Testing hypotheses in an $I(2)$ model with piecewise linear trends. An analysis of the persistent long swings in the Dmk/$ rate

Søren Johansen,* Economics Department, University of Copenhagen and CREATES, University of Aarhus
Katarina Juselius, Economics Department, University of Copenhagen
Roman Frydman, Department of Economics, New York University
Michael Goldberg, Department of Economics, University of New Hampshire

May 27, 2009

Abstract

This paper discusses the $I(2)$ model with breaks in the deterministic component and illustrates with an analysis of German and US prices, exchange rates, and interest rates in 1975–1999. It provides new results on the likelihood ratio test of overidentifying restrictions on the cointegrating relations when they contain piecewise linear trends. One important aim of the paper is to demonstrate that a structured $I(2)$ analysis is useful for a better understanding of the empirical regularities underlying the persistent swings in nominal exchange rates, typical in periods of floating exchange rates.

JEL:C32, C52, F41

Keywords: PPP, Long swings puzzle, Cointegrated VAR, Test of Overidentification

*Corresponding author: Søren Johansen, Department of Economics, University of Copenhagen, Studiestræde 6, 1455 Copenhagen K, Denmark. Tlf: +45 35323071. Email: soren.johansen@econ.ku.dk

1 Introduction

Over the past three decades, floating currencies have shown a tendency to undergo persistent swings away from purchasing power parity (PPP) for extended periods of time that are followed by periods in which exchange rates move persistently back toward this benchmark.[1] These long swings in real exchange rates have puzzled international macroeconomists for a long time. The literature has attempted to account for such fluctuations with an REH sticky-price monetary model, such as Dornbusch (1976) or one of its New Open Economy Macroeconomic (NOEM) formulations, see Lane (2001) for a review article of the NOEM literature.

These models typically imply that the nominal exchange rate and relative goods prices are unit-root processes, while the real exchange rate is stationary. This theory has led empirical researchers to make use of an $I(1)$ model to analyze the data. Most report that the $I(1)$ null hypothesis cannot be rejected for the nominal exchange rate and relative goods prices, but that it can for the real exchange rate in favor of stationarity. They also report that the real exchange rate, though stationary, is very persistent, i.e., a near-$I(1)$ process.

However, assuming that an $I(1)$ model is appropriate, rather than *testing* whether the $I(2)$ condition can be rejected on open-economy macroeconomic data sets, is problematic. When data are approximately $I(2)$ rather than $I(1)$ a conclusion that the real exchange rate is stationary (or stationary but highly persistent) reached in the $I(1)$ model is likely to be misleading. This is because when an $I(1)$ model is applied to $I(2)$ data, cointegration is not from $I(1)$ to $I(0)$, but from $I(2)$ to $I(1)$. Thus, the 'consensus' conclusion in the literature that real exchange rates are stationary but highly persistent, might have to be replaced by the conclusion that the *change* in real exchange rate is stationary but highly persistent. Indeed, the results of Juselius's (2008) cointegrated VAR analysis of German and US goods prices and exchange rates over the post Bretton Woods period show that this is the case. When the $I(1)$ model is applied, nonetheless, the choice of $r = 1$ (suggested by the trace test) leaves an unrestricted characteristic root of 0.99 in the model, essentially demonstrating that the estimated long-run relation is in fact $CI(2,1)$ rather than $CI(1,1)$. The paper concludes that an empirical understanding of currency swings is not likely to be reached without allowing for an $I(2)$ component and including interest rates.

That interest rates are important for a full understanding of the long swings puzzle also follows from the monetary model of Frydman and Goldberg (2007, 2008), which replaces the Rational Expectations Hypothesis (REH) with an Imperfect Knowledge Economics (IKE) representation of forecasting behavior. In Frydman et al. (2009), we show that this IKE model implies that relative goods prices, as well as nominal and real exchange rates, are $I(2)$ and that there is a cointegrating relationship between the real exchange rate and the real interest rate differential. All of these results are

[1]The PPP puzzle is the inability of one model to account for both the high persistence and the high volatility of real exchange rates. For reviews of this literature, see Rogoff (1996), Taylor and Taylor (2004), Sarno and Taylor (2003), and Mark (2001).

found in this paper. They are strengthened by other studies that find $I(2)$ trends in other data sets that include exchange rates, goods prices, and money supplies, see Juselius (1994), Kongsted (2003, 2005), Kongsted and Nielsen (2004), and Bacchiocchi and Fanelli (2005).

Thus, there are two competing economic theories attempting to explain the pronounced persistence in exchange data, one claiming real exchange rates are stationary, though highly persistent, i.e. near $I(1)$, the other claiming that the *change* of real exchange rates is $I(1)$, but highly persistent, i.e. the *level* of the real exchange is near $I(2)$. The cointegrated VAR model seems ideal for discriminating between these two views because hypotheses of $I(0)$, $I(1)$, and $I(2)$ can be formulated within the unrestricted VAR without having to impose any of them from the outset.

The paper tests both the $I(1)$ and the $I(2)$ condition in a cointegrated VAR model that describes German and US prices, exchange rates and interest rates in the period 1975-1999. Since the $I(2)$ condition cannot be rejected, the data are structured into three different levels of persistence. This enables us to confront the basic assumptions underlying the competing theories with the data.

The reunification of Germany in 1991 was a major institutional event which is likely to have caused a structural break in the data and the test for overidentifying restrictions on the cointegration structure is developed for the $I(2)$ model with breaks in the deterministic component. The test procedures discussed in this paper build on previous work in Johansen (1992, 1995, 1997, 2006a), Rahbek et al. (1999), Paruolo (2000, 2002), Nielsen and Rahbek (2007).

The paper is organized as follows. Section 2 motivates the need for why the $I(2)$ model needs deterministic components with breaks. Section 3 discusses the Maximum Likelihood parametrization of the $I(2)$ model, while Section 4 shows how to test structural hypotheses in that model. Section 5 estimates an unrestricted VAR model with deterministic components containing breaks and tests the final model specification. Section 6 discusses the choice of the two reduced rank indices determining the number of $I(1)$ and $I(2)$ trends in the model. Section 7 reports a number of test results based on non-identifying hypotheses as a general description of $CI(2,1)$ relationships in the data. Section 8 reports an overidentified long-run structure of polynomially cointegrated relations and Section 9 uses the MA representation to discuss whether changes in real exchange rate and long interest rate differential can be considered stationary but highly persistent. Section 10 concludes.

2 Deterministic components and the $I(2)$ model

A proper specification of deterministic terms in the $I(2)$ model is necessary for the model to fit the data and hence yield statistically good estimates. From the outset, an $I(2)$ model generates a linear deterministic trend from the initial values, but no trend in the long-run relations. Therefore, if a trend is needed in the cointegrating relations it has to be explicitly modelled. See Rahbek et al. (1999) for a discussion.

<Figure 1 here>

Figure 1, upper panel, shows the graphs of the price differential and the nominal exchange rate and illustrates the tendency in the latter to undergo long swings. Three features stand out: (1) the downward sloping stochastic trend in price differentials; (2) the big persistent swings in the nominal exchange rate evolving around a similar downward sloping trend as in relative prices; and (3) an indication of a change in the slope of the relative price trend around 1991 (possibly also around 1980-81). The graph in Figure 1, lower panel, of the real exchange rate, $ppp = p_1 - p_2 - s$, shows that the downward sloping trend in relative prices has been approximately canceled by the similar trend in the nominal exchange rate. The long persistent swings remain essentially unchanged.

The change in the relative price trend mentioned in (3) is likely to be associated with the German re-unification in 1991:1. This was a very significant event which is likely to have strongly affected German, but not US, prices. The merging of East and West German prices may have produced a step effect in German prices which, as long as it is purely technical, should be removed prior to the VAR analysis. The possibility of an additive step effect has been tested and the effect removed prior to the VAR analysis using a procedure in Nielsen (2004). In addition, the unification is likely to have produced dynamic effects on prices, exchange rates and interest rates to be accounted for by breaks in the deterministic component within the dynamics of the model. We allow for the possibility of such breaks by a broken linear trend in the long-run relations, a step dummy in the growth rates, and an impulse dummy to account for a blip in the acceleration rates. For a detailed description of the role of deterministic trends in the $I(2)$ model, see Section 3 and Juselius (2006, Chapter 16.). These effects are tested in Section 7.

<Figure 2 here>

As appears from the graphs of the data in Figure 2, there are a number of additional outlying observations, the majority of them belonging to the short-term interest rates in the period of monetary targeting 1980–1982. Hansen and Johansen (1999) tested the hypothesis of constant parameters for this highly volatile period and it was strongly rejected. Because of this, we exclude the observations from 1980:2–1982:3 from the model analysis[2].

To summarize, the baseline VAR model is specified with a linear deterministic trend allowing for a shift in the slope in 1991:1, a step and an impulse dummy at the same date as well as a few additional impulse dummies to be subsequently defined.

[2] As a sensitivity check, the model has been estimated without the hole in 1980–1982 and without correcting for outliers. The main conclusions hold, but the results seem less reliable.

3 The $I(2)$ model with piecewise linear deterministic terms

To ensure that the piecewise linear deterministic components do not cumulate to quadratic or cubic trends we need to discuss how to restrict these components to enter the $I(0)$, $I(1)$, and $I(2)$ directions of the VAR model. Without loss of generality, this discussion can be based on the VAR(3) model formulated in acceleration rates, changes and levels:

$$\begin{aligned}
\Delta^2 x_t &= \Gamma_1 \Delta^2 x_{t-1} + \Gamma \Delta x_{t-1} + \Pi x_{t-1} \qquad (1)\\
&\quad + \Phi_s D_{s,t} + \Phi_p D_{p,t} + \Phi_{tr} D_{tr,t} + \mu_0 + \mu_1 t + \mu_2 t_{b,t} + \varepsilon_t, \\
t &= 1975{:}7,...,1998{:}12, \qquad (2)
\end{aligned}$$

where $x_t' = [pp_t, s_{12,t}, \Delta p_{2,t}, b_{1,t}, b_{2,t}, s_{1,t}, s_{2,t}]$, with $pp_t = (p_1 - p_2)_t$ describing the log of relative prices, $s_{12,t}$ the Dmk/\$ rate, $b_{1,t}, b_{2,t}$ the long-term bond rates, $s_{1,t}, s_{2,t}$ the short-term interest rates.[3] We denote by $t_{b,t} = (t - b + 1)^+$ a broken linear trend, and $D_{s,t}$ is a vector of step dummies $(\ldots, 0, 0, 1, 1, \ldots)$, $D_{p,t}$ is a vector of permanent impulse dummies $(\ldots, 0, 0, 1, 0, \ldots)$, $D_{tr,t}$ is a vector of transitory impulse dummies $(\ldots, 0, 0, 1, -1, 0, 0, \ldots)$, and all parameters are unrestricted.

