix. Thus, $\mu = [\mu_1' \ \mu_2']'$ with $\mu_1 = [v' \ (\mathrm{vec}\varphi)']'$. We also assume that no overidentifying restrictions are imposed on cointegrating vectors so that $\beta' = [I_r \ -A(\phi)] = [I_r \ -A]$ with $\phi = \mathrm{vec}A$. The cointegrating vectors are thus identified by normalizing restrictions. A convenient feature of test statistic S in this case is that it is invariant to a particular normalization. To see this, suppose that $\beta' = [I_r \ -A]$ is transformed as $\beta' \to \xi\beta'$ and that the parameters v and φ are transformed similarly as $v \to \xi v$ and $\varphi \to \xi\varphi$. Here ξ is any nonsingular $r \times r$ matrix and, if the parameter matrix α is transformed as $\alpha \to \alpha\xi^{-1}$, we obtain a reparameterization of the original model. The ML residuals $\hat{\varepsilon}_t$ are clearly invariant to these transformations. From the definitions it can further be seen that these transformations do not change the value of $\hat{G}_{3t}$ in test statistic S and the same is also true for $\hat{G}_{2t}$ where the transformation amounts to multiplying $u_{t-1}(\hat{\theta}_1)$ by ξ. Finally, consider the sequence $\hat{G}_{1t}$ which in the present context becomes

$$\hat{G}_{1t} = - \begin{bmatrix} y_{2,t-1} \otimes \hat{\alpha}' \\ \hat{\alpha}' \\ f_t(\hat{\mu}_2) \otimes \hat{\alpha}' \\ \left(\dfrac{\partial \, vec \, f_t(\hat{\mu}_2)'}{\partial \mu_2} \right) \otimes \hat{\varphi}' \hat{\alpha}' \end{bmatrix}.$$

From this and the definition of test statistic S it can readily be seen that the desired invariance also holds for the sequence $\hat{G}_{1t}$ and hence for test statistic S as a whole.

Although our original formulation of test statistic S applies even when overidentifying restrictions are imposed on cointegrating vectors or the parameters v and φ in (4.8) it seems convenient to consider such restrictions only after broader aspects of cointegrating relations have been specified. This means that in practice these overidentifying restrictions are specified only after the cointegrating rank and the form of the deterministic sequence $g_t(\mu)$ or the function $f(x;\mu_2)$ in (4.8) have been specified. If this approach is adopted test statistic S is only applied in cases where the above mentioned invariance property holds so that one does not need to worry about the possible effect of incorrect normalization on the test. In the special case where the null model can be estimated by the reduced rank regression method this particularly means that the resulting ML estimators can be used directly to obtain test statistic S. This happens, for example, when $g_t(\mu) = v$.

5 Empirical application

As an empirical application of the ideas put forward in the previous sections, the relationship between own-yield of broad money and a set of interest rates as an opportunity cost of money is studied.[1] It is often suggested (see, e.g., Rasche (1992)) that in an economy with perfect capital markets a system with a set of opportunity costs of money and own-yield of money should contain one common trend. This means that, if interest rates are believed to be *I(1)* processes, they should be cointegrated with cointegrating rank one less than the number of considered series.[2]

We use monthly Finnish data from the Bank of Finland data base covering the period 1980–1995. In addition to own-yield of money *(IOWN)*, our data set contains three interest rates which are 1-month money market rate *(I1M)*, 3-months money market rate *(I3M)* and 5-years bond rate *(IBOND)*. Graphs of these series[3] are depicted in figure 1. Recently Luukkonen et al. (1999) analyzed the same data set by using a standard cointegrated VAR model with an intercept term restricted to the cointegrating space. Their results were rather surprising. The own yield of money was not cointegrated with any of the three interest rates and only two cointegrating relations were found instead of the expected three.

