Exponential or Logistic Smooth Transition Cointegration in the Term Structure of Interest Rates? A Case of a Transition Economy

Joanna Bruzda

Abstract

In the paper, on an example of the Polish economy, the expectations hypothesis of the term structure of interest rates is examined in the smooth transition (STR) cointegration framework. The analysis is based on weekly observations on interest rates for instruments with maturities ranging from 1 week till 10 years like WIBOR (the Warsaw Inter-Bank Offered Rate), Treasury bill rates and government bond rates. The study can be thought of as an attempt to discriminate between three competing explanations of the lack of empirical support of the expectation hypothesis of the term structure: nonstationary risk premia, presence of transaction costs and policymakers’ actions. In the investigation two approaches to testing for exponential and first- as well as second-order logistic smooth transition cointegration are utilized: one based on Taylor series approximations to the transition functions under study and second taking advantage of inf statistics computed for different values of nuisance parameters in the range of stationarity of an adjustment process. It turns out that the tests for STR cointegration considered here are able to find long-term relationships, where linear-based tests fail. They generally support a nonlinear adjustment hypothesis, where the observed nonlinearity seems to result from policymakers’ actions rather than from the presence of transaction costs.

Keywords: smooth transition cointegration, STAR processes, term structure of interest rates

JEL: E4, C5

1 Introduction

In the theory of the term structure of interest rates an important role is played by the expectations hypothesis. The hypothesis implies that the annual \( n \)-period interest rate, \( R_t^{(n)} \), is the weighted average of the expected future one-period interest rates plus a risk premium.
Let us consider investing \( A \) in a zero coupon bond with \( n \) years to maturity. The future value of the investment is:

\[
FV = A(1 + R_i^{(n)})^n.
\]

Next, consider the alternative strategy of reinvesting \( A \) in a series of one-period investments with the one-year spot yields, \( r_{t+i} \), \( i = 0, 1, ..., n-1 \). Expected future value of this investment is given by:

\[
E_t(FV) = A(1 + r_t)(1 + E_t r_{t+1})(1 + E_t r_{t+2})... (1 + E_t r_{t+n-1}).
\]

*Pure expectations hypothesis* assumes that future values (1) and (2) are equal, i.e.

\[
(1 + R_i^{(n)})^n = \prod_{i=0}^{n-1} (1 + E_t r_{t+i}).
\]

Taking logarithms of (3) and using the approximation \( \ln(1 + z) = z \) for \( z \) close to zero, we arrive at the postulated approximate linear relationship:

\[
R_i^{(n)} = \frac{1}{n} \sum_{i=0}^{n-1} E_t r_{t+i}.
\]

A more general relationship between interest rates for bonds with different maturities is given by:

\[
R_i^{(n)} = \frac{1}{n} \sum_{i=0}^{n-1} E_t r_{t+i} + L_t^{(n)},
\]

(see Hall et al., 1992), where \( L_t^{(n)} \) is the term premium including agents’ liquidity preferences and a risk premium. Under different assumptions concerning the term premium different term structure hypotheses are obtained with the simplest *pure expectations hypothesis* saying that the premium is equal zero. Some elementary algebraic transformations allow to write (5) as:

\[
R_i^{(n)} - r_t = \frac{1}{n} \sum_{i=0}^{n-1} \sum_{j=0}^{l} E_t \Delta r_{t+j+1} + L_t^{(n)},
\]

where \( \Delta r_{t+j+1} = r_{t+j+1} - r_t \). If the time premium is stationary and the interest rates are \( \text{I}(1) \) processes\(^3\), the relation (5) implies that \( R_i^{(n)} \) and \( r_t \) are cointegrated with the cointegrating vector \([1, -1]\).

Empirical studies usually provide mixed results on cointegration analysis of the term structure of interest rates, which depend considerably on specifications underlying econometric models, on sample periods used, maturities under scrutiny, and kinds of instruments\(^4\). Explanations for the lack of empirical support for the expectation hypothesis of the term structure fall into two main categories: nonlinearity of adjustment.

---

\(^2\) For an overview of theories of the term structure of interest rates see, for example, Cuthbertson, Nitzsche (2004).

\(^3\) Many empirical studies document that interest rates are integrated of order one, see, for example, Engle, Granger (1987), Hall et al. (1992), Cuthbertson et al. (1998).

\(^4\) For a discussion see Pagan et al. (1996).
processes and nonstationary risk premia. The nonlinear adjustment to the term structure relationship is usually motivated by the presence of transaction costs. This transaction costs hypothesis calls for a three-regime dynamics for the adjustment process with a unit root imposed in the middle regime. Such a dynamics can be modelled with the help of either a three-regime SETAR process or an exponential (or, alternatively, second-order logistic) smooth transition autoregressive process. However, it is usually pointed out that temporal as well as cross-sectional aggregation causes transitions between regimes being smooth rather than sharp, which supports the STAR specifications. Additionally, as the transition parameter becomes large, the second-order logistic STAR model approaches a restricted three-regime SETAR model, with the restriction that the outer regimes are identical. Due to this the family of second-order LSTAR processes nests the three-regime SETAR process as a limiting case.

However, there is also another explanation of the nonlinear adjustment to the term structure relationship referring to policymakers’ actions. The second explanation, advocated by Bohl and Siklos (2004), is based on the observation that policymakers raise interest rates gradually but reduce them quickly once the threat of higher future inflation has dissipated. Increasing short-term interest rates gradually may influence expectations on inflation and, hence, long-term interest rates. Due to this small changes in the interest rates, which central bank controls directly, can have a relatively larger effect on long-term interest rates. This calls for a two-regime specification of the adjustment process. The one admitted here takes the form of a first-order logistic STAR model.

The influence of transaction costs and the interest rate spread asymmetry caused by policymakers’ actions in a natural way give rise to relax the linearity assumption in cointegration analysis by considering non-linear adjustment processes. There are numerous econometric studies providing evidence on non-linear adjustments to long-term relationships between interest rates. Anderson (1997) made the first attempt to relax the implicit assumption of linearity of an adjustment process to the term structure relationship. In her paper she examined the expectations hypothesis in the US economy with the help of the Balke and Fomby (1997) procedure and univariate threshold error correction models as well as smooth transition error correction models pointing at the smooth transition specification as a better way to describe an influence of transaction costs in aggregated quantities. Subsequent papers (Enders, Siklos, 2001, Bohl, Siklos, 2004, Kuo, Enders, 2004, McMillan, 2004) examined the term structures in the US, German, Japan and British economies making use of the Enders and Siklos (2001) methodology and documenting threshold or, alternatively, smooth transition asymmetries. Another strand of research focuses on threshold vector equilibrium correction models (see Hansen, Seo, 2002, Seo, 2003 and Clements, Galvão, 2004) or smooth transition vector equilibrium correction models (Sarno, Thornton, 2003). All these studies document existence of non-linear adjustments to long-term relationships for yields with different maturities and their empirical evidence is based on: lower residual variance of the estimated non-linear error correction models as compared to linear models, a better post-sample performance, and results of linearity tests against threshold alternatives.

In this paper I take a similar viewpoint and attempt to verify the expectations hypothesis of the term structure of interest rates in the Polish economy in the smooth transition cointegration framework. The study can be thought of as an attempt to discriminate between three competing explanations of the lack of empirical support of the ex-
pection hypothesis: nonstationary risk premia, presence of transaction costs and policymaker’s actions. In the investigation two approaches to testing for exponential and first- as well as second-order logistic smooth transition cointegration are utilized: one based on Taylor series approximations to the transition functions under study and second taking advantage of \( \inf t \) statistics computed for different values of nuisance parameters in the range of stationarity of an adjustment process. The rest of the paper is organised as follows. Section 2 describes the methodology used in the study. Section 3 discusses the results of simulation analysis concerning the power and size of STR cointegration tests, while section 4 presents the empirical results. A brief summary is given in section 5.

