DISCUSSION PAPER SERIES

No. 4958

FORECASTING THE SPOT EXCHANGE RATE WITH THE TERM STRUCTURE OF FORWARD PREMIA: MULTIVARIATE THRESHOLD COINTEGRATION

Michel R van Tol and Christian C Wolff

FINANCIAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP4958.asp**

ISSN 0265-8003

FORECASTING THE SPOT EXCHANGE RATE WITH THE TERM STRUCTURE OF FORWARD PREMIA: MULTIVARIATE THRESHOLD COINTEGRATION

Michel R van Tol, Maastricht University
Christian C Wolff, LIFE, Maastricht University and CEPR

Discussion Paper No. 4958
March 2005

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 4958

March 2005

ABSTRACT

Forecasting the Spot Exchange Rate with the Term Structure of Forward Premia: Multivariate Threshold Cointegration

In this paper we develop a multivariate threshold vector error correction model of spot and forward exchange rates that allows for different forms of equilibrium reversion in each of the cointegrating residual series. By introducing the notion of an indicator matrix to differentiate between the various regimes in the set of nonlinear processes we provide a convenient framework for estimation by OLS. Empirically, out-of sample forecasting exercises demonstrate its superiority over a linear VECM, while being unable to out-predict a (driftless) random walk model. As such we provide empirical evidence against the findings of Clarida and Taylor (1997).

JEL Classification: C51, C53 and F31
Keywords: foreign exchange, multivariate threshold cointegration and TAR models

Michel R van Tol
Faculty of Economics and Business Administration
Maastricht University
P.O. Box 616
6200 MD Maastricht
THE NETHERLANDS
Email: m.tol@berfin.unimaas.nl

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=162193

Christian C Wolff
Limburg Institute of Financial Economics
Maastricht University
P.O. Box 616
6200 MD Maastricht
THE NETHERLANDS
Tel: (31 43) 388 3838
Fax: (31 43) 388 4875
Email: c.wolff@berfin.unimaas.nl

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=105187

Submitted 24 February 2005

Forecasting the Spot Exchange Rate with the Term Structure of Forward Premia: Multivariate Threshold Cointegration

Michel R. van Tol [a], Christian C.P. Wolff [b,*]

[a] *Maastricht University*
[b] *Maastricht University and Centre for Economic Policy Research (CEPR)*

Abstract

In this paper we develop a multivariate threshold vector error correction model of spot and forward exchange rates that allows for different forms of equilibrium reversion in each of the cointegrating residual series. By introducing the notion of an indicator matrix to differentiate between the various regimes in the set of nonlinear processes we provide a convenient framework for estimation by OLS. Empirically, out-of sample forecasting exercises demonstrate its superiority over a linear VECM, while being unable to out-predict a (driftless) random walk model. As such we provide empirical evidence against the findings of Clarida and Taylor (1997).

J.E.L. Classification: C51; C53; F31

Keywords: Multivariate threshold cointegration; TAR models; Foreign exchange

1. Introduction

Since Meese and Rogoff's (1983a,b) illustration that extant structural exchange rate models are unable to out-predict the naïve alternative of a driftless random walk, researchers have been applying a plethora of new econometric techniques in an attempt at uncovering empirical regularities that have thus far eluded them. Despite the rapid evolution of these techniques research subsequent to the Meese and Rogoff papers has failed to convincingly overturn their findings. On occasion statistical models are able to capture the dynamics of exchange rate movements succinctly but are usually later shown to be too sample specific and therefore poor predictors out-of-sample. Thus, when an apparently theoretically sound model produces out-of-sample forecasts that significantly out-predict the random walk model, at various horizons, it piques the curiosity of any who have an interest in the field.

Just such results arose in an interesting investigation of the forecastability of spot exchange rates by Clarida and Taylor (1997). They develop a framework for the prediction of future spot rates by extracting information from the term structure of forward exchange premiums. They do so by showing that while both the spot and forward rates are integrated processes an economically justifiable combination of these processes is stationary – implying the presence

* Corresponding author. Address: Faculty of Economics and Business Administration, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Email: c.wolff@berfin.unimaas.nl

of the property of cointegration and the existence of a vector error correction model (VECM) that may be used for forecasting spot rates.

The theoretical framework on which their empirical study hinges is interesting and leads naturally to the idea of threshold cointegrated spot and forward rates, which is developed in this paper. Specifically, a multivariate threshold cointegration framework for spot and forward exchange rates is developed and implemented empirically. An important component of the empirical implementation is our analysis of out-of-sample forecasts.

Linear cointegration implicitly presupposes that reversion to long-run equilibrium is both continuous and symmetric, while economic conjectures abound as to why this need not necessarily be the case; the law of one price being a pertinent example. Consequentially, Balke and Fomby (1997) detail veridical equilibrium reversion processes that more accurately reflect the empirical findings insinuated by the studies of a plenitude of researchers from the many, diverse, sub-domains in economics. By allowing for the presence of thresholds in the disequilibrium process of a cointegrated system they further the general tendency of developing nonlinear generalizations of linear techniques; a deliberate break from Slutzky's (1927) linear paradigm.

Section 2 details and presents a general multivariate threshold cointegration framework. Section 3 discusses a two-step estimation procedure whereby the threshold cointegrated system is estimated conditionally on parameter estimates of the threshold autoregressions which we hypothesis govern the sequence of disequilibria. Section 4 presents the theoretical framework of Clarida and Taylor (1997) which we extend to allow for the incorporation of threshold effects in the forward premia of a number of exchange rates. Section 5 delineates a comprehensive study of the stationarity and linearity properties of the forward premia. Once nonlinearity has been determined we estimate the multivariate threshold model these properties imply and utilize the model in producing out-of-sample forecasts which are juxtaposed against the forecasts of a series of alternative forecasting models. Section 6 concludes and provides a number of suggestions for future research.

2. A Multivariate Threshold Cointegration Framework

Curiously, while a plenitude of researchers have developed statistical tests and procedures for the case of threshold cointegration in a bivariate system (see Balke and Fomby (1997), Enders and Granger (1998), Lo and Zivot (2001), Hansen and Seo (2002) and Seo (2003a, 2003b)) multivariate threshold cointegration has, as yet, been left untended. In this paper we introduce the notion of multivariate threshold cointegration whereby each time-series of cointegrating residuals in a cointegrated system are governed by separate threshold autoregressive models. Additionally, we present a simple and convenient formulation of the multivariate threshold VECM it implies.

To see how multivariate threshold cointegration, of rank *r*, can arise consider a *K*-dimensional $VAR(p)$ model

$$\mathbf{x}_t = \boldsymbol{\mu} + \boldsymbol{\Gamma}_1 \mathbf{x}_{t-1} + \cdots + \boldsymbol{\Gamma}_p \mathbf{x}_{t-p} + \boldsymbol{\eta}_t \tag{1}$$

where $\mathbf{x}_t = (x_{1,t}, x_{2,t}, \ldots, x_{K,t})'$, for $t = 1,2,\ldots,N$, with $\boldsymbol{\eta}_t$ an innovation process i.e. $E[\boldsymbol{\eta}_t] = \mathbf{0}$, $E[\boldsymbol{\eta}_t \boldsymbol{\eta}_t'] = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}$, and $E[\boldsymbol{\eta}_t \boldsymbol{\eta}_s'] = \mathbf{0}$ for $s \neq t$. Suppose that the process is nonstationary with $\left|\mathbf{I}_2 - \boldsymbol{\Gamma}_1 z - \cdots - \boldsymbol{\Gamma}_p z^p\right| = (1-\lambda_1 z)\cdots(1-\lambda_n z) = 0$ for $z = 1$, where λ_i are the reciprocals of the roots of the determinantal polynomial. Logically, at least one of them must equal unity; while all other roots are assumed to lie outside the unit circle, i.e. all $\lambda_i \neq 1$ are inside the complex unit circle. Then since $\left|\mathbf{I}_2 - \mathbf{\Gamma}_1 - \cdots - \mathbf{\Gamma}_p\right| = 0$ the matrix $\mathbf{\Pi} = \mathbf{I}_2 - \mathbf{\Gamma}_1 - \cdots - \mathbf{\Gamma}_p$ is singular and if $rank(\mathbf{\Pi}) = r$, it may be decomposed as $\mathbf{\Pi} = \boldsymbol{\alpha\beta}$, where $\boldsymbol{\alpha}$ is $(K \times r)$ and $\boldsymbol{\beta}$ is $(r \times K)$. The matrix $\boldsymbol{\alpha}$ is typically referred to as the loading matrix while $\boldsymbol{\beta}$ is the matrix of cointegrating vectors. If $r = 0$, $\Delta\mathbf{x}_t$ will have a stable $VAR(p-1)$ representation and, if $r = K$, $\mathbf{x}_t$ is a stationary $VAR(p)$ process.

If $0 < rank(\mathbf{\Pi}) = r < K$, model (1) has the following error correction representation

$$\Delta\mathbf{x}_t = \boldsymbol{\mu} - \boldsymbol{\alpha\beta}\mathbf{x}_{t-1} + \mathbf{\Psi}_1\Delta\mathbf{x}_{t-1} + \cdots + \mathbf{\Psi}_{p-1}\Delta\mathbf{x}_{t-p+1} + \boldsymbol{\eta}_t \tag{2}$$

where $\mathbf{\Psi}_i = -\left(\mathbf{\Gamma}_{i+1} + \cdots + \mathbf{\Gamma}_p\right)$ for $i = 1,\ldots,p-1$[1]. Let, allowing for the simplifying assumption that each equilibrium error process is independent of all others, the *r*-vector $\mathbf{z}_t = \left[z_{1,t}, z_{2,t},\ldots,z_{r,t}\right]'$, where $z_{i,t} = \beta_i x_t$ with β_i denoting the *ith* row vector of $\boldsymbol{\beta}$, define *r*-discrepancies from long-run equilibria and, additionally, allow each $z_{i,t}$ to have a *k*-regime TAR representation[2], i.e.

$$z_{i,t} = \sum_{j=1}^{k}\left[\phi_{i,0}^{(j)} + \phi_{i,1}^{(j)} z_{it-1} + \cdots + \phi_{i,p^{(j)}}^{(j)} z_{i,t-p} + \varepsilon_{i,t}^{(j)}\right] I\left(\theta_i^{(j-1)} < v_{i,t-d} \le \theta_i^{(j)}\right) \tag{3}$$

where the positive integers *j*, *d*, and p denote regime index, threshold delay and autoregressive lag-order, respectively; $\varepsilon_{i,t}^{(j)}$ is assumed to be a martingale difference sequence with respect to the past history of z_t; $\phi_{i,l}^{(j)}$ denotes the autoregressive coefficient of the *lth* lag in regime *j*; $I(A)$ is an indicator function such that it equals unity when event *A* occurs and is zero otherwise; $v_{i,t-d}$ is the delay variable which is evaluated relative to the threshold vector $\boldsymbol{\theta}_i = \left[\theta_i^{(0)}, \theta_i^{(1)},\ldots,\theta_i^{(k)}\right]$, where $\theta_i^{(k)} = -\theta_i^{(0)} = \infty$. If $v_{i,t-d} \equiv z_{i,t-d}$ then (3) is typically referred to as a self-exciting TAR (SETAR) model.