There is a potential explanation for the previous result. The own-yield of money is a weighted average of the yields of various components of broad money. The Ministry of Finance restricted banks' certificates of deposits (CD) issue until the beginning of 1987 but gradually lifted the quota due to the plans of the Bank of Finland of starting the open-market operations in March 1997. The CDs were chosen as the material for open-market operations. The open-market operations were started because the Bank of Finland had a long-run objective to improve the functioning of the money markets and to provide central bank financing in a deregulated economic environment. After that period the CDs' stock grew rapidly and, consequently, the gap between the own-yield of moneyand the opportunity cost of money diminished substantially. Due to the long-run objectives of the Bank of Finland, the change in the gap can be considered as exogenous with respect to the level of interest rates. The standard model of Luukkonen et al. (1999) assumed that this gap is constant and not changing gradually in time. Thus, the incapability of the employed model to allow for the effects of changing regulations might be the reason for the previous unexpected results.

---

[1] The computation were done with Gauss 3.2 with CML library. We also thank Bent Nielsen letting us use his DisCo program.

[2] Here it is implicitly assumed that the expectation hypothesis of the term structure holds and that a possible risk premium is stationary.

[3] Due to the devaluation speculations in August-September 1986, the short-term interest rates rose temporarily two percentage points. We adjust the data for this spike, which was particularly acute in the 1-month and 3-month money market rates.

Figure 1. **Data**
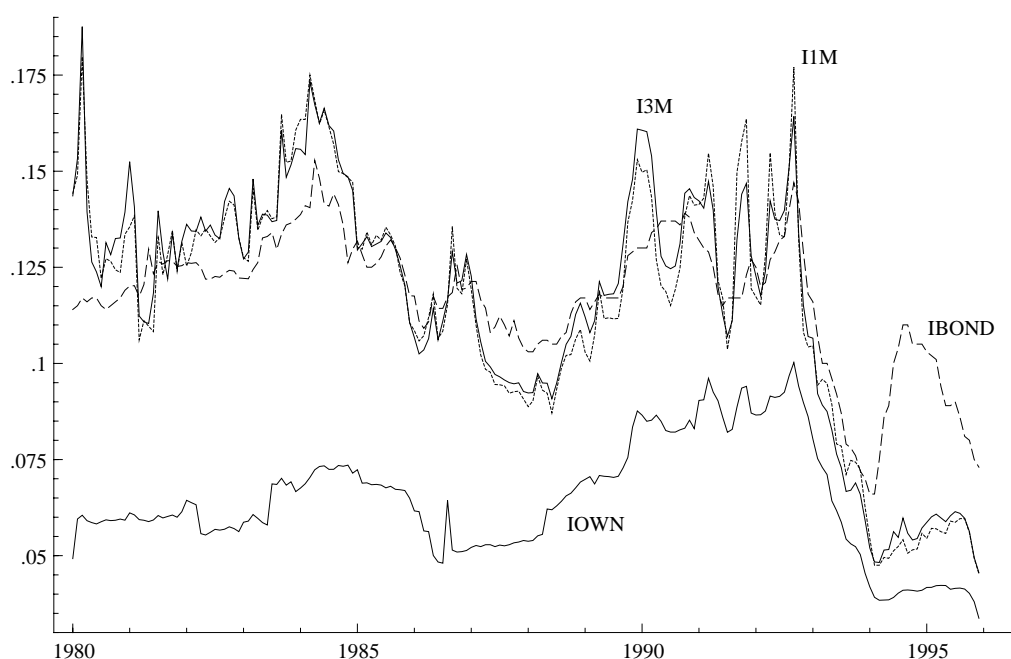
The left panel of figure 2 depicts the first three cointegrating relations of the standard reduced rank VAR(4) model of Luukkonen et al. (1999). The cointegrating relations are ordered according to the canonical correlations. The third cointegrating relation exhibits non-stationary features which is in line with the cointegrating rank tests of Luukkonen et al. (1999). However, as our previous discussion suggests, this nonstationarity might be modelled by a smooth function of time which could pick up the 'gradualism' in the deregulation and banks' adjustment to the new procedures of the Bank of Finland. In order to formally test for this idea, we use the test procedure developed in the previous section. Thus, we continue the analysis by assuming that the cointegrating rank is three and the lag length of the VAR model is four. On the basis of the graph of the third cointegrating relation in the left panel of figure 2 a function of the logistic form might be appropriate. This would mean choosing $w_t = t/T$ in test statistic S. However, in addition to this choice, we also apply the test with $w_t = [t/T\ (t/T)^2]'$. The outcome of these two tests is reported in the two middle columns of table 1. When the test is applied with $w_t = t/T$ the null hypothesis of no structural change is rejected even at the one per cent significance level but with the other choice of w_t a rejection at seven per cent significance level is only possible. The difference between the outcomes of these two tests also suggests that a function of the logistic form is more appropriate than, for instance, a bell shaped function like the one in (2.6).