2 Testing for smooth transition cointegration

It is now well recognised that standard unit root and cointegration tests, like the ADF test and the Engle-Granger and Johansen procedures, are not appropriate for investigating nonlinear processes (see, for example, Pippenger, Goering, 1993, 2000, Dufrenot, Mignon, 2002). In the interplay between nonstationarity, cointegration and non-linearity two problems arise. Firstly, the standard linear tests lack their power in the case of stationary nonlinear processes. Secondly, even a more dangerous problem seem to be serious size distortions in the case of nonstationary and nonlinear processes, which can be wrongly recognised as stationary ones. It is quite obvious that misspecifying a stable nonlinear process as a nonstationary one will lead to misleading impulse responses and worse forecasting abilities. On the other hand, allowing for a nonlinear adjustment should enable to find long-term relationships, where linear cointegration tests fail. Several authors have addressed these questions and suggested unit root tests against specific nonlinear alternatives (see, for example, Enders, Granger, 1998, Caner, Hansen, 2001, Eklund, 2003, Kapetanios, Shin, Snell, 2003, Kiliç, 2003, Bec, Ben Salem Carrasco, 2004, Park, Shintani, 2005) as well as tests of the hypothesis of no cointegration against an alternative assuming a particular nonlinear stationary adjustment process (see Enders, Siklos, 2001, Kapetanios, Shin, Snell, 2006). This paper contributes to the second strand of research by suggesting tests of the hypothesis of no linear cointegration against exponential and first- as well as second-order logistic smooth transition autoregressive adjustment processes. \( F \) and \( t \) tests for exponential smooth transition cointegration have been suggested by Kapetanios, Shin and Snell (2006) (hereafter KSS). The \( F \) tests considered here are an attempt to generalise the KSS approach by providing a specification procedure, which can be helpful to distinguish between linear, LSTR and ESTR (or 2LSTR) cointegration.\(^5\) The modelling technique resembles the well-known Teräsvirta procedure (see Teräsvirta, 1994) to test for LSTAR and ESTAR linearity in stationary time series. Moreover, a different approach to test for STR cointegration is considered. The second method is based on a grid search over possible values of parameters of a transition function under study in a way similar to unit root tests suggested in Kiliç (2003) and Park, Shintani (2005).

\(^5\) ESTR, LSTR and 2LSTR cointegration denote, respectively, exponential, first-order logistic and second-order logistic smooth transition cointegration.
Similarly to KSS the starting point here is a particular specification of the nonlinear vector error correction model considered by Saikkonen (2005). Let us consider a \((k + 1)\)-dimensional vector of \(I(1)\) processes, \(Z_t\), in the form:

\[
\Delta Z_t = a \beta' Z_{t-1} + g(\beta' Z_{t-1}) + \sum_{j=1}^{p} \Gamma_j \Delta Z_{t-j} + \epsilon_t, \tag{7}
\]

where \(\epsilon_t\) is a \((k + 1)\)-dimensional white noise with \(E\epsilon_t = 0\) and \(E\epsilon_t \epsilon'_t = I_{k+1}\). \(a\) and \(\beta\) are \((k + 1) \times r\) parameter matrices of rank \(r\), \(\Gamma_j\) are \((k + 1) \times (k + 1)\) parameter matrices and \(g: \mathbb{R}^r \rightarrow \mathbb{R}^k\) is a nonlinear function to be described in detail shortly. Further it is assumed that the following assumptions are fulfilled:

(i) The distribution of \(\epsilon_t\) is absolutely continuous with respect to the Lebesgue measure and has a density that is bounded away from zero on compact subsets of \(\mathbb{R}^k\).

(ii) Define \(A(z) = (1 - z)I_k - a \beta' z - \sum_{j=1}^{p} \Gamma_j (1 - z) z^j\). If \(\det A(z) = 0\) then \(|z| > 1\) or \(z = 1\), whereas the number of unit roots is equal \(k + 1 - r\).

(iii) The function \(g\) is (Borel) measurable, locally bounded, and asymptotically no greater than a linear function, i.e. \(g(x) = O(\|x\|)\) as \(\|x\| \to \infty\), where \(\|\|\) denotes the Euclidean norm.

Under assumptions (i)–(iii), Saikkonen (2005) proves that there exists a choice of initial values \(Z_{-p}, Z_{-p+1}, \ldots, Z_0\) such that \(\beta' Z_t\) and the components of \(\Delta Z_t\) are both strictly and covariance stationary.

Further in the paper I assume that there exists at most one cointegrating vector, i.e. \(r \leq 1\), and concentrate on conditional modeling of the scalar process \(Y_t\) given the \(k\)-dimensional process \(X_t\) and the past values of \(Z_t\), where \(Z_t = (Y_t, X_t)'\). Let us define \(u_t = Y_t - \beta' X_t\), where \(\beta\) is a \(k \times 1\) vector of cointegrating parameters. Conformably with the partition of \(Z_t\) the vector \(a\) is decomposed as \(a = (\phi, a_x)'\), where it is assumed that \(a_x = 0\) and \(\phi < 0\). Further I focus on the following functional forms for the function \(g\):

- exponential function (see KSS):
  \[
g(u_{t-1}) = -\rho u_{t-1} e^{-\gamma (u_{t-1} - c)^2}, \quad \gamma > 0, \tag{8}
\]

- first-order logistic function:
  \[
g(u_{t-1}) = -\rho u_{t-1} \frac{b}{1 + e^{-\gamma (u_{t-1} - c)}}, \quad 0 < b < 1, \quad \gamma \neq 0, \tag{9}
\]

- second-order logistic function:
  \[
g(u_{t-1}) = \rho u_{t-1} \left[\frac{1}{1 + e^{-\gamma (u_{t-1} - c)^2}} - 1\right], \quad \gamma > 0, \quad c > 0. \tag{10}
\]
Clearly, the functions (8)–(10) fulfill assumption (iii) above. Next, let us define the partition \( \mathbf{p} = (\mathbf{p}_\rho, \mathbf{p}_\omega)' \) and assume \( \mathbf{p}_\omega = \mathbf{0} \). Furthermore, it is assumed that there is no cointegration among the components of \( \mathbf{X}_t \). The last assumption together with \( \mathbf{a}_x = \mathbf{0} \) and \( \mathbf{p}_\omega = \mathbf{0} \) implies weak exogeneity of \( \Delta \mathbf{X}_t \) for structural parameters in the equation for \( \Delta Y_t \).\(^6\) Conformably with the partition of \( \mathbf{Z}_t \) the error terms are decomposed as \( \varepsilon_t = (\varepsilon_{yt}, \varepsilon_{xt})' \) and the parameter matrices as \( \Gamma_j = (\Gamma_{yj}', \Gamma_{xj}')' \), \( j = 1, \ldots, p \). Then the model (7) can be rewritten in the form:

\[
\Delta Y_t = \varphi u_{t-1} + g(u_{t-1}) + \sum_{j=1}^{p} \gamma_{yj} \Delta Z_{t-j} + \varepsilon_{yt}, \tag{11}
\]

\[
\Delta \mathbf{X}_t = \sum_{j=1}^{p} \Gamma_{xj} \Delta Z_{t-j} + \varepsilon_{xt}. \tag{12}
\]

If the covariance matrix of \( \varepsilon_t \) is decomposed as:

\[
\Sigma = \begin{pmatrix} \mathbf{\sigma}_{yy} & \mathbf{\sigma}_{yx} \\ \mathbf{\sigma}_{xy} & \mathbf{\Sigma}_{xx} \end{pmatrix},
\]

then \( \varepsilon_{yt} \) may be expressed in the form: \( \varepsilon_{yt} = \omega e_{xt} + e_t \), where \( \omega = \mathbf{\sigma}_{yx} \mathbf{\Sigma}_{xx}^{-1} \).

\( e_t \sim i.i.d.(0, \mathbf{\sigma}^2_e) \) is uncorrelated with \( \varepsilon_{xt} \) and

\[
\mathbf{\sigma}^2_e = E(\varepsilon_{yt} - \mathbf{\sigma}_{yx} \mathbf{\Sigma}_{xx}^{-1} \varepsilon_{xt}) (\varepsilon_{yt} - \mathbf{\sigma}_{yx} \mathbf{\Sigma}_{xx}^{-1} \varepsilon_{xt}) = \mathbf{\sigma}_{yy} - \mathbf{\sigma}_{yx} \mathbf{\Sigma}_{xx}^{-1} \mathbf{\sigma}_{xy}. \tag{14}
\]