The hypothesis that x_t is $I(1)$ is formulated as a reduced rank hypothesis

$$\Pi = \alpha\beta' \text{ , where } \alpha, \beta \text{ are } p \times r, \qquad (3)$$

implicitly assuming that Γ is unrestricted. The hypothesis that x_t is $I(2)$ is formulated as an additional reduced rank hypothesis

$$\alpha_\perp' \Gamma \beta_\perp = \xi\eta', \text{ where } \xi, \eta \text{ are } (p - r) \times s_1, \qquad (4)$$

where $\alpha_\perp$ denotes a $p \times (p-r)$ matrix of rank $p-r$ for which $\alpha_\perp'\alpha = 0$, and the notation $\overline{\alpha_\perp} = \alpha_\perp(\alpha_\perp'\alpha_\perp)^{-1}$ is used. Condition (3) is associated with the variables in levels and (4) with the variables in differences. The intuition is that the differenced process also contains unit roots when data are $I(2)$.

Based upon (4) we define the orthogonal decompositions

$$(\beta, \beta_{\perp 1} = \overline{\beta_\perp}\eta, \beta_{\perp 2} = \beta_\perp \eta_\perp) \text{ and } (\alpha, \alpha_{\perp 1} = \overline{\alpha_\perp}\eta, \alpha_{\perp 2} = \alpha_\perp \xi_\perp)$$

of dimensions r, s_1, and s_2 respectively. The moving average representation of the solution of the $I(2)$ model under assumptions (3) and (4), and the condition $|\alpha_{\perp 2}' \Psi \beta_{\perp 2}| \neq 0$,

[3] Annual interest rates in percentages have been transformed to monthly rates by dividing by 1200 to achieve comparability with monthly log changes.

(where $\Psi = \Gamma\overline{\beta}\overline{\alpha}'\Gamma + I_p - \Gamma_1$) is given by

$$\begin{aligned} x_t &= C_2 \sum_{j=1}^{t}\sum_{i=1}^{j}(\varepsilon_i + \Phi_s D_{s,i} + \Phi_p D_{p,i} + \Phi_{tr} D_{tr,i} + \mu_0 + \mu_1 i + \mu_2 t_{91:1,i}) \\ &\quad + C_1 \sum_{j=1}^{t}(\varepsilon_j + \Phi_s D_{s,j} + \Phi_p D_{p,j} + \Phi_{tr} D_{tr,j} + \mu_0 + \mu_1 j + \mu_2 t_{91:1,j}) \\ &\quad + C^*(L)(\varepsilon_t + \Phi_s D_{s,t} + \Phi_p D_{p,t} + \Phi_{tr} D_{tr,t} + \mu_0 + \mu_1 t + \mu_2 t_{91:1,t}) + A + Bt, \end{aligned} \tag{5}$$

where A and B are functions of initial values (x_0, x_{-1}, x_{-2}) and the coefficient matrices are complicated functions of the parameters, satisfying the relations

$$C_2 = \beta_{\perp 2}(\alpha'_{\perp 2}\Psi\beta_{\perp 2})^{-1}\alpha'_{\perp 2}, \tag{6}$$

$$\beta' C_1 + \overline{\alpha}'\Gamma C_2 = 0, \quad C_1\alpha + C_2\Gamma\overline{\beta} = 0, \tag{7}$$

$$\beta'_{\perp 1} C_1 = -\overline{\alpha_{\perp 1}}'(I - \Psi C_2), \tag{8}$$

$$(\beta, \beta_{\perp 1})' B = 0, \quad \beta' A + \bar{\alpha}'\Gamma B = 0, \tag{9}$$

see Johansen (1992) for more details. It follows from (5), (6), (7), and (8) that the processes

$$\Delta^2 x_t, \quad \beta' x_t + \bar{\alpha}'\Gamma\Delta x_t, \quad (\beta, \beta_{\perp 1})'\Delta x_t \tag{10}$$

are trend stationary. This means that x_t is $I(2)$, and that the long-run (polynomially cointegrating) relations are $\beta' x_t + \bar{\alpha}'\Gamma\Delta x_t$. Finally $(\beta, \beta_{\perp 1})'\Delta x_t$ are the medium-run relations between the differences of x_t.

To facilitate the interpretation of the $I(2)$ trends and how they load into the variables, we denote $\breve{\beta}_{\perp 2} = \beta_{\perp 2}(\alpha'_{\perp 2}\Psi\beta_{\perp 2})^{-1}$, so that

$$C_2 = \breve{\beta}_{\perp 2}\alpha'_{\perp 2}. \tag{11}$$

It appears that C_2 has a similar reduced rank representation as C_1 in the $I(1)$ model, so that it is straightforward to interpret $\alpha'_{\perp 2}\sum_{j=1}^{t}\sum_{i=1}^{j}\varepsilon_i$ as a measure of the s_2 second order stochastic trends which load into the variables x_t with the weights $\breve{\beta}_{\perp 2}$.

From (5) it follows that an unrestricted constant will cumulate twice to a quadratic trend, and an unrestricted trend to a cubic trend and similarly for the step dummy and the broken trend. Thus, the coefficients of the deterministic components need to be appropriately restricted in the model equations to avoid undesirable effects in the process.

Since the parameters in model (1) are restricted by the two reduced rank conditions, (3) and (4), Johansen (1997) proposed a parametrization where the individual parameters are variation independent. The parametrization below supplemented with deterministic terms is used in the empirical analysis:

$$
\begin{aligned}
\Delta^2 x_t &= \alpha \left[\rho' \begin{pmatrix} \tau \\ \tau_{01} \\ \tau_0 \end{pmatrix}' \begin{pmatrix} x_{t-1} \\ t_{91:1,t-1} \\ t-1 \end{pmatrix} + \begin{pmatrix} \psi \\ \psi_{01} \\ \psi_0 \end{pmatrix}' \begin{pmatrix} \Delta x_{t-1} \\ D_{s91:1,t-1} \\ 1 \end{pmatrix} \right] \\
&+ \alpha_{\perp\Omega} \kappa' \begin{pmatrix} \tau \\ \tau_{01} \\ \tau_0 \end{pmatrix}' \begin{pmatrix} \Delta x_{t-1} \\ D_{s91:1,t-1} \\ 1 \end{pmatrix} + \Phi_p D_{p,t} + \Phi_{tr} D_{tr,t} + \varepsilon_t,
\end{aligned} \tag{12}
$$

where $\alpha_{\perp\Omega} = \Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}$, $t_{91:1,t}$ is a linear trend starting in 1991:1, $Ds_{91:1,t}$ is a step dummy starting in 1991:1, $D_{p,s}$ is a vector of impulse dummies, and $D_{tr,t}$ is a transitory dummy, see Table 1 for a complete description.

In the following we use $d_t = (t_{91:1t}, t)'$ and write (12) in a more compact form as

$$
\Delta^2 x_t = \alpha(\rho'\tilde{\tau}'\tilde{x}_{t-1} + \tilde{\psi}'\Delta\tilde{x}_{t-1}) + \alpha_{\perp\Omega}\kappa'\tilde{\tau}'\Delta\tilde{x}_{t-1} + \Phi_p D_{p,t} + \Phi_{tr} D_{tr,t} + \varepsilon_t, \tag{13}
$$

where $\tilde{\tau} = (\tau', \tau_{01}, \tau_0)' = (\tau', \tau_d')'$, $\tilde{\psi} = (\psi', \psi_{01}, \psi_0)' = (\psi', \psi_d')'$, and $\tilde{x}_t = (x_t', d_t')'$. Let $\alpha_\Omega' = (\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}$, then $\alpha_\Omega'\alpha_{\perp\Omega} = 0$, and $\alpha_\Omega'\alpha = I_r$, and compared to the parametrization (1) we have $\tau = (\beta, \beta_{\perp 1})$, $\psi' = \alpha_\Omega'\Gamma$, and $\kappa' = \alpha_\perp'\Gamma\bar{\tau}$.

In the polynomially cointegrating relation $\rho'\tilde{\tau}'\tilde{x}_{t-1} + \tilde{\psi}'\Delta x_{t-1}$, the term $\tilde{\tau}'\Delta\tilde{x}_{t-1}$ is already trend stationary so we define the coefficient $\tilde{\delta}' = \tilde{\psi}'\tilde{\tau}_\perp(\tilde{\tau}_\perp'\tilde{\tau}_\perp)^{-1}\tilde{\tau}_\perp'$ as the coefficient to $\Delta\tilde{x}_{t-1}$ needed to render $\tilde{\beta}'\tilde{x}_t + \tilde{\delta}'\Delta\tilde{x}_t$ trend stationary.

We let p_1 be the dimension of $\tilde{x}_t$. An advantage of this parametrization is that the parameters $(\alpha, \tilde{\tau}, \tilde{\psi}, \rho, \kappa, \Phi_p, \Phi_{tr}, \Omega)$ are variation independent. We next give expressions for the main components of the deterministic trends in x_t and $\tau' x_t$ and the trend stationary relations $\Delta^2 x_t$, $\beta' x_t + \psi'\Delta x_t$, and $\tau'\Delta x_t$. The proof of the next Lemma is essentially contained in Rahbek et al. (1999) who investigated the case of a restricted linear term. We need to pay special attention to the dummy variables which generate many different trends, because they are not linearly independent when shifted in time.

Lemma 1 *The deterministic term in the model equation (12):*

$$
\mu_t = \alpha\rho'\tau_d' d_{t-1} + (\alpha\psi_d' + \alpha_{\perp\Omega}\kappa'\tau_d')\Delta d_{t-1} + \Phi_p D_{p,t} + \Phi_{tr} D_{tr,t},
$$

implies that x_t has a deterministic trend which consists of $d_t = (t_{91:1,t}, t)'$, a linear trend from the initial values, and broken linear trends from the impulse dummies plus terms that are bounded, and a sum of impulse dummies with exponentially decreasing coefficients. Hence the trend in $\Delta^2 x_t$ is a sum of impulse dummies with exponentially decreasing coefficients.

Proof. It follows from (5) that the initial values generate a linear trend $A + Bt$, and the term $\Phi_p D_{p,t} + \Phi_{tr} D_{tr,t}$ generates broken linear trends, step dummies, and a linear combination of impulse dummies with exponentially decreasing coefficients. The remaining terms are generated by d_t and Δd_t and give

$$[C_2\alpha_{\perp\Omega}\kappa'\tau_d' + C_1\alpha\rho'\tau_d']\sum_{j=1}^{t} d_{j-1} + C_1[\alpha\psi_d' + \alpha_{\perp\Omega}\kappa'\tau_d']d_{t-1}$$
$$+\sum_{i=0}^{\infty} C_i^*[\alpha\rho'\tau_d' d_{t-1-i} + (\alpha\psi_d' + \alpha_{\perp\Omega}\kappa'\tau_d')\Delta d_{t-1-i}],$$

where we have used that $C_2\alpha = 0$, see (6).

We want to prove that

$$C_2\alpha_{\perp\Omega}\kappa' + C_1\alpha\rho' = 0,$$

by multiplying it by $(\beta, \beta_{\perp1}, \beta_{\perp2})'$, so that $\sum_{j=1}^{t} d_{j-1}$ does not appear.