Conditional on the existence of *k* non-trivial thresholds, in (3), a general multivariate threshold error correction representation may be formulated as

$$\Delta\mathbf{x}_t = \boldsymbol{\mu} + \boldsymbol{\alpha}^{(1)}\mathbf{I}^{(1)}\mathbf{z}_{t-1}^{(1)} + \boldsymbol{\alpha}^{(2)}\mathbf{I}^{(2)}\mathbf{z}_{t-1}^{(2)} + \cdots + \boldsymbol{\alpha}^{(k)}\mathbf{I}^{(k)}\mathbf{z}_{t-1}^{(k)} + \sum_{i=1}^{p-1}\mathbf{\Psi}_i\Delta\mathbf{x}_{t-i} + \boldsymbol{\eta}_t \tag{4}$$

or, more compactly, as

$$\Delta\mathbf{x}_t = \boldsymbol{\mu} + \sum_{j=1}^{k}\boldsymbol{\alpha}^{(j)}\mathbf{I}^{(j)}\mathbf{z}_{t-1}^{(j)} + \sum_{i=1}^{p-1}\mathbf{\Psi}_i\Delta\mathbf{x}_{t-i} + \boldsymbol{\eta}_t \tag{5}$$

where $\boldsymbol{\alpha}^{(j)}$ is the $(K \times r)$ loading matrix associated with equilibrium errors $\mathbf{z}_t^{(j)}$ when $\theta_i^{(j-1)} < v_{i,t-d} \le \theta_i^{(j)}$, for all $i = 1,\ldots,r$; $\mathbf{I}^{(j)}$ is a $(r \times rk)$ block diagonal indicator matrix, the *jth* element of the main diagonal being equal to unity when $\theta_i^{(j-1)} < v_{i,t-d} \le \theta_i^{(j)}$ and is zero otherwise with all off-diagonal elements set equal to zero; $\mathbf{z}_{t-1}^{(j)}$ is a lagged-value of the

[1] See Hylleberg and Mizon (1989) for other representations of cointegrated processes.

[2] In the context of the framework that we develop the equilibrium error processes may be assigned any functional form; including an allowance for the interplay between two or more disequilibrium processes.

disequilibrium term being explicitly dependent on the functional form present in the *jth* regime of the *ith* cointegrating residual[3]. See Appendix for further clarification.

Stationarity of (5) is ensured if $\mathbf{z}_{i,t}^{(1)}$ and $\mathbf{z}_{i,t}^{(k)}$ are stationary for all *i*; Tjøstheim (1990) has shown that a sufficient condition for stationarity of (3), and hence (5), is that the roots of the AR processes in the outer-regimes are less than $|1|$[4]. Unit root testing in the presence of threshold-nonlinearity requires a model under the alternative that takes account of this characteristic; Pippenger and Goering (1993) were the first to show the severely detrimental effect threshold boundaries have on conventional unit root tests. Consequently, Enders and Granger (1998), Berben and van Dijk (1999) and Kapetanios and Shin (2002) have developed tests specifying single and double-threshold models under the alternative. In Van Tol and Wolff (2004) we developed an alternative unit root test, with attractive power properties.

In this section we developed a general multivariate threshold error correction model with the capacity of allowing differing threshold processes to define each cointegrating residual series; including the special case of linear cointegration, when trivial threshold effects are present in each series, as well as, no cointegration when, for example, each disequilibrium term is an element of the inner-regime of a Band-TAR process admitting a unit root process there within. Additionally, the formulation is conducive to standard estimation procedures, the specifics of which are detailed in the next section.

3. Multivariate Threshold VECM Estimation

3.1. TAR Model Parameter Estimation

The general multivariate threshold VECM developed in the previous section convolutes around two estimation problems. The first relates to specifying and estimating TAR models for each series of cointegrating errors while the second entails estimation of the autoregressive parameters of the VECM, conditional on the TAR model estimates. Disentangling these estimations seems the most tractable approach; direct estimation being in an embryonic stage of development relates primarily to bivariate cointegrated systems. Resultantly, the estimation procedures which we detail in this section relate firstly to TAR model estimation and then the conditionally estimated parameters of model (5).

The range of possible specifications of model (3) is richly diverse. Consider the following two "first-difference" reparameterizations of (3)

$$\begin{aligned}\Delta z_t = & \left[\phi_0^{(1)} + \phi_1^{(1)} z_{t-1}\right] I_t\left(z_{t-1} < \theta^{(1)}\right) + \left[\phi_0^{(2)} + \phi_1^{(2)} z_{t-1}\right] I_t\left(\theta^{(1)} \le z_{t-1} < \theta^{(2)}\right) + \\ & \left[\phi_0^{(3)} + \phi_1^{(3)} z_{t-1}\right] I_t\left(z_{t-1} \ge \theta^{(2)}\right) + \varepsilon_t\end{aligned} \tag{6}$$

$$\begin{aligned}\Delta z_t = & \left[\phi_1^{(1)} (z_{t-1} - \theta^{(1)})\right] I_t\left(z_{t-1} < \theta^{(1)}\right) + \left[\phi_0^{(2)} + \phi_1^{(2)} z_{t-1}\right] I_t\left(\theta^{(1)} \le z_{t-1} < \theta^{(2)}\right) \\ & \left[\phi_1^{(3)} (z_{t-1} - \theta^{(2)})\right] I_t\left(z_{t-1} \ge \theta^{(2)}\right) + \varepsilon_t\end{aligned} \tag{7}$$

[3] Note that the Threshold VECM, or TVECM, typically discussed in the literature relates only to cointegration in a bivariate system and, hence, defines a model that is very different from model (6.5). A bivariate TVECM (see Lo and Zivot (2001) for a detailed discussion) has the advantage of allowing the lagged changes in a VECM to be determined by the threshold variable of the TAR model governing the series of cointegrating residuals; while the multivariate setting does not possess this capacity.

[4] See Van Tol (2005) for a more detailed discussion of the conditions required to ensure stationarity of a TAR process.

where, for convenience, it is assumed that $\varepsilon_t \sim IIN\left(0, \sigma_\varepsilon^2\right)$ while Δ denotes the first-difference operator i.e. $\Delta z_t = z_t - z_{t-1}$. Notationally, the remainder may be inferred from (3). Equations (6) and (7) are generalizations of models that are typically referred to as Equilibrium-TAR and Band-TAR models; see Balke and Fomby (1997). Here we refrain from instituting the requirement that the thresholds are symmetrically dispersed around the unconditional mean of the cointegrating error process, nor do we require that the AR processes in the outer-regimes are equivalent. Enders and Granger (1998) introduced the idea of using momentum as a delay variable by substituting z_{t-d} in a SETAR model with Δz_{t-d}; terming the resultant model a Momentum-TAR, or M-TAR, model. Their findings, when specifying an M-TAR model under the alternative of a unit root test, indicate the potential contribution that the model could have when allowing discrepancies from equilibrium in a threshold cointegrated system to be modeled by an M-TAR model. Subjecting (3) to the requirement that the piecewise linear autoregressive function be continuous everywhere, results in an important TAR sub-class, the continuous TAR or CTAR model, introduced into the literature by Chan and Tsay (1998)[5].

Balke and Fomby (1997) further restrict specifications (6) and (7) by allowing the inner-regime to be driven by a unit root process. The consequences of doing so for (5) are interesting since then the process will follow a $VAR(p)$ model when the cointegrating errors are within the corridor formed by the threshold vector of (4). When in the outer-regimes, however, $\mathbf{\Pi} = \boldsymbol{\alpha}^{(j)}\boldsymbol{\beta}$ will be of reduced-rank implying that an equilibrium reverting mechanism governs the fluctuating cointegrating errors. Hence, (5) switches between a $VAR(p)$ and a VECM specification depending on the magnitude of the delay variable relative to the locations of the thresholds in (3).