Substituting for \( \varepsilon_{yt} \) in (11) we obtain the following conditional nonlinear error correction (NEC) model:

\[
\Delta Y_t = \varphi u_{t-1} + g(u_{t-1}) + \omega \Delta \mathbf{X}_t + \sum_{j=1}^{p} \psi_{yj} \Delta Z_{t-j} + e_t, \tag{15}
\]

where \( \psi_{yj} = \gamma_{yj} - \omega \Gamma_{xj}, j = 1, \ldots, p \). With \( g \) defined according to (8)–(10) the model (15) becomes the conditional smooth transition regression error correction model (STR ECM), which constitutes the basic object of interest in further methodological considerations in this section. The model (12) is the marginal VAR model. Let us rewrite (15) as:

\[
\Delta Y_t = \varphi u_{t-1} + g(u_{t-1}) + \omega_1 \Delta \mathbf{X}_t + \omega_1 \Delta \mathbf{X}_{t-1} + \ldots + \omega_p \Delta \mathbf{X}_{t-p} +
\]

\[
+ \delta_1 \Delta Y_{t-1} + \ldots + \delta_p \Delta Y_{t-p} + e_t, \tag{16}
\]

where \( \psi_j = (\delta_j, \omega_j)' \), \( j = 1, \ldots, p \). If the structural parameters in (16) fulfill the common factor (COMFAC) restrictions in the form \( \omega = \mathbf{\beta}'_x \), \( \omega_1 = -\delta_1 \mathbf{\beta}'_x \), \ldots, \( \omega_p = -\delta_p \mathbf{\beta}'_x \), then the conditional model may be written as:

\(^6\) In fact to validate an asymptotic inference about no cointegration in the NEC framework a more stringent condition of strict exogeneity is required. Alternatively, the NEC test equations are augmented with leads of \( \Delta \mathbf{X}_t \) – see KSS.
The equation (17) is a nonlinear variant of the test equation in the augmented Dickey-Fuller test. Further both the ECM-based tests and the residual-based tests are considered, which utilize the equations (16) and (17), respectively. It is to be expected that, when the COMFAC restrictions are not fulfilled, tests based on (16) should have better power properties as compared to the nonlinear Engle-Granger procedure.

In further considerations it is admitted that \( \varphi = \rho \), which results in some meaningful specifications. In the exponential and second-order logistic cases it is equivalent to imposing, respectively, a unit root and ‘nearly unit root’ dynamics in the middle regime, while in the first-order logistic case the model (15) with \( g \) defined in (9) is clearly over-parametrized and the assumption \( \varphi = \rho \) reduces the redundant parameter. Then, for global stationarity of the error correcting mechanism of the ESTR i 2LSTR error correction models we need \( \rho < 0 \), while the appropriate condition for stationarity in the first-order logistic case is \( \frac{b-2}{1-b} < \rho < 0 \).

Figure 1 presents example nonlinear functions \( f(u_{t-1}) = \varphi u_{t-1} + g(u_{t-1}) \) describing the error correcting mechanisms considered here. The functions correspond to the exponential, first-order logistic and second-order logistic cases, respectively, and are as follows:

\[
\begin{align*}
f(u_{t-1}) &= -0.1 u_{t-1} \left( 1 - e^{-0.1 u_{t-1}^2} \right), \quad (18) \\
f(u_{t-1}) &= -0.1 u_{t-1} \left( 1 - \frac{0.8}{1 + e^{-0.5 u_{t-1}}} \right), \quad (19) \\
f(u_{t-1}) &= -0.1 u_{t-1} \frac{1}{1 + e^{-0.5 u_{t-1}^2}}. \quad (20)
\end{align*}
\]

The conditions for global stationarity of the analysed error correcting mechanisms suggest that in testing for STR cointegration the pragmatic residual-based two-step approach can be applied, where in the first step the adjustment process \( u_t \) is replaced with the OLS residuals \( \hat{u}_t = y_t - \hat{\beta} x_t \), while in the second step the following hypotheses are set: \( H_0: \rho = 0 \) against \( H_1: \rho < 0 \). However, such a test, utilizing the standard \( t \) statistic for the parameter \( \rho \), cannot be performed directly due to the presence of nuisance parameters, which are identified only under the alternative. To overcome this problem two approaches are usually admitted: one making use of a Taylor series approximation to the transition function of interest and second based on a grid search over possible values.

\[\Delta Y_t = \rho Y_{t-1} \left( 1 - \frac{b}{1 + e^{-\gamma V_{t-1}}} \right) + \varepsilon_t \text{ under } \gamma \to \infty \]
of the unidentified parameters. I what follows I describe in detail the two testing frame-
works.

![Fig. 1 Exponential and logistic error correcting mechanisms](image)

A Taylor series approximation in testing for ESTR cointegration has been suggested in KSS, where the authors consider the conditional ESTR ECM model in the form:

$$\Delta Y_t = \rho u_{t-1} \left( 1 - e^{-\gamma (u_{t-1} - c)^2} \right) + \omega \Delta X_t + \sum_{j=1}^{\rho} \psi_{yj} \Delta Z_{t-j} + e_t,$$

$$\rho < 0, \gamma > 0,$$

which is the model (15) with \( g \) given in (8) and \( \varphi = \rho \), and set the null and alternative hypotheses as \( H_0: \gamma = 0 \) and \( H_1: \gamma > 0 \). Then, the first-order Taylor series expansion of the function \( F(x) = 1 - e^{-x} \) around 0, where \( x = \gamma (u_{t-1} - c)^2 \), gives the approximation \( T_1(x) = x \), what leads to the following test equation:

$$\Delta Y_t = \alpha_1 u_{t-1} + \alpha_2 u_{t-1}^2 + \alpha_3 u_{t-1}^3 + \omega \Delta X_t + \sum_{j=1}^{\rho} \psi_{yj} \Delta Z_{t-j} + e_t,$$

(22)

where \( \alpha_1 = c^2 \gamma \rho \), \( \alpha_2 = -2c \gamma \rho \) and \( \alpha_3 = \gamma \rho \). To test for ESTR cointegration an \( F \)-type statistic for \( \alpha_1 = \alpha_2 = \alpha_3 = 0 \) is suggested. The statistic is computed as follows:

$$F = \frac{(SSR_0 - SSR_1) / q}{SSR_1 / (T - q - pk - k)},$$

(23)

where \( SSR_0 \) and \( SSR_1 \) stand for the sums of squared residuals in the restricted and unrestricted models and \( q \) is the number of restrictions. Assuming \( c = 0 \) in (22) one arrives at the test equation:

$$\Delta Y_t = \alpha_3 u_{t-1}^3 + \omega \Delta X_t + \sum_{j=1}^{\rho} \psi_{yj} \Delta Z_{t-j} + e_t.$$

(24)

A test for ESTR cointegration based on (23) consists in testing for negativity of the parameter \( \alpha_3 \) with the usual \( t \) statistics.

Here I discuss analogous \( F \) tests resulting from considering (15) with \( \varphi = \rho \) and \( g \) given in (9) and (10). The conditional LSTR ECM model has the form:
\[ \Delta Y_t = \rho u_{t-1} \left( 1 - \frac{b}{1 + e^{-\gamma(u_{t-1} - c)}} \right) + \omega \Delta X_t + \sum_{j=1}^{p} \psi_{yj} \Delta Z_{t-j} + e_t, \quad (25) \]

Testing for linearity of (25) consists in testing \( H_0: \gamma = 0 \) against the two-sided alternative. As previously, to overcome the problem of unidentified parameters a Taylor series approximation is used. The first-order Taylor approximation around 0 to the function \( F(x) = 1 - \frac{b}{1 + e^{-x}} \), where \( x = \gamma(u_{t-1} - c) \), is given by \( T_1(x) = (1 - \frac{b}{2}) - \frac{b}{4} x \), while the third-order approximation has the form \( T_3(x) = (1 - \frac{b}{2}) - \frac{b}{4} x + \frac{b}{48} x^3 \). The two approximations give the following auxiliary equations:

\[ \Delta Y_t = \alpha_1 u_{t-1} + \alpha_2 u_{t-1}^2 + \omega \Delta X_t + \sum_{j=1}^{p} \psi_{yj} \Delta Z_{t-j} + e_t, \quad (26) \]

where \( \alpha_1 = (1 - \frac{b}{2}) \rho + \frac{bc\rho}{4} \) and \( \alpha_2 = -\frac{b\rho}{4} \), and

\[ \Delta Y_t = \alpha_1 u_{t-1} + \alpha_2 u_{t-1}^2 + \alpha_3 u_{t-1}^3 + \alpha_4 u_{t-1}^4 + \omega \Delta X_t + \sum_{j=1}^{p} \psi_{yj} \Delta Z_{t-j} + e_t. \quad (27) \]

Then, to test jointly for no cointegration and linearity the following hypotheses are set:

\( H_0: \alpha_1 = \alpha_2 = 0 \) in the case of the regression (26) and \( H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0 \) in the more general setting (27). As in the ESTR cointegration case, to verify these hypotheses the \( F \) statistic (23) is used with \( k = 2 \) and 4, respectively.