1. Multiplying byβ' we find the result from $\beta' C_2 = 0$ and $\beta' C_1\alpha = 0$, see (6) and (7).

2. Multiplying by $\beta_{\perp1}'$ and useing $\beta_{\perp1}' C_2 = 0$ we see that the first term is zero, and from the expression $\beta_{\perp1}' C_1 = -\bar{\alpha}_{\perp1}'(I_p - \Psi C_2)$, see (8), we find that the next term is zero.

3. Finally we multiply by $\beta_{\perp2}'$ and show that

$$\beta_{\perp2}' C_2\alpha_{\perp\Omega}\kappa' + \beta_{\perp2}' C_1\alpha\rho' = 0.$$

From (7) we find

$$C_1\alpha = -C_2\Gamma\bar{\beta} = -C_2\alpha_{\perp\Omega}\kappa'\tau'\bar{\beta},$$

so we get

$$\beta_{\perp2}' C_2\alpha_{\perp\Omega}\kappa' + \beta_{\perp2}' C_1\alpha\rho' = \beta_{\perp2}' C_2\alpha_{\perp\Omega}\kappa'(I_{r+s_1} - \tau'\bar{\beta}\rho').$$

The following identity holds

$$I_{r+s_1} - \tau'\bar{\beta}\rho' = I_{r+s_1} - \tau'\tau\rho(\rho'\tau'\tau\rho)^{-1}\rho' = \rho_\perp(\rho'(\tau'\tau)^{-1}\rho_\perp)^{-1}\rho_\perp(\tau'\tau)^{-1}. \quad (14)$$

To see this, set $\beta = \tau\rho$ and multiply by the full rank matrix $(\rho_\perp, \tau'\tau\rho)$.

Thus we find

$$\beta_{\perp2}' C_2\alpha_{\perp\Omega}\kappa' + \beta_{\perp2}' C_1\alpha\rho' = \beta_{\perp2}' C_2\alpha_{\perp\Omega}\kappa'\rho_\perp(\rho'(\tau'\tau)^{-1}\rho_\perp)^{-1}\rho_\perp(\tau'\tau)^{-1}.$$

Finally we use that $\alpha_{\perp2} = \alpha_\perp\xi_\perp$ where $\xi = \kappa'\rho_\perp$ and find

$$C_2\alpha_{\perp\Omega}\kappa'\rho_\perp = \beta_{\perp2}(\alpha_{\perp2}'\Psi\beta_{\perp2})^{-1}\alpha_{\perp2}'\Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\kappa'\rho_\perp,$$

but

$$\alpha_{\perp2}'\Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\kappa'\rho_\perp = \xi_\perp'\alpha_\perp'\Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\kappa'\rho_\perp = \xi_\perp'\xi = 0.$$

Thus the trend in x_t has the form indicated in the Lemma. ∎

Corollary 2 *The deterministic term in* Δx_t^2, $\tilde{\beta}'\tilde{x}_t+\tilde{\psi}'\Delta\tilde{x}_t$, *and* $\tilde{\tau}'\Delta\tilde{x}_t$ *consists of a sum of impulse dummies with exponentially decreasing weights, which we denote generically by* R_t. *It follows that*

$$\begin{aligned} E(\beta' x_t + \psi'\Delta x_t) &= -\rho'\tau_d' d_t - \psi_d'\Delta d_t + R_t, \qquad (15)\\ E(\tau'\Delta x_{t-1}) &= -\tau_d'\Delta d_t + R_t. \qquad (16)\end{aligned}$$

Therefore

$$E(\tau' x_t) = -\tau_d' d_t + \text{bounded terms.} \qquad (17)$$

Proof. It follows from Lemma 1 that $E(\Delta^2 x_t) = R_t$ and from the model equation (12) we find using $\alpha_\Omega = (\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}$ that

$$\begin{aligned} \alpha_\Omega' E(\Delta^2 x_t) &= E(\rho'\tilde{\tau}'\tilde{x}_{t-1} + \tilde{\psi}'\Delta\tilde{x}_{t-1}) + \alpha_\Omega'\Phi_p D_{p,t} + \alpha_\Omega'\Phi_{tr} D_{tr,t},\\ \alpha_\perp' E(\Delta^2 x_t) &= \kappa' E(\tilde{\tau}'\Delta\tilde{x}_{t-1}) + \alpha_\perp'\Phi_p D_{p,t} + \alpha_\perp'\Phi_{tr} D_{tr,t}.\end{aligned}$$

From the first relation it is seen that

$$E(\beta' x_t + \psi'\Delta x_t) = -\rho'\tau_d' d_t - \psi_d'\Delta d_t + R_t,$$

so that

$$\rho E(\tau'\Delta x_{t-1}) = E(\beta'\Delta x_t) = -\rho'\tau_d'\Delta d_t + R_t. \qquad (18)$$

From the second relation it follows that

$$\kappa' E(\tau'\Delta x_{t-1}) = -\kappa'\tau_d'\Delta d_{t-1} + R_t.$$

Writing $\kappa'\tau' = \kappa'(\bar{\rho}\rho' + \bar{\rho}_\perp\rho_\perp')\tau'$ we find

$$\begin{aligned} \kappa'\bar{\rho}_\perp\rho_\perp' E(\tau'\Delta x_{t-1}) &= \kappa' E(\tau'\Delta x_{t-1}) - \kappa'\bar{\rho}\rho' E(\tau'\Delta x_{t-1})\\ &= -\kappa'\bar{\rho}_\perp\rho_\perp'\tau_d'\Delta d_{t-1} + R_t,\end{aligned}$$

so that

$$\rho_\perp' E(\tau'\Delta x_{t-1}) = -\rho_\perp'\tau_d'\Delta d_{t-1} + R_t,$$

which together with (18) gives (16) and by summation we find (17). ■

Thus the main terms in the trend of the polynomially cointegrating relations is d_t and Δd_t, whereas $\tilde{\tau}'\Delta x_t$ has only Δd_t.

The four deterministic terms modelled by the parameters $(\tau_{01}, \tau_0, \psi_{01}, \psi_0)$ in (12) are exactly the (main) deterministic terms that appear in the trend stationary processes.

4 Testing hypotheses in the $I(2)$ model

We discuss in this section hypotheses on the parameters α, $\tilde{\beta}$, and $\tilde{\tau}$, in the parametrization (13).

4.1 Hypotheses on α

We discuss two types of hypotheses, the hypothesis of no levels feed-back and the hypothesis of a unit vector in α.

First, let $x_t = (x'_{1t}, x'_{2t})'$ be a decomposition of the variables into two sets of $p - m$ and m variables, and decompose $\alpha = (\alpha'_1, \alpha'_2)'$ similarly. The hypothesis on no levels feed-back from x_2 to x_1

$$\alpha = \begin{pmatrix} \alpha_1 \\ 0 \end{pmatrix} = \begin{pmatrix} I_{p-m} \\ 0 \end{pmatrix} \alpha_1, \tag{19}$$

or $\alpha_2 = 0$, means that the acceleration $\Delta^2 x_{2t}$ does not react immediately to a disequilibrium error in the polynomial cointegration relations $\beta' x_{t-1} + \delta' \Delta x_{t-1}$. Expressed differently this means that the error term ε_{2t} cumulates to common trends and in this sense the variables in x_{2t} are pushing variables with long-run impact. The hypothesis of weak exogeneity of x_{2t} is a restriction on the rows of $(\alpha, \alpha_{\perp 1})$, and that is not tested here, see however, Paruolo and Rahbek (1999).

Second, the hypothesis that a unit vector, e_1, is in α, as formulated by

$$\alpha = (e_1, e_{1\perp}\phi). \tag{20}$$

An equivalent way of saying this is that the first row of $\alpha_\perp$ is zero, $e'_1 \alpha_\perp = 0$, so that

$$\alpha_\perp = e_{1\perp}\psi.$$

This has the interpretation that the errors of the first equation are not cumulating and in this sense the variable is purely adjusting Juselius (2006 p. 200).

Both hypotheses are restrictions on the coefficient of the (asymptotically) stationary polynomial cointegration relations, $\beta' x_{t-1} + \delta' \Delta x_{t-1}$, and therefore the likelihood ratio test statistics are asymptotically χ^2 with degrees of freedom mr and $p + r - 1$ respectively, corresponding to the number of restricted parameters.

4.2 Tests on $\tilde{\beta}$

We consider in Section 8 test for linear restrictions on each $\tilde{\beta}$ vector

$$\tilde{\beta} = (h_1 + H_1\varphi_1, \ldots, h_r + H_r\varphi_r), \tag{21}$$

where h_i is $p_1 \times 1$ and linearly independent of H_i which is $p_1 \times m_i$ of rank m_i, both known, and φ_i is an unknown parameter of dimension $m_i \times 1$.

We then need the result

Lemma 3 *In the $I(2)$ model with piecewise linear deterministic trends, and $\tilde{\beta}$ identified by the restrictions (21), the asymptotic distribution of the maximum likelihood estimator of $(\varphi_1, \ldots, \varphi_r)$ is mixed Gaussian, so that the asymptotic distribution of the likelihood ratio test for the hypothesis (21) is $\chi^2(\sum_{i=1}^r (p_1 - r - m_i))$.*

The proof is given in Appendix by establishing that the asymptotic distribution of the score function is mixed Gaussian with a non-singular conditional variance. Related results are derived previously. In Johansen (1997) the asymptotic distribution of $\hat{\beta}$ were given for the case of no restriction and no deterministic components, and Rahbek et al. (1999) show the result for the case of no restrictions when the process has a linear trend. General conditions on β were dealt with by Boswijk (2000) and Johansen (2006a) but for the model with no deterministic components. Here we give a direct proof for the model at hand when $\tilde{\beta}$ is identified by the linear restrictions (21).

The result will be applied to simplify the estimated polynomial cointegration relations. The hypotheses do not involve the coefficient $\tilde{\delta}$ in $\tilde{\beta}'\tilde{x}_t + \tilde{\delta}'\Delta\tilde{x}_t$, because the asymptotic theory for such hypotheses has not been worked out.

4.3 Tests on $\tilde{\tau}$

By decomposing $\tilde{\tau} = \tilde{\tau}\rho\bar{\rho}' + \tilde{\tau}\rho_\perp\bar{\rho}_\perp' = \tilde{\beta}\bar{\rho}' + \tilde{\eta}\bar{\rho}_\perp'$ it was proved above that $\widehat{\tilde{\beta}} - \tilde{\beta}$, suitably normalized, is asymptotically Gaussian given $\beta_{\perp 2}'C_2\int_0^u W(s)ds$ and $\beta_{\perp 1}'C_1W(u)$, see (30). Similarly, one can show that $\widehat{\tilde{\eta}} - \tilde{\eta}$, suitably normalized, is asymptotically mixed Gaussian given $\beta_{\perp 2}'C_2W(u)$. The two limits are mixed Gaussian but not jointly mixed Gaussian, so that hypotheses that involve both $\tilde{\beta}$ and $\tilde{\eta}$ (and ρ) need not give rise to an asymptotic χ^2 distribution. Johansen (2006a, Theorem 5) gave a sufficient condition for asymptotic mixed Gaussian inference in a submodel of the $I(2)$ model. The condition formulates a separation between the parameters, which ensures that the usual conditioning argument leading to asymptotic χ^2 distributions holds. In the case of deterministic terms we have exactly the same structure of asymptotic distributions so the same theorem holds, and the investigation of hypotheses on τ is valid also for $\tilde{\tau}$.