Generally, TAR models may be categorized as belonging to one of two sub-classes: those that have regimes where the AR process is a functional of a threshold parameter and those that do not. Fitting SETAR models with regimes that are independent of the values of the threshold vector is typically done by sequential conditional least squares. In this case the continuous threshold parameter space $\Theta \subseteq \Re^{k-1}$ is restricted to the set of points

$$\mathbf{\Theta} = \left\{\boldsymbol{\theta} \mid z_{\left(\lfloor N^* \pi_0 \rfloor\right)} \le \theta^{(i)} \le z_{\left(\lfloor (1-\pi_1) N^* \rfloor\right)}\right\} \tag{8}$$

where $z_{(i)}$ denotes the *ith* order statistic of the delay variable; (π_0, π_1) are trimming parameters which ensure that an adequate number of observations are present in the outer-regimes to ensure the accurate estimation of the parameters of that regime; $N^* = N - \max(d, L)$ is the effective sample size where L is an *a priori* choice of the maximum allowable dimension of the lag space; and $\lfloor \cdot \rfloor$ denotes a floor function here defined as the integer part operator. Parameter estimation is done by computing the residual variance conditional on the *gth* candidate threshold vector as

$$\tilde{\sigma}_g^2 = \tfrac{1}{N^*} \sum_{j=1}^{k} \left(\Delta \mathbf{z}_g^{(j)} - \mathbf{F0}_g^{(j)} \tilde{\mathbf{\Phi}}_g^{(j)}\right)' \left(\Delta \mathbf{z}_g^{(j)} - \mathbf{F0}_g^{(j)} \tilde{\mathbf{\Phi}}_g^{(j)}\right) \tag{9}$$

[5] Instituting the restrictions $p = 1$; $\phi_0^{(2)} = 0$; and $\phi_1^{(2)} = 1$ results in three-regime SETAR models that allow for unit root behavior in the inner-regime and autoregressive behavior in the outer-regimes. It is easily shown that restricting (7), in this manner, will result in a model that meets the "continuous everywhere" requirement implying that it belongs to the CTAR sub-class.

where $\Delta\mathbf{z}_g^{(j)}$ denotes the dependent variable in regime *j*, in (6) for example; and $\mathbf{F0}_g^{(j)}$ is the $\left(N^{*(j)} \times L\right)$ matrix of regressor values, where $N^{*(j)}$ is the effective sample size in regime *j*[6]; and the vector of autoregressive coefficient estimates are defined as $\widetilde{\mathbf{\Phi}}_g^{(j)} = \left[\widetilde{\phi}_1^{(j)}, \widetilde{\phi}_2^{(j)}, \ldots, \widetilde{\phi}_L^{(j)}\right]'$ where $\widetilde{\phi}_i^{(j)}$ is the coefficient estimate of the *ith* lag in regime *j*. Letting *G* denote the total number of candidate threshold vectors the estimated threshold vector is given by $\hat{\mathbf{\theta}} = \mathbf{\theta}_{\hat{g}}$ where $\hat{g} = \arg\min_{1 \le g \le G} \widetilde{\sigma}_g^2$, while $\hat{\mathbf{\Phi}}^{(j)} = \hat{\mathbf{\Phi}}_{\hat{g}}^{(j)} = \left(\mathbf{F0}_{\hat{g}}^{(j)'} \mathbf{F0}_{\hat{g}}^{(j)}\right)^{-1} \left(\mathbf{F0}_{\hat{g}}^{(j)'} \Delta\mathbf{z}\right)$.

Allowing $p < L$ requires the use of information criteria, such as the Akaike Information Criterion (*AIC*) measure of Tong (1983), to determine the appropriate lag-order in each regime[7]. Similarly, the normalized AIC (*N AIC*), as suggested by Tong and Lim (1980), is typically utilized to determine the value of the unknown threshold delay. Chan (1993) has shown that $\hat{\mathbf{\theta}}$ is asymptotically independent of $\hat{\mathbf{\Phi}}^{(j)}$, for all *j*, and that the latter estimates are normally distributed, as is the case when the threshold vector is known. This is a consequence of his illustration that, for arbitrary *L*, the conditional LS estimates of the parameters of a $TAR(p;d;k)$ model are strongly consistent.

If the regressor matrix may be written as the first-degree polynomial matrix $\mathbf{F}_g^{(j)} = \mathbf{F0}_g^{(j)} + \mathbf{U}_g^{(j)}$, where $\mathbf{U}_g^{(j)} = [c, c, \ldots, c]' \left[\theta_g^{(i)}, \theta_g^{(i)}, \ldots, \theta_g^{(i)}\right]$ for *c*, being a scalar, and $\theta_g^{(i)}$, being the threshold on which the AR process in regime *j* is explicitly dependent, then the continuous threshold parameter space will have to be searched to obtain the optimal threshold vector location. The task is relatively simple when the threshold vector is a scalar and may be computed efficiently by utilizing the fitting procedure proposed by Coakley *et al.* (2003). The fitting approach that they advocate involves subjecting an interpolated residual sum of squares function to the standard optimality condition. The support points required by the interpolating algorithm are computed via Givens transformations of QR factorizations of the (augmented) regressor matrix $\left[\mathbf{F0}_g^{(j)} \mid \Delta\mathbf{z}_g^{(j)}\right]$. When generalizing the fitting approach to allow for multiple thresholds in a TAR process, as we did in Van Tol (2005), the LS fitting problem becomes decidedly less trivial; involving the need for a multidimensional interpolation-based optimization routine defined over a set of non-overlapping threshold hypercubes.

3.2. Conditional Multivariate VECM Parameter Estimation

Given the *r* TAR model estimates, discussed in the previous section, parameter estimation of the multivariate threshold VECM is fairly straightforward. The estimation procedure results in sets of threshold vector estimates, here, denoted by $\hat{\mathbf{\theta}}_i = \left\lfloor \hat{\theta}_i^{(0)}, \hat{\theta}_i^{(1)}, \ldots, \hat{\theta}_i^{(k)} \right\rfloor$, for each of the *r* cointegrating residual series. Conditional on these estimates let $\mathbf{I}_M^{(j)}$ denote a $N^* \times r$ matrix of indicator vectors, for $j = 1,2,\ldots,k$, then denoting the *ith* element of the *tth* row vector by $\mathbf{I}_M^{(j)}(t,i)$ we have

$$\mathbf{I}_M^{(j)}(t,i) = \begin{cases} 1 & if \quad \hat{\theta}_i^{(j-1)} < v_{i,t-d} \le \hat{\theta}_i^{(j)} \\ 0 & otherwise \end{cases} \tag{10}$$

[6] Frequently augmented with a unit vector when including a constant term. For reasons of notational convenience we abstain from including it here.

[7] Alternatively, the more parsimonious Schwarz Bayesian criterion (SBC) or the Hannan-Quinn criterion (HQC) may be used (see Kapetanios (1999)).

Then $\mathbf{I}_M^{(j)} \circ \mathbf{z}'$, where $\mathbf{z}$ is a $r \times N^*$ matrix in which the *ith* element of each column vector is defined as in (A4)[8] for $t = 1,2,\ldots,N^*$ and $\circ$ denotes a Hadamard product[9].

Subsequently, if we denote by $\mathbf{Z}$ the horizontally concatenated matrix of regressor values i.e. $\left[1 \,|\, \mathbf{I}_M^{(1)} \circ \mathbf{z}_{i,t-1}^{(1)}{}' \,|\, \mathbf{I}_M^{(2)} \circ \mathbf{z}_{i,t-1}^{(2)}{}' \,|\, \mathbf{I}_M^{(3)} \circ \mathbf{z}_{i,t-1}^{(3)}{}' \,|\, \Delta\mathbf{x}_{t-1} \,|\cdots|\, \Delta\mathbf{x}_{t-p+1}\right]$ OLS estimates of the coefficients, $\mathbf{B}$, are obtained as $\hat{\mathbf{B}} = \mathbf{YZ}(\mathbf{Z'Z})^{-1}$ where $\mathbf{Y} = \left[\Delta\mathbf{x}_1, \Delta\mathbf{x}_2, \ldots, \Delta\mathbf{x}_{N^*}\right]$ and an estimate of the residual covariance matrix is given by $\Sigma_\eta = \mathbf{Y}\left(I_{N^*} - \mathbf{Z}(\mathbf{Z'Z})^{-1}\mathbf{Z}\right)\mathbf{Y}'$ with I_{N^*} an identity matrix of order N^*. Finally, Lütkepohl (1993) provides a detailed discussion on VECM order selection and diagnostic tests of model adequacy applicable to the case at hand. In the section that follows we introduce the empirical framework of Clarida and Taylor (1997) and detail the applicability of the underlying notions of model (5) to that context.

4. Forecasting Spot Exchange Rates in a Multivariate Threshold Cointegration Framework

4.1. The Clarida-Taylor Empirical Framework

Clarida and Taylor (1997) develop an empirical framework, drawing upon a similar framework employed by Hall *et al.* (1992) in a study of the term structure of treasury bill yields, in which a 5-variable system of spot and forward exchange rates are hypothesized to be propelled by a single common trend. In order to be able to compare results, we obtained the Harris Bank database, maintained by Richard Levich[10], which allows us to investigate discretely sampled, weekly data on spot and 4-, 13-, 26-, and 52-week forward dollar exchange rates for Germany, Japan, and the United Kingdom. Our data span the period 1977:1 through 1993:52 and are identical to those used in Clarida and Taylor (1997) and thus should not in principle be a source of any statistical discrepancy. All estimations are carried out using data for the period 1977:1 through 1990:26 while we reserve all complementary data for the purpose of evaluating forecasts.

Below we briefly delineate the Clarida and Taylor (1997) framework after which we introduce an alternative formulation of the short-run dynamics of the process.

Let

$$\mathbf{y}_t = \left[s_t, f_{h(1),t}, f_{h(2),t}, \ldots, f_{h(4),t}\right]' \quad (11)$$

where s_t and $f_{h(\tau),t}$ denote the logarithm of the spot and forward exchange rates at horizon $h(\tau)$, respectively. Assuming the empirically pervasive finding of a unit in the spot rate it is allowed to evolve according to

$$s_t = w_t + v_t \quad (12)$$

[8] See Appendix.

[9] If *A* and *B* are matrices of the same order, say $m \times n$, with elements a_{ij} and b_{ij}, respectively, then the $m \times n$ matrix $A \circ B$ defines a Hadamard product when its *ijth* element is $a_{ij} b_{ij}$.

[10] We thank Richard Levich for kindly making the data available.

where $v_t \sim I(0)$ and $w_t = \gamma + w_{t-1} + e_t$. Letting $E(s_{t+h(\tau)} \mid \Omega_t)$ denote the mathematical expectation of the spot rate at horizon $h(\tau)$, conditional on the information that is available at time *t*, Ω_t, then any deviation from the risk-neutral efficient markets hypothesis maybe defined as

$$\delta_{h(\tau),t} \equiv f_{h(\tau),t} - E(s_{t+h(\tau)} \mid \Omega_t) \tag{13}$$

Combining (12) and (13) results in the following expression of the forward rate at horizon $h(\tau)$

$$f_{h(\tau),t} = h(\tau)\gamma + w_t + E(v_{t+h(\tau)} \mid \Omega_t) + \delta_{h(\tau),t} \tag{14}$$

and subsequently to

$$f_{h(\tau),t} - s_t = h(\tau)\gamma + E(v_{t+h(\tau)} - v_t \mid \Omega_t) + \delta_{h(\tau),t} \tag{15}$$

which is the forward premium at horizon $h(\tau)$. Clearly, if the assumption in (13) holds each of the *j* forward premia are stationary implying the existence of a single common trend.