In the case of the second-order logistic error correcting mechanism (10) the resulting conditional 2LSTR ECM is the following:

\[ \Delta Y_t = \rho u_{t-1} \left( \frac{1}{1 + e^{-\gamma(u_{t-1} - c)^2}} \right) + \omega \Delta X_t + \sum_{j=1}^{p} \psi_{yj} \Delta Z_{t-j} + e_t, \quad (28) \]

\[ \rho < 0, \quad \gamma > 0, \quad c > 0. \]

A test for linearity of (28) consists in testing \( H_0: \gamma = 0 \) against \( H_1: \gamma > 0 \). This time the first-order Taylor series approximation around 0 to the function \( F(x) = \frac{1}{1 + e^{-x}} \), where \( x = \gamma(u_{t-1} - c)^2 \), given by \( T_1(x) = \frac{1}{2} + \frac{1}{4} x \), leads to the test equation in the form:

\[ \Delta Y_t = \alpha_1 u_{t-1} + \alpha_3 u_{t-1}^3 + \omega \Delta X_t + \sum_{j=1}^{p} \psi_{yj} \Delta Z_{t-j} + e_t, \quad (29) \]

\[ A \text{ positive value of } \gamma \text{ means that the adjustment process is more mean-reverting in the regime defined by } u_{t-1} \leq c \quad \text{and more persistent if } u_{t-1} > c, \quad \text{while if } \gamma \text{ is negative an inverse situation takes place.} \]
where \( \alpha_1 = \frac{q}{2} - \frac{e^2}{4} \) and \( \alpha_3 = \frac{mp}{4} \). In testing for the joint hypothesis of no cointegration and linearity one may set the null hypothesis as \( H_0: \alpha_1 = \alpha_3 = 0 \) and use the statistic (23) with \( k = 2 \).

To accommodate for deterministic components in the analysed processes one can consider regressions based on the de-meaned and de-trended data, i.e.

\[
Y_t^* = \beta'_x X_t^* + u_t^*,
\]

(30)

\[
Y_t^+ = \beta'_x X_t^+ + u_t^+,
\]

(31)

where the superscripts ‘*’ and ‘+’ denote the de-meaned and de-trended data, respectively (comp. KSS). Then, all the processes in the test equations are replaced with their de-meaned (or de-trended) counterparts, for example:

\[
\Delta Y_t^* = \alpha_1 u^*_{t-1} + \alpha_2 u^2_{t-1} + \alpha_3 u^3_{t-1} + \alpha_4 u^4_{t-1} + \omega \Delta X_t^* + \sum_{j=1}^{\infty} \psi_{ij} \Delta Z_{t-j} + \epsilon_t.
\]

(32)

Beside considering tests in the NEC framework it is also possible to suggest the corresponding residual-based tests, in which the following test equations (or their de-meaned and de-trended versions) are utilized:

\[
\Delta u_t = \alpha_1 u_{t-1} + \alpha_2 u^2_{t-1} + \epsilon_t,
\]

(33)

\[
\Delta u_t = \alpha_1 u_{t-1} + \alpha_2 u^2_{t-1} + \alpha_3 u^3_{t-1} + \epsilon_t,
\]

(34)

\[
\Delta u_t = \alpha_1 u_{t-1} + \alpha_2 u^2_{t-1} + \alpha_3 u^3_{t-1} + \alpha_4 u^4_{t-1} + \epsilon_t,
\]

(35)

\[
\Delta u_t = \alpha_1 u_{t-1} + \alpha_3 u^3_{t-1} + \epsilon_t.
\]

(36)

Then, in testing the null hypothesis of no cointegration and linearity the \( F \) statistic in the form:

\[
F = \frac{\text{SSR}_0 - \text{SSR}_1}{\text{SSR}_0/(n - q)}.
\]

(37)

can be used, where \( \text{SSR}_0 = \sum_{t=1}^{n} \Delta u_t^2 \), \( \text{SSR}_1 \) is the sum of squared residuals of the appropriate test equation and \( q \) is the number of restrictions. In the case of autocorrelated errors the test equations (33)–(34) can be augmented with lagged differences with the number of lags chosen according to standard model selection criteria.

The tests considered so far make it possible to suggest a specification procedure to distinguish between linear, ESTR and LSTR cointegration in a way similar to the well-known Teräsvirta procedure to test for STAR nonlinearity in the univariate context (see Teräsvirta, 1994). The specification technique takes advantage of the general to specific modelling and starts with the most general \( F_4 \) tests of the hypothesis \( H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0 \) concerning parameters in the test equation (35) or (27). If the null is rejected, the general conclusion is that the adjustment process under scrutiny is not a linear unit root process.\(^{10}\) The conclusion, especially combined with results of

\(^{10}\) Taking into consideration that the \( F_4 \) tests may have low power as they are testing four coefficients, when the null is not rejected it might be interesting to execute also the subsequent \( F_3 \) and \( F_2 \) tests as well as the standard Engle-Granger test.
some co-mixing tests, can be thought of as an evidence of (linear or nonlinear) cointegration. The next step consists in testing for the significance of the last parameter $\alpha_4$ in the test equation with the ordinary $t$ statistic, which under stationarity of the underlying process and normally and identically distributed errors has the usual Student’s $t$ distribution. The test of the hypothesis $H_0$: $\alpha_4 = 0$ against the two-sided alternative will provide two main indications. If the null is rejected, first of all we conclude that LSTR cointegration takes place. Secondly, the sign of the parameter $\alpha_4$ will give additional information about the LSTAR adjustment process, i.e. if it is negative, the parameter $\gamma$ is positive and the process under scrutiny is more mean-reverting in the regime defined by $u_{t-1} \leq c$, while if it is positive, the process is more persistent in this regime. If the null is accepted, we turn to testing the conditional hypothesis $H_{03}$: $\alpha_3 = 0 | \alpha_4 = 0$ against $H_1$: $\alpha_3 < 0 | \alpha_4 = 0$. Rejection of $H_{03}$ after acceptance of $H_{04}$ may be treated as an indication for the presence of ESTR (or 2LSTR) cointegration. If the null $H_{03}$ is accepted, the subsequent conditional hypothesis should be tested, i.e. $H_{02}$: $\alpha_2 = 0 | \alpha_3 = \alpha_4 = 0$. After $H_{03}$ have been accepted, a rejection of $H_{02}$ may be treated as an evidence of LSTR cointegration. This time, if the sign of the parameter $\alpha_2$ is negative, one can conclude that the LSTAR adjustment process is more mean-reverting in the regime defined by $u_{t-1} > c$, while if it is positive, the parameter $\gamma$ is positive too and the process behaves ‘more stationary’ if $u_{t-1} \leq c$. If the null $H_{02}$ is accepted, the most likely dynamics is the linear one and one can conclude that linear cointegration takes place.

As was mentioned earlier, a second approach to overcome the nuisance parameter problem is based on a grid search over possible values of the parameters of the transition function under study. For simplicity I concentrate on testing in the NEG framework and in the case of the ESTAR and first-order LSTAR adjustments assume $c = 0$. Then, to test for LSTR cointegration the following test equation can be utilized:

$$\Delta u_t = \rho u_{t-1} \left( 1 - \frac{b}{1 + e^{-\gamma u_{t-1}}} \right) + \varepsilon_t,$$

which is used to compute the test statistic:

$$\inf_{b, \gamma} t = \inf_{(b, \gamma) \in \text{Bd}} \hat{t}_{\rho=0}(b, \gamma).$$

In fact low-order polynomial regressions are rather poor proxies for non-linearity of the STAR type. Due to this the alternative hypothesis nests approximations to many other stationary and nonstationary nonlinear processes and the inference about LSTR and ESTR cointegration is possible only if the STAR family of models is assumed a priori as the DGP of the adjustment process.