We consider first the same restriction on all vectors in $\tilde{\tau}$, that is,

$$\tilde{\tau} = H\phi, \tag{22}$$

where H is $p_1 \times m$ is known and ϕ is an $m \times (r + s_1)$ matrix of unknown parameters. An equivalent formulation is $R'\tilde{\tau} = 0$, where $R = H_\perp$.

The other hypothesis corresponds to (20), that is,

$$\tilde{\tau} = (b, b_\perp\varphi), \tag{23}$$

where b is $p_1 \times 1$ and known and φ is a $(p_1 - 1) \times (r + s_1 - 1)$ matrix of unknown parameters.

The test statistic for the first test is asymptotically distributed with degrees of freedom $(p_1 - m)(r + s_1)$, and, in general, the test for the second one is also asymptotically distributed as χ^2 with $s_2 - 1$ degrees of freedom. There is, however, one case when the asymptotic distribution is not a χ^2 distribution. This is when the vector b is a vector in $\tilde{\beta}$, that is, when the hypothesis $\tilde{\beta} = (b, b_\perp\xi)$ is satisfied, see the discussion in Johansen (2006a) of this hypothesis.

This problem can be avoided by first testing the hypothesis $\tilde{\beta} = (b, b_{\perp}\xi)$ and, if accepted, then b is a vector in $\tilde{\tau}$. If it is rejected, we can test $\tilde{\tau} = (b, b_{\perp}\varphi)$ and apply the χ^2 distribution because we have checked that $b \notin sp(\tilde{\beta})$.

<Table 1 here>

5 Misspecification testing of the baseline VAR

The VAR model in (12) was specified to account for the German reunification as explained in Section 2[4]. In addition there are a number of outlying observations that need to be accounted for. Since the VAR estimates have been shown to be reasonably robust to moderate excess kurtosis (long tails) as long as the error distribution is symmetrical (Gonzalo, 1994) only extraordinarily large shocks producing skewed residuals have been corrected for[5]. The dummies and their estimated effects are reported in Table 1, which shows that the very large shocks were associated with large and unexplainable changes (given our data and our model) in the short-term interest rates and the US bond rate. The dummy variable, $Dtax$, measures the impact on German prices from a number of excise taxes in 1991:7, 1991:1, and 1993:1 to finance the re-unification. All dummy variables, except the one in 1984:1 which is a transitory dummy $(\ldots, 0, 1, -1, 0, \ldots)$, are impulse dummies $(\ldots, 0, 1, 0, \ldots)$.

With this model specification, Table 2 shows that the model passes most of the specification tests. However, multivariate normality and no ARCH are rejected. The former is mostly due to excess kurtosis in the nominal exchange rate and the interest rates. The latter is mostly due to some ARCH effects in the bond rates. Because, the cointegrated VAR results should be reasonably robust to excess kurtosis and ARCH effects as long as they are moderately sized, we continue with this model specification.

<Table 2 here>

6 Determining the two reduced rank indices

The number of stationary polynomially cointegrating relations, r, and the number of $I(1)$ trends, s_1, among the $p-r$ common stochastic trends are determined by the likelihood ratio test. The asymptotic distribution was found by Nielsen and Rahbek (2007) for the case of no deterministics. and Kurita (2007) found the limit distributions when there are broken linear trends. Since our model has a broken linear trend restricted to be in the cointegration relations, and a shift dummy restricted to the differences,

[4] All calculations have been performed in the computer package CATS in RATS (Dennis et al. 2005.)

[5] Note that, contrary to the static regression model, the dummies do not eliminate the corresponding observation. They only account for the unanticipated shock at the time it occurred, essentially saying that an event outside the chosen information set had caused the shock. Next period it is no longer unanticipated and the dynamics of the model should take account of the observed data.

the asymptotic distributions of the likelihood ratio tests have been simulated with a program kindly made available by Heino Bohn Nielsen.

<Table 3 here>

Table 3 reports the tests of the joint hypothesis (r, s_1, s_2) for all values of r, s_1 and s_2, where s_2 is the number of $I(2)$ trends and $s_1 = p - r - s_2$. The test procedure starts with the most restricted model $(r = 0, s_1 = 0, s_2 = 7)$ in the upper left hand corner, continues to the end of the first row $(r = 0, s_1 = 7, s_2 = 0)$, and proceeds similarly row-wise from left to right until the first acceptance.

<Table 4 here>

The first non rejection is for $\{r = 2, s_1 = 3, s_2 = 2\}$ with a p-value of 0.09. This case implies seven unit roots in the model which is in conflict with the results in Table 4 which suggest that the unrestricted VAR contains at most six (near) unit roots. The following two cases $\{r = 2, s_1 = 4, s_2 = 1\}$ and $\{r = 3, s_1 = 2, s_2 = 2\}$ are strongly accepted and imply six unit roots in the model. To further check these two cases, Table 4 reports the roots for the case of $r = 2, 3$ assuming no $I(2)$ trends, $s_2 = 0$, and $s_2 = 1, 2$ respectively. The case $\{r = 3, s_1 = 4, s_2 = 0\}$ would leave two near unit roots (0.93 and 0.91) in the model, whereas $\{r = 2, s_1 = 5, s_2 = 0\}$ would leave one near unit root (0.90) in the model. Imposing either $\{r = 2, s_1 = 4, s_2 = 1\}$ or $\{r = 3, s_1 = 2, s_2 = 2\}$ removes all large roots from the model. Therefore, the choice seems to be between two (three) polynomial cointegration relations $(\tilde{\beta}' x_t + \tilde{\delta}' \Delta \tilde{x}_t)$ and three (two) medium-run relations in differences $(\tilde{\beta}'_{\perp 1} \Delta \tilde{x}_t)$. Checking the $t-$values of $\hat{\alpha}_3$ shows five highly significant coefficients (with t-values in the range of 15.4 to 3.4), which suggests that the third polynomial cointegration relation is indeed stationary. This is further supported by the graphs in Figure 5. Based on this, we will continue with the case $\{r = 3, s_1 = 2, s_2 = 2\}$.

How can we understand the finding of two $I(2)$ trends? The persistent downward sloping trend of the price differential and the persistent long swings in the nominal exchange rate (see Figure 1, upper panel) may very well have the property of a near $I(2)$ process corresponding to a double root of (1.0 and 0.93) and (1.0 and 0.91) respectively. The second root may not be exactly $I(2)$ but persistent enough for the test not to reject it as a unit root.

Altogether, the results seem to suggest that the persistent movements in the data are (near) $I(2)$. Treating the process as $I(1)$ is likely to yield unreliable inference (see Johansen, 2006b and Juselius, 2008) as one (or two) very persistent components in the data would be treated as stationary.

7 Testing non-identifying hypotheses

The three groups of tests in this section provide an approximate description of the properties of the data. To illustrate what the tests find, Figure 3, upper panel, shows the graphs of $\Delta(pp)_t$ and $\Delta s_{12,t}$ filtered through a twelve-month moving average.[6] It is notable that the inflation rate differentials exhibit more persistent behavior than the changes in exchange rates. The middle panel shows the short-term interest spread and the bottom panel the long-term spread. In both cases, the spreads exhibit pronounced persistence.

<Figure 3 here>

7.1 Same restriction on all $\tilde{\tau}$

There are a number of interesting hypotheses that can be formulated as the same restrictions on $\tilde{\tau}$, described in Section 4.3, expressed either as $\tilde{\tau} = H\varphi$ or $R'\tilde{\tau} = 0$. We test the following four hypotheses:

1. $\mathcal{H}_1 : R_1'\tilde{\tau} = 0$, where $R_1' = [1, 1, 0, 0, 0, 0, 0, 0, 0]$, i.e. we test whether we can impose the *ppp* restriction on all $\tilde{\tau}$ vectors. If not rejected, it would imply that the nominal to real transformation $x_t' = [ppp_t, \Delta p_{1,t}, \Delta p_{2,t}, b_{1,t}, b_{2,t}, s_{1,t}, s_{2,t}]$ would be econometrically valid (Kongsted, 2005) in the sense of transforming an $I(2)$ vector to an $I(1)$ without loss of information. The hypothesis is rejected based on $\chi^2(5) = 16.64\,[0.01]$.

2. $\mathcal{H}_2 : R_2'\tilde{\tau} = 0$, where $R_2' = [0, 0, 0, 0, 0, 0, 0, 1, 0]$, i.e. we test whether the broken trend is long-run excludable from $\tilde{\tau}$. Based on $\chi^2(5) = 9.99\,[0.08]$, the hypothesis was borderline not rejected. Thus, there is only weak evidence that the direction of the trend in relative prices and/or nominal exchange rates changed at the time of the re-unification of Germany.

3. $\mathcal{H}_3 : R_3'\tilde{\tau} = 0$, where $R_3' = [0, 0, 0, 0, 0, 0, 0, 0, 1]$, i.e. we test whether the trend is long-run excludable from $\tilde{\tau}$. The hypothesis was rejected based $\chi^2(5) = 13.75\,[0.02]$. Thus, there is some evidence that the trend is econometrically needed as a local approximation of the downward sloping trend in relative prices and in the nominal exchange rate.

4. $\mathcal{H}_4 : R_4'\tilde{\tau} = 0$, where
$$R_4' = \begin{bmatrix} 0, 0, 0, 0, 0, 0, 0, 0, 1 \\ 0, 0, 0, 0, 0, 0, 0, 1, 0 \end{bmatrix},$$
i.e. we test whether the trends can be left out of the long-run relations. The hypothesis was rejected based $\chi^2(10) = 24.50\,[0.01]$. Thus, the hypothesis that the downward sloping trend strongly visible in Figure 1 is stochastic rather than deterministic is rejected. However, the hypotheses of no trends were only borderline rejected, which supports our prior assumption that a deterministic trend should only be considered a local approximation.

[6] The series are plotted as a 12 months moving average to filter out the strong seasonal and other high frequency components in prices. The data used in the analysis is of course non-filtered.

Imposing the *ppp* transformation of the data $[ppp_t, \Delta p_{1,t}, \Delta p_{2,t}, b_{1,t}, b_{2,t}, s_{1,t}, s_{2,t}]$, despite not being supported by the test, is likely to suppress some information in the data. The estimated long-run structure in Section 8 shows that one relation is consistent with the *ppp* transformation, whereas the other two are not. Also, test results of $\mathcal{H}_2 - \mathcal{H}_4$ show that the linear trend is significant in the long-run relations. The hypotheses that the slope changed at the reunification seems to have empirical support, though not very significantly so. The estimated results in Section 8 show that the trend effects in the long-run relations are absolutely tiny but, nevertheless, highly significant. Altogether, the above tests seem to confirm our model specification.

