Testing uncovered interest rate parity and term structure using three-regime threshold unit root VECM: an application to the Swiss “Isle” of interest rates

KRISHNAKUMAR, Jaya, NETO, David


DOI : 10.1111/j.1468-0084.2011.00644.x
Testing Uncovered Interest Rate Parity and Term Structure Using a Three-regime Threshold Unit Root VECM: An Application to the Swiss ‘Isle’ of Interest Rates*

JAYA KRISHNAKUMAR and DAVID NETO

Department of Econometrics, University of Geneva 40, Bd. du Pont d’Arve, CH-1211, Geneva 4, Switzerland (e-mails: jaya.krishnakumar@unige.ch; david.neto@unige.ch)

Abstract

In this article, a three-regime multivariate threshold vector error correction model with a ‘band of inaction’ is formulated to examine uncovered interest rate parity (UIRP) and expectation hypothesis of the term structure (EHTS) of interest rates for Switzerland. Combining both UIRP and EHTS in a model that allows for nonlinearities, we investigate whether the Swiss advantage is disappearing with respect to Europe. Our results favour threshold cointegration and show that both hypotheses hold, at least in one of the three regimes of the process for Switzerland/Germany. The same is not true between Switzerland and the United States.

I. Introduction

Since the Bretton Woods agreement, Swiss interest rates have been lower than Organization for Economic Co-operation and Development countries, and in particular its European neighbours. Would this advantage disappear if one day Switzerland decides to join the European Union? Several factors account for this advantage. The main ones are no doubt the bank secrecy principle (which explains why some investors accept a lower return), and the credibility of the Swiss Central Bank in controlling inflation (which has made investments in Switzerland quite safe). However, the growing macroeconomic stability in Europe (including European Monetary Union and bordering countries) seems to produce a standardization of interest rates, with the result that the Swiss advantage in terms of interest rates, that is, the ‘Swiss isle of interest rates’, may begin to disappear with respect to Europe. Indeed, in a floating exchange rate regime, the uncovered interest rate parity (UIRP) should hold if the exchange rate is stable. While the Swiss franc appreciated against the US dollar and against the UK pound sterling during these last few years, it seems to stabilize against the Euro currency. Considering these stylized facts, it would be interesting

*We are grateful to two anonymous referees and the associate editor for many valuable comments. This work was supported by a grant from the Swiss National Fund for Scientific Research.

JEL classification numbers: C12, C32, E43.
to test if the Swiss ‘isle’ still holds towards its European neighbours or only vis-à-vis the rest of the world.

The area of international finance is particularly rich in such relationships that have important implications for policy design and evaluation. Purchasing power parity (PPP), expectation hypothesis of term structure (EHTS) of interest rates or the aforementioned UIRP, are examples of ‘laws’ that a policy maker would be interested in knowing whether they are verified or not in practice. These hypotheses are also important in macroeconomic theory. Many models, like for instance the Dornbusch’s (1976) overshooting model, which explains exchange rate dynamics, have UIRP as an assumption. This hypothesis also appears as a long-term equilibrium in the Mundell–Fleming model. However, many empirical findings often fail to provide evidence of UIRP. Indeed, for many economies, the spot exchange rate seems to be negatively correlated with the lagged forward premium contrary to the expected result under the UIRP condition (Engel, 1996). This result is commonly referred to as the ‘forward premium anomaly’. Several arguments have been put forth to explain this anomaly: time-varying risk premia, peso problem, and nonlinearities because of transaction costs, limits to speculation or central bank interventions. The latter one – presence of nonlinearities – has recently been investigated by Baillie and Kılıç (2006) through an asymmetric and nonlinear relation between exchange rate change and lagged forward premium. Through a logistic smooth transition regression (LSTR) model, the authors find what while one regime is consistent with the anomaly, another one seems to exhibit a higher probability that the parity holds. These nonlinearities can also lead to a rejection of the EHTS.

Enders and Siklos (2001) propose the idea of asymmetric cointegration as an extension of the threshold cointegration concept put forward earlier by Balke and Fomby (1997). They also propose a methodology to test no-cointegration against threshold cointegration but do not derive any asymptotic results. They investigate the term structure of interest rates through particular variants of the threshold autoregressive model (SETAR and momentum TAR) and find threshold effects in the error correction dynamics. The theoretical gap is filled by Hansen and Seo (2002) for a threshold vector error correction model (TVECM) involving one cointegrating relationship and they apply their no-threshold cointegration test to the term structure of interest rates finding evidence of threshold effect.

More recently, Bekaert, Wei and Xing (2007) carry out a joint test of UIRP and EHTS and find evidence against the latter but inconclusive for the former. UIRP is seen to be currency dependent but not horizon dependent. Brüggemann and Lütkepohl (2005) find evidence for UIRP and EHTS for the United States and the Euro zone. To verify if the Swiss ‘isle’ holds, we focus on the simultaneous verification of EHTS and UIRP using a cointegrated system. As Brüggemann and Lütkepohl (2005) and Bekaert et al. (2007), we find it interesting to examine both hypotheses simultaneously since if UIRP holds at a short horizon, it should also hold at a long horizon as long as EHTS holds. Our aim therefore is to test these two hypotheses in a cointegrated system, which allows the disequilibria to be nonlinear. The easiest modelling route is probably a TVECM as many theoretical results (on estimation and inference) are available in the literature. Gonzalo and Pitarakis (2006) provide some stochastic properties of TVECM, in particular conditions for stability of such processes. Seo (2006) develops a test for linear cointegration with the equilibrium error specified as a SETAR process and examines the law of one price when threshold
dynamics is introduced. Kapetanios and Shin (2006) investigate a test similar to that of Seo (2006) for a three-regime SETAR model where the corridor regime follows a random walk, already used by Taylor (2001) in a PPP investigation framework. Note that two pitfalls are combined when one wishes to test for a unit root in a threshold model where the transition variable is the lagged dependent variable itself. In addition to the well-known issue of the non-identification of the threshold parameter under the null hypothesis (see Hansen, 1996), the transition variable is also non-stationary under the null. Thus, although the model of Caner and Hansen (2001) and Gonzalo and Pitarakis (2006) covers a large set of processes, it does not include this particular case considered in Seo (2006) and in Kapetanios and Shin (2006). However, as discussed in Caner and Hansen (2001), one could always take the lagged difference of the threshold variable to ensure its stationarity. Enders and Siklos (2001) note that this choice could often be suitable in economic models.

Except for Gonzalo and Pitarakis (2006), the aforementioned studies on TVECM assume a single cointegrating relationship. This underlying assumption simplifies both the estimation procedure and inference. Although determining the number of cointegrating relationships for a set of integrated variables and estimating them has been an important area of research in standard cointegration theory (one can cite Phillips and Durlauf, 1986; Johansen, 1988; Phillips, 1991; among others), there are few developments in the threshold cointegrated context. Gonzalo and Pitarakis (2006) propose a methodology to test for cointegrating rank in a TVECM by directly estimating the unknown ranks of the coefficient matrices using a model selection approach introduced by Gonzalo and Pitarakis (1998, 2002). In a recent paper, Krishnakumar and Neto (2009) consider estimation and inference procedures for a TVECM with more than one cointegrating relation and develop a test for the cointegrating rank.

In this article, we apply this test for testing EHTS along with UIRP between Switzerland and Germany, the biggest European neighbour of Switzerland as well as between Switzerland and the United States. For this purpose, we formulate a cointegrated system in which the errors (disequilibria) are assumed to follow a three-regime multivariate threshold model comprising a unit-root middle regime corresponding to an ‘inaction band’ following the terminology of Taylor (2001), and no unit roots in the two regimes outside of the band.

Our article is organized as follows. In section II, we give a formal specification of the two economic laws to be tested. These laws give rise to three possible relationships with four interest rates. In section III, we go on to present the general formulation of the TVECM. We successively consider the tests of no-cointegration and of the number of cointegrating relations (reduced rank tests). Section IV presents the empirical results. The article ends with some concluding remarks in section V.