16

Figure 2.   **Cointegrating relations of the model with and
                    without nonlinear terms**
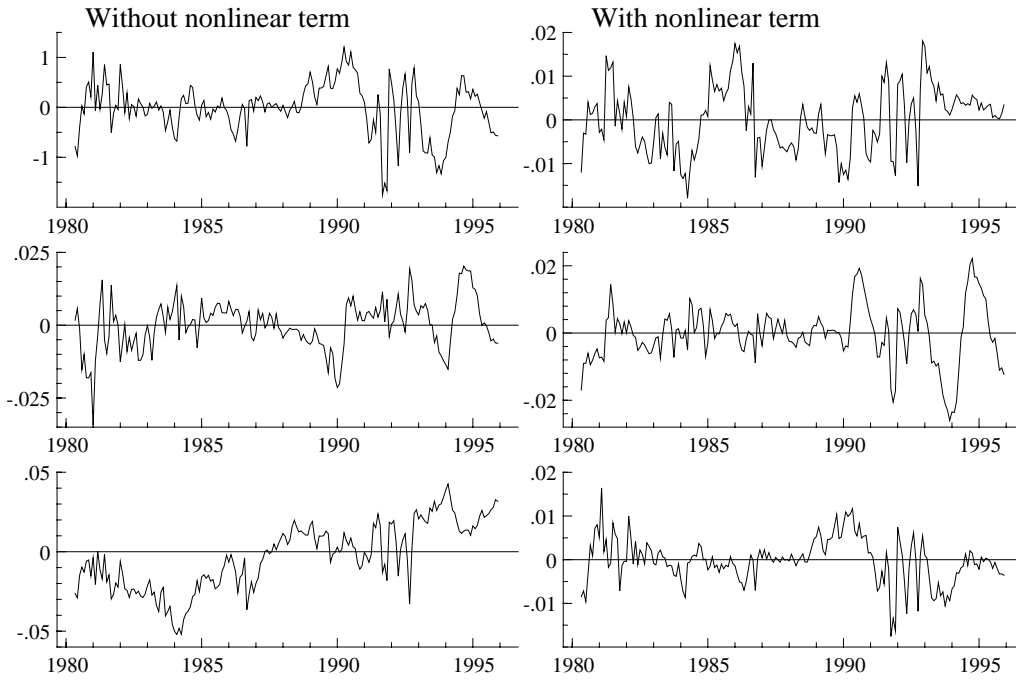
The first column presents cointegrating relations based on unrestricted reduced rank estimation of the VAR(4) model of Luukkonen et al. (1999). The cointegrating relations are ordered according to the canonical correlations so that the first graph corresponds the largest canonical correlation. The second column presents cointegrating relations based on the estimates of model (5.1).