The assumption can be easily relaxed leading to a bit more complicated and by far more computationally intensive testing procedure. However, it is worth noticing that this assumption seems to be quite natural in cointegration analysis, where one can expect that an adjustment process behaves differently for negative and positive deviations from a long-term value or, alternatively, that the adjustment is symmetric around 0 but not proportional.
The statistic (39) takes the lowest possible value over a space of relevant values for $b$ and $\gamma$. The parameter $b$ controls the discrepancy between the negative and positive regime, while $\gamma$ decides upon the smoothness of the transition between these regimes. A natural candidate for the set $B$ is a set of equally spaced points in the interval $(0,1)$, for example $B = \{0.01, 0.02, \ldots, 0.99\}$. With such a choice of $B$ admitting only positive values of $\gamma$ means that a more mean-reverting dynamics in the negative regime is expected. If one expects the opposite, he or she may be interested in performing a ‘symmetrical’ test, admitting a less persistent behaviour in the positive regime. To this end, we need to revert the values of $\gamma$ (i.e. assume that $\gamma < 0$). As in practice we usually will not know, which of the two presuppositions is more adequate, we may be interested in executing both tests at once, and base our inference on the minimal value of the two inf $t$ statistics. This is equivalent to admitting both negative and positive values of $\gamma$ for example we may assume $\Gamma = \{-5, -4.95, \ldots, -0.05, 0.05, \ldots, 4.95, 5\}$. As our space for the grid search includes values that approximate the linear autoregression case, we may expect some power gains over the standard Engle-Granger test.\footnote{In fact such power gains for a wide range of economically meaningful values of parameters of an LSTAR adjustment process have been documented in Bruzda (2007). A bootstrap test based on (38) outperforms also the $F$ tests based on the equations (33) and (35), especially for medium and large sample sizes.} It is worth stressing the test based on (39) tests directly against LSTR cointegration, as the grid search is performed over parameter values lying in the range of stationarity of an LSTAR adjustment process. In practice, to implement the test it is advisable to re-parametrize the logistic transition function to make the test statistic approximately scale-free.\footnote{See Teräsvirta (1994), van Dijk, Teräsvirta, Franses (2002), Park, Shintani (2005).} To this end, I suggest using the sample standard deviation of the transition variable as the scaling factor\footnote{Note that this sample standard deviation is not an estimate of the appropriate characteristic of the transition variable as under the null hypothesis this variable is nonstationary.} and base the inference on the equation:

$$\Delta u_t = \rho u_{t-1} \left[ 1 - \frac{b}{1 + e^{-\gamma \left( \frac{u_{t-1}}{s_f} \right)}} \right] + \varepsilon_t,$$

where $s_f = \sqrt{\frac{\sum_{t=1}^{n} \varepsilon_t^2}{n}}$.

In testing for ESTR cointegration the test equation in the re-parametrized form becomes:

$$\Delta u_t = \rho u_{t-1} \left[ 1 - e^{-\frac{u_{t-1}}{s_f} \gamma} \right] + \varepsilon_t,$$

where $s_f$ is defined above, and the test statistic is:

$$\inf_{\gamma \in \Gamma} \hat{\tau} = \inf_{\rho \in \Gamma} \hat{\tau} \rho = 0 (\gamma).$$
This time, however, the set of possible values of the nuisance parameter $\gamma$ comprises positive values only. For example, in the simulation experiments in the next section the set $\Gamma = \{0.05, 0.1, \ldots, 4.95, 5\}$ was used.

In the case of the second-order logistic transition function the following re-parametrized version of the appropriate test equation can be used:

$$
\Delta u_t = \rho u_{t-1} \left\{ \frac{1}{1 + e^{-b \left( \frac{u_{t-1}}{c} \right)^2 - 1}} \right\} + \varepsilon_t, \quad (43)
$$

where $b = \gamma c^2$. Then, to test for 2LSTR cointegration the following statistic is computed:

$$
\inf_{b, c} t = \inf_{(b, c) \in B \times C} \hat{\rho}_{0}(b, c), \quad (44)
$$

where the set B contains some arbitrary positive values (for example, in the subsequent sections I apply $B = \{5, 10, 15, 25, 30\}$), while C is chosen in a data-dependent way (for example, in the analysis below I admit 10 equally-spaced points between 0 and the 0.85 quantile of the empirical distribution of the transition variable).

Results obtained by Bec, Ben Salem and Carrasco (2004) as well as Park and Shintani (2005) in the univariate context suggest that it could be possible to base our statistical inference, utilizing the suggested inf $t$ tests, on an asymptotic distribution of these test statistics. However, taking into account that such an asymptotic distribution depends on random limits of the sets of possible values of the nuisance parameters and that in small and medium-sized samples size distortions of the inf $t$ tests can be substantial \(^{17}\), it seems that designing a bootstrap version of the tests may constitute a better solution. In what follows, I utilize a residual-based block bootstrap method, in which the following triangular system is considered:

$$
\Delta y_t = \Psi \Delta x_t + f(u_{t-1}) + \eta_t, \quad (45)
$$

$$
\Delta x_t = v_t, \quad (46)
$$

where $u_t = y_t - \beta x_t$ and $f(u_{t-1}) = \phi u_{t-1} + g(u_{t-1})$. Having computed residuals $\hat{u}_t$, the equation (45) may be fitted to the data as to minimize the residual sum of squares conditionally on a grid search over possible values of the parameters of the transition function under study. Then, to allow for the possible temporal dependence in $\eta_t$ and $\nu_t$, as well as for the dependence between these two processes, blocks of the pairs $(\hat{\eta}_t, \hat{\nu}_t)$ are resampled and bootstrap samples are constructed by imposing the null hypothesis, i.e. through the equations:

$$
\Delta y_t^* = \hat{\Psi} \Delta x_t^* + \hat{\eta}_t^*, \quad (47)
$$

$$
\Delta x_t^* = v_t^*, \quad (48)
$$

\(^{16}\) Compare Bec, Ben Salem, Carrasco (2004) in a slightly different context.

\(^{17}\) Compare Park, Shintani (2005).
where \( \tilde{\eta}_t^\ast \) and \( \tilde{v}_t^\ast \) are resampled from \( \tilde{\eta}_t \) and \( v_t \), respectively. Both – fixed and stationary – block bootstrap may be applied. If deterministic components are present in the analysed processes, the bootstrap samples should be de-meaned and de-trended as well.\(^{18}\) This bootstrap method corresponds to the sampling scheme \( S_2 \) according to the classification in Li, Maddala (1997), in which the so-called unrestricted residuals are utilized. Alternatively, the scheme \( S_3 \) may be applied, in which the residuals \( \tilde{\eta}_t \) are obtained by imposing the null hypothesis in the equation (45), i.e. by estimating \( \Delta Y_t = \Psi \Delta X_t + \eta_t \).\(^{19}\) In the next section the two bootstrap schemes are examined in the context of the ESTR and 2LSTR inf \( t \) tests. Additionally, their overall performance is compared with the appropriate results of the \( F \) tests for STR cointegration. Only a limited number of experiments is presented.

### 3 Simulation analysis

In the first part of the analysis the size of the bootstrap ESTR and 2LSTR cointegration inf \( t \) tests was examined. To this end, the following data generating process (DGP) was used:

\[
\begin{align*}
\Delta Y_t &= \lambda \Delta X_t + \epsilon_t, \\
\Delta X_t &= v_t, \\
u_t &= Y_t - \beta X_t, \\
\begin{bmatrix} \epsilon_t \\
v_t \end{bmatrix} &\sim n.i.d. \left( 0, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right).
\end{align*}
\] (49)

where \( \lambda = 1; 0 \), \( \beta = 1 \), \( \sigma_1 = 1 \), \( \sigma_2 = 1; 2 \). In each case series of length 200 were simulated with initial values taken randomly from the \( N(0, 1) \) distribution and the first 100 observations were rejected. The de-meaned one-regressor case was exclusively considered here. The number of simulations was set to 1000 and the number of bootstrap replications was 200. In the case of the ESTR cointegration tests the equation (41) and its scale-adjusted NEC counterpart were utilized, while the 2LSTR cointegration was examined on the base of the equation (43) and the corresponding NEC model. The definitions of the sets \( \Gamma, B \) and \( C \) were as discussed in the previous section, i.e. \( \Gamma = \{0.05, 0.1, \ldots, 4.95, 5\} \), \( B = \{5, 10, 15, 25, 30\} \), while \( C \) comprised 10 equally-spaced points between 0 and the 0.85 quantile of the empirical distribution of the transition variable.

\(^{18}\) See Kapetanios (2003).

\(^{19}\) There is also a third method, applicable to the first-order logistic case, which makes use of residuals obtained by imposing the constraint in the form \( b = 0 \). Simulation analyses presented in Bruzda (2007) show that both the scheme \( S_2 \) and the third ‘restricted nonlinearity’ method, when applied to the LSTR cointegration tests in the NEG and NEC frameworks, are generally superior to \( S_3 \) in terms of power and do not cause significant size distortions. Additionally, the ‘restricted nonlinearity’ method produces results that are less sensitive to the parameters of the first-order logistic function.
The results of this experiment for $\lambda = 0$ are given in Table A1 in the Appendix. They clearly point at the unrestricted residuals method as a suitable one for practical applications. This method does not produce substantial size distortions, except maybe for the case of the ESTR cointegration test in the NEG framework, where a slight tendency to underreject the null was observed. On the other hand, the restricted residuals method significantly underrejects the null hypothesis, especially in the case of the ESTR cointegration tests (much better but still unsatisfactory results were obtained for the 2LSTR cointegration tests).