II. The economic relations

EHTS of interest rates

Let \( R_t^{(n)} \) be the continuously compounded \( n \)-period interest rate and let \( R_t^{(n_0)} \) be a shorter-term interest rate, say \( n_0 \)-period interest rate. The EHTS of interest rates leads to the following relation:
where \( n_0/n \) is an integer, \( E_t \) denotes the conditional expectation on the information available at time \( t \) and \( a_t^{(n_0,n)} \) is a premium, which could vary over time with the maturities \( n \) and \( n_0 \). This premium may incorporate risk considerations or investors’ liquidity preferences. Campbell and Shiller (1987) show that if the short-term interest rate is \( I(1) \), then EHTS implies that the long-term rate will also be \( I(1) \) and the spread will be stationary, that is, if \( R_t^{(n)} \sim I(1) \) and \( R_t^{(n_0)} \sim I(1) \), then \( S_t^{(n_0,n)} = R_t^{(n)} - R_t^{(n_0)} \sim I(0) \).

We can see this statement by rearranging equation (1) as:

\[
R_t^{(n)} - R_t^{(n_0)} = \frac{n_0}{n} \sum_{j=0}^{n_0-1} \sum_{i=1}^{n-j} E_t R^{(n_0)}_{t+j} + a_t^{(n_0,n)},
\]

where \( \Delta R_t^{(n)} = R_t^{(n)} - R_t^{(n_0)} \). The right side of this equation is stationary if \( R_t^{(n)} \) is \( I(1) \) and \( a_t^{(n_0,n)} \) is stationary or constant. Thus there is a cointegration relationship between these two rates \( (R_t^{(n)}, R_t^{(n_0)})' \), with the cointegration vector given by \( (1, -1)' \). Hall, Anderson and Granger (1992) extend this argument to a set of \( N \) maturities: \( n_0 \leq n_1 \leq \cdots \leq n_j \leq \cdots \leq n_{N-1} \). They state that if the short-term rate \( R_t^{(n_0)} \) is \( I(1) \), then all the others will also be \( I(1) \), and their spreads with the short-term rate should all be \( I(0) \). In other words, there will be \((N-1)\) cointegrating vectors with the parameter values summing up to 0. The cointegration space is generated by the rows of the \((N-1 \times N)\) matrix as shown:

\[
\begin{bmatrix}
1 & -1 \\
1 & -1 \\
\vdots & \ddots \\
1 & -1
\end{bmatrix},
\]

where the coefficient 1 corresponds to \( R_t^{(n_0)} \) and blank elements are zeros.

**Uncovered interest rate parity**

The UIRP hypothesis at the \( n_0 \)-period horizon states that

\[
\frac{1}{n_0} (E_t e_{12,t+n_0} - e_{12,t}) = R_t^{(n_0)} - R_t^{(n_1)} + b_t^{(n_0)},
\]

where \( R_t^{(n_0)} \) and \( R_t^{(n_1)} \) are time-\( t \) continuously compounded \( n_0 \)-period interest rates of countries 1 and 2, respectively, \( e_{12} \) is the logarithm of the spot exchange rate denominated in the currency of country 1 and \( b_t^{(n_0)} \) a stationary risk premium.

Following Bekaert et al. (2007), the UIRP at a short horizon \( n_0 \) and the EHTS at a long horizon \( n \) imply UIRP at the long horizon \( n \). Let us write down the EHTS at the long horizon \( n \) for both countries:

\[\text{Note that in this general case, there are } N - 1 \text{ premia } a_t^{(n_0,n)} \text{ for } j = 0, \ldots, N - 1.\]
Combining the above two equations with the short-term UIRP [equation (2)], we get the following relation:

\[ R_{1,t}^{(n)} - R_{2,t}^{(n)} = \frac{n_0}{n} \sum_{i=0}^{n/n_0-1} E_i (R_{1,t+n_0i}^{(n)} - R_{2,t+n_0i}^{(n)}) + a_{1,t}^{(n_0,n)} - a_{2,t}^{(n_0,n)} + 1 \sum_{r=1}^{2N-1} \Delta \epsilon_{12,t+n_0r} + a_{1,t}^{(n_0,n)} - a_{2,t}^{(n_0,n)} + b_r^{(n_0,n)}. \]  

(4)

Few works investigate this tight connection between EHTS and UIRP hypothesis. Only Bekaert et al. (2007) and Wolters (2002) have recently pointed it out. The implication of equation (4) is that if the exchange rate changes are weakly exogenous, then \( R_{1,t}^{(n)} - R_{2,t}^{(n)} \) will be I(0). Thus, for a system with \( N \) maturities per country [equations (1) and (2)], there should be at most \( 2N-1 \) cointegration relations according to UIRP and EHTS. In this article, we will test the aforesaid relations for two countries and two maturities (short- and long-term).

III. The econometric framework

The model

The following general notation is used. We denote the projection matrix associated with a \((p \times r)\) matrix \( A \) as \( P_j \), and \( M_j = I - P_j \), the orthogonal complement of \( A \) as \( A_\perp \) which is a \((p \times p-r)\) matrix of full column rank such that \( A'A_\perp = 0 \). We denote as \( t_T \) the \( T \)-vector of ones and we will use \( I(1) \) and \( I(0) \) to represent time series that are integrated of orders 1 and 0, respectively.

Let us consider the \( p \)-dimensional finite-order VAR model for an \( I(1) \) vector \( X_t \):

\[ A(L)X_t = \epsilon_t, \quad t = 1, 2, \ldots, T, \]

where \( A(L) = I_p - A_1 L - \cdots - A_K L^K \), the first \( K \) data points \( X_1, X_2, \ldots, X_0 \) are fixed, and where \( \epsilon_t \) are i.i.d. \( N(0, \Omega) \). The error correction form of the previous VAR is given by:

\[ \Delta X_t = \Pi \Delta X_{t-1} + A_1 \Delta X_{t-2} + \cdots + A_{K-1} \Delta X_{t-K+1} + \epsilon_t. \]

The hypothesis of cointegration implies a matrix \( \Pi \) of reduced rank such that \( \Pi = \Gamma A' \), where \( \Gamma \) and \( A \) are \((p \times r)\) matrices with \( r < p \), in particular we set \( A' = [I_r - \beta] \) where \( \beta \) is the \((r \times p)\) matrix which stacks the cointegrated vectors. The \( r \) equilibrium errors are denoted as follows:

\[ z_{r,t} = A'X_t, \]

(5)

whereas the \( p-r \) elements of \( X_t \) are merely:

\[ \Delta X_{p-r,t} = z_{p-r,t}. \]

This is the triangular representation widely used in the empirical economic literature (see Philips, 1991; Saikkonen, 1992).
Three-regime threshold unit root VECM

While the traditional cointegration theory of Engle and Granger (1987) assumes linearity in adjustment to the long-run equilibrium, many economic situations are characterized by nonlinear adjustments that might be caused by transaction costs or other institutional market interventions. A simple and convenient way to capture such features is to use a threshold model for the equilibrium error processes (many threshold models are available and have been used in the literature, see e.g. Balke and Fomby, 1997; Enders and Siklos, 2001; Taylor, 2001; among others). Therefore, we formulate a stationary multivariate threshold model for the error process, similar to the one specified by Tsay (1998) with three regimes:

\[
z_{r,t} = \begin{cases} 
\phi_1 z_{r,t-1} + \epsilon_t & \text{if } s_{t-d} \in (-\infty, \theta_1] = A_1, \\
\phi_2 z_{r,t-1} + \epsilon_t & \text{if } s_{t-d} \in (\theta_1, \theta_2] = A_2, \text{ with } \theta_1 = -\theta_2, \\
\phi_3 z_{r,t-1} + \epsilon_t & \text{if } s_{t-d} \in [\theta_2, +\infty) = A_3.
\end{cases}
\]

This model assumes that the threshold variable \( s_{t-d} \) is a strictly stationary and ergodic univariate sequence, whose distribution \( F \) is continuous everywhere and symmetric. The delay \( d \) is strictly positive to ensure the predeterminedness of the process. If \( d = 0 \) then the variable \( s \) has to be assumed exogenous. \( \theta_1 \in \Theta \), where \( \Theta \) is a closed and bounded subset of the sample space of the variable \( s_{t-d} \). Moreover, we assume the error term \( \epsilon_t \) to be an \( r \)-vector of i.i.d sequences with zero mean, constant variance and finite \( 2\delta \) moments for some \( \delta > 2 \).