<Figure 4 here>

7.2 A known vector in $\tilde{\tau}$

Next, we shall test five hypotheses formulated as a known vector b in $\tilde{\tau}$. If not rejected, they imply that the variable in question (conditional on Δx_{t-1}) is at most $I(1)$. If in addition, it is not a vector in $\tilde{\beta}$, then the test implies it is $I(1)$. This hypothesis was rejected for all variables tested below, which implies that our tests are tests of $I(1)$.

1. $\mathcal{H}_5 : \tilde{\tau} = (d_1, d_{1\perp}\varphi)$ where $d_1 = [1, 0, 0, 0, 0, 0, 0, 0, 0]$, i.e. we test whether relative prices is a unit vector in $\tilde{\tau}$. This hypothesis is strongly rejected based on $\chi^2(4) = 55.56\,[0.00]$, implying that this variable can be considered $I(2)$.

2. $\mathcal{H}_6 : \tilde{\tau} = (d_2, d_{2\perp}\varphi)$ where $d_2 = [0, 1, 0, 0, 0, 0, 0, 0, 0]$, i.e. we test whether nominal exchange rate is a unit vector in $\tilde{\tau}$. This hypothesis is rejected based on $\chi^2(4) = 9.76\,[0.04]$, implying that this variable can be considered $I(2)$.

3. $\mathcal{H}_7 : \tilde{\tau} = (d_3, d_{3\perp}\varphi)$ where $d_3 = [1, -1, 0, 0, 0, 0, 0, 0, 0]$, i.e., we test whether the real exchange rate is a unit vector in $\tilde{\tau}$. This hypothesis is not rejected based on $\chi^2(4) = 4.90\,[0.30]$, which implies that ppp_t can be considered $I(1)$.

4. $\mathcal{H}_8 : \tilde{\tau} = (d_4, d_{4\perp}\varphi)$ where $d_4 = [0, 0, 0, 1, -1, 0, 0, 0, 0]$, i.e., we test whether the bond rate spread is a unit vector in $\tilde{\tau}$. This hypothesis is not rejected based on $\chi^2(4) = 3.65\,[0.46]$, which implies that $b_{1,t} - b_{2,t}$ can be considered $I(1)$.

5. $\mathcal{H}_9 : \tilde{\tau} = (d_5, d_{5\perp}\varphi)$ where $d_5 = [0, 0, 0, 0, 0, 1, -1, 0, 0]$, i.e., we test whether the short spread is a unit vector in $\tilde{\tau}$. This hypothesis is borderline not rejected based on $\chi^2(4) = 8.43\,[0.08]$, which implies that this variable can be considered $I(1)$ but with a large second root.

As discussed in the introduction, the monetary model with IKE implies that Δppp_t is stationary but highly persistent, or near $I(1)$, whereas REH implies that Δppp_t is white noise or alternatively that ppp_t is stationary. The graphs of $\Delta s_{12,t}$ and Δpp_t in Figure 3, illustrate that both are highly persistent processes, which is supported by the test results of $\mathcal{H}_5$ and $\mathcal{H}_6$. The graph of Δppp_t in the lower panel looks almost indistinguishable from the graph of $\Delta s_{12,t}$[7] (though with opposite sign). The fact that

[7]This is mostly because $\Delta s_{12,t}$ has a much larger variance than Δpp_t.

$\mathcal{H}_7$ could not be rejected must, therefore, imply that Δppp_t (though stationary) is a highly persistent process.[8] This is also confirmed by the estimated Moving Average (MA) representation in Section 9. Hypothesis $\mathcal{H}_8$ could also not be rejected, implying that the ppp and $b_{1,t} - b_{2,t}$ are of a similar order of integration and, hence, could be cointegrated (as they turned out to be). As the monetary model with IKE implies cointegration between the two variables, this is of some interest.

7.3 Hypotheses on α

Finally we shall test two hypotheses described in Section 4.1 both formulated on α. The first one is a test of a zero row in α and implies no long-run levels feedback of the variable in question. The second is a test of a unit vector in α and implies pure adjustment of the variable in question.

1. $\mathcal{H}_{10} : e_i'\alpha = 0$, where e_i is the i^{th} unit vector. This hypothesis was not rejected for nominal exchange rate, $\chi^2(3) = 5.24\,[0.15]$, and the US bond rate, $\chi^2(3) = 1.27\,[0.74]$. The joint hypothesis is also accepted based on $\chi^2(6) = 6.578[0.362]$.

2. $\mathcal{H}_{11} : \alpha = (e_i, e_{i\perp}\phi)$. This hypothesis was accepted for the US inflation rate based on $\chi^2(3) = 2.41\,[0.66]$. Thus, US prices seem purely adjusting.

The finding that there are little long-run levels feed-back on the US bond rate and the nominal exchange rate and that prices are strongly adjusting are hard to reconcile with the REH theories, but again perfectly consistent with IKE behavior.

8 Testing identifying restrictions on the long-run structure

The decomposition $\tilde{\tau} = (\tilde{\beta}, \tilde{\beta}_{\perp 1})$ defines three stationary polynomially cointegrating relations, $\tilde{\beta}_i' x_t + \tilde{\delta}\tilde{\psi}'\Delta\tilde{x}_t$, $i = 1, 2, 3$ and five stationary cointegration relations between the differenced variables, $\tilde{\tau}'\Delta\tilde{x}_t$. How to test over-identifying restrictions on $\tilde{\beta}$ was discussed in Section 4.2, whereas overidentifying tests on $\tilde{\beta}_{\perp 1}$ have not yet been derived. Even though unrestricted estimates of $\tilde{\beta}_{\perp 1}$ can be calculated, we will not discuss them here as they may not be economically meaningful.

To obtain standard errors of the estimated $\tilde{\beta}$ coefficients we need to impose identifying restrictions on each of the polynomially cointegrating relations reported above. The asymptotic distribution of the test of identifying restrictions on $\tilde{\beta}$ is given in Lemma 1.

<Table 5 here>

[8] Applying an autoregressive model to Δppp_t, produced a root of approximately 0.85-0.90, depending on the length of the lag structure. However, the AR model was not a good description of the time-series behavior of Δppp_t, and should therefore only be seen as indicative.

When interpreting the $\tilde{\beta}$ relations below we shall only include the first two elements of $\tilde{\delta}'\Delta x_t$, corresponding to the inflation rate differentials and the depreciation/appreciation rate, as the pp and the s_{12} were the only variables tested to be $I(2)$ in the previous section.

The first cointegrating relation involves the long-term real interest rate spread and ppp. Similar relationships have previously been found in Juselius (1995, 2006, Chapter 21) and Juselius and MacDonald (2004, 2007). The close co-movements between the two series, illustrated in Figure 1, are quite remarkable.[9]

$$\tilde{\beta}_1'x + \tilde{\delta}_1'\Delta x = (b_1 - 0.92\Delta p_1) - (b_2 - 0.92\Delta p_2) + 0.15\Delta s_{12} - 0.01ppp + 0.000007t. \quad (24)$$

The second is a relation between the US term spread and US inflation relative to German inflation. It can be interpreted as expected inflation, measured by the term spread, as a function of actual inflation rates and the change in the Dmk/$ rate:

$$\tilde{\beta}_2'x + \tilde{\delta}_2'\Delta x = (b_2 - s_2) + 0.60\Delta p_2 - 0.51\Delta p_1 - 0.17\Delta s_{12} + 0.001s_{12} + 0.000019t_{91.1}. \quad (25)$$

The third relation, essentially a relation for German inflation rate, is similar to the relation found in Juselius and MacDonald (2007). It shows that the German inflation rate has been (almost) homogeneously related to US inflation rate, German short-term interest rate, and the change in the Dmk/$ rate:

$$\tilde{\beta}_3'x + \tilde{\delta}_3\Delta x = 1.31\Delta p_1 - 0.31\Delta p_2 - 0.74s_1 - 0.07\Delta s_{12} - 0.01pp - 0.00002t_{91.1}. \quad (26)$$

All three relations contain a tiny, but significant, trend effect which is not straightforward to interpret. It seems, however, likely that the linear trend effect in the relations is a proxy for some information not included in the analysis. For example, the small trend effect in (24) might very well account for a productivity differential between the two economies. In (25) the re-unification trend might be a proxy for a change in the market's re-assessment of the riskiness of the nominal Dmk/$ rate. In (26) the trend together with the pp may imply that German inflation rate, in addition to following the US inflation rate, the short-term interest rate, and the change in the Dmk/$ rate, has adjusted in the long run to the deviation of relative prices from trend. Figure 5 shows that the three polynomially cointegrating relations are very stationary.

<Figure 5 here>

The estimated α coefficients show that relative prices and the US inflation rate adjust very significantly to all three cointegration relations, whereas the nominal exchange rate adjusts to the first two relations, though less significantly so. Of the two long-term rates, the US bond rate is not significantly adjusting (except with a tiny coefficient to the first relation) consistent with the test of a zero row in α in Section 7.3, whereas German bond rate is very significantly adjusting to the second relation. Of the two short term interest rate, the German rate is very significantly adjusting to all three relations, whereas the US rate is essentially adjusting to the first relation.

[9] REH monetary models imply that the long-term real interest rate spread and ppp are separately $I(0)$. By sharp contrast, the cointegrating relationship between these two variables is consistent with the monetary model under IKE. See Frydman et al. (2009).

9 The (near) $I(2)$ trends and how they load into the data

Table 6 reports the weights with which the two (near) $I(2)$ stochastic trends have affected the variables of the system. As the weights of the $I(1)$ trends are complicated functions of all estimated matrices (see Section 3), they will not be reported. The discussion will, therefore, focus on the $I(2)$ trends and how the estimated weights are helpful in understanding more fully the test results of the previous section.

<Table 6 here>

We note that the first $I(2)$ trend, essentially measuring the twice cumulated shocks to the bond spread and the German term spread, loads into pp and s_{12} with coefficients of the same sign, while not exactly the same magnitude. The coefficients to the interest rates are very close to zero. We interpret this $I(2)$ trend as describing the long-run downward sloping trend visible in both relative prices and nominal exchange rates The second $I(2)$ trend, essentially measuring the twice cumulated shocks to the US short-term interest rate, loads primarily into s_{12} and into pp with a coefficient of opposite sign, reflecting the tendency of the nominal exchange rate to move away from relative prices for extended periods of time. It also loads into the remaining variables with coefficients which might be large enough to suggest some significant effects. Based on this, we can now get an expression for the (near) $I(2)$ properties of the ppp:

$$pp - s_{12} = (1.556 - 2.432)\alpha'_{\perp 2,1}\sum\sum\hat{\varepsilon}_s - (0.693 + 3.356)\alpha'_{\perp 2,2}\sum\sum\hat{\varepsilon}_s$$

This expression suggests that the real exchange rate is indeed near $I(2)$ in the sense of being strongly influenced by the second $I(2)$ trend. The fact that it could not be rejected as an $I(1)$ process in Section 7.2, $\mathcal{H}_7$, suggests that the loading -0.876 is not statistically significant from zero and that the loadings to the second stochastic trend have large standard errors, consistent with the high volatility characterizing exchange rate movements in currency markets. This volatility is also visible in the relatively larger loadings to the nominal exchange rate compared to relative prices. Also, the finding that the $I(1)$ hypothesis was rejected for both $s_{12,t}$ and pp_t, but not for ppp_t suggests that the loadings 2.43 and 1.56 are both significant, whereas 0.876 is not.