The authors subsequently assume that the equilibrium reversion process implicit in (13) is both symmetric and continuous while this need not necessarily be the case. It is conceivable that economic agents are unresponsive to random fluctuations in $\delta_{h(\tau),t}$, possibly for a prolonged period, yet have an economic incentive to act when these exceed some threshold level. Hence, the forward premia may be locally non-stationary while maintaining the global property of stationarity; in the sense of Balke and Fomby (1997).

Assuming that deviations from RNEMH are stationary and equilibrium reversion is guided by the following mechanism

$$\delta_{h(\tau),t} = \begin{cases} \alpha^{(3)} + \beta^{(3)}\left(\delta_{h(\tau),t-1} - \theta^{(2)}\right) + \varepsilon_t & if \quad \delta_{h(\tau),t-1} > \theta^{(2)} \\ \delta_{h(\tau),t-1} + \varepsilon_t & if \quad \theta^{(1)} < \delta_{h(\tau),t-1} \le \theta^{(2)} \\ \alpha^{(1)} + \beta^{(1)}\left(\delta_{h(\tau),t-1} - \theta^{(1)}\right) + \varepsilon_t & if \quad \delta_{h(\tau),t-1} \le \theta^{(1)} \end{cases} \tag{16}$$

which is a Band-TAR model the following model is implied for the forward premia

$$f_{h(\tau),t} - s_t = h(\tau)\gamma + \begin{cases} \alpha^{(3)} + \beta^{(3)}\left(\delta_{h(\tau),t-1} - \theta^{(2)}\right) + E(v_{t+h(\tau)} - v_t \mid \Omega_t) + \varepsilon_t & if \quad \delta_{h(\tau),t-1} > \theta^{(2)} \\ \delta_{h(\tau),t-1} + E(v_{t+h(\tau)} - v_t \mid \Omega_t) + \varepsilon_t & if \quad \theta^{(1)} < \delta_{h(\tau),t-1} \le \theta^{(2)} \\ \alpha^{(1)} + \beta^{(1)}\left(\delta_{h(\tau),t-1} - \theta^{(1)}\right) + E(v_{t+h(\tau)} - v_t \mid \Omega_t) + \varepsilon_t & if \quad \delta_{h(\tau),t-1} \le \theta^{(1)} \end{cases} \tag{17}$$

if we additionally assume that $\varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right)$ then if $\eta_t \equiv \varepsilon_t + E(v_{t+h(\tau)} - v_t \mid \Omega_t)$, $\eta_t \sim N(0, \sigma_\eta^2)$ where $0 < \sigma_\eta^2 < \infty$ we may simplify model (17) as

$$f_{h(\tau),t} - s_t = h(\tau)\gamma + \begin{cases} \alpha^{(3)} + \beta^{(3)}\left(\delta_{h(\tau),t-1} - \theta^{(2)}\right) + \eta_t & \textit{if} \quad \delta_{h(\tau),t-1} > \theta^{(2)} \\ \delta_{h(j),t-1} + \eta_t & \textit{if} \quad \theta^{(1)} < \delta_{h(\tau),t-1} \le \theta^{(2)} \\ \alpha^{(1)} + \beta^{(1)}\left(\delta_{h(\tau),t-1} - \theta^{(1)}\right) + \eta_t & \textit{if} \quad \delta_{h(\tau),t-1} \le \theta^{(l)} \end{cases} \tag{18}$$

Specifications (16) through (18) are stationary processes admitting a unit root when the process is in the inner-corridor and autoregressive equilibrium reversion when in an outer-regime. Interestingly, the thresholds of this model act as attractors hence we, implicitly, assume that the equilibrating process is towards the inner-band instead of towards a particular equilibrium level. Note that specification (18) allows for asymmetry in both the dispersion of the thresholds around the unconditional mean of the process, as well as, differing speeds of equilibrium reversion; the model is therefore more general than that originally proposed by Balke and Fomby (1997).

4.2. Nonlinearity and Stationarity of Forward Premiums

Postulating that each of the τ forward premiums possesses the properties of stationarity and nonlinearity requires an investigation into these properties. Since the testing strategy for linearity and a nonlinear unit root conventionally require the use of the supremum-Wald statistic[11] it is convenient to state each hypothesis separately and then detail the testing procedure. A general formulation of the linearity hypothesis may be stated as

$$H_0 : \mathbf{\Phi}^{(j_1)} = \mathbf{\Phi}^{(j_2)} \text{ vs. } H_1 : \mathbf{\Phi}^{(j_1)} \neq \mathbf{\Phi}^{(j_2)} \tag{19}$$

where j_1 and j_2 represent the indices of all possible subsets of contiguous regimes with $j_1, j_2 \in \{1,2,\dots,k\}$, while the unit root hypothesis may be stated as

$$H_0 : \gamma_1^{(j_1)} = \gamma_1^{(j_2)} = 0 \text{ vs. } H_1 : \gamma_1^{(j_1)} < 0 \quad \textit{and} \quad \gamma_1^{(j_2)} < 0 \tag{20}$$

where, *here*, j_1 and j_2 represent the indices of the outer-regimes with $j_1, j_2 \in \{1,2,\dots,k\}$ and $\gamma_1^{(j_r)} = \phi_1^{(j_r)} - 1$, given the general TAR formulation in (3).

If the threshold locations are known, testing hypotheses (19) and (20) is achieved by means of computing the Wald statistic

$$\boldsymbol{W} = \frac{N^*}{2}\left(\frac{\tilde{\sigma}^2 - \hat{\sigma}^2}{\hat{\sigma}^2}\right) \tag{21}$$

where $\tilde{\sigma}^2$ is the estimated residual variance under the null and $\hat{\sigma}^2$ is computed as is done in (9), conditional on the known threshold vector. If $\boldsymbol{\theta}$ is unknown it is appropriate to utilize the supremum-Wald test statistic, given the estimates $\hat{d}$ and $\hat{p}$, which may be formulated as

$$\boldsymbol{W}^{\sup}\left(\hat{d}, \hat{p}\right) = \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \boldsymbol{W}\left(\boldsymbol{\theta}, \hat{d}, \hat{p}\right) \tag{22}$$

[11] Andrews and Ploberger (1994) also suggest the use of the average Wald and average exponential Wald tests. We refrain from utilizing these tests here due to the arbitrariness of the weightings implicit in these tests.

where

$$W(\boldsymbol{\theta}, \hat{d}, \hat{p}) = \frac{N^*}{2}\left(\frac{\tilde{\sigma}_n^2(\hat{p}) - \hat{\sigma}^2(\boldsymbol{\theta}, \hat{d}, \hat{p})}{\hat{\sigma}^2(\boldsymbol{\theta}, \hat{d}, \hat{p})}\right) \quad (23)$$

and $\hat{\sigma}^2(\boldsymbol{\theta}, \hat{d}, \hat{p})$ is the residual variance given some candidate threshold vector; defined such that the elements are equal to an order statistic of delay variable for models like (6). As alluded to previously, estimating the F-statistic for model (7) is more complex as it requires that each non-overlapping threshold hypercube[12] be searched by subjecting the RSS function, which is rational and continuous there within[13], to the standard optimality condition. Since the threshold parameter is unidentified under the null the asymptotic distribution of the sup-Wald test statistic is non-standard, consequently when testing (19), significance levels are computed using 200 replications of Hansen's (1996, 1999) fixed-regressor bootstrap. Significance of the unit root hypothesis is determined by comparing the value computed in (22) with the critical values computed by various authors given different stationary (SE)TAR model specifications under the alternative hypothesis.

Tables I and II present the results of the linearity and unit root testing procedures, respectively. The results in Table I present a very clear image of the type of nonlinearity present in each of the forward premiums; being unable to reject the linearity hypothesis when testing for double-threshold nonlinearity while being able to do so for all but one premium series when testing for single-threshold nonlinearity. Hence, we may conclude, with the exception of the 12 month dollar-yen forward premium, that each premium is decidedly nonlinear and that this linearity is of the single-threshold type. Note that these findings indicate the absence of the inner-regime in model (18) yet does not violate our basic premise that the forward premiums are nonlinear in nature.

Table II presents the results of a battery of unit root tests; including the augmented Dickey-Fuller, Enders-Granger (specifying stationary two-regime SETAR and MTAR models under the alternative), Berben-van Dijk (specifying a two-regime CTAR under the alternative), Kapetanios-Shin (specifying a stationary three-regime SETAR model under the alternative) and the lag-augmented double-threshold momentum-TAR models that we proposed as alternative data-generating mechanisms in the unit root tests that we developed in the previous chapter[14,15]. The results generally favor the conclusion that each of the forward premium series are stationary processes. Additionally, note that the forward premiums tend to become less stationary as the horizon increases.

Given the evidence in Table I that statistically discernable single-threshold nonlinearity is present in each of the series it is appropriate that in judging whether the processes are stationary most weight is applied to the single-threshold tests. There, only the 12 month dollar-yen forward premium is unable to cause rejection of the unit root hypothesis; since there is insufficient evidence of nonlinearity the appropriate test is the ADF test, this test indicates an inability to reject the unit root hypothesis[16].