Turning to power evaluation of the ESTR and 2LSTR cointegration tests I consider the following data generating processes:

\[
\begin{align*}
\Delta Y_t &= \lambda \Delta X_t + \rho u_{t-1} \left[ 1 - e^{-\gamma u_{t-1}^2} \right] + \varepsilon_t, \\
\Delta X_t &= v_t, \\
u_t &= Y_t - \beta X_t,
\end{align*}
\]

\[
\begin{pmatrix}
\varepsilon_t \\
v_t
\end{pmatrix} \sim n.i.i.d. \left( 0, \begin{bmatrix}
\sigma_1^2 & 0 \\
0 & \sigma_2^2
\end{bmatrix} \right),
\]

where $\lambda = 1; 0$, $\gamma = 0; 1$, $\beta = 1$, $\sigma_1 = 1$, $\sigma_2 = 1$, and $\rho = -0.025; 0; 0.5$, and

\[
\begin{align*}
\Delta Y_t &= \lambda \Delta X_t + \rho u_{t-1} \left[ \frac{1}{1 + \left( \frac{u_{t-1}}{c} \right)^2} \right] + \varepsilon_t, \\
\Delta X_t &= v_t, \\
u_t &= Y_t - \beta X_t,
\end{align*}
\]

\[
\begin{pmatrix}
\varepsilon_t \\
v_t
\end{pmatrix} \sim n.i.i.d. \left( 0, \begin{bmatrix}
\sigma_1^2 & 0 \\
0 & \sigma_2^2
\end{bmatrix} \right),
\]

where $\lambda = 1; 0$, $\beta = 1$, $\sigma_1 = \sigma_2 = 1$, $\rho = -0.5; 0; 0.5$, $b = 5; 15; 25$ and $c = 1; 2.5$. As previously, series of length 100 were examined. They were obtained after simulating series of length 200 and dropping first 100 initial observations. The de-meaned one-regressor case was only considered. In the case of the DGP (50) the residual- and ECM-based bootstrap tests for ESTR and 2LSTR cointegration as well as the $F$ tests for the STR cointegration in the NEG and NEC frameworks were investigated. Additionally, the appropriate results for the Enders-Siklos (ES) threshold cointegration test and the Engle-Granger (EG) linear cointegration test are presented. In the case of the DGP (51) the residual- and ECM-based bootstrap tests and the Engle-Granger procedure were exclusively examined. The number of overall replications was set to 1000 for the DGP (50) and 500 for the DGP (51), while that of bootstrap replications was 200. If $\lambda = 0$, the common factor restriction is not fulfilled and it might be expected that the ECM-based tests will perform better than the residual-based tests. The results for $\lambda = 0$

---

\(\lambda\) is a parameter in the data generating process that controls the degree of cointegration. Its value affects the performance of the cointegration tests. When $\lambda = 0$, the common factor restriction is not fulfilled, and the ECM-based tests are expected to perform better than the residual-based tests.
are presented in Tables A2–A4 in the Appendix, while for $\lambda = 1$ (as well as for time series of length 250 and for the de-trended case) are available upon request. In the experiment concerning the DGP (50) simulated small-sample critical values for the $F$ tests were used. The implementation of the bootstrap $t$ tests was as in the size examination. The main results of power evaluation can be summarized as follows:

- As for the bootstrap tests the overall impression is that the unrestricted residuals method provides much better results as compared with the restricted residuals method, especially when the adjustment processes are more persistent.
- As expected, the $t$ and $F$ tests based on NEC models provide substantial power gains over the NEG tests, when the COMFAC restriction is violated. However, further results, not presented here, indicate that there exist some (slight) power gains even if the common factor restriction is fulfilled. The power superiority of the NEC tests increases with the increase in the variance of the innovation in $X_t$.
- The most surprising finding is that in majority of cases (with the only exception being the ‘nearly unit root’ case with increased innovation variance, when tested in the NEG framework) the $t$ tests for 2LSTR cointegration outperform the $t$ ESTR cointegration test even if the true DGP was (50), i.e. even if an ESTAR adjustment was assumed. Except for the case mentioned above, the 2LSTR tests dominate also the $F$ tests and the Enders-Siklos and Engle-Granger procedures.
- The $t$ tests for ESTR cointegration provide slightly better results than the $F$ tests if $\rho = -0.1$. In the other cases the two families of tests give similar results or even the $F$ tests outperform the $t$ tests occasionally. However, it is worth mentioning that further results indicate that with the increase of the sample size the $t$ tests eventually dominate the $F$ tests.
- It seems that the ES and EG tests provide results which are more or less similar to that obtained with the $F$ tests. Among the $F$ tests best power is usually associated with more parsimonious specifications.
- In the case of the DGP (51) one can observe that the power performance of all tests depends crucially on the value of the parameter $c$ and becomes worse with the increase of $c$.

Two practical indications result form the above simulation analysis. Firstly, it seems that for medium and large-sized data samples the best way to test for smooth transition cointegration are the bootstrap $t$ tests considered here. Secondly, if one is interested in testing for cointegration in the presence of transaction costs, a better way to perform such an analysis are tests for 2LSTR cointegration than tests for ESTR cointegration. The second indication is additionally supported by the observation, that a limiting case of the second-order LSTAR process is the SETAR process suggested by Balke and Fomby (1997) to examine the hypothesis of transaction costs. Due to this 2LSTR cointegration nests the three-regime threshold cointegration as its limiting case. In the next section this findings are confronted with some empirical data.
4 Empirical results

In the empirical investigation weekly observations (Monday to Monday) on the following interest rates were analysed: WIBOR (Warsaw Inter-Bank Offered Rate) for 1-week, 1-month and 1-year deposits, 52-week Treasury bill rates and yields on 2-year and 10-year benchmark government bonds (see Fig. 2). The period of analysis was set arbitrary from January 2001 till March 2005 (218 observations). Further in the text the following abbreviations are used: W1W, W1M, W1Y – WIBOR for 1-week, 1-month and 1-year deposits, respectively, TBILLS – Treasury bill rates and O2Y, O10Y – yields on the appropriate benchmark bonds.

Fig. 2 Weekly observations on Polish interest rates for instruments with different maturities

In the preliminary analysis the 6 series were tested for unit roots with the help of the ADF and KPSS tests. All the series turned out to be integrated of order one. Further the Engle-Granger test for cointegration was executed on OLS residuals of the following equations:

\[ r_t = \alpha_0 + \alpha_1 R_t + \eta_t, \]  
\[ r_t = \alpha_0 + \alpha_0 t + \alpha_1 R_t + \eta_t. \]

where \( r_t \) denotes a short-term interest rate and \( R_t \) stands for a long-term interest rate.

The results are given in Table 1. This univariate linear testing strategy was able to find long-term relationships in 3 out of 14 cases for deposits with relatively close maturities.

Table 1. The results of the Engle-Granger cointegration tests

<table>
<thead>
<tr>
<th>Short- vs. long-term interest rates</th>
<th>W1M</th>
<th>W1Y</th>
<th>TBILLS</th>
<th>O2Y</th>
<th>O10Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equation (52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1W</td>
<td>-4,538** (6)</td>
<td>-2,208 (8)</td>
<td>-2,344 (5)</td>
<td>-2,375 (5)</td>
<td>-2,796 (5)</td>
</tr>
<tr>
<td>W1M</td>
<td>-2,094 (3)</td>
<td>-2,491 (0)</td>
<td>-2,486 (4)</td>
<td>-2,632 (0)</td>
<td>-2,958 (3)</td>
</tr>
<tr>
<td>W1Y</td>
<td>-3,062 (4)</td>
<td>-2,958 (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBILLS</td>
<td>-3,332* (1)</td>
<td>-1,599* (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2Y</td>
<td>-3,022 (4)</td>
<td>-2,952 (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equation (53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1W</td>
<td>-4,272** (6)</td>
<td>-2,380 (8)</td>
<td>-2,595 (5)</td>
<td>-2,430 (5)</td>
<td>-1,739 (8)</td>
</tr>
<tr>
<td>W1M</td>
<td>-1,959 (3)</td>
<td>-1,784 (3)</td>
<td>-2,296 (4)</td>
<td>-1,978 (0)</td>
<td>-2,491 (3)</td>
</tr>
<tr>
<td>W1Y</td>
<td>-2,993 (4)</td>
<td>-2,491 (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBILLS</td>
<td>-3,492 (1)</td>
<td>-2,752 (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Critical values at the 5% and 10% level of significance for the equations without trend are: –3.365 and –3.065, while for the equations with trend are: –3.824 and –3.529 (see McKinnon, 1991); ‘*’ (‘**’) denotes rejection at the 5% (10%) critical level; in brackets are given augmentations in the Dickey-Fuller test equations.