These last assumptions are standard in the literature, which deals with threshold models and they are important for the asymptotic theory. They have been widely discussed in Caner and Hansen (2001) and Gonzalo and Pitarakis (2006). However, because we deal with three regimes and not with two as in the aforementioned references, we added an extra assumption on the distribution of \( s_{t-d} \), which is its symmetry. Without being really restrictive, this assumption facilitates the derivation and tabulation of the no-cointegration test statistic given in section III. In particular, it allows us to restrict the number of nuisance parameters to one in the limiting distribution of the test statistic by writing \( F(\theta_1) = 1 - F(\theta_2) \).

Our framework excludes cross-effects between the \( r \) disequilibria including such effects only in the short-term dynamics of the VECM. This means that we assume that the matrices \( \phi_{j}, j = \{1, 2, 3\} \), are diagonal. This assumption is necessary for stating the stationarity conditions for \( z_{r,t} \). Indeed, since \( z_{r,t} \) must be \( I(0) \) under cointegration, if the matrices \( \phi_j \) are diagonal then the conditions given by Chan et al. (1985) for a univariate process apply individually on each element of \( z_{r,t} \); denoting the \( k \)th diagonal element of \( \phi_j \) as \( \phi_{jk} \) for \( k = \{1, \ldots, r\} \), we should have \( |\phi_{1k}| < 1, |\phi_{2k}| < 1 \) and \( \phi_{1k}\phi_{2k} < 1 \). In what follows we let \( \phi_2 = I_r \), such that the second regime is a ‘band of inaction’, or formally speaking there is no-cointegration inside this band. Such models are particularly relevant for describing exchange rate dynamics, PPP and law of one price (Taylor, 2001; Kapetanios and Shin, 2006). Therefore, \( z_t \) can be viewed as a threshold unit root (TUR) process introduced by González and Gonzalo (1998). Caner and Hansen (2001) investigate inference in such threshold processes with a unit root calling them partial unit root processes. Following the terminology of Caner and Hansen (2001), we can say that our variables are ‘partially cointegrated’ when the disequilibrium of the long-run relationship follows a TUR process. Let us rewrite the model (6) with \( \phi_2 = I_r \) as:

---

2To our knowledge, the conditions given by Chan et al. (1985) for a univariate TAR(1) process are the only ones available to date and it is beyond the scope of this study to examine a generalization to a multivariate setting.
\[(\Phi_1(L)z_{r,t})\|_{[A_i]} + (1 - L)\|_{[A_i]} + (\Phi_3(L)z_{r,t})\|_{[A_i]} = \epsilon_t, \quad (7)\]

where \(\Phi_j(L) = I_r - \phi_j L\), for \(j = \{1, 3\}\).

We now present the TVECM in the form of a proposition (with proof in appendix A).

**Proposition 1.** The TVECM associated with the system described by [equations (5) and (7)], can be written as:

\[\Delta X_t = \Pi^{(1)} X_{t-1} \|_{[A_i]} + \Pi^{(3)} X_{t-1} \|_{[A_i]} + \Lambda(L) \Delta X_{t-1} + \epsilon_t, \quad (8)\]

where \(\Pi^{(j)} = \Gamma^{(j)} A'\),

\[\Gamma^{(j)} = \left(\begin{array}{c} -\Phi_j(1) \\ 0 \end{array}\right), \quad \text{for} \quad j = \{1, 3\}, \]

\[\Lambda(L) = \begin{pmatrix} 0 & \beta g(L) \\ 0 & g(L) \end{pmatrix}, \]

with \(g(L) = \sum_{i=1}^{q} g_i L_i^{-1}\), and where \(g_i\) are \((p - r \times p - r)\) coefficient matrices.

**Proof.** See appendix A.

The presence of unit roots in the threshold model (6) for the disequilibria induces some consequences on the ranks of \(\Pi^{(j)}\) matrices in the TVECM. In our model, we have the limiting case where the \(r\) disequilibria are random walks in the middle regime, that is, \(\phi_2 = I_r\), such that the second regime is a ‘band of inaction’ (see Taylor, 2001). Therefore, all the loading parameters of the TVECM are 0 in this middle band. However, if only some elements in the diagonal of \(\phi_2\) were equal to 1, then the rank of \(\Pi^{(2)}\) would be less than \(r\).

As an example, consider a four-dimensional integrated system with \(r = 3\) cointegration relationships, that is, \(X_t = (X_{1,t}, X_{2,t}, X_{3,t}, X_{4,t})', A' = [I_3, -\beta]\), where \(\beta\) is \((3 \times 1)\), \(\Delta X_{4,t} = z_{4,t}\) with \(z_{4,t}\) a (finite autocorrelated) stationary process and where the vector \(z_{r,t} = (z_{1,t}, z_{2,t}, z_{3,t})'\) is described by a trivariate threshold process as follows:

\[\Delta z_{1,t} = \rho_{11} z_{1,t-1} \|_{[A_i]} + \rho_{13} z_{3,t-1} \|_{[A_i]} + \epsilon_{1,t}, \]
\[\Delta z_{2,t} = \rho_{21} z_{1,t-1} \|_{[A_i]} + \rho_{23} z_{3,t-1} \|_{[A_i]} + \epsilon_{2,t}, \]
\[\Delta z_{3,t} = \rho_{31} z_{1,t-1} \|_{[A_i]} + \rho_{33} z_{3,t-1} \|_{[A_i]} + \epsilon_{3,t}, \]

with \(\rho_{jk} = \phi_{jk} - 1\) and stationarity conditions given by \(\rho_{1k} < 0\), \(\rho_{3k} < 0\) and \((\rho_{3k} + 1) < 1\) for \(k = \{1, 2, 3\}\).

Now, if we assume \(\rho_{33} = 0\) (i.e. \(\phi_{33} = 1\)), then \(\Pi^{(3)} = \Gamma^{(3)} A'\) will look like this:

\[\Pi^{(3)} = \begin{pmatrix} \gamma_{11}^{(3)} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \gamma_{33}^{(3)} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \gamma_{11}^{(3)} & 0 & 0 & \gamma_{11}^{(3)} \beta_{11} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma_{33}^{(3)} & \gamma_{33}^{(3)} \beta_{31} \end{pmatrix}, \]

such that \(rk(\Pi^{(3)}) = 2\) whereas \(rk(\Pi^{(1)}) = 3\). Therefore, we say that there is one ‘not-activated’ relation of cointegration in the third regime.
Another remark must be made concerning the short-term coefficients $\Lambda_t$ in equation (8). Here they do not switch with the regime. This is because the threshold model specified for the disequilibrium error is autoregressive of order 1 [cf. equation (6)]. In general, one lag seems to adequately describe the disequilibrium dynamics, with the coefficient estimates of higher lags turning out to be insignificant. Now, it is straightforward to show that when the order of autoregression in equation (6) is greater than 1, then the threshold effect also concerns the short-run parameters.