Given the strong evidence of single-threshold nonlinearity in the forward premiums Balke and Fomby (1997) suggest a two-step approach for examining threshold cointegration

[12] This construct is extensively detailed in Van Tol (2005).
[13] The proof of this statement can be found in the appendix of Coakley *et al.* (2003).
[14] For details on the relative strengths of each of these tests we refer readers to that chapter.
[15] For critical values and additional details with respect to these unit root tests we refer readers to Dickey and Fuller (1979); Enders and Granger (1997); Berben and van Dijk (1999); and Kapetanios and Shin (2002).
[16] The implications of this finding in the out-of-sample forecasting exercises is very strong, generating forecast errors that, in magnitude, far exceed those of a simple (driftless) random walk.

utilizing standard time-series methods, which they argue work "reasonably well". Pippenger and Goering (1993, 2000) have, however, poignantly illustrated the detrimental effect that the presence of thresholds, in the equilibrium reversion process, has on both linear unit root and cointegration tests. For the bivariate case Seo and Hansen (2002) and Seo (2003) propose algorithms for the determination of maximum likelihood estimates of a complete threshold cointegration model, including the cointegrating vector. Unfortunately, no test exists for the multivariate case and as such we assume a known matrix of cointegrating vectors. Implicitly, the linearity and unit root tests have assumed the existence of the following cointegration matrix

$$\boldsymbol{\beta} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix} \tag{24}$$

Clarida and Taylor (1997) are unable to reject the set of economically motivated restrictions implied by (24); based on a likelihood ratio test of the hypothesis that exactly four linearly independent forward premiums comprise the basis of the cointegration space[17]. Hence, we are in position whereby we assume support for the empirical framework set forth by the authors, by means of assumption (24), and have shown overwhelming evidence that the cointegrating residuals, or forward premiums, are well defined as single-threshold SETAR models and are stationary; with the exception of a single forward premium series. Resultantly, we are in a position to estimate the multivariate TVECM these findings imply.

4.4. Multivariate Threshold Vector Error Correction Model Results

Since we have found overwhelming support for the hypothesis that equilibrium reversion in each series of forward premiums is asymmetric we desire to incorporate these findings into a model of the form represented by (5). Tables IV, V and VI report model estimates[18] when a TAR model of the following form is estimated on the forward premiums in the first-step (Threshold estimates and Box-Ljung statistics are reported at the top of each table)

$$\left(s_t - f_{h(j)}\right)_t = \begin{cases} \phi_1^{(2)}\left(\left(s_t - f_{h(j)}\right)_{t-1} - \theta\right) + \varepsilon_t & if \quad f_{h(j),t-1} > \theta \\ \phi_1^{(1)}\left(\left(s_t - f_{h(j)}\right)_{t-1} - \theta\right) + \varepsilon_t & if \quad f_{h(j),t-1} \le \theta \end{cases} \tag{25}$$

here the threshold functions as the attractor. Also, we have restricted the lag-order and threshold delay variables to 1, for reasons of simplicity. Subsequently, conditional on the

[17] Our reassessment of their findings in Van Tol (2005) does however indicate that these findings are not as strong as they claim.

[18] Note that we report the full set of parameter estimates. Firstly, imposing the restriction that the coefficient matrices on all matrices other than the first-order matrix did not result in a significant change in the maximized likelihood; sequential application of LR tests having led to this conclusion. Secondly, we found that removing all insignificant coefficient estimates led to models that appeared rather arbitrary while repetitive removal of the "least" significant regressor tended to destroy the model in it's entirety; this left the model for the UK without any right-side variables, for example – implying that the model degenerated into a random walk model – despite the appeal of this development we find it uninteresting to allow this to occur and hence continue to leave all regressors in the model. Note that Clarida and Taylor (1997) found most regressors to be significant at the 5% level.

estimates of the thresholds for each forward premium series we construct the indicator matrix of (10) and estimate a model of the form

$$\Delta \mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\alpha}^{(1)}\mathbf{I}^{(1)} \circ \left(s - f_{h(j)}\right)_{t-1}^{(1)} + \boldsymbol{\alpha}^{(2)}\mathbf{I}^{(2)} \circ \left(s - f_{h(j)}\right)_{t-1}^{(2)} + \boldsymbol{\Psi}_1 \Delta \mathbf{y}_{t-1} + \boldsymbol{\eta}_t \tag{26}$$

There are two reasons for including just a single lagged difference: the first relates to maintaining enough degrees of freedom for accurate model parameter estimation while the second relates to the fact that we intend to juxtapose our findings with that of Clarida and Taylor (1997) who, due to the instability of the maximum-likelihood algorithm that they employ, estimate a "linear" VECM with only a single lagged-difference; the authors cite their forecasting results as evidence that a single lag is adequate.

Table III presents estimation results for the dollar-sterling system. The first interesting finding is that the likelihood ratio test of the linearity restriction is unable to reject the null. This is curious as the linearity tests of Table I strongly rejected that the forward premiums are linear AR processes, further evidence being provided by the Q-statistics of the SETAR estimates at the top of the table. One possible explanation could be that the nonlinearity in the forward premiums is mitigated by the inclusion of lagged differences. There is, otherwise, no clearly discernable pattern in the estimated coefficients. The linearity hypothesis is firmly rejected for the dollar-mark and dollar-yen systems in Tables IV and V, where, additionally, the different model adequacy measures indicate that the estimated models are well specified.

4.5. Out-of-Sample Forecasting Results

Utilizing models (25) and (26) we generate dynamic forecasts of the spot exchange rate at 1-, 3-, 6-, and 12-month horizons over the period 1990:27 through 1993:52. In assessing the out-of-sample forecasting performance of these models we compute similar forecasts using a "linear" VECM, including a single lagged-difference as a regressor; an unrestricted forth order VAR model; a driftless random walk model; a forward premium regression; and a standard forward rate model. While the RMSE and MAE of the multivariate threshold VECM is presented in levels that of the alternative forecasting models are presented as a ratio. The ratio is computed by dividing the RMSE/MAE of model (26) by that of the alternative model. Hence, values below unity indicate the relative superiority of the approach that we advocate.

The forecasting results, which are presented in Tables VI, VII and VIII, show that in two of the three systems model (26) out-predicts the "linear" VECM by up to approximately 9%. The proposed model is unable to systematically out-predict the (driftless) random walk model; despite producing lower forecast errors at the 1-month horizon for the dollar-yen system the model tends to veer off strongly at longer forecast horizons, having a ratio of 36 and 38 at the 52-week horizon. Hence, we find strong evidence that not only is the linear cointegration framework adopted by Clarida and Taylor (1997) severely misspecified, due to the clear presence of a single-threshold boundary in the forward premiums, but also that their claims of out-predicting the random walk model at various horizons is overstated and possibly a spurious result despite the elegance of their empirical framework.

Interestingly, when inspecting graphs of the forward premium series, illustrated in Figure 1, we find that the forward premiums of the dollar-mark and dollar-yen systems exhibit behavior during the out-of-sample period that significantly deviates from that of the estimation period. These observations make the findings of Clarida and Taylor (1997) all the more surprising.

5. Conclusion

In this paper we formulated a convenient specification of a multivariate threshold vector error correction model. The model allows each series of cointegrating residuals to be governed by a nonlinear equilibrium reverting process while obeying long-run equilibrium constraints.

We applied these concepts and constructs to the task of forecasting spot exchange rates utilizing the empirical framework of Clarida and Taylor (1997) in which they show that, under the assumption of stationary deviations from the risk-neutral efficient markets hypothesis, the forward premiums comprise a basis for the cointegration space of a system consisting of spot and forward foreign exchange rates. We were able to establish that the forward premiums exhibit a strong tendency towards equilibrium reversion and that this reversion is asymmetric in nature. Combining this result with the findings of Pippenger and Goering (2000), who show that standard cointegration tests lack power when the disequilibrium process displays asymmetric reversion, forms a strong argument against the overly-supportive claims of Clarida and Taylor (1997). Out-of-sample forecasting exercises illustrated the relative superiority of the multivariate threshold VECM over the linear VECM, yet, despite this improvement in forecasting performance, we remain incapable of systematically out-predicting a simple (driftless) random walk model. Suggestions for future research include developing procedures capable of incorporating asymmetric equilibrium reversion of the cointegrating residuals when estimating the matrix of cointegrating vectors.

Appendix

To further clarify the formulation of model (5) we briefly digress to discuss an example. Suppose a trivariate cointegrated system, propelled by a single common trend, with each time-series of cointegrating residuals being governed by a double-threshold Band-TAR model. A Band-TAR model describes the following time-series process

$$z_t = \begin{cases} \sum_{i=1}^{p} \phi_i^{(1)}\left(z_{t-i} - \theta^{(1)}\right) + \varepsilon_t & if \quad z_{t-d} < \theta^{(1)} \\ \mu + \sum_{i=1}^{p} \phi_i^{(2)} z_{t-i} + \varepsilon_t & if \quad \theta^{(1)} \le z_{t-d} \le \theta^{(2)} \\ \sum_{i=1}^{p} \phi_i^{(3)}\left(z_{t-i} - \theta^{(2)}\right) + \varepsilon_t & if \quad z_{t-d} > \theta^{(2)} \end{cases} \tag{A1}$$

or, alternatively, using the indicator function $I^{(j)}(A)$, which is equal to unity when event A occurs and is zero otherwise, (A1) may be written as

$$\begin{aligned} z_t = {} & I^{(1)}\left(z_{t-d} < \theta^{(1)}\right)\left[\sum_{i=1}^{p} \phi_i^{(1)}\left(z_{t-i} - \theta^{(1)}\right)\right] + I^{(2)}\left(\theta^{(1)} \le z_{t-d} \le \theta^{(2)}\right)\left[\mu + \sum_{i=1}^{p} \phi_i^{(2)} z_{t-i}\right] + \\ & I^{(3)}\left(z_{t-d} > \theta^{(2)}\right)\left[\sum_{i=1}^{p} \phi_i^{(3)}\left(z_{t-i} - \theta^{(2)}\right)\right] + \varepsilon_t \end{aligned} \tag{A2}$$