Taking into consideration that linear cointegration tests lack their power, when adjustment processes are nonlinear, in the next step I turned to tests for STR cointegration. Two kinds of tests were performed: the \( \inf t \) tests and the \( F \) tests. As the Engle-Granger tests based on (53) provided generally worse results as compared to the case without deterministic trend, the de-meaned case was exclusively considered. In order to retain similarity with the Engle-Granger procedure only NEG tests were performed. The results of the nonlinear cointegration analysis are given in Tables 2–3.

Table 2. The results of the \( \inf t \) tests for STR cointegration

<table>
<thead>
<tr>
<th>Short- vs. long-term interest rates</th>
<th>W1M</th>
<th>W1Y</th>
<th>TBILLS</th>
<th>O2Y</th>
<th>O5Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSTR cointegration tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESTR cointegration tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1M</td>
<td>-5.127 [-5.511; -5.239]</td>
<td>-5.127 [-5.511; -5.239]</td>
<td>-5.127 [-5.511; -5.239]</td>
<td>-5.127 [-5.511; -5.239]</td>
<td>-5.127 [-5.511; -5.239]</td>
</tr>
<tr>
<td>2LSTR cointegration tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\('*‘ (‘**’) denotes rejection at the 5% (10%) critical level; in square brackets bootstrap 5% and 10% critical values are reported; the block length in the overlapping block bootstrap was set arbitrary to 20, non-
augmented test equations were used and the number of bootstrap replications was set to 500; the definitions of the sets of possible values of the nuisance parameter were as follows: in the LSTR cointegration test $\Gamma = \{-5, -4.95, \ldots, -0.05, 0.05, \ldots, 4.95, 5\}$, $B = \{0.1, 0.2, \ldots, 0.9\}$, in the ESTR cointegration test $\Gamma = \{0.025, 0.05, \ldots, 4.975, 5\}$ and in the 2LSTR cointegration tests $B = \{5, 10, 15, 25, 30\}$, while $C$ comprised 10 equally-spaced points between 0 and the 0.85 quantile of the empirical distribution of the transition variable.

Table 3. The results of the $F$ tests for STR cointegration

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$F_{NEG4}$</th>
<th>$t$ (p-value)</th>
<th>$F_{NEG3}$</th>
<th>$t$ (p-value)</th>
<th>$F_{NEG2}$</th>
<th>$t$ (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>W1W–W1M</td>
<td>7.372**</td>
<td>-1.427 (0.155)</td>
<td>9.105**</td>
<td>-0.572 (0.568)</td>
<td>13.538**</td>
<td>3.218 (0.002)</td>
</tr>
<tr>
<td>W1W–W1Y</td>
<td>1.321</td>
<td>1.764</td>
<td>2.654</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1W–TBILLS</td>
<td>1.845</td>
<td>2.468</td>
<td>3.498</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1W–O2Y</td>
<td>2.156</td>
<td>2.871</td>
<td>4.215</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1W–O10Y</td>
<td>3.663**</td>
<td>1.134 (0.258)</td>
<td>4.449**</td>
<td>-0.933 (0.352)</td>
<td>6.243**</td>
<td>-2.129 (0.035)</td>
</tr>
<tr>
<td>W1M–W1Y</td>
<td>3.093*</td>
<td>-1.389 (0.166)</td>
<td>3.466</td>
<td></td>
<td>3.331</td>
<td></td>
</tr>
<tr>
<td>W1M–TBILLS</td>
<td>5.477**</td>
<td>-2.501 (0.013)</td>
<td>5.090**</td>
<td>-2.723 (0.007)</td>
<td>3.810</td>
<td></td>
</tr>
<tr>
<td>W1M–O2Y</td>
<td>2.879</td>
<td>3.737</td>
<td>4.809*</td>
<td>-1.835 (0.068)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1M–O10Y</td>
<td>2.335</td>
<td></td>
<td>3.773</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1Y–O2Y</td>
<td>3.664**</td>
<td>1.022 (0.308)</td>
<td>4.536**</td>
<td>-0.641 (0.522)</td>
<td>6.618**</td>
<td>-1.932 (0.055)</td>
</tr>
<tr>
<td>W1Y–O10Y</td>
<td>3.034</td>
<td>2.972</td>
<td>4.364</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBILLS–O2Y</td>
<td>4.506**</td>
<td>0.763 (0.446)</td>
<td>5.825**</td>
<td>-0.444 (0.658)</td>
<td>8.672**</td>
<td>-2.446 (0.015)</td>
</tr>
<tr>
<td>TBILLS–O10Y</td>
<td>3.113*</td>
<td>1.495 (0.136)</td>
<td>3.386</td>
<td></td>
<td>4.978*</td>
<td>0.141 (0.888)</td>
</tr>
<tr>
<td>O2Y–O10Y</td>
<td>2.941</td>
<td>3.311</td>
<td>4.970*</td>
<td>0.902 (0.368)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* (**) denotes rejection at the 5% (10%) critical level with the help of simulated critical values for $n = 250$ and the number of regressors $k = 1$: 3,541 (5%), 3,068 (10%) for the $F_4$ test; 4,418 (5%), 3,818 (10%) for the $F_3$ test; 5,652 (5%), 4,766 (10%) for the $F_2$ test; the critical values were obtained by simulating independent random walks, computing OLS residuals for the de-meaned data and performing the appropriate tests; 5000 replications were executed; augmentation was the same as in the linear cointegration tests (see Table 1); if the null hypothesis of the $F$ test was rejected, the standard $t$ test for significance of the last parameter in the test equation was also performed; numbers in brackets are $p$-values.

The results of the STR cointegration tests provide much stronger evidence in favour of the expectation hypothesis in the term structure of interest rates as compared to the linear cointegration tests. In particular, the LSTR and 2LSTR inf $t$ tests point quite strongly at, respectively, 8 and 9 long-term relationships among 14 examined regressions. The two tests gave almost the same results, what may seem surprising, as they assume two different kinds of dynamics of adjustment processes: one corresponding to an asymmetric 2-regime behaviour and second admitting a symmetric 3-regime dynamics. The ESTR inf $t$ tests gave much worse results as compared to the 2LSTR cointegration tests. This finding, however, confirms the results of the simulation analysis presented in the previous section, which document a better power performance of the 2LSTR inf $t$
test over the ESTR test. The general $F_4$ tests reject their null in 7 out of 14 cases, but the joint confirmation of the presence of cointegration obtained with the $\inf t$ LSTR (or 2LSTR) and the $F$ cointegration tests takes place in 4 cases only. This may result from a strong ARCH effect, which is observed in the data. The ARCH effect is likely to distort results of the $F$ tests. When the $F$ tests reject their null hypotheses, they seem to favour a 2-regime LSTAR specification instead of an ESTAR model for the adjustment processes (see, for example, the ‘W1W–W1M’ and ‘TBILLS–O2Y’ cases in Table 3). However, it is worth stressing that the discrimination between the LSTR and ESTR cointegration on the base of the $F$ tests is not an easy task, as the tests have usually lower power than the corresponding $\inf t$ tests.\textsuperscript{21}

5 Conclusions

The joint study of nonstationarity and nonlinearity is very important for different types of empirical analyses. As standard unit root and cointegration tests lack their power and show serious size distortions when nonlinearities are at play, there is a need to develop new tests, which might help to find long-term relationships, where linear-based tests fail. The STR cointegration tests considered here inscribe to this strand of research. They can be viewed as supplementary to two- and three-regime threshold cointegration tests, having two properties worth stressing. Firstly, the tests are more general as they nest threshold cointegration as a limiting case. Secondly, in aggregate quantities smooth behaviour seems to be more likely than sharp and, due to this, the STR cointegration tests might be somewhat more adequate to study macroeconomic and financial phenomena.