The expression for the long-term bond spread is given by:

$$b_1 - b_2 = (0.008 - 0.000)\alpha'_{\perp 2,1}\sum\sum\hat{\varepsilon}_s + (0.091 - 0.084)\alpha'_{\perp 2,2}\sum\sum\hat{\varepsilon}_s,$$

showing that the loadings to both $I(2)$ trends are tiny and, as the testing of $\mathcal{H}_7$ showed, not significant.

Finally an expression for the short-term spread is given by:

$$s_1 - s_2 = (0.017 - 0.000)\alpha'_{\perp 2,1}\sum\sum\hat{\varepsilon}_s + (0.162 - 0.119)\alpha'_{\perp 2,2}\sum\sum\hat{\varepsilon}_s,$$

showing that the loadings to the $I(2)$ trends are tiny also in this case, but that the loadings to both trends are larger than for the bond spread, probably explaining why $\mathcal{H}_8$ was almost rejected.

10 Conclusions

This paper has discussed a number of likelihood ratio tests in the $I(2)$ model. Using these procedures we have been able to investigate the empirical regularities behind the long swings in the Dmk/$ rate. This has been done by structuring the data according to different levels of persistence using the $I(2)$ model. We have argued that to ignore such trends when they are present in the data is likely to impede a full understanding of the data. Moreover, the $I(2)$ framework enabled us to present some empirical regularities in characterizing the long swings properties of real and nominal exchange rates. The finding that the $I(1)$ or near $I(2)$ hypothesis cannot be rejected for these variables indicates a rejection of the monetary model under REH in favor of its IKE counterpart. The fact that price inflation was found to be 'purely' adjusting, whereas there was little evidence of long-run feed-back on nominal exchange rates, is also in conflict with the assumptions of the REH models. In striking contrast, these results accord well with the IKE monetary model of currency swings, helping us to resolve one of the core anomalies in international macroeconomics, the PPP puzzle (Frydman et al. 2008, 2009). Thus, the common practice of not testing for double unit roots in the data may lead economists to draw erroneous inferences from their 'statistical' analyses.

From a more general point of view, we find that the general-to-specific approach of a cointegrated VAR model is potentially very important as a way of making abductive inference in economics (Hoover, 2006). This is because it allows us to systematically search for an econometric model that is as simple as possible (but not more so), without distorting some of the information in the data. Thus, this approach should allow us to ask if the standard theory is too restrictive and, if so, what theoretical structure we should be looking for in explaining regularities in the data. Instead of leaving the investigation to the interplay between theory and the data, the specific-to-general approach imposes constraints according to the 'favored' model and is therefore likely to be detrimental to understanding the data properly from a theoretical and empirical perspective (see Juselius and Franchi, 2007).

11 Appendix: Proof of Lemma 3

When discussing the asymptotic distribution of the estimator of $\tilde{\beta} = (\beta', \beta_d')' = \rho'(\tau', \tau_{01}', \tau_0')$ it is convenient to introduce the true value denoted by β^0, and normalize the parameter and the estimator on the $p_1 \times r$ matrix $c' = (\bar{\beta}^0, 0)'$ as $\tilde{\beta}_c = \tilde{\beta}(c'\tilde{\beta})^{-1}$. Then the restrictions on $\tilde{\beta}_c$ are not linear, but an expansion of $\tilde{\beta}_c$ around $\tilde{\beta}_c^0$ shows that

$$\begin{aligned}
\tilde{\beta}_c - \tilde{\beta}_c^0 &= c_\perp(\tilde{\beta}_\perp^{0\prime} c_\perp)^{-1} \tilde{\beta}_\perp^{0\prime}(\tilde{\beta} - \tilde{\beta}^0) + O(|\tilde{\beta}_c - \tilde{\beta}^0|^2) \\
&= \begin{pmatrix} P_{\beta_\perp^0} & 0 \\ -\beta_d \bar{\beta}^{0\prime} & I_2 \end{pmatrix} (\tilde{\beta} - \tilde{\beta}^0) + O(|\tilde{\beta}_c - \tilde{\beta}^0|^2),
\end{aligned}$$

where we have used the expressions

$$c_{\perp}=\left(\begin{array}{cc}\beta_{\perp}^{0} & 0\\ 0 & I_{2}\end{array}\right),\quad \tilde{\beta}_{\perp}^{0}=\left(\begin{array}{cc}\beta_{\perp}^{0} & -\bar{\beta}^{0}\beta_{d}^{0\prime}\\ 0 & I_{2}\end{array}\right). \tag{27}$$

In the following we shall not write the superscript 0 to indicate the true value.

We decompose $\tilde{\tau}=\tilde{\tau}\rho\bar{\rho}'+\tilde{\tau}\bar{\rho}_{\perp}\rho_{\perp}'=\tilde{\beta}\bar{\rho}'+\tilde{\eta}\rho_{\perp}'$, and replace the variation independent parameters $(\rho,\tilde{\tau})$ by the variation independent parameters $(\tilde{\beta},\tilde{\eta},\rho)$. The likelihood function is denoted $\ell_{T}(\theta)$ and the score function with respect to φ_{i} is therefore

$$\begin{aligned}\frac{\partial}{\partial\varphi_{i}}\ell_{T}(\theta) &= \alpha_{i}'\Omega^{-1}\sum_{t=1}^{T}\varepsilon_{t}(x_{t-1}'P_{\beta_{\perp}}-d_{t-1}'\beta_{d}\bar{\beta}';d_{t-1}')H_{i}\\ &\quad+\bar{\rho}_{i}'\kappa\alpha_{\perp\Omega}'\sum_{t=1}^{T}\varepsilon_{t}(\Delta x_{t-1}'P_{\beta_{\perp}}-\Delta d_{t-1}'\beta_{d}\bar{\beta}';\Delta d_{t-1})H_{i}.\end{aligned}$$

The second term is always dominated by the first, so we focus on the first. We find using $P_{\beta_{\perp}}=P_{\beta_{\perp1}}+P_{\beta_{\perp2}}$ and $\beta_{\perp1}=\bar{\tau}\rho_{\perp}$. The (main) trend in $\beta_{\perp1}'x$ is $-\rho_{\perp}'(\tau'\tau)^{-1}\tau_{d}'=\nu d_{t}$, say, see (17)

$$\begin{aligned}& H_{i}'\left(\begin{array}{cc}P_{\beta_{\perp}} & -\bar{\beta}\beta_{d}'\\ 0 & I_{2}\end{array}\right)\left(\begin{array}{c}x_{t-1}\\ d_{t-1}\end{array}\right)\\ =\;& H_{i}'\left(\begin{array}{c}P_{\beta_{\perp2}}\\ 0\end{array}\right)x_{t-1}+H_{i}'\left(\begin{array}{c}-\bar{\beta}\beta_{d}'\\ 0\end{array}\right)d_{t-1}+H_{i}'\left(\begin{array}{c}P_{\beta_{\perp1}}\\ 0\end{array}\right)x_{t-1}\\ =\;& H_{i}'\left(\begin{array}{c}\overline{\beta_{\perp2}}\\ 0\end{array}\right)\beta_{\perp2}'x_{t-1}+H_{i}'\left(\begin{array}{c}-\bar{\beta}\beta_{d}'+\overline{\beta_{\perp1}}v\\ I_{2}\end{array}\right)d_{t-1}+H_{i}'\left(\begin{array}{c}\overline{\beta_{\perp1}}\\ 0\end{array}\right)(\beta_{\perp1}'x_{t-1}-vd_{t-1})\\ =\;& M_{i2}\beta_{\perp2}'x_{t-1}+M_{id}d_{t-1}+M_{i1}(\beta_{\perp1}'x_{t-1}+\rho_{\perp}'(\tau'\tau)^{-1}\tau_{d}'d_{t-1}),\end{aligned} \tag{28}$$

say, where the terms are of the order $T^{3/2},T$, and $T^{1/2}$ respectively, because the cumulated random walk in $\beta_{\perp2}'x_{t-1}$ dominates the first term, the second, d_{t-1}, is of the order of T and in the last term the process has been detrended so it is dominated by the random walk.

Because $\tilde{\beta}$ is identified we have that $H_{i}'\tilde{\beta}_{\perp}$ has full rank m_{i}, because if $\phi'H_{i}'\tilde{\beta}_{\perp}$ were zero, then $H_{i}\phi$ would be a linear combination of $\beta_{j},\ j=1,...,r$ and $h_{i}+H_{i}(\varphi_{i}+\phi)$ would satisfy the same restrictions as β_{i}. The assumption of identification then shows that $\phi=0$. This implies that also

$$\operatorname{rank}\left(H_{i}'\left(\begin{array}{cc}P_{\beta_{\perp}} & -\bar{\beta}\beta_{d}'\\ 0 & I_{2}\end{array}\right)\right)=\operatorname{rank}(M_{i2},M_{id},M_{i1})=m_{i},$$

and we exploit this as follows. Let $\operatorname{rank}(M_{i2})=m_{2i}$ so that $M_{i2}=v_{i}w_{i}'$ where v_{i} is $m_{i}\times m_{2i}$ and w_{i} is $s_{2}\times m_{2i}$ are of rank m_{2i}. Similarly let $\bar{v}_{i\perp}'M_{id}=b_{i}c_{i}'$ be of rank m_{di}, then the matrix $\bar{b}_{i\perp}'\bar{v}_{i\perp}'M_{i1}$ has rank $m_{1i}=m_{i}-m_{2i}-m_{di}$, as is seen from the display

$$\left(\begin{array}{c}\bar{v}_{i}'\\ \bar{b}_{i}'\bar{v}_{i\perp}'\\ \bar{b}_{i\perp}'\bar{v}_{i\perp}'\end{array}\right)(M_{i2},M_{id},M_{i1})=\left(\begin{array}{ccc}\bar{v}_{i}'M_{i2} & \bar{v}_{i}'M_{id} & \bar{v}_{i}'M_{i1}\\ 0 & \bar{b}_{i}'\bar{v}_{i\perp}'M_{id} & \bar{b}_{i}'\bar{v}_{i\perp}'M_{i1}\\ 0 & 0 & \bar{b}_{i\perp}'\bar{v}_{i\perp}'M_{i1}\end{array}\right). \tag{29}$$

The asymptotic behavior of the various processes in (28) are summarized in

$$\begin{aligned} T^{-3/2}\bar{v}_i' M_{i2}\beta_{\perp 2}' x_{[Tu]} &\xrightarrow{d} \bar{v}_i' M_{i2}\beta_{\perp 2}' C_2 \int_0^u W(s)ds = G_{i2}(u), \\ T^{-1}\bar{b}_i'\bar{v}_{i\perp}' M_{id} d_{[Tu]} &\rightarrow \bar{b}_i'\bar{v}_{i\perp}' M_{id} d(u)' = G_{id}(u), \\ T^{-1/2}\bar{b}_{i\perp}'\bar{v}_{i\perp}' M_{i1}(\beta_{\perp 1}' x_{[Tu]} + \rho_\perp'(\tau'\tau)^{-1}\tau_d' d_{[Tu]}) &\xrightarrow{d} \bar{b}_{i\perp}'\bar{v}_{i\perp}' M_{i1}\beta_{\perp 1}' C_1 W(u) = G_{i1}(u), \end{aligned} \tag{30}$$

and we define $G_i(u) = (G_{2i}(u)', G_{di}(u)', G_{1i}(u)')'$, and $d(u) = \lim_{T\to\infty} T^{-1}d_{[Tu]}$ which is assumed to exist, that is, the broken linear trend breaks at a given fraction of the sample.