Hence, defining

$$\mathbf{I} = \left[I^{(1)}, I^{(2)}, I^{(3)}\right] \tag{A3}$$

as a $(1 \times k)$ indicator vector, where $I^{(j)} = I^{(j)}(A)$, we can reformulate (A2) compactly as

$$z_t = \mathbf{I}\mathbf{f}\left(z_{t-}\right) \tag{A4}$$

where

$$\mathbf{f}\left(z_{t-}\right) = \begin{bmatrix} \sum_{i=1}^{p} \phi_i^{(1)}\left(z_{t-i} - \theta^{(1)}\right) + \varepsilon_t \\ \mu + \sum_{i=1}^{p} \phi_i^{(2)} z_{t-i} + \varepsilon_t \\ \sum_{i=1}^{p} \phi_i^{(3)}\left(z_{t-i} - \theta^{(2)}\right) + \varepsilon_t \end{bmatrix} \tag{A5}$$

is a vector allowing for the incorporation of the different functional forms present in each of the *k*-regimes of the model. Denoting the *ith* disequilibrium process $z_{i,t}$, for $i = 1,2$, given the presence of a single common trend (see Stock and Watson (1988)), we can simply extend the formulation in (A4) to allow for multiple cointegrating error series. Let $\mathbf{I}_i$ denote vector (A3) and $\mathbf{f}_i\left(z_{t-}\right)$ denote vector (A5) for the *ith* disequilibrium term, then

$$\mathbf{z}_t = \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_1\mathbf{f}_1\left(z_{t-}\right) \\ \mathbf{I}_2\mathbf{f}_2\left(z_{t-}\right) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \mathbf{f}_1\left(z_{t-}\right) \\ \mathbf{f}_2\left(z_{t-}\right) \end{bmatrix}$$

and since $\begin{bmatrix} \mathbf{I}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix} = \begin{bmatrix} I_1^{(1)} & I_1^{(2)} & I_1^{(3)} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_2^{(1)} & I_2^{(2)} & I_2^{(3)} \end{bmatrix}$ it is useful to decompose the matrix according to which regime $z_{i,t}$ is in i.e.

$$\begin{bmatrix} \mathbf{I}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \mathbf{f}_1(z_{t-}) \\ \mathbf{f}_2(z_{t-}) \end{bmatrix} = \begin{bmatrix} I_1^{(1)} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_2^{(1)} & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{f}_1(z_{t-}) \\ \mathbf{f}_2(z_{t-}) \end{bmatrix} + \begin{bmatrix} 0 & I_1^{(2)} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_2^{(2)} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{f}_1(z_{t-}) \\ \mathbf{f}_2(z_{t-}) \end{bmatrix} + \begin{bmatrix} 0 & 0 & I_1^{(3)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_2^{(3)} \end{bmatrix} \begin{bmatrix} \mathbf{f}_1(z_{t-}) \\ \mathbf{f}_2(z_{t-}) \end{bmatrix} \tag{A6}$$

Hence, denoting the matrix corresponding with regime *j* in (A6) by $\mathbf{I}^{(j)}$ and $\mathbf{z}_t = \begin{bmatrix} \mathbf{f}_1(z_{t-}) \\ \mathbf{f}_2(z_{t-}) \end{bmatrix}$ we have that (A6) may be conveniently formulated as $\sum_{j=1}^{k} \mathbf{I}^{(j)} \mathbf{z}_t$ and the error correction term $\boldsymbol{\beta} \mathbf{x}_{t-1}$, in (2), may be substituted with $\sum_{j=1}^{k} \mathbf{I}^{(j)} \mathbf{z}_{t-1}$ leading to the formulation in model (5).

Table I
Linearity Tests of the Dollar-Sterling, Dollar-Mark and Dollar-Yen Forward Premiums.

*indicates rejection of the linearity hypothesis. p-values are computed using Hansen's (1996, 1999) fixed-regressor bootstrap with 200 replications.

Horizon	Single-Threshold non-linearity		Double-Threshold non-linearity	
	Supremum-Wald	Significance level	Supremum-Wald	Significance Level
Linearity Tests of the Forward Premia: Dollar-Sterling				
1-month	6.904	1.000*	80.931	0.000
3-month	1.355	0.635*	3.324	0.140*
6-month	0.835	0.335*	9.006	0.000
12-month	0.910	0.335*	6.190	0.010
Linearity Tests of the Forward Premia: Dollar-Mark				
1-month	15.712	1.000*	0.000	0.000
3-month	3.469	0.880*	59.000	0.000
6-month	3.373	0.815*	3.648	0.000
12-month	2.930	0.765*	8.258	0.020
Linearity Tests of the Forward Premia: Dollar-Yen				
1-month	3.293	0.850*	7.824	0.000
3-month	0.832	0.095*	1.495	0.000
6-month	0.955	0.085*	3.737	0.000
12-month	0.402	0.000	1.437	0.000

Table II
Linear and Nonlinear Unit Root Tests of the Dollar-Sterling, Dollar-Mark and Dollar-Yen Forward Premia

* indicates rejection of the unit root hypothesis at the 5% significance level. Reported Values for the ADF test are pseudo t-statistics while the nonlinear unit root tests specifying single- and double threshold models under the alternative are supremum-Wald statistics. In all cases the tests were performed on "demeaned" series. The ADF, Enders-Granger and van Tol-Wolff tests allow a lag-augmented data generating mechanism with *L* set to 3. The van Tol-Wolff tests additionally have a maximum threshold delay value of 3. The trimming parameters and the minimal fraction of observations in each regime are set to 15%.

	Linear	Single-Threshold Models			Double-Threshold Models		
Horizon	ADF (AR)	Enders-Granger (SETAR)	Enders-Granger (MTAR)	Berben – van Dijk (CTAR)	Kapetanios-Shin (SETAR)	Van Tol-Wolff (EQ-MTAR)	Van Tol-Wolff (Band-MTAR)
Unit Root Tests of the Forward Premia: Dollar-Sterling							
1-month	-5.390*	21.257*	28.677*	56.441*	114.194*	96.886*	98.162*
3-month	-4.060*	12.479*	14.329*	9.628*	15.684*	16.253*	16.549*
6-month	-3.363*	6.242*	12.112*	6.555*	13.648*	16.607*	17.359*
12-month	-3.409*	6.806*	7.953*	7.024*	11.264*	9.237	10.198*
Unit Root Tests of the Forward Premia: Dollar-Mark							
1-month	-4.987*	16.146*	15.464*	28.704*	30.478*	21.955*	32.683*
3-month	-4.098*	31.288*	22.175*	101.600*	97.816*	82.073*	87.282*
6-month	-3.060*	6.873*	11.399*	6.983*	8.226	4.103	8.446
12-month	-3.005*	7.896*	14.033*	7.731*	9.203*	4.374	12.487*
Unit Root Tests of the Forward Premia: Dollar-Yen							
1-month	-3.720*	15.771*	11.395*	11.132*	8.883	6.779	10.391*
3-month	-3.051*	7.056*	6.139*	6.376*	3.892	3.276	3.231
6-month	-2.713	4.895*	9.168*	4.855*	4.124	6.078	9.996*
12-month	-2.405	3.882	6.780*	4.291	4.023	3.541	11.293*

Table III
FIML Error Correction Model for the Five-Variable System: Dollar-Sterling

Sample period is 1977:1 to 1990:26. The Q-statistics are Box-Ljung statistics computed at 13 autocorrelations of the residual series; H is Hosking's (1980) multivariate portmanteau statistic computed at 13 autocorrelations; All distributions are distributed as central χ^2 under the null hypothesis, with the degrees of freedom indicated. Figures in parentheses are marginal significance levels.

SETAR Model Parameter Estimates of the Forward Premiums

	$s_t - f_{4,t}$	$s_t - f_{13,t}$	$s_t - f_{26,t}$	$s_t - f_{52,t}$
$\hat{\theta}$	-0.0016	-0.0035	0.0122	0.0257
$Q(13)$	154.830 (0.000)	27.902 (0.009)	44.950 (0.000)	24.028 (0.031)

Multivariate Threshold Vector Error Correction Model Estimates

Explanatory Variables	Model for Δs_t Coeff.	t	Model for $\Delta f_{4,t}$ Coeff.	t	Model for $\Delta f_{13,t}$ Coeff.	t	Model for $\Delta f_{26,t}$ Coeff.	t	Model for $\Delta f_{52,t}$ Coeff.	t
Δs_{t-1}	-0.312	-0.533	0.415	0.709	-0.287	-0.487	-0.540	-0.902	-0.702	-1.133
$\Delta f_{4,t-1}$	0.548	0.850	-0.523	-0.811	0.733	1.128	0.784	1.189	0.728	1.066
$\Delta f_{13,t-1}$	0.462	0.619	0.644	0.863	-0.164	-0.218	0.517	0.677	0.416	0.527
$\Delta f_{26,t-1}$	-1.295	-1.828	-1.120	-1.582	-0.787	-1.104	-1.524	-2.105	-0.816	-1.089
$\Delta f_{52,t-1}$	0.579	1.737	0.562	1.686	0.483	1.438	0.741	2.175	0.355	1.006
Lower Regime $(s - f_4)_{t-1}$	-0.970	-0.622	-0.380	-0.244	-1.256	-0.799	-1.485	-0.930	-1.480	-0.896
Lower Regime $(s - f_{13})_{t-1}$	0.099	0.105	-0.103	-0.109	0.601	0.633	0.112	0.116	0.382	0.383
Lower Regime $(s - f_{26})_{t-1}$	0.172	0.203	0.145	0.171	-0.044	-0.052	0.556	0.640	0.243	0.271
Lower Regime $(s - f_{52})_{t-1}$	0.015	0.053	0.024	0.085	0.023	0.082	-0.130	-0.456	-0.016	-0.056
Upper Regime $(s - f_4)_{t-1}$	0.779	0.531	1.682	1.147	0.878	0.594	0.905	0.604	0.894	0.576
Upper Regime $(s - f_{13})_{t-1}$	-0.187	-0.182	-0.494	-0.480	0.120	0.116	-0.525	-0.499	-0.328	-0.302
Upper Regime $(s - f_{26})_{t-1}$	0.553	0.588	0.661	0.704	0.680	0.718	1.428	1.487	1.301	1.309
Upper Regime $(s - f_{52})_{t-1}$	-0.339	-0.900	-0.390	-1.035	-0.491	-1.293	-0.714	-1.853	-0.683	-1.714
Constant	0.000	0.056	0.000	-0.007	-0.003	-0.417	0.002	0.275	0.001	0.067
	$Q(13) = 9.63$ (0.72)		$Q(13) = 9.62$ (0.73)		$Q(13) = 10.35$ (0.67)		$Q(13) = 8.90$ (0.78)		$Q(13) = 9.146$ (0.762)	
			$H(325) = 533.07$ (0.00)				$\chi^2 = 16.96$ (1.00)			

Table IV
FIML Error Correction Model for the Five-Variable System: Dollar-Mark

Sample period is 1977:1 to 1990:26. The Q-statistics are Box-Ljung statistics computed at 13 autocorrelations of the residual series; H is Hosking's (1980) multivariate portmanteau statistic computed at 13 autocorrelations; All distributions are distributed as central χ^2 under the null hypothesis, with the degrees of freedom indicated. Figures in parentheses are marginal significance levels.