There is a final practical consideration that is worth mentioning here as we will be dealing with interest rates in our empirical work. It is generally agreed that interest rates do not remain because of the presence of a premium in spite of rises in capital movements over the last decades. Including a drift, say a $p$-vector $\mu$, in model (8) is compatible with two schemes that are discussed in more detail in Krishnakumar and Neto (2009). If $\Gamma^{(j)} \mu \neq 0$, then the intercept will be both in the error correction part and as a constant in the cointegration relationships (denoted as $\beta_0$). The TVECM will then be written as:

$$
\Delta X_t = P(t^{(1)} \mu_{1}^{(1)} A_1) + P(t^{(3)} \mu_{3}^{(3)} A_3) + \Gamma^{(1)} (\beta_0 + A' X_{t-1}) \|_{(A_1)} \\
+ \Gamma^{(3)} (\beta_0 + A' X_{t-1}) \|_{(A_3)} + \Lambda(L) \Delta X_{t-1} + \epsilon_t,
$$

Whereas, if $\Gamma^{(j)} \mu = 0$, for $j = 1$ and 3, then the TVECM will be:

$$
\Delta X_t = \Gamma^{(1)} \beta_0 \|_{(A_1)} + \Gamma^{(3)} \beta_0 \|_{(A_3)} + A' X_{t-1} \|_{(A_1)} + A' X_{t-1} \|_{(A_3)} + \Lambda(L) \Delta X_{t-1} + \epsilon_t,
$$

where $\beta_0 = C \beta_0$, with a suitable matrix $C$ (see Krishnakumar and Neto, 2009). Our empirical model will therefore be a more general specification that is, model (8) with an intercept $\mu$ such that $\mu = \mu^{(1)} \|_{(A_1)} + \mu^{(3)} \|_{(A_3)}$.

**Estimation**

Since the threshold does not affect the coefficients of the lagged variations in our TVECM, under the normality of $\epsilon_t$, we can concentrate out the short-term parameters $\Lambda_t$ to simplify the notation. Let us denote $\Psi (\epsilon_t)$ as $\Omega$. This ‘concentration’ proceeds as follows: regressing $\Delta X_t$ and $X_{t-1}$ on $\Delta X_{t-1}, \ldots, \Delta X_{t-q-1}$, and using the residuals of the two previous regressions, denoted as $\Delta X_t^*$ and $X_{t-1}^*$, to focus only on the error correction part, we can write the TVECM [equation (8)] as:

$$
\Delta X_t^* = D_i(\theta_1) + (X_{t-1}^* \otimes \Gamma(\theta_1)) \text{vec}(B) + \epsilon_t^*, 
$$

where $D_i(\theta_1) = \mu^{(1)} \|_{(A_1)} + \mu^{(3)} \|_{(A_3)}$, $B = \left( \begin{array}{cc} A' & 0 \\ 0 & A' \end{array} \right)$, $\Gamma(\theta_1) = (\Gamma^{(1)} \Gamma^{(3)})$, $X_{t-1}^* = (X_{t-1}^* \|_{(A_1)}$, $X_{t-1}^* \|_{(A_3)})$.

Let us write the long-run parameters as $\text{vec}(B) = H \text{vec}(\beta) + h$, where $H$ and $h$ are respectively a $(4rp \times r(p - r))$ selection matrix and a $4rp$-dimensional vector. The vector $h$ selects the ones in $A$. The log-likelihood function is given by:
When the threshold is unknown, it has to be searched over a grid on $[\tau, 1 - \tau]$. In practice, this grid is given by the values taken by the selected transition variable. However, the choice of $\tau$ is somewhat arbitrary; empirical investigations take $\tau$ to be 5%, 10% or 15%. In general, it is chosen to ensure that there are enough observations in each

For a given threshold

Step 1. Matrices $\Pi^{(1)}$ and $\Pi^{(3)}$ can be consistently estimated from a regression of $\Delta X_t^*$ on $X^*_{t-1}$ from $\Delta X_t^* = D_t(\theta_1) + \Pi^{(1)} X^*_{t-1} + \Pi^{(3)} X^*_{t-1} + \varepsilon_t^*$. The triangular normalisation implies that the first $r$ columns of $\Pi^{(1)}$ and $\Pi^{(3)}$ are equal to $\Gamma^{(1)}$ and $\Gamma^{(3)}$, respectively. Thus, one can use these columns of $\hat{\Pi}^{(1)}$ and $\hat{\Pi}^{(3)}$ as consistent estimates of the loading factors $\Gamma^{(1)}$ and $\Gamma^{(3)}$.

Step 2. Using $\hat{\mu}^{(1)}$, $\hat{\mu}^{(3)}$, $\hat{\Gamma}^{(1)}$, $\hat{\Gamma}^{(3)}$ and $\hat{\Omega}(\theta_1) = T^{-1} \sum_t \hat{\varepsilon}_t^* \hat{\varepsilon}_t^{*'}$ from step 1, we obtain a feasible estimator of $\beta$ by concentrated maximum likelihood.

Step 3. Using $\hat{\beta}$, we can re-estimate $\Gamma$, $\mu^{(1)}$ and $\mu^{(3)}$ by running regression (9): $\Delta X_t^* = D_t(\theta_1) + \Gamma(\theta_1) B X^*_{t-1} + \varepsilon_t^*$.

For an unknown threshold

When the threshold is unknown, it has to be searched over a grid on $[\tau, 1 - \tau]$. In practice, this grid is given by the values taken by the selected transition variable. However, the choice of $\tau$ is somewhat arbitrary; empirical investigations take $\tau$ to be 5%, 10% or 15%. In general, it is chosen to ensure that there are enough observations in each
regime and that the limits for the test statistics used in inference are non-degenerate (see Andrews, 1993; Hansen, 1996). In practice, various values must be tried in empirical applications to check the robustness of the results to the selected value of \( \tau \).

The concentrated likelihood function is \( L_T(\theta_1) = -\frac{T}{2} \log |\hat{\Omega}(\theta_1)| - \frac{T}{2}, \) and so an estimator of \( \theta_1 \) is given by \( \hat{\theta}_1 = \arg \min \log |\hat{\Omega}(\theta_1)| \) (see Tsay, 1989; Enders and Siklos, 2001; Hansen and Seo, 2002).

Testing no-cointegration

Testing the no-cointegration hypothesis consists in testing \( H_0 : \Gamma = 0 \) in:

\[
\Delta X = t^{(1,3)}_T \mu^{(1,3)} + z^{(1,3)} \Gamma' + \Delta X_{-1} A + \varepsilon,
\]

where \( \Delta X \) and \( \Delta X_{-1} \) are \((T \times p)\) and \((T \times pq)\) matrices that stack \( \Delta X_1 \) and \((\Delta X_{-1}, \ldots, \Delta X_{-q})\) over \( t \), \( t^{(1,3)}_T = (t^{(1)}_T, t^{(3)}_T) \), \( \mu^{(1,3)} = (\mu^{(1)}, \mu^{(3)}) \) and \( z^{(1,3)} = (z^{(1)}, z^{(3)}) \), where \( t^{(j)}_T = t_T \mathbb{1}(A_j) \) for \( j = 1,3 \) and the \((T \times r)\) matrices \( z^{(j)} \) stack the \( r \)-vectors \( z_{r,T}^j \mathbb{1}(A_j) \) over \( t \). \( \mathbb{1}(\cdot) \) is a \((qp \times p)\) matrix.

If we assume \( \Gamma^{(j')} \mu \neq 0 \), then the intercept is present in the TVECM and \( \mu^{(1,3)} \) will not be 0. The threshold parameter \( \theta_1 \) is identified under the null hypothesis as the intercept switches with the regimes. However, it still appears as a nuisance parameter [with \( F(\theta_1) \) continuous on the interval \([0,1]\)] and it becomes impossible to provide a tabulation of the distribution of this statistic. In this case, a supremum version of the Wald statistic can be used to test \( H_0 \) (see Davies, 1987; Andrews and Ploberger, 1994; Hansen, 1996; Gonzalo and Pitarakis, 2006; for a discussion on such a statistic). We have:

\[
\sup_{\theta_1 \in [\tau,1-\tau]} \mathcal{W}_{0,T}(F(\theta_1)),
\]

where

\[
\mathcal{W}_{0,T}(F(\theta_1)) = \text{tr}\left\{ \Omega^{-1}(z^{(1,3)} \mathbb{M} e^{*})' (z^{(1,3)} \mathbb{M} e^{(1,3)})^{-1}(z^{(1,3)} \mathbb{M} e^{*}) \right\},
\]

where \( \mathbb{M} = I_T - \mathbb{X}(\mathbb{X}' \mathbb{X})^{-1} \mathbb{X}' \) with \( \mathbb{X} = (t^{(1,3)}_T, \Delta X_{-1}) \).