We now define the normalizing matrices

$$A_{iT}^{-1} = (T^{3/2}v_i, Tv_{i\perp}b_i, T^{1/2}v_{i\perp}b_{i\perp})' \text{ and } A_{iT} = (T^{-3/2}\bar{v}_i, T^{-1}\bar{v}_{i\perp}\bar{b}_i, T^{-1/2}\bar{v}_{i\perp}\bar{b}_{i\perp}),$$

and find that the triangular structure (29) implies that the limit of the normalized score function is

$$T^{-1/2}A_{iT}'H_i'\sum_{t=1}^T \begin{pmatrix} P_{\beta_\perp} & -\bar{\beta}\beta_d' \\ 0 & I_2 \end{pmatrix}\begin{pmatrix} x_{t-1} \\ d_{t-1} \end{pmatrix}\varepsilon_t'\Omega^{-1}\alpha_i \xrightarrow{d} \int_0^1 G_i(dW)'\Omega^{-1}\alpha_i.$$

Thus the normalized score function with respect to the parameters $\varphi_1, \ldots, \varphi_r$ is asymptotically mixed Gaussian, because the processes $G_i(u)$ depend on $C_2W(u)$ and $\beta_{\perp 1}'C_1W(u)$, both of which are functions of $(\alpha_{\perp 1}, \alpha_{\perp 2})'W(u)$, see (7), and therefore independent of $\alpha'\Omega^{-1}W(u)$.

The i,jth block of the asymptotic conditional variance is given by the $m_i \times m_j$ matrix

$$\alpha_i'\Omega^{-1}\alpha_j\int_0^1 G_i(u)G_j(u)'du. \tag{31}$$

Similarly the main term of the information with respect to φ_i and φ_j is given by

$$\alpha_i'\Omega^{-1}\alpha_j H_j'\begin{pmatrix} P_{\beta_\perp} & -\bar{\beta}\beta_d' \\ 0 & I_2 \end{pmatrix}\sum_{t=1}^T\begin{pmatrix} x_{t-1} \\ d_{t-1} \end{pmatrix}\begin{pmatrix} x_{t-1} \\ d_{t-1} \end{pmatrix}'\begin{pmatrix} P_{\beta_\perp} & 0 \\ -\beta_d\bar{\beta}' & I_2 \end{pmatrix}H_i,$$

and normalized by $T^{-1/2}A_{iT}'$ and $T^{-1/2}A_{iT}$, the limit is given by (31). The asymptotic distribution of the estimator of φ_i is found from the usual expansion of the likelihood function and is given by

$$\{T^{-1/2}A_{iT}^{-1}(\hat{\varphi}_i - \varphi_i^0)\} \xrightarrow{d} \{\alpha_i'\Omega^{-1}\alpha_j\int_0^1 G_i(u)G_j(u)'du\}^{-1}\{\int_0^1 G_i(dW)\Omega^{-1}\alpha_i\}.$$

12 Acknowledgement

We gratefully acknowledge detailed and valuable comments and suggestions from two anonymous referees and from the editor in charge, Peter Boswijk. Support from Center for Research in Econometric Analysis of Time Series, CREATES, funded by the Danish National Research Foundation is gratefully acknowledged by the first author.

13 References

Bacchiocchi, E. and L. Fanelli, 2005, Testing the purchasing power parity through $I(2)$ cointegration techniques. Journal of Applied Econometrics 20, 749–770.

Boswijk, H. P., 2000, Mixed Normality and Ancillarity in $I(2)$ Systems. Econometric Theory 16, 878–904.

Dennis, J., Johansen, S. and K. Juselius, 2005, CATS for RATS: Manual to Cointegration Analysis of Time Series. Estima, Illinois.

Dornbusch, R., 1976, Expectations and Exchange Rate Dynamics. Journal of Political Economy, December, 1161–1174.

Frydman, R. and M. Goldberg, 2007, Imperfect Knowledge Economics: Exchange rates and Risk. Princeton University Press, Princeton.

Frydman, R. and M. Goldberg, 2008, Macroeconomic Theory for a World of Imperfect Knowledge. Capitalism and Society. 3, 1–76.

Frydman, R., Goldberg, M., Johansen S. and K. Juselius, 2008, A Resolution of the Purchasing Power Parity Puzzle: Imperfect Knowledge and Long Swings. Discussion paper 08-31, University of Copenhagen.

Frydman, R., Goldberg, M., Johansen S. and K. Juselius, 2009, Resolving the Purchasing Power Parity Puzzle: Endogenous Monetary Policy or Imperfect Knowledge? Discussion paper, University of Copenhagen, and working paper, Center on Capitalism and Society, Columbia University.

Gonzalo, J., 1994, Five alternative methods of estimating long-run equilibrium relationships. Journal of Econometrics 60, 203–233.

Hansen, H. and S. Johansen, 1999, Some tests for parameter constancy in the cointegrated VAR. Econometrics Journal 2, 306–333.

Hoover, K., 2006, The past as the future: The Marshallian approach to Post-Walrasian macro, in: D. Colander, (Ed.), Post Walrasian Economics: Beyond the Dynamic Stochastic General Equilibrium Model, Chapter 12. Cambridge University Press, Cambridge, Mass., pp. 239–257.

Johansen, S., 1992, A Representation of Vector Autoregressive Processes Integrated of Order 2. Econometric Theory 8, 188–202.

Johansen, S., 1995, Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. 2.ed. 1996, Oxford University Press, Oxford.

Johansen, S., 1997, Likelihood Analysis of the $I(2)$ Model. Scandinavian Journal of Statistics 24, 433–462.

Johansen, S., 2006a, Statistical analysis of hypotheses on the cointegrating relations in the $I(2)$ model. Journal of Econometrics 132, 81–115.

Johansen, S., 2006b, Confronting the economic model with the data, in: D. Colander, (Ed.), Post Walrasian Economics: Beyond the Dynamic Stochastic General Equilibrium Model., Chapter 12. Cambridge University Press, Cambridge, Mass., pp. 287–300.

Juselius, K., 1994, On the Duality Between Long-run Relations and Common Trends in the $I(1)$ and the $I(2)$ case. An Application to Aggregate Money Holdings.

Econometric Reviews 13, 151–178.

Juselius, K., 1995, Do the purchasing power parity and the uncovered interest rate parity hold in the long run? - An example of likelihood inference in a multivariate time-series model. Journal of Econometrics 69, 211–240.

Juselius, K., 2006, The Cointegrated VAR Model: Methodology and Applications. Oxford University Press, Oxford.

Juselius, K., 2008, The long swings puzzle. What the data tell when allowed to speak freely, in: T. C. Mills, and K. Patterson, (Eds.), Palgrave Handbook of Econometrics, Vol. 2 Empirical Econometrics, Chapter 8. Palgrave Macmillan, Basingstoke, United Kingdom.

Juselius, K. and M. Franchi, 2007, Taking a DSGE Model to the Data Meaningfully. Economics 4, pp. 1–38.

Juselius, K., and R. MacDonald, 2004, The International Parities Between USA and Japan. Japan and the World Economy 16, 17–34.

Juselius, K., and R. MacDonald, 2007, International Parity Relationships between Germany and the United States: A Joint Modelling Approach, in: A. M. Zumaquero, (Ed.), International Macroeconomics: Recent Development. Nova Science Publishers, 79-103.

Kongsted, H. C., 2003, An $I(2)$ cointegration analysis of small country import price determiniation. Econometrics Journal 12, 163–178.

Kongsted, H. C., 2005, Testing the nominal-to-real transformation. Journal of Econometrics 124, 205–225.

Kongsted, H. C. and H. B. Nielsen, 2004, Analyzing $I(2)$ Systems by Transformed Vector Autoregressions. Oxford Bulletin of Economics and Statistics 66, 379–397.

Kurita, T., 2007, $I(2)$ Cointegration analysis in the presence of deterministic shifts. Working Paper, Faculty of Economics, Fukuoka University.

Lane, P. R., 2001, The New Open Economy Macroeconomics: A Survey. Journal of International Economics 54, 235–266.

Mark, N. C., 2001, International Macroeconomics and Finance. Blackwell Publishing, Malden, Mass.

Nielsen, H. B., 2004, Cointegration analysis in the presence of outliers. The Econometrics Journal 7, 249–271.

Nielsen, H. B. and A. Rahbek, 2007, The likelihood ratio test for cointegration ranks in the $I(2)$ model. Econometric Theory 23, 615–637.

Paruolo P., 2000, Asymptotic efficiency of the two stage estimator in $I(2)$ systems. Econometric Theory 16, 524–550.

Paruolo P., 2002, Asymptotic inference on the moving average impact matrix in cointegrated $I(2)$ VAR systems. Econometric Theory 18, 673–690.

Paruolo, P. and A. Rahbek, 1999, Weak exogeneity in $I(2)$ VAR systems. Journal of Econometrics 93, 281–308.

Rahbek, A., Kongsted, H. C. and C. Jørgensen, 1999, Trend stationarity in the $I(2)$ cointegration model. Journal of Econometrics 90, 265–289.

Rogoff, K., 1996, The purchasing power parity puzzle. Journal of Economic Literature 34, 647–68.

Sarno, L. and M. P. Taylor, 2003, An empirical investigation of asset price bubbles in Latin American emerging financial markets. Applied Financial Economics 13, 635–643.

Taylor, A. and M. P. Taylor, 2004, The purchasing power parity debate. Journal of Economic Perspectives 4, 135–158.