SETAR Model Parameter Estimates of the Forward Premiums

	$s_t - f_{4,t}$	$s_t - f_{13,t}$	$s_t - f_{26,t}$	$s_t - f_{52,t}$
$\hat{\theta}$	-0.0048	-0.0146	-0.0272	-0.0475
$Q(13)$	143.590 (0.000)	97.150 (0.000)	30.930 (0.003)	46.678 (0.000)

Multivariate Threshold Vector Error Correction Model Estimates

		Model for Δs_t		Model for $\Delta f_{4,t}$		Model for $\Delta f_{13,t}$		Model for $\Delta f_{26,t}$		Model for $\Delta f_{52,t}$	
	Explanatory Variables	Coeff.	t	Coeff.	t	Coeff.	t	Coeff.	t	Coeff.	t
	Δs_{t-1}	-0.796	-0.829	-0.416	-0.434	-1.241	-1.303	-0.546	-0.572	-0.708	-0.735
	$\Delta f_{4,t-1}$	0.993	1.013	0.542	0.553	2.021	2.078	0.757	0.777	0.719	0.730
	$\Delta f_{13,t-1}$	-0.148	-0.364	-0.103	-0.254	-1.055	-2.616	-0.040	-0.098	-0.023	-0.057
	$\Delta f_{26,t-1}$	0.479	0.832	0.470	0.818	0.802	1.405	-0.004	-0.008	0.482	0.834
	$\Delta f_{52,t-1}$	-0.510	-1.333	-0.472	-1.237	-0.516	-1.360	-0.159	-0.418	-0.457	-1.190
Lower Regime	$(s - f_4)_{t-1}$	1.427	0.658	1.993	0.921	2.341	1.088	1.344	0.624	1.458	0.670
Lower Regime	$(s - f_{13})_{t-1}$	2.676	1.873	2.522	1.767	3.334	2.352	2.044	1.440	1.985	1.385
Lower Regime	$(s - f_{26})_{t-1}$	-1.345	-1.433	-1.214	-1.296	-1.83	-1.961	-0.724	-0.777	-0.670	-0.711
Lower Regime	$(s - f_{52})_{t-1}$	0.188	0.566	0.144	0.434	0.266	0.806	0.036	0.109	0.068	0.205
Upper Regime	$(s - f_4)_{t-1}$	1.267	0.749	1.557	0.922	0.563	0.335	1.105	0.657	1.159	0.683
Upper Regime	$(s - f_{13})_{t-1}$	-1.436	-1.447	-1.476	-1.489	-0.486	-0.494	-1.293	-1.312	-1.142	-1.147
Upper Regime	$(s - f_{26})_{t-1}$	0.600	0.865	0.537	0.776	0.024	0.035	0.627	0.910	0.339	0.487
Upper Regime	$(s - f_{52})_{t-1}$	0.040	0.162	0.601	0.243	0.160	0.651	0.010	0.039	0.125	0.501
		0.000	0.171	0.000	0.204	0.000	0.167	0.000	-0.412	0.000	0.183
	Constant	$Q(13) = 13.75$		$Q(13) = 13.95$		$Q(13) = 13.21$		$Q(13) = 13.50$		$Q(13) = 13.39$	
				(0.38)		(0.43)		(0.41)			
		(0.39)		$H(325) = 457.92$ (0.00)				$\chi^2 = 88.57$ (0.00)		(0.418)	

Table V
FIML Error Correction Model for the Five-Variable System: Dollar-Yen

Sample period is 1977:1 to 1990:26. The Q-statistics are Box-Ljung statistics computed at 13 autocorrelations of the residual series; H is Hosking's (1980) multivariate portmanteau statistic computed at 13 autocorrelations; All distributions are distributed as central χ^2 under the null hypothesis, with the degrees of freedom indicated. Figures in parentheses are marginal significance levels.

SETAR Model Parameter Estimates of the Forward Premiums

	$s_t - f_{4,t}$	$s_t - f_{13,t}$	$s_t - f_{26,t}$	$s_t - f_{52,t}$
$\hat{\theta}$	-0.0014	-0.0167	-0.0322	-0.0490
$Q(13)$	35.678 (0.001)	13.289 (0.426)	23.085 (0.041)	18.070 (0.155)

Multivariate Threshold Vector Error Correction Model Estimates

Explanatory Variables	Model for Δs_t Coeff.	t	Model for $\Delta f_{4,t}$ Coeff.	t	Model for $\Delta f_{13,t}$ Coeff.	t	Model for $\Delta f_{26,t}$ Coeff.	t	Model for $\Delta f_{52,t}$ Coeff.	t
Δs_{t-1}	-1.021	-1.158	-0.632	-0.716	-0.921	-1.046	-0.974	-1.103	-0.600	-0.669
$\Delta f_{4,t-1}$	2.507	2.088	1.938	1.613	2.589	2.161	2.334	1.940	1.942	1.590
$\Delta f_{13,t-1}$	-1.340	-1.403	-1.237	-1.294	-1.849	-1.940	-1.217	-1.271	-1.746	-1.797
$\Delta f_{26,t-1}$	0.461	0.647	0.552	0.774	0.781	1.098	0.361	0.505	1.172	1.615
$\Delta f_{52,t-1}$	-0.545	-1.813	-0.556	-1.846	-0.540	-1.780	-0.447	-1.485	-0.716	-2.339
Lower Regime $(s - f_4)_{t-1}$	-0.094	-0.074	0.366	0.288	-0.300	-0.237	-0.234	-0.184	0.112	0.087
Lower Regime $(s - f_{13})_{t-1}$	0.998	0.682	0.695	0.475	1.237	0.847	0.556	0.379	0.721	0.485
Lower Regime $(s - f_{26})_{t-1}$	-1.175	-1.206	-1.112	-1.142	-1.287	-1.325	-0.810	-0.831	-1.224	-1.236
Lower Regime $(s - f_{52})_{t-1}$	0.297	1.378	0.319	1.478	0.336	1.562	0.277	1.280	0.434	1.979
Upper Regime $(s - f_4)_{t-1}$	-1.287	-0.718	-0.613	-0.342	-1.155	-0.645	-0.898	-0.499	-0.970	-0.532
Upper Regime $(s - f_{13})_{t-1}$	1.266	1.110	1.187	1.041	1.764	1.550	1,307	1.144	1.371	1.182
Upper Regime $(s - f_{26})_{t-1}$	-0.491	-0.617	-0.571	-0.717	-0.788	-0.992	-0.437	-0.548	-0.775	-0.958
Upper Regime $(s - f_{52})_{t-1}$	0.031	0.142	0.048	0.220	0.068	0.316	-0.004	-0.017	0.143	0.649
Constant	-0.002	-0.277	0.001	0.083	-0.002	-0.350	-0.003	-0.407	0.001	0.059
	$Q(13) = 12.48$ (0.49)		$Q(13) = 12.77$ (0.47)		$Q(13) = 12.90$ (0.46)		$Q(13) = 12.77$ (0.466)		$Q(13) = 14.45$ (0.34)	
			$H(325) = 403.00$ (0.00)				$\chi^2 = 45.12$ (0.267)			

Table VI
Results of the Forecasting Exercises: Dollar-Sterling

Notes: Forecast period is 1990:27 to 1993:52. For the VECM the RMSE and the MAE is expressed in levels. For the alternative forecasts, the RMSE or the MAE is expressed as the inverse of its ratio to the corresponding figure for the VECM. Thus a figure less than 1 indicates the relative superior performance by the VECM.

	TVECM (level)	VECM (ratio)	VAR (ratio)	Random Walk (ratio)	Forward Premium Regression (ratio)	Forward Rate (ratio)
RMSE						
4-week horizon	0.0402	1.007	1.007	1.036	0.997	1.039
13-week horizon	0.0860	0.994	0.966	1.143	0.997	1.170
26-week horizon	0.1280	1.012	0.965	1.245	0.940	1.323
52-week horizon	0.4903	2.593	0.431	4.222	2.367	4.947
MAE						
4-week horizon	0.0298	1.015	1.016	1.032	1.011	1.033
13-week horizon	0.0637	1.005	0.970	1.161	1.016	1.181
26-week horizon	0.0911	1.020	0.913	1.295	0.946	1.371
52-week horizon	0.1875	1.156	0.247	2.189	1.009	2.536

Table VII
Results of the Forecasting Exercises: Dollar-Mark

Notes: Forecast period is 1990:27 to 1993:52. For the VECM the RMSE and the MAE is expressed in levels. For the alternative forecasts, the RMSE or the MAE is expressed as the inverse of its ratio to the corresponding figure for the VECM. Thus a figure less than 1 indicates the relative superior performance by the VECM.

	TVECM (level)	VECM (ratio)	VAR (ratio)	Random Walk (ratio)	Forward Premium Regression (ratio)	Forward Rate (ratio)
RMSE						
4-week horizon	0.0350	0.998	1.008	1.026	0.975	1.027
13-week horizon	0.0780	0.990	1.015	1.183	0.971	1.194
26-week horizon	0.1115	0.974	1.043	1.331	0.917	1.357
52-week horizon	0.3051	0.950	1.877	3.732	1.867	3.921
MAE						
4-week horizon	0.0274	0.998	1.007	1.016	0.981	1.017
13-week horizon	0.0625	0.995	1.022	1.178	0.991	1.191
26-week horizon	0.0856	0.979	1.042	1.372	0.902	1.426
52-week horizon	0.2303	0.911	1.670	3.987	1.586	4.111

Table VIII
Results of the Forecasting Exercises: Dollar-Yen

Notes: Forecast period is 1990:27 to 1993:52. For the VECM the RMSE and the MAE is expressed in levels. For the alternative forecasts, the RMSE or the MAE is expressed as the inverse of its ratio to the corresponding figure for the VECM. Thus a figure less than 1 indicates the relative superior performance by the VECM.