According to the assumptions made on the threshold variable in the previous section (and summarized in appendix B, assumption A0), it is always possible to trim the subset \( \Theta \) such that \( P(s_{t-\tau} \leq \theta) = \tau > 0 \) and \( P(s_{t-\tau} \leq \theta) = 1 - \tau, \) with \( \theta < \theta_1 < \theta_2 < \hat{\theta} \). Let \( U_t = F(s_t) \) with \( U_t \) following a uniform distribution \( U_{[0,1]} \). As \( \mathbb{1}(s_{t-\tau} \in A(\theta_1)) = \mathbb{1}(F(s_{t-\tau}) \in A(F(\theta_1))) \), we can use \( \mathbb{u} = F(\theta_1) = 1 - F(\theta_2) \) since we assumed that the threshold variable admits a symmetric density [assumption A0, item (ii) in appendix B] and \( \theta_1 = -\theta_2 \) in equation (6). The following theorem gives the asymptotic distribution of our test statistic. In what follows, integrals are taken over the unit interval and ‘\( \Rightarrow \)’ indicates convergence in distribution.
Table 1 gives the critical values of the limiting distribution of \( \sup \mathcal{W}_0 \) for various dimensions of \( p \). The asymptotic distribution has been calculated by Monte Carlo simulations. The stochastic integrals have been evaluated at 10,000 points over the argument \( u \) and 100 steps over the argument \( u = F(\theta_t) \). The critical values have been computed as the corresponding empirical quantiles from 10,000 replications. These values are also reported for various ranges \( [\tau, 1 - \tau] \). Caner and Hansen (2001) discuss the inconsistency of the tests when one chooses \( \tau = 0 \) in threshold models. Indeed, because the critical value of the statistic increases as \( \tau \) decreases, the rejection of the null requires a larger value of the statistic as \( \tau \) tends to 0. It follows that \( \tau \) should be set in the interior of \( (0, 1) \).

Theorem 1. Let assumptions A0 and A1 (in appendix B) hold. Then, under \( H_0 \), the limiting distribution of the sup-Wald statistic is:

\[
\sup \mathcal{W}_0 \Rightarrow \sup_{u \in [\tau, 1 - \tau]} \text{tr} \left\{ Q(u, u)' S(u)^{-1} Q(u, u) \right\} \equiv \sup_{u \in [\tau, 1 - \tau]} (2u)^{-1} \text{tr} \left\{ Q(u, u)' \left( \int \tilde{B}(u) \tilde{B}(u)' \right)^{-1} Q(u, u) \right\},
\]

with \( Q(u, u) = (\int dB(u, u) \tilde{B}(u)', \int dB(u, u) \tilde{B}(u)') \), \( S(u) = \text{diag}(u, u) \otimes \int \tilde{B}(u) \tilde{B}(u)' \), where \( B(u, u) \sim N(0, ru) \) is an \( r \)-vector two-parameter Brownian motion, such that \( B(u, 1) = B(u) \), and where \( \tilde{B}(u) \) denotes the demeaned Brownian motion \( \tilde{B}(u) = B(u) - \int B(u) \).

Proof. See appendix B.
Testing reduced rank

The simplest situation is when there are the same number of cointegrating relationships active in both regimes. Formally it means \( rk(\Pi^{(1)}) = rk(\Pi^{(3)}) = r \leq p - 1 \) [where \( rk(\cdot) \) stands for the rank function]. However, as it has been mentioned, the presence of a unit root in \( \Phi_1(L) \) or \( \Phi_3(L) \) in equation (7) implies \( rk(\Pi^{(1)}) \neq rk(\Pi^{(3)}) \). In our example in the previous section, the presence of one unit root in the dynamics of \( z_{t,r} \) yielded a column of zeros in \( \Pi^{(3)} \), that is, one zero in the diagonal of \( \Gamma^{(3)} \) and we referred to this case as one in which one cointegration relationship was inactive in the third regime.

Krishnakumar and Neto (2009) propose a methodology to test \( H_0 : r = r_0 \) against the alternative \( H_1 : r = r_0 + 1 \) (for \( r \geq 1 \)) inspiring from Lütkepohl and Saikkonen (2000). Through Monte Carlo experiments, they show that, while their rank test works as well as the standard likelihood ratio (LR) test in absence of a unit root in the threshold model (6) [i.e. \( rk(\Pi^{(1)}) = rk(\Pi^{(3)}) \)], it does a better job when there are unit roots [i.e. \( rk(\Pi^{(1)}) \neq rk(\Pi^{(3)}) \)]. This rank test does not require the triangular representation which allowed us to obtain a closed form solution of the parameter estimators. Let us briefly summarize it. Noting that \( P_\ell + P_{\perp} = I_p \), inserting it in the TVECM equation (without considering the lagged variables to lighten the notations) and splitting the two regimes, we can write for a given matrix \( A \) (and \( A_{\perp} \)):

\[
\Delta X_t = (\mu^{(1)} + \kappa_1 z_{t-1}^{(1)} + \lambda_1 v_{t-1}^{(1)}) \mathbb{1}_{\{A_{\perp}\}} + (\mu^{(3)} + \kappa_3 z_{t-1}^{(3)} + \lambda_3 v_{t-1}^{(3)}) \mathbb{1}_{\{A_{\perp}\}} + \varepsilon_t,
\]

where \( \kappa_j = \Pi^{(j)} A(A')^{-1} \), \( \lambda_j = \Pi^{(j)} A_{\perp}(A_{\perp}')^{-1} \) and \( v_j^{(j)} = A_{\perp}' X_t(j) \). Then, under \( H_0 : r = r_0 = rk(\Pi^{(j)}) \), we have \( \lambda_j = 0 \), and \( \kappa_j = \Gamma^{(j)} \), whereas, under the alternative, some columns of \( A_{\perp} \) will be cointegrating vectors so that \( \lambda_j \neq 0 \). Thus, one can base the rank test on testing \( H_0 : \lambda_j = 0 \). For an easier implementation of the test, following Lütkepohl and Saikkonen (2000), we multiply each equation by \( \Gamma^{(j)} \) to give:

\[
\Delta \tilde{X}_t = \tilde{\mu}^{(j)} + \tilde{\kappa}_j z_{t-1}^{(j)} + \tilde{\lambda}_j v_{t-1}^{(j)} + \tilde{\varepsilon}_t,
\]

where for any variable \( x \), we denote \( \tilde{x} = \Gamma^{(j)'} x \) and for any parameter \( a \), we denote \( \tilde{a} = \Gamma^{(j)'} a \). Thus, the null becomes \( H_0 : \tilde{\lambda}_j = 0 \). Piling up the observations, we have the following regression model for both regimes:

\[
\Delta \tilde{X}_j = (\tilde{\mu}^{(j)} \otimes I_T) + \tilde{z}^{(j)} \tilde{\kappa}_j + \tilde{v}^{(j)} \tilde{\lambda}_j + \tilde{\varepsilon}_j,
\]

where \( \Delta \tilde{X}_j \) and \( \tilde{\varepsilon}_j \) are \((T \times 4)\) matrices which stack the observations on \( \Delta \tilde{X}_j \) and \( \tilde{\varepsilon}_j \), respectively. The \((T \times r_0)\) and \((T \times 4 - r_0)\) matrices \( \tilde{z}^{(j)} \) and \( \tilde{v}^{(j)} \) stack \( z_{t-1}^{(j)} \) and \( v_{t-1}^{(j)} \), respectively. Then the test statistic \( \mathcal{W}_T^{(j)}(F(\theta), r_0) \) is obtained as a Wald-statistic for \( \lambda_j = 0 \) in equation (14).