Thus, we augment the model of Luukkonen et al. (1999) by including a logistic function of time in the cointegrating relations. This leads to the model

$$\Delta y_t = \alpha\{y_{1,t-1} - Ay_{2,t-1} - v - \varphi f(\gamma(t/T - \tau))\} + \Gamma q_t + \varepsilon_t \qquad (5.1)$$

where q_t contains lagged differences (lag length is four), $y_{1,t} = [IOWN_t\ IBOND_t\ I3M_t]'$ and $y_{2,t} = IIM_t$. Moreover, A, v and φ are 3×1 parameter vectors and $f(\gamma(t/T-\tau)) = \{1 + exp[-\gamma(t/T - \tau)]\}^{-1}$. Note that we assume that the form of nonlinearity, ie the function $f(\cdot)$, is the same in each cointegrating relation.

Table 1.   **Testing continuous structural change**

Degree of polynomial	Degrees of freedom	Standard model		Model (5.1)	
		Test statistic	p value	Test statistic	p value
1	3	10.55	0.01	4.41	0.22
2	6	11.49	0.07	5.43	0.49

The right column of figure 2 shows the cointegrating relations based on the estimates of model (5.1). According to visual inspection all the three cointegrating relations look stationary which supports the prior belief of three cointegrating relations. Unfortunately there seems to be no formal test for cointegration available in the case of smooth deterministic trends. However, since the estimated

logistic function turned out to be fairly steep[4], we approximated it by a step dummy and applied the LR test for cointegration with critical values computed by the Disco program of Nielsen (1993). Because of this approximation and also because the date where the value of this step dummy is changed is determined by the ML estimate of the parameter τ in (5.1) these critical values should only be treated as approximations of their correct asymptotic counterparts. Using this procedure the p value obtained for the null hypothesis that the cointegrating rank is at most two was approximately 0.07. Thus, even this approximate test was not strongly against the prior belief of three cointegrating relations.

Before discussing ML estimates obtained for the above model we test the adequacy of the specified nonlinearity. Thus, we use the general form of test statistic S and, in the same way as above, apply it both with $w_t = t/T$ and $w_t = [t/T \; (t/T)^2]$'. The results, reported in the last two columns of table 1, imply that the deterministic nonlinearity can be captured by a single logistic function so that the specified model is satisfactory in this respect.

As to the ML estimates of the parameters of model (5.1), we first note that the t values obtained for the last two components of the parameter vector φ were only 0.19 and 0.05. This suggests that the nonlinearity is only present in the first cointegrating relation which describes the relationship between own-yield of money and 1-month money market rate. The p value of the LR test for the joint hypothesis $\varphi_2 = \varphi_3 = 0$ was 0.98 supporting this view. Thus, it is reasonable to consider the corresponding restricted version of model (5.1). The ML estimates obtained for the long run parameters of this model are as follows (standard errors in parenthesis).

---

[4] See the discussion below.

18

$$
\Delta \begin{bmatrix} IOWN \\ IBOND \\ I3M \\ I1M \end{bmatrix}_t = \begin{bmatrix} -0.137 & 0.045 & 0.077 \\ (0.038) & (0.049) & (0.103) \\ 0.056 & -0.170 & 0.176 \\ (0.125) & (0.030) & (0.040) \\ 0.212 & 0.046 & 0.00145 \\ (0.084) & (0.100) & (0.056) \\ 0.330 & 0.036 & 0.390 \\ (0.074) & (0.156) & (0.186) \end{bmatrix}
$$

$$
\left\{ \begin{bmatrix} IOWN \\ IBOND \\ I3M \end{bmatrix}_{t-1} - \begin{bmatrix} 0.633 \\ (0.038) \\ 0.464 \\ (0.049) \\ 0.963 \\ (0.024) \end{bmatrix} I1M_{t-1} - \begin{bmatrix} -0.0224 \\ (0.005) \\ 0.064 \\ (0.0058) \\ 0.0056 \\ (0.0029) \end{bmatrix} \right.
$$

$$
\left. - \begin{bmatrix} 0.0268 \\ (0.0021) \\ 0 \\ \\ 0 \end{bmatrix} \frac{1}{1 + \exp\left[-\underset{(51.1)}{80.1}(t/T - \underset{(0.010)}{0.448}) \right]} \right\} + \hat{\Gamma} q_t + \hat{\varepsilon}_t
$$