	TVECM (level)	VECM (ratio)	VAR (ratio)	Random Walk (ratio)	Forward Premium Regression (ratio)	Forward Rate (ratio)
RMSE						
4-week horizon	0.0268	0.9934	0.9946	0.9738	0.9739	0.9743
13-week horizon	0.0599	0.9758	1.0704	1.1024	1.0438	1.1036
26-week horizon	0.1255	0.9879	1.7635	1.8125	1.5189	1.7941
52-week horizon	3.156	0.9447	12.711	36.652	26.300	35.189
MAE						
4-week horizon	0.0209	0.9920	0.9928	0.9869	0.9677	0.9882
13-week horizon	0.0483	0.9738	1.0598	1.1415	1.0569	1.1375
26-week horizon	0.0969	0.9738	1.6784	1.8763	1.3920	1.8365
52-week horizon	2.3698	0.9420	12.316	38.710	23.612	36.916

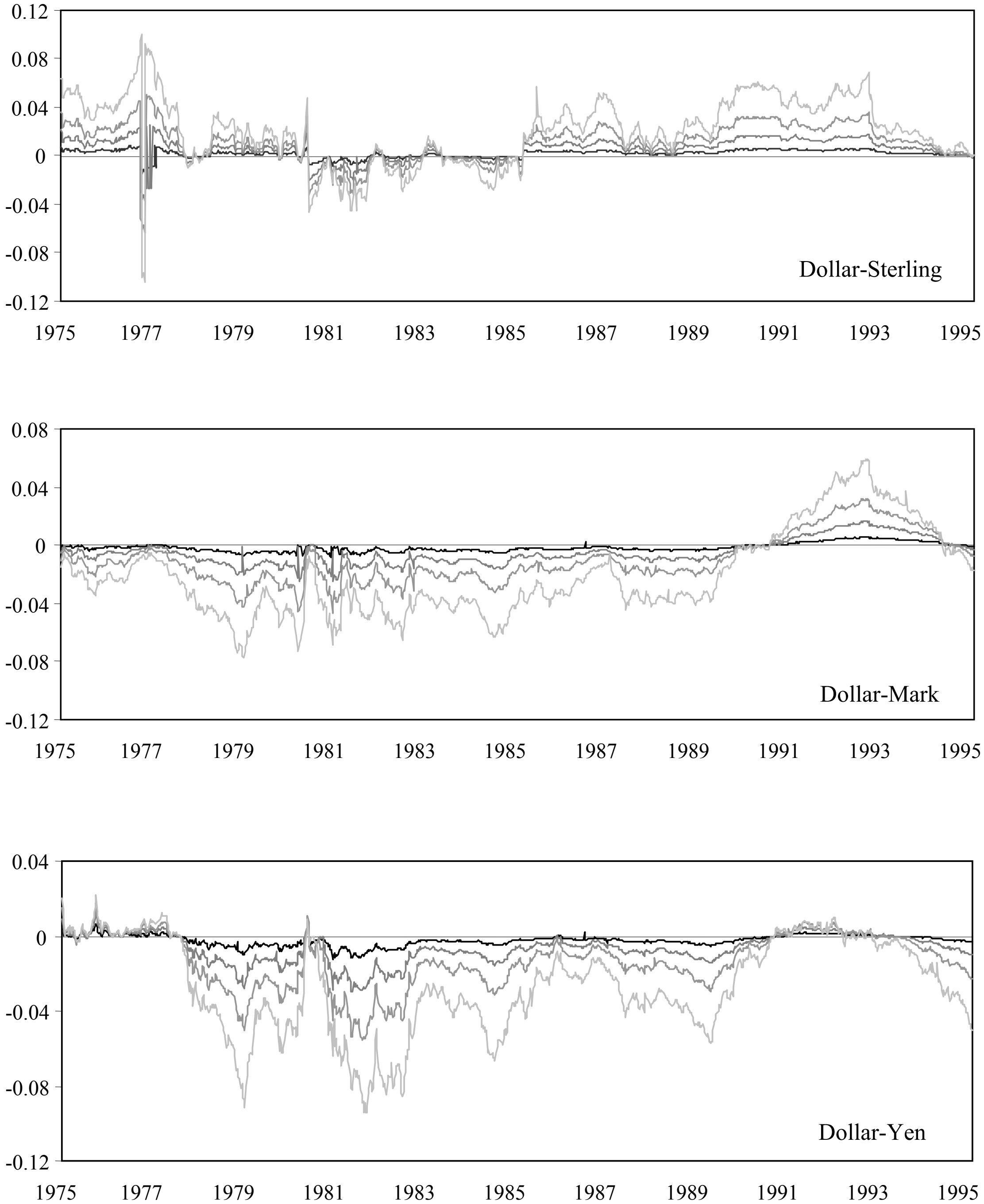

Figure 1. Dollar-Sterling, Dollar-Mark and Dollar-Yen 1-, 3-, 6- and 12-month Forward premia. The dark solid line is the 1-month forward premium while the 3-, 6- and 12-month forward premiums are represented by increasingly less dark solid lines.

References

Andrews, D.W.K. and W. Ploberger, 1994, Optimal Tests when a Nuisance Parameter is Present only under the Alternative, *Econometrica* **62**, 1383-1414.

Balke, N.S. and T.B. Fomby, 1997, Threshold Cointegration, *International Economic Review* **38**, 627-646.

Berben, R. and D. van Dijk, 1999, Unit Root Tests and Asymmetric Adjustment: A Reassessment, *Econometric Institute Research Report El-9902/A.*

Chan, K.S., 1993, Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model, *The Annals of Statistics* **21**, no.1, 520-533.

Chan, K.S. and R.S. Tsay, 1998, Limiting Properties of the Least Squares Estimator of a Continuous Threshold Autoregressive Model, *Biometrika*, **85**, 413-426.

Clarida, R.H., and M.P. Taylor, 1997, The Term Structure of Forward Exchange Premiums and the Forecastability of Spot Exchange Rates: Correcting the Errors, *The Review of Economics and Statistics* **3**, 353-362.

Coakley, J., A.M. Fuertes and M.T. Perez, 2003, Numerical Issues in Threshold Autoregressive Modeling of Time Series, *Journal of Economic Dynamics and Control* **27**, 2219-2242.

Hall, A., H. Anderson, and C.W.J. Granger, 1992, A Cointegration Analysis of Treasury Bill Yields, *The Review of Economics and Statistics* **74**, no. 1, 116-126.

Hansen, B.E., 1996, Inference when a Nuisance Parameter is not Identified under the Null Hypothesis, *Econometrica* **64**, 413-430.

Hansen, B.E., 1999, Testing for Linearity, *Journal of Economic Surveys* **13**, No. 5, 551-576.

Hansen, B.E. and B. Seo, 2002, Testing for Two-Regime Threshold Cointegration in Vector Error Correction Models, Journal of Econometrics, *Journal of Econometrics* **110**, 293-318.

Hosking, J.R.M., 1980, The Multivariate Portmanteau Statistic, *Journal of the American Statistical Association* **75**, 602-608.

Hylleberg, S. and G.E. Mizon, 1989, Cointegration and Error Correction Mechanisms, *Economic Journal* **99**, 113-125.

Kapetanios, G., 1999, Model Selection in Threshold Models, *Working Paper,* Department of Applied Economics, University of Cambridge.

Kapetanios, G. and Y. Shin, 2002, Unit Root Tests in Three-Regime SETAR Models, *Working Paper.*

Lo, M.C., and E. Zivot, 2001, Threshold Cointegration and Nonlinear Adjustment to the Law of One Price, *Macroeconomic Dynamics* **5**, 533-576.

Lütkepohl, H., 1993, Introduction to Multiple Time Series Analysis, 2nd edition, Springer-Verlag.

Meese, R.A. and K. Rogoff, 1983a, Empirical Exchange Rates of the Seventies, *Journal of International Economics*, **14**, 3-24.

Meese, R.A. and K. Rogoff, 1983b, The Out-of-Sample Failure of Empirical Exchange Rate models, in J. Frenkel, ed Exchange Rates and International Macroeconomics, University of Chicago Press, Chicago.

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.

Seo, M., 2003a, Nonlinear Mean Reversion in the Term Structure of Interest Rates, *Journal of Economic Dynamics and Control* **27**, 2243-2265.

Seo, M., 2003b, Bootstrap Testing for the Presence of Threshold Cointegration in a Threshold Vector Error Correction Model, *Unpublished Working Paper.*

Slutzky, E., 1927, The Summation of Random Causes as the Source of Cyclic Processes, *Econometrica* **5**, 105-146.

Stock, J., and M. Watson, 1988, Testing for Common Trends, *Journal of the American Statistical Association* **83**.

Tjøstheim, D., 1990, Non-linear Time Series and Markov Chains, *Advances in Applied Probability* **22**, 587-611.

Tong, H., 1983, Threshold Models in Non-linear Time Series Analysis," *Lecture Notes in Statistics* **21**, Springer-Verlag.

Tong, H., 1990, Non-linear time series: A dynamical system approach, Oxford Statistical Science series 6, Clarendon Press Oxford 1990.

Tsay, R.S., 1989, Testing and Modeling Threshold Autoregressive Processes, *Journal of the American Statistical Association* **84**, 231-240.

Tsay, R.A., Testing and Modeling Multivariate Threshold Models," *Unpublished Working Paper*.

Van Tol, M.R., 2005, Beyond the Threshold: Theoretical and Empirical Nonlinear Time-Series Econometrics of Foreign Exchange Markets, *Doctoral Dissertation*, Maastricht University.

Van Tol, M.R., and C.C.P. Wolff, 2004, Unit Root Testing in the Presence of Double-Threshold Nonlinearity, *Working Paper*, Maastricht University.

Wolff, C.C.P., 1987, Time-Varying Parameters and the Out-of-Sample Forecasting Performance of Structural Exchange Rate Models, *Journal of Business and Economic Statistics* **5**, 87-97.

Wolff, C.C.P., 1988, Exchange Rates, Innovations and Forecasting, *Journal of International Money and Finance* **7**, 49-61.