Once again the threshold parameter \( \theta_1 \) appears as a nuisance parameter with \( F(\theta_1) \) continuous on \([0, 1]\). Like for the no-cointegration test statistic, we use the supremum Wald statistic over a grid of possible values of the threshold. Thus, for a given \( A \), for regime \( j \), this statistic, denoted as \( sup \mathcal{W}^{(j)} \), is defined as:

\[
\sup \mathcal{W}^{(j)}(r_0) = \sup_{F(\theta_1) \in [0, 1]} \mathcal{W}_T^{(j)}(F(\theta_1), r_0).
\]
The limiting distribution of this statistic is non-standard and it is given and tabulated for various dimensions of $p - r$ in Krishnakumar and Neto (2009). Critical values compatible with an intercept in both the error correction term and the cointegration relationships are reproduced in Table A1 in appendix B.

Residual-based bootstrap for finite sample distribution

It is well known that the distribution of this kind of statistics needs to be approximated for finite samples (Caner and Hansen, 2001; Seo, 2006). For this purpose, we follow the residual-based bootstrap methodology that Seo (2006) uses for a similar statistic, which consists in resampling independently and with replacement the residuals of regression (8) under the restriction $\Gamma = 0$, that is, $\Pi^{(1)} = \Pi^{(3)} = 0$. In a first step we estimate the parameters and the errors of equation (8) and then get the bootstrapped sample, denoted as $\Delta X_t^b$, from:

$$\Delta X_t^b = \tilde{\mu}^{(1)} + \tilde{\mu}^{(3)} + \tilde{\Lambda}(L)\Delta X_{t-1}^b + \tilde{\epsilon}_t^b,$$

where $\tilde{\epsilon}_t^b$ denotes the bootstrapped sample of residuals from $(\hat{\epsilon}_t - T^{-1} \sum_{j=1}^{T} \hat{\epsilon}_j)_{t=1}^T$.

Seo (2006) shows how the power of a no-cointegration test can be significantly improved with the bootstrap critical values for a sample size of 100. Since our empirical investigation deals with a bit less than $T = 200$ observations, we performed our bootstrap experiments for this sample size. The bootstrap critical values at 1%, 5% and 10% are given in the next section for the no-cointegration test and for the rank test.

IV. Empirical results

The data

To carry out a joint test of the EHTS and UIRP hypothesis for Switzerland and United States on one hand, and Switzerland and Germany on the other, we use monthly interest rate series over the period January 1993 to October 2008, representing 190 observations. Several pairs of short-/long-term rates are examined: for short-term interest rates we use 1-, 3- and 6-month money market rates (denoted as $R_{us}$, $R_{sw}$ and $R_{gm}$) and for the long-term interest rates we have the 10-year government bond rates (denoted as $r_{us}$, $r_{sw}$ and $r_{gm}$). Note that we do not use money market rate as a long-run rate because the longest one available for Switzerland is the 1-year rate, and using such a maturity would yield a particularly high risk premium for the couples 1/3/6 months and 1 year. The data used are taken from Datastream and FRED II databases. Moreover, we start our investigation in 1993 to avoid some structural changes (in particular the European integration) and the inversion of the Swiss interest rate curve from the end of the 80s to the first half of the 90s (see Figure 1). Some of our series are depicted in Figure 2.

It is generally agreed that interest rates are $I(1)$ without linear trend. Hence the unit root test results, which confirm the difference-stationary nature of each variable, are not reported here to save space (see Figure 2). So the dimension of our $x$ vector is $p = 4$, we have: $X_t = (R_{sw,t}, R_{us,t}, R_{us,t}, r_{us,t})'$ or $X_t = (R_{sw,t}, r_{sw,t}, R_{gm,t}, r_{gm,t})'$. According to the theory outlined in section II, we expect to find at most three cointegrating relationships in each regime of the TVECM. The optimal lag of TVECM is selected using the Akaike information criterion (AIC). Our results indicate the presence of only one lag on $\Delta X_t$. 

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011
Estimation and testing

Our model does not allow several transition variables (e.g. one for each cointegration relation) but only one for the full system. While the estimation of a TVECM with more than one transition variable can be envisaged without any major difficulty (see van Tol and Wolff, 2005), inference based on such a model requires fundamental modifications. Our inference is based on Tsay’s (1998) formulation of the multivariate threshold model, and on the TVECM used in Gonzalo and Pitarakis (2006) which assume a stationary and ergodic transition variable, crucial for the derivation of the asymptotic properties. Hence we keep the same transition variable for all our cointegrating relations.

The choice of the (unique) transition variable is somewhat tricky. One could try several variables depending on the economic context and select the ‘best’ one based on an information criterion such as the Bayesian information criterion (BIC) or the AIC.

Here our transition variable is chosen to be the smallest deviation from the variation of the spreads: \( \min(\Delta S_t) = s_t \), where \( S_t = J'X_t \) and the matrix \( J \) is given by: \( J = [(1, -1, 0, 0); (0, 0, 1, -1)]' \). After examination of the BIC\(^3\) for several values of the delay \( d \), we finally select \( d = 1 \). However, note that the values obtained are all close to each other up to \( d = 3 \).

Our choice for the transition variable is justified by the fact that taking a function of both country spreads incorporates money market information from both economies. Moreover, as it has been mentioned, UIRP should hold at a long horizon as long as EHTS holds, therefore a transition variable built from rate spreads would be a good candidate for a transition variable. Note that we take the changes of \( S_t \) rather than \( S_t \) itself to ensure stationarity of our transition variable even under the null of no-cointegration (see Figure 3). Indeed, the limiting distribution of the test statistics would change in case the transition variable is not stationary. Seo (2006) provides a no-cointegration test for a two-regime TVECM with only

---

\(^3\)We preferred BIC to AIC because it penalizes more the additional parameters.
Figure 2. Plots of interest and their spreads
Figure 3. Plots of (some) changes in spreads and their empirical distributions.
one cointegrating relation when the transition variable is non-stationary under the null. The four graphs at the bottom of Figure 3 check the symmetry of the empirical distribution of our transition variable (to save some space, we only plot here the couple 1 month/10 years, however, for the other maturities the empirical distributions are similar). We compare the kernel estimates with the Gaussian curve which is symmetric [recall that only symmetry is required in our framework, see assumption A0, item (ii) in appendix B]. This assumption does not seem to be violated for our sample size.

Table 2 reports the results of the no-cointegration [equation (11)] and reduced rank tests described in the previous section. However, as it has been pointed out in section III, our sample is probably not large enough to use the asymptotic critical values of these test statistics (presented in Table 1 and Table A1 in appendix B). Therefore, their distributions have been approximated for a sample size of 200 observations, which roughly corresponds to the size of our sample and for \( \tau = 0.15 \). The bootstrap critical values for the no-cointegration test, obtained from 1,000 bootstrap replications, are as follows. At 5%: 20.58 for \( p = 2 \); 27.27 for \( p = 3 \); and 33.11 for \( p = 4 \). At 1%: 26.09, 33.52 and 39.70, for \( p = 2, 3, 4 \), respectively. At 10% we have 17.55, 23.87 and 30.34. The same experiments are performed for cointegrating rank tests and the following bootstrap critical values are obtained. At 5%: 8.65 for \( p = r = 1 \); 21.29 for \( p = r = 2 \); and 37.28 for \( p = r = 3 \). At 1%: 15.88, 28.33 and 42.91, respectively, for the same values of \( p = r \). At 10%: 6.76, 18.52 and 33.62, respectively. To conduct the cointegrating rank tests, the TVECM model is estimated under the triangular normalization over 134 grid points on the empirical distribution of our threshold variable, that is, \( F_T(s_t - d) \). This grid set corresponds to 70% of the estimated values of the transition variable, which are taken as possible values of the threshold. The discarded values are the largest and smallest 15% of \( s_t - d \), that is, \( \tau = 0.15 \). For each estimated model, we computed the statistics \( \text{sup} W_T^{(1)} \) and \( \text{sup} W_T^{(3)} \). Since the interest rates are highly heteroskedastic, standard deviations of parameter estimates are computed using the Newey–West method (Newey and West, 1987) and heteroskedasticity-consistent statistic values are obtained.