The empirical analysis presented in the paper shows that the expectation hypothesis of the term structure of interest rates in the Polish financial market is fulfilled for pairs of yields on instruments with relatively close maturities. The nonlinear cointegration tests considered here provide much stronger evidence in favour of the expectation hypothesis in the term structure as compared to the linear cointegration tests. They seem to favour an LSTAR adjustment rather than an ESTAR model for equilibrium errors. This observation stays in accordance with the hypothesis saying that a nonlinear dynamics of interest rate spreads results from policymakers’ actions.

References


\textsuperscript{21} See Section 3 about the power properties of the $F$ tests in the case of an ESTAR adjustment, as well as Bruzda (2007) about the power performance of these tests in the case of first-order logistic STAR adjustments.


Bruzda J. (2007), Bootstrap LSTR cointegration tests. Simulation analysis and an application to money demand modelling, unpublished manuscript, Nicolaus Copernicus University, Torun.


Appendix

Table A1 Size of bootstrap ESTR and 2LSTR cointegration tests (in %)

<table>
<thead>
<tr>
<th>$\sigma_\epsilon$</th>
<th>Tests for ESTR cointegration</th>
<th>Tests for 2LSTR cointegration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>inf $\hat{b}_{2LSTR}$</td>
<td>inf $\hat{b}_{ESTR}$</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Unrestricted residuals method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>8.2</td>
<td>3.3</td>
</tr>
<tr>
<td>2</td>
<td>8.4</td>
<td>3.3</td>
</tr>
<tr>
<td>Restricted residuals method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3.9</td>
<td>1.1</td>
</tr>
<tr>
<td>2</td>
<td>4.1</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Table A2 Power of bootstrap ESTR and 2LSTR cointegration tests (in %); DGP (50)

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\sigma_\epsilon$</th>
<th>$\gamma$</th>
<th>Tests for ESTR cointegration</th>
<th>Tests for 2LSTR cointegration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>inf $\hat{b}_{2LSTR}$</td>
<td>inf $\hat{b}_{ESTR}$</td>
<td>inf $\hat{b}_{2LSTR}$</td>
<td>inf $\hat{b}_{ESTR}$</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>5</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Unrestricted residuals method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.025</td>
<td>1</td>
<td>1</td>
<td>9.3</td>
<td>4.7</td>
</tr>
<tr>
<td>-0.025</td>
<td>1</td>
<td>0.1</td>
<td>10.3</td>
<td>5.3</td>
</tr>
<tr>
<td>-0.025</td>
<td>2</td>
<td>1</td>
<td>7.0</td>
<td>3.6</td>
</tr>
<tr>
<td>-0.025</td>
<td>2</td>
<td>0.1</td>
<td>7.3</td>
<td>4.4</td>
</tr>
<tr>
<td>-0.1</td>
<td>1</td>
<td>1</td>
<td>27.3</td>
<td>14.6</td>
</tr>
<tr>
<td>-0.1</td>
<td>1</td>
<td>0.1</td>
<td>20.8</td>
<td>11.5</td>
</tr>
<tr>
<td>-0.1</td>
<td>2</td>
<td>1</td>
<td>21.1</td>
<td>11.7</td>
</tr>
<tr>
<td>-0.1</td>
<td>2</td>
<td>0.1</td>
<td>16.9</td>
<td>7.9</td>
</tr>
<tr>
<td>-0.5</td>
<td>1</td>
<td>1</td>
<td>99.9</td>
<td>99.9</td>
</tr>
<tr>
<td>-0.5</td>
<td>1</td>
<td>0.1</td>
<td>98.9</td>
<td>94.8</td>
</tr>
<tr>
<td>-0.5</td>
<td>2</td>
<td>1</td>
<td>100</td>
<td>99.9</td>
</tr>
<tr>
<td>-0.5</td>
<td>2</td>
<td>0.1</td>
<td>100</td>
<td>99.5</td>
</tr>
<tr>
<td>Restricted residuals method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table A3 Power of $F$ tests for STR cointegration (in %); DGP (50)

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$\sigma$</th>
<th>$\gamma$</th>
<th>$F_{\text{NEC}}$</th>
<th>$F_{\text{NEG}}$</th>
<th>$F_{\text{ES}}$</th>
<th>$F_{\text{CUC}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.025</td>
<td>1</td>
<td>1</td>
<td>8.9</td>
<td>9.7</td>
<td>10.0</td>
<td>10.1</td>
</tr>
<tr>
<td>0.025</td>
<td>1</td>
<td>0.1</td>
<td>9.4</td>
<td>9.9</td>
<td>10.5</td>
<td>10.9</td>
</tr>
<tr>
<td>0.025</td>
<td>2</td>
<td>1</td>
<td>7.5</td>
<td>6.7</td>
<td>7.2</td>
<td>6.5</td>
</tr>
<tr>
<td>0.025</td>
<td>2</td>
<td>0.1</td>
<td>7.6</td>
<td>3.4</td>
<td>3.9</td>
<td>0.8</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>1</td>
<td>24.4</td>
<td>14.8</td>
<td>14.4</td>
<td>12.1</td>
</tr>
<tr>
<td>0.025</td>
<td>1</td>
<td>0.1</td>
<td>18.6</td>
<td>11.0</td>
<td>7.7</td>
<td>10.7</td>
</tr>
<tr>
<td>0.025</td>
<td>1</td>
<td>0.1</td>
<td>16.8</td>
<td>9.7</td>
<td>2.8</td>
<td>18.8</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>0.1</td>
<td>13.7</td>
<td>6.8</td>
<td>1.0</td>
<td>13.3</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>0.1</td>
<td>98.9</td>
<td>95.3</td>
<td>67.4</td>
<td>95.3</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>1</td>
<td>95.2</td>
<td>99.7</td>
<td>93.0</td>
<td>99.7</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>0.1</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>1</td>
<td>99.9</td>
<td>95.2</td>
<td>93.0</td>
<td>99.9</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>0.1</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>1</td>
<td>93.0</td>
<td>99.9</td>
<td>99.3</td>
<td>99.7</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>0.1</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>1</td>
<td>99.9</td>
<td>95.2</td>
<td>93.0</td>
<td>99.9</td>
</tr>
<tr>
<td>0.025</td>
<td>0.1</td>
<td>0.1</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Table A4 Power of bootstrap ESTR and 2LSTR cointegration tests (in %); DGP (51)

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$c$</th>
<th>$b$</th>
<th>Tests for ESTR cointegration</th>
<th>Tests for 2LSTR cointegration</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>1</td>
<td>5</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td>0.01</td>
<td>1</td>
<td>15</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td>0.01</td>
<td>1</td>
<td>25</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.5</td>
<td>1</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.5</td>
<td>15</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.5</td>
<td>25</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.5</td>
<td>1</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.5</td>
<td>15</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.5</td>
<td>25</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.5</td>
<td>1</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.5</td>
<td>15</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.5</td>
<td>25</td>
<td>$F_{\text{NEC}}$</td>
<td>$F_{\text{NEG}}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>-0.5</td>
<td>1</td>
<td>15</td>
<td>97.4</td>
<td>86.0</td>
</tr>
<tr>
<td>-0.5</td>
<td>1</td>
<td>25</td>
<td>97.0</td>
<td>85.8</td>
</tr>
<tr>
<td>-0.5</td>
<td>2.5</td>
<td>5</td>
<td>81.4</td>
<td>50.8</td>
</tr>
<tr>
<td>-0.5</td>
<td>2.5</td>
<td>15</td>
<td>81.8</td>
<td>54.8</td>
</tr>
<tr>
<td>-0.5</td>
<td>2.5</td>
<td>25</td>
<td>81.8</td>
<td>54.8</td>
</tr>
<tr>
<td>-0.1</td>
<td>1</td>
<td>5</td>
<td>22.4</td>
<td>12.0</td>
</tr>
<tr>
<td>-0.1</td>
<td>1</td>
<td>15</td>
<td>24.4</td>
<td>14.0</td>
</tr>
<tr>
<td>-0.1</td>
<td>2.5</td>
<td>5</td>
<td>17.8</td>
<td>8.8</td>
</tr>
<tr>
<td>-0.5</td>
<td>1</td>
<td>5</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>-0.5</td>
<td>1</td>
<td>15</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>-0.5</td>
<td>1</td>
<td>25</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>-0.5</td>
<td>2.5</td>
<td>5</td>
<td>99.6</td>
<td>97.6</td>
</tr>
<tr>
<td>-0.5</td>
<td>2.5</td>
<td>15</td>
<td>99.8</td>
<td>97.6</td>
</tr>
<tr>
<td>-0.5</td>
<td>2.5</td>
<td>25</td>
<td>99.8</td>
<td>97.6</td>
</tr>
</tbody>
</table>