The estimation results are interesting. The relationship between 1-month money market rate *(I1M)* and bond rate *(IBOND)* differs significantly from a one-to-one relationship.[5] This one-to-one relationship is preserved between 1-month and 3-months money market rates. The constant term in both relationships significantly differs from zero suggesting, for example, that the 3-months money market rate is on the average a half percentage point above the 1-month money market rate. This is consistent with the common observation that the yield-curve is, on the average, upward-sloping.

The relationship between own-yield of money and 1-month money market rate exhibits structural change in the intercept term. On the average, the own-yield of money has moved along with 1-month money market rate, but the responses have been smaller than unity (0.633). The nonlinear term suggests that the 'adjusted gap' between these two variables has narrowed by 2.7 percentage point. The estimate of τ is 0.448 which suggests that the mid-point of this transition

---

[5] One-to-one relationship suggested by, among others, Campbell and Shiller (1987) and Stock and Watson (1988) – and recently studied, e.g., by Lanne (1997) – is based on the cointegration implications of the expectation hypothesis of the term structure of interest rates. The one-to-one relationship between the own-yield and the opportunity cost of money is suggested by, e.g., Ripatti (1998).

period has been in the April of 1987. That is close to the date when the Bank of Finland started its market operations (March 1987). The shift to the new level has been quite rapid $(\gamma = -80)$ – graphical investigation suggests that the major part of the shift occurred within twelve months surrounding the mid-point (April 1987). It is quite natural that the transition started half a year before the mid-point, because the banks' CD quotas were extended already before the Bank of Finland started its open-market operations in order to enlarge the money markets. It is also natural to believe that it took some time after the start of the open-market operations to saturate banks' needs to issue CDs. The uncertainty in the estimation of the parameter γ is likely to arise from this rapid shift, which means that we do not have very many observations from the 'transition period'. Note that this also means that it is not reasonable to interpret this results as an indication of insignificance.[6] The estimate of the parameter α is reasonable. The residuals of the model exhibit no autocorrelation, but they suffer from non-normality and, in some cases, heteroscedasticity. Trying to improve the model in this respect is outside the scope of this paper, however.

6   Conclusion

This paper has argued that the puzzle of 'missing cointegrating vector', which is not uncommon in empirical analysis of cointegrated systems, can in some cases be explained by structural changes and solved by including nonlinear deterministic trends in cointegrating relations. This idea was implemented in the paper by extending the conventional cointegrated VAR model to a fairly general form. The case of continuous deterministic trends was emphasized because gradual structural changes are often conceivable and may not be well modeled by conventional dummy variables which have so far been mainly used in these contexts.

The usefulness of the proposed approach was demonstrated in the paper by an empirical example on interest rate data. In this example a previous model was augmented by including a logistic trend term in the cointegrating relations and the puzzle of 'missing cointegrating vector' could thereby be solved. According to our general idea, we could also pinpoint reasons for this augmentation and interpret the parameter estimates in the logistic trend in a reasonable way. However, even in our model no one-to-one relationship between interest rates was found.

Finally, in the same way as in Lin and Teräsvirta (1994), it is also here worth mentioning that in the case of a nonlinear time trend the ML estimation required in our approach is demanding and can cause problems. The employed estimation algorithm may converge slowly and find a local optimum. Good starting values are therefore an important prerequisite to succesful parameter estimation.

---

[6] See Lin and Teräsvirta (1994) for a similar result and further discussions.

References

Andrews, D.K.W. (1993) **Tests for Parameter Instability and Structural Change with Unknown Change Point.** Econometrica 61, 821–856.

Andrews, D.W.K. – McDermott, C.J. (1995) **Nonlinear Econometric Models with Deterministically Trending Variables.** Review of Economic Studies 62, 343–360.