The first line in Table 2 shows that the threshold cointegration hypothesis cannot be rejected for all studied cases. Moreover, while the estimates for the Swiss-US couple lead to only one cointegrating relationship, three cointegrating relationships seem to activated at least in one or both regimes for the couple Switzerland/Germany. We can attribute this result to the growing economic integration in Europe over the period. Table 2 only reports the long-run parameter estimates of these cases so that we can examine if the data are consistent with UIRP and EHTS. Noting that the three cointegrating relations are specified by a triangular TVECM, we have the following estimated relationships for parity and rate spreads: \( r_{sw,t} + \beta_2 r_{*t} \), \( R_{sw,t} + \beta_3 r_{*t} \), where (*) stands for \((us)\) or \((gm)\), and \( R_{sw,t} + \beta_4 r_{sw,t} \), where \( \beta_4 = -\frac{\beta_1}{\beta_3} \). Estimates of \( \beta_4 \) are given in Table 2 with their standard deviations obtained using the \( \delta \)-method.

\[ V(g(\hat{\beta})) = \frac{\partial g(\hat{\beta})}{\partial \beta} V(\hat{\beta}) \frac{\partial g(\hat{\beta})}{\partial \beta}^T. \]

\[ V(g(\hat{\beta})) = \frac{\partial g(\hat{\beta})}{\partial \beta} V(\hat{\beta}) \frac{\partial g(\hat{\beta})}{\partial \beta}^T. \]

\[ V(g(\hat{\beta})) = \frac{\partial g(\hat{\beta})}{\partial \beta} V(\hat{\beta}) \frac{\partial g(\hat{\beta})}{\partial \beta}^T. \]

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011
TABLE 2

<table>
<thead>
<tr>
<th></th>
<th>Germany–Switzerland</th>
<th>US–Switzerland</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 month/10 years</td>
<td>3 months/10 years</td>
</tr>
<tr>
<td>( H_0 : r_0 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r_0 = 1 )</td>
<td>( \sup W^{(0)} )</td>
<td>81.991</td>
</tr>
<tr>
<td>( r_0 = 2 )</td>
<td>( \sup W^{(1)} )</td>
<td>29.864*,**</td>
</tr>
<tr>
<td>( r_0 = 3 )</td>
<td>( \sup W^{(3)} )</td>
<td>70.385</td>
</tr>
<tr>
<td>( r_0 = 3 )</td>
<td>( \sup W^{(3)} )</td>
<td>33.97</td>
</tr>
</tbody>
</table>

Main parameter estimates

\[
\begin{align*}
\hat{\theta}_1 &= 0.042, \\
\hat{\beta}_1 &= -0.349, \\
\hat{\beta}_2 &= -0.666, \\
\hat{\beta}_3 &= -0.671, \\
\hat{\beta}_4 &= -0.524, \\
\text{LR} &= -91.88, \\
\text{P-value} &= 0.999
\end{align*}
\]

Notes: The null hypothesis cannot be rejected at *5% and **10% levels. These results are for \( \tau = 0.15 \). LR, likelihood ratio.

from our TVECM can be performed by testing \( H_0 : \beta_1 = \beta_2 = \beta_4 = -1 \). Using \( \beta_4 = -\frac{\hat{\theta}_1}{\hat{\beta}_1} \), and combining it with \( H_0 \), that is, \( \beta_2 = -1 \), we get \( \beta_1 = -1 \) as well. Thus, the test can be carried out directly from the TVECM for:

\[
H_0 : \begin{pmatrix} 1 & 0 & 0 & \beta_1 \\ 0 & 1 & 0 & \beta_2 \\ 0 & 0 & 1 & \beta_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix},
\]

using an LR statistic. The LR statistics and the corresponding \( p \)-values [which correspond here to a \( \chi^2_{(3)} \) distribution] for the selected estimates are reported in the last row of Table 2. The values conclude that the null hypothesis of the consistency of EHTS and UIRP hypothesis cannot be rejected. The estimates of the loadings illustrate the asymmetry of the error correction dynamics outside the ‘band of inaction’. The results for the couple (1 month/10 years) are reproduced below as an illustration:

\[
\hat{\Gamma}^{(1)} = \begin{pmatrix} -0.056 & 0 & 0 \\ 0 & -0.205 & 0 \\ 0 & 0 & -0.059 \end{pmatrix}, \quad \hat{\Gamma}^{(3)} = \begin{pmatrix} -0.107 & 0 & 0 \\ 0 & -0.241 & 0 \\ 0 & 0 & -0.113 \end{pmatrix}.
\]
As $\tilde{\Gamma}^{(j)} = (-\tilde{\Phi}_j(1)\theta^{(1 \times 3)})'$, we can write these results in terms of threshold model parameters [equation (7)] as follows:

$$\tilde{\Phi}^{(1)} = \begin{pmatrix} 0.944 & 0 & 0 \\ 0 & 0.795 & 0 \\ 0 & 0 & 0.941 \end{pmatrix}, \quad \tilde{\Phi}^{(3)} = \begin{pmatrix} 0.893 & 0 & 0 \\ 0 & 0.758 & 0 \\ 0 & 0 & 0.886 \end{pmatrix}.$$

These estimates suggest that the speed of adjustment is greater in the lower regime. The fact that the estimates of $\phi_{11}^{(1)}$ and $\phi_{33}^{(1)}$ are close to 1 highlights a high persistence in this regime. To conclude, our data confirm both EHTS and UIRP when we consider the big European neighbour country (Germany). This result is probably because of the growing standardization of European economies (including European Monetary Union and bordering countries) over the last 20 years. However, the two hypotheses do not hold when Swiss rates are compared with US rates.

V. Conclusion

This article examines the EHTS of interest rate and UIRP for Switzerland/United States and Switzerland/Germany in a TVEC framework. It is well known that it is difficult to get empirical evidence for these long-run relationships even though they are still used as fundamental laws in economic theory. We focus here on the Swiss case, which is perhaps a tricky one because Swiss rates have been particularly low since the Second World War. The Swiss advantage can be mainly explained by the bank secrecy principle and the Central Bank’s credibility. However, the European economic integration leading to a standardization of European rates, plus the recent European debate over the legitimacy of the bank secrecy principle in Switzerland, may act against the Swiss advantage. It is in this context that we test whether UIRP still holds for Switzerland.

Many empirical studies show that such relationships do not hold if some nonlinearities, resulting from transaction costs and/or Central Bank interventions, are not taken into account. As a threshold model is well known to be able to capture regime changes and persistence, we first look at estimation and inference for a multivariate three-regime TVECM allowing for more than one cointegrating relationship, which includes one regime of inaction between two symmetric bands. For this particular three-regime framework, we follow the methodology of Caner and Hansen (2001) to derive a no-cointegration test and compute the critical values. Then we develop a reduced rank test for detecting the number of cointegration relationships in each regime and compute the critical values. Applying these tests to our data set, we find that both UIRP and EHTS hypotheses are jointly accepted (cannot be rejected) when we consider Switzerland and Germany, whereas these hypotheses do not hold when US rates are considered. Moreover, we also find asymmetry in the activation of the cointegrating relationships. Indeed only one regime is consistent with both hypotheses. In conclusion, even if the Swiss advantage in terms of interest rates has not yet completely vanished, it seems to be weaker towards Europe.

Final Manuscript Received: January 2011
References


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


© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011


### Appendix A: Proof of proposition 1

Let us consider the cointegrated triangular system (5). The VECM associated with it is given by:

\[ \Delta X_t = \Gamma A' X_{t-1} + \epsilon_t, \]

where \( \epsilon_t = \left( \begin{array}{c} I_r \\ \beta \\ 0 \end{array} \right) z_t, \) with \( z_t = (z_{r,t} z_{p-r,t}) \), \( z_{p-r,t} \) a white noise, \( \Gamma' = [-I_r, 0] \), and \( A' = [I_r - \beta] \) (see Phillips, 1991; Saikkonen, 1992). Rewriting \( \epsilon_t \) as:

\[ \left( \begin{array}{c} \epsilon_{r,t} \\ \epsilon_{p-r,t} \end{array} \right) = \left( \begin{array}{c} z_{r,t} + \beta z_{p-r,t} \\ z_{p-r,t} \end{array} \right), \]

and combining the VECM and the threshold model (7), we have:

\[ \Delta X_t = \Gamma A' X_{t-1} + \left( \begin{array}{c} I_r \\ 0 \\ I_{p-r} \end{array} \right) \left( \begin{array}{c} \phi_1 z_{p-r,t-1} L(A_{1}) + z_{p-r,t-1} L(A_{1}) + \phi_3 z_{p-r,t-1} L(A_{1}) + \epsilon_t \\ \phi_3 z_{p-r,t-1} L(A_{1}) + \epsilon_t \end{array} \right), \]

Rearranging this last equation, we obtain:

\[ \Delta X_t = \Gamma A' X_{t-1} + \left( \phi_1 L(A_{1}) + I_r (1 - L(A_{1}) - L(A_{1})) + \phi_3 L(A_{1}) \right) A' X_{t-1} + \begin{pmatrix} \epsilon_t + \beta z_{p-r,t} \\ \epsilon_t \end{pmatrix}. \]

Using \( \Gamma' = [-I_r, 0] \), we have:

\[ \Delta X_t = \left( \phi_1 - I_r \right) L(A_{1}) + \phi_3 L(A_{1}) A' X_{t-1} + \epsilon_t. \]

Denoting \( \Gamma^{12} = (\phi_1 - I_r, 0) \), the expression of the TVECM, \( \Delta X_t = \Pi^{1} X_{t-1} L(A_{1}) + \Pi^{12} X_{t-1} L(A_{1}) + \epsilon_t \), is then immediate.
Let us assume now that \( z_{p-r,t} \) has a finite \( q \)-autoregressive representation such that 
\[ G(L)z_{p-r,t} = \zeta_t, \]
where \( \zeta_t \) is an i.i.d process with 0 mean, 
\[ G(L) = I_{p-r} - L g(L), \]
\[ g(L) = \sum_{i=1}^{q} g_i L^{-i} \]
and where \( g_i \) are coefficient matrices of dimension \((p - r \times p - r)\). Substituting in our VECM, we obtain:
\[ \Delta X_t = \Pi^{(1)} X_{t-1}^\perp (A_1) + \Pi^{(2)} X_{t-1}^\perp (A_2) + \left( \frac{\beta g(L)z_{p-r,t-1}}{g(L)z_{p-r,t-1}} \right) + \left( \frac{\beta \zeta_t + \epsilon_t}{\zeta_t} \right). \]

Denoting by \( \epsilon_t \) the last error vector and keeping in mind that \( \Delta X_{p-r,t-1} = z_{p-r,t-1} \), we get the expression given in the body of proposition 1.

**Appendix B: Proof of theorem 1**

For the derivation of the limiting distribution of the sup-Wald statistic, we need the following assumptions.

**Assumption A0.** (i) \( s_{t-d} \) is a strictly stationary and ergodic sequence, whose distribution 
\( F \) is continuous everywhere. (ii) \( s_{t-d} \) admits a symmetric density function. (iii) The delay \( d \) is strictly positive to ensure the predeterminedness of the process. (iv) If \( d = 0 \) then the variable \( s_t \) has to be assumed exogenous. (v) \( \theta_1 \in \Theta \), where \( \Theta \) is a closed and bounded subset of the sample space of the variable \( s_{t-d} \).

**Assumption A1.** Consider the sequence \( \{ e_t, U_t \} \), where \( e_t \) is a \( p \)-dimensional vector and \( U_t \) is a univariate process whose marginal distribution is \( U_{[0,1]} \), and let \( \mathcal{F}_{t-1} \) be the natural filtration. We assume (i) \( e_k \), \( k = 1, \ldots, p \) and \( U_t \) are strictly stationary, ergodic and strong mixing with mixing coefficients \( a_t \) satisfying \( \sum_{t=1}^{\infty} a_t^{1/2 - 1/\delta} < \infty \) for some \( \delta > 2 \), (ii) \( e_k \) are independent of \( \mathcal{F}_{t-1} \) for \( k = 1, \ldots, p \), and (iii) \( \mathbb{E}(e_k) = 0, \mathbb{E}|e_k|^4 < \infty \), for \( k = 1, \ldots, p \).

**Theorem A1.** Consider the \( p \)-dimensional \( I(1) \)-vector process \( X_t \), such that \( \Delta X_t = e_t \), where \( e_t \) satisfies assumption A1 with covariance matrix \( \Omega_e \). Then, as \( T \to \infty \) and under assumptions A0 and A1, we have:

(i) \[ B_T(u) = T^{-1/2} \sum_{t=1}^{T} \mathbb{1}_{\{U_{t-d} \leq F(\theta_1)\}} e_t \Rightarrow \Omega_e^{1/2} B(u, u), \text{ on } (u, u) \in [0,1]^2, \]

(ii) \[ T^{-1} \sum_{t=1}^{T} X_{t-1}^\perp (U_{t-d} \leq F(\theta_1)) e'_t \Rightarrow \int B(u) dB(u, u) \Omega_e^{1/2}. \]

(iii) \[ T^{-2} \sum_{t=1}^{T} \mathbb{1}_{\{U_{t-d} \leq F(\theta_1)\}} X_{t-1}^\perp X^\perp_{t-1} \Rightarrow \mu \Omega_e^{-1/2} \left( \int B(u) dB(u, u) \right) \Omega_e^{1/2}. \]

**Proof of theorem A1.** These results are direct multivariate extensions of the results of Caner and Hansen (2001, theorems 1–3, pp. 1560–1561) for the univariate case.

Let us now turn to our statistic in equation (11). To derive its limiting distribution, we rewrite it as follows:
\[ \mathcal{W}_0(F(\theta_1)) = \text{tr} \left\{ \frac{1}{T} (z^{(1,3)} M \eta)^T z^{(1,3)} (z^{(1,3)} M \eta)^{-1} z^{(1,3)} M \eta \right\}, \]
with $\eta = \varepsilon \Omega^{-1/2}$. Since under $H_0$, $z^{(1,3)}$ is $I(1)$, we have:

$$
\frac{1}{T}(z^{(1,3)'M\eta}) = \frac{1}{T}z^{(1,3)'\eta} \Rightarrow \left( \int dB(u, u)\tilde{B}(u)', \int dB(u, u)\tilde{B}(u)' \right),
$$

$$
\frac{1}{T^2}(z^{(1,3)'Mz^{(1,3)}}) = \frac{1}{T^2}z^{(1,3)'z^{(1,3)}} \Rightarrow \begin{pmatrix} \mathbf{u} & 0 \\ 0 & \mathbf{u} \end{pmatrix} \otimes \int \tilde{B}(u)\tilde{B}(u)' du.
$$

<p>| TABLE A1 |</p>
<table>
<thead>
<tr>
<th>Critical values for sup-Wald statistic, for model with $\Gamma_{j\mu}=0$ (Krishnakumar and Neto, 2009)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$p - m$</strong></td>
</tr>
<tr>
<td>$3$</td>
</tr>
<tr>
<td>$4$</td>
</tr>
<tr>
<td>$5$</td>
</tr>
<tr>
<td>$6$</td>
</tr>
</tbody>
</table>

**Note:** Calculated from 10,000 simulations.