Andrews, D.K.W – Ploberger, W. (1994) **Optimal Tests when a Nuisance Parameter is Present only Under the Alternative.** Econometrica 62, 1383–1414.

Campbell, J.Y. – Shiller, R.J. (1987) **Cointegration test of present value models.** Journal of Political Economy 95(5), 1052–1088.

Dahlhaus, R. (1996a) **Fitting Time Series Models to Nonstationary Processes.** The Annals of Statistics. 25, 1–37.

Dahlhaus, R. (1996b) **Asymptotic Statistical Inference for Nonstationary Processes with Evolutionary Spectra.** In Athens Conference on Applied Probability and Time Series Analysis 2, P.M. Robinson and M. Rosenblatt (eds.), Springer, New York.

Davies, R.B. (1977) **Hypothesis Testing when the Nuisance Parameter is Present Only Under the Alternative.** Biometrika 64, 247–254.

Davies, R.B. (1987) **Hypothesis Testing When the Nuisance Parameter is Present Only Under the Alternative.** Biometrika 74, 33–43.

Eitrheim, Ø. – Teräsvirta, T. (1996). **Testing the Adequacy of Smooth Transition Autoregressive Models.** Journal of Econometrics 74, 59–75.

Hansen, B.E. (1996) **Inference When a Nuisance Parameter is Not Identified Under the Null Hypothesis.** Econometrica 64, 413–430.

Johansen, S. (1995) **Likelihood Based Inference in Cointegrated Vector Autoregressive Models.** Oxford University Press, Oxford.

Lanne, M. (1997) **Essays on Inference in Time Series Models with Near Unit Roots.** University of Helsinki.

Leybourne, S. – Newbold, P. – Vougas, D. (1998) **Unit Roots and Smooth Transitions.** Journal of Time Series Analysis 19, 83–97.

Lin C.-F. – Teräsvirta, T. (1994) **Testing the Constancy of Regression Parameters Against Continuous Structural Change.** Journal of Econometrics 62, 211–228.

Lütkepohl, H. (1996) **Handbook of Matrices**. John Wiley, Chichester.

Luukkonen, R. – Ripatti, A. – Saikkonen, P. (1999) **Testing for a Valid Normalization of Cointegrating Vectors in Vector Autoregressive Processes.** Journal of Business & Economic Statistics (forthcoming).

Nielsen, B. (1993) **DisCo, version 1.1, A Program for Simulating Asymptotic Tables of the Cointegration Rank Test.** Institute of Mathematical Statistics, University of Copenhagen.

Park, J.Y. (1990) **Testing for Unit Roots and Cointegration by Variable Addition.** In Rhides, G.F. and Fomby, T.B. (eds.), Advances in Econometrics. Greenwich, JAI Press.

Park, J. –Hahn, S. (1999) **Cointegrating Regressions with Time Varying Coefficients.** Econometric Theory 15, 664–703.

Phillips, P.C.B. – Hansen, B.E. (1990) **Statistical Inference in Instrumental Variables Regression with I(1) Processes.** Review of Economic Studies 57, 99–125.

Rasche, R.H. (1992) **Money Demand and the Term Structure; Some New Ideas on an Old Problem.** Econometrics and economic theory paper 9211, Michigan State University.

Ripatti, A. (1998) **Stability of the Demand for M1 and Harmonized M3 in Finland.** Empirical Economics, 23, 317–337.

Saikkonen, P. (2000a) **Consistent Estimation in Cointegrated Vector Autoregressive Models with Nonlinear Time Trends in Cointegrating Relations.** Econometric Theory (forthcoming).

Saikkonen, P. (2000b) **Statistical Inference in Cointegrated Vector Autoregressive Models with Nonlinear Time Trends in Cointegrating Relations.** Econometric Theory (forthcoming).

Stock, J.H. – Watson, M.W. (1988) **Testing for Common Trends.** Journal of the American Statistical Association 83, 1097–1107.