Table 1: Estimated outlier coefficients

	Dummy variables in the model								
	$Dtax$	$Dp_{79:10}$	$Dp_{82:08}$	$Dtr_{84:01}$	$Dp_{85:02}$	$Dp_{86:02}$	$Dp_{88:08}$	$Dp_{89:02}$	$Dp_{97:07}$
$\Delta^2 pp$	0.01 [9.77]	*	*	*	*	*	*	*	0.01 [3.49]
$\Delta^2 s_{12}$	*	*	*	0.06 [2.67]	*	-0.07 [-2.35]	*	*	*
$\Delta^2 \Delta p_2$	*	*	-0.00 [-2.75]	*	*	-0.00 [-3.33]	*	*	*
$\Delta^2 b_1$	*	*	*	*	0.00 [2.36]	*	*	*	*
$\Delta^2 b_2$	*	0.00 [2.31]	-0.00 [-2.64]	*	0.00 [4.08]	-0.00 [-3.96]	*	*	*
$\Delta^2 s_1$	*	*	*	*	0.00 [2.79]	*	0.00 [5.55]	0.00 [4.40]	*
$\Delta^2 s_2$	*	0.00 [5.18]	-0.00 [-12.67]	-0.00 [-13.41]	0.00 [3.27]	*	*	*	*

t-values in brackets, * indicates a t-value < 2.0

Table 2: Misspecification tests

Multivariate tests (p-values in brackets)							
Residual autocorrelation LM(1)				$\chi^2(49)$	=		69.98 (0.03)
Residual autocorrelation LM(2)				$\chi^2(49)$	=		22.38 (1.00)
Test for normality				$\chi^2(14)$	=		**81.64** (0.00)
Test for ARCH LM(1)				$\chi^2(784)$	=		**942.43** (0.00)
Univariate tests:							
	$\Delta^2 pp$	$\Delta^2 s_{12}$	$\Delta^2 \Delta p_2$	$\Delta^2 b_1$	$\Delta^2 b_2$	$\Delta^2 s_1$	$\Delta^2 s_2$
ARCH	0.49 [0.78]	2.26 [0.32]	5.62 [0.06]	6.95 [0.03]	5.30 [0.07]	1.96 [0.37]	0.40 [0.82]
Skew.	0.01	0.22	0.10	0.19	−0.10	0.38	−0.04
Kurt.	2.88	4.16	3.63	3.63	3.90	4.38	4.87
Norm.	0.02 [0.99]	**13.38** [0.00]	5.43 [0.07]	5.65 [0.06]	**9.32** [0.01]	**16.10** [0.00]	**29.38** [0.00]

Table 3: The likelihood ratio test statistics for cointegration rank indices

$p-r$	r	$s_2=7$	$s_2=6$	$s_2=5$	$s_2=4$	$s_2=3$	$s_2=2$	$s_2=1$	$s_2=0$
7	0	1737.63 [0.00]	1402.12 [0.00]	1137.46 [0.00]	908.43 [0.00]	715.21 [0.00]	528.18 [0.00]	379.58 [0.00]	280.64 [0.00]
6	1		1115.86 [0.00]	885.59 [0.00]	659.57 [0.00]	470.53 [0.00]	319.31 [0.00]	190.76 [0.00]	160.82 [0.00]
5	2			646.64 [0.00]	455.20 [0.00]	271.26 [0.00]	**141.15** [0.09]	**94.27** [1.00]	**93.63** [1.00]
4	3				265.40 [0.00]	134.94 [0.00]	**69.81** [0.71]	**39.63** [1.00]	**40.67** [1.00]
3	4					64.60 [0.01]	32.60 [0.84]	15.58 [1.00]	22.18 [1.00]
2	5						13.04 [0.73]	2.03 [1.00]	8.06 [1.00]
1	6							0.91 [1.00]	0.17 [1.00]

Table 4: The 7 largest estimated characteristic roots

r	s_1	s_2	1	2	3	4	5	6	7
7	0	0	1.0	0.98 (0.04)	0.98 (−0.04)	0.97	0.89 (0.06)	0.89 (−0.06)	0.41
3	4	0	1.0	1.0	1.0	1.0	0.93	0.91	0.43
3	2	2	1.0	1.0	1.0	1.0	1.0	1.0	0.38
2	5	0	1.0	1.0	1.0	1.0	1.0	0.90	0.59
2	4	1	1.0	1.0	1.0	1.0	1.0	1.0	0.51

Table 5: An identified long-run structure in β

The structure : $\tilde{\beta} = (h_1 + H_1\varphi_1, \ldots, h_r + H_r\varphi_r)$, $\chi^2(10) = 9.19[0.51]$									
	pp	s_{12}	Δp_2	b_1	b_2	s_1	s_2	$t_{91.1}^{1)}$	$t^{1)}$
$\tilde{\beta}_1'$	**−0.01** [−23.49]	**0.01** [23.49]	0.00 [NA]	**1.00** [NA]	**−1.00** [NA]	0.00 [NA]	0.00 [NA]	0.00 [NA]	**0.07** [5.60]
$\tilde{\beta}_2'$	0.00 [NA]	**0.001** [6.32]	**0.09** [10.25]	0.00 [NA]	**1.00** [NA]	0.00 [NA]	**−1.00** [NA]	**0.19** [4.36]	0.00 [NA]
$\tilde{\beta}_3'$	**−0.01** [−5.51]	0.00 [NA]	**1.00** [NA]	0.00 [NA]	0.00 [NA]	**−0.74** [−16.48]	0.00 [NA]	**−0.20** [−4.83]	0.00 [NA]
$\tilde{\delta}_1'$	**−0.92**	**0.15**	0.03	0.03	0.04	0.04	0.04	0.00	−0.01
$\tilde{\delta}_2'$	**−0.51**	**−0.17**	0.02	0.02	0.02	0.03	0.02	0.00	−0.00
$\tilde{\delta}_3'$	**1.31**	**−0.07**	−0.04	−0.04	−0.05	−0.06	−0.05	0.00	0.00
α_1'	**0.39** [3.99]	**−4.47** [−3.19]	**−0.59** [−8.62]	0.00 [0.67]	**−0.03** [−2.86]	**0.03** [4.59]	**−0.04** [−4.86]		
α_2'	**0.29** [4.53]	**2.88** [3.11]	**−0.29** [−6.48]	**−0.03** [−6.58]	0.00 [0.30]	**−0.06** [−13.55]	0.00 [0.66]		
α_3'	**−0.29** [−3.00]	−1.47 [−1.06]	**−0.48** [−6.98]	0.01 [1.44]	0.00 [0.33]	**0.02** [2.67]	**−0.02** [−2.39]		

1) The trend has been multiplied by 10000.

Table 6: The common stochastic trends and their loadings

$$\begin{bmatrix} pp_t \\ s_{12,t} \\ \Delta p_{2,t} \\ b_{1t} \\ b_{2,t} \\ s_{1,t} \\ s_{2,t} \end{bmatrix} = \begin{bmatrix} 1.556 & -0.693 \\ 2.432 & 3.356 \\ 0.007 & 0.074 \\ 0.008 & 0.091 \\ 0.000 & 0.084 \\ 0.017 & 0.162 \\ 0.000 & 0.119 \end{bmatrix} \begin{bmatrix} \alpha'_{\perp 2,1} \sum\sum \hat{\varepsilon}_s \\ \alpha'_{\perp 2,2} \sum\sum \hat{\varepsilon}_s \end{bmatrix} +$$

$$+ \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \\ c_{51} & c_{52} & c_{53} & c_{54} \\ c_{61} & c_{62} & c_{63} & c_{64} \\ c_{71} & c_{72} & c_{73} & c_{74} \end{bmatrix} \begin{bmatrix} \alpha'_{\perp 2,1} \sum \hat{\varepsilon}_i \\ \alpha'_{\perp 2,2} \sum \hat{\varepsilon}_i \\ \alpha'_{\perp 1,1} \sum \hat{\varepsilon}_i \\ \alpha'_{\perp 1,2} \sum \hat{\varepsilon}_i \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \\ b_{51} & b_{52} \\ b_{61} & b_{62} \\ b_{71} & b_{72} \end{bmatrix} \begin{bmatrix} t_{91.1} \\ t \end{bmatrix}$$

where

$$\begin{bmatrix} \alpha'_{\perp 2,1} \\ \alpha'_{\perp 2,2} \end{bmatrix} = \begin{bmatrix} \underset{[-0.13]}{-0.00} & \underset{[1.06]}{0.01} & \underset{[-0.30]}{-0.00} & \underset{[NA]}{\mathbf{1.00}} & \underset{[-8.03]}{\mathbf{-0.47}} & \underset{[-7.99]}{\mathbf{-0.47}} & \underset{[NA]}{\mathbf{-0.00}} \\ \underset{[0.29]}{0.01} & \underset{[0.24]}{0.00} & \underset{[-0.78]}{-0.03} & \underset{[NA]}{\mathbf{0.00}} & \underset{[-0.77]}{-0.12} & \underset{[-0.05]}{-0.01} & \underset{[NA]}{\mathbf{1.00}} \end{bmatrix}$$

t-ratios are given in [ ] and standard errors are calculated using Paruolo (2002).

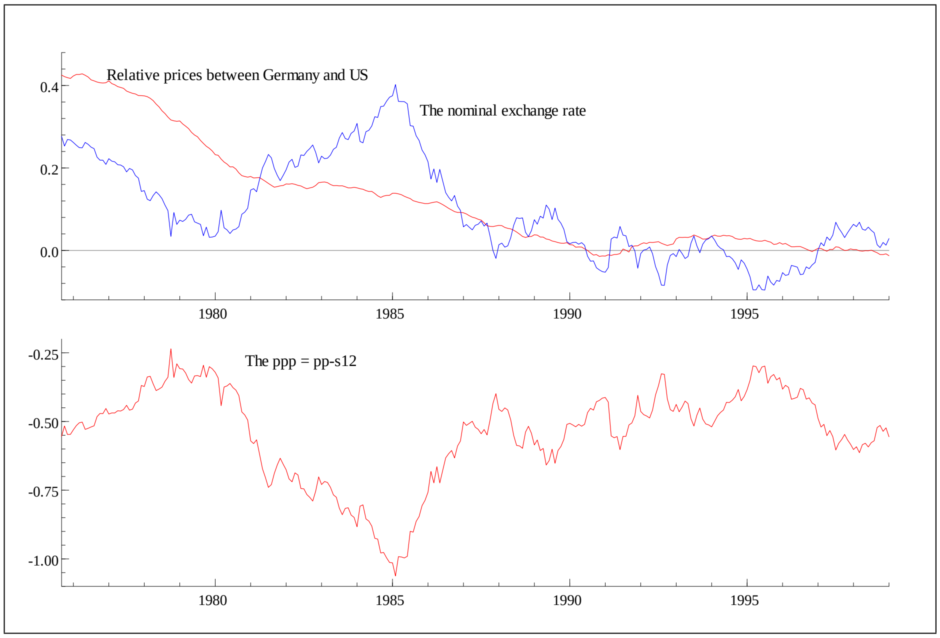

Figure 1: The graphs of the (mean and range adjusted) German-US price differential, pp, and the nominal exchange rate, s_{12} (upper panel), and the $ppp = pp - s_{12}$ (lower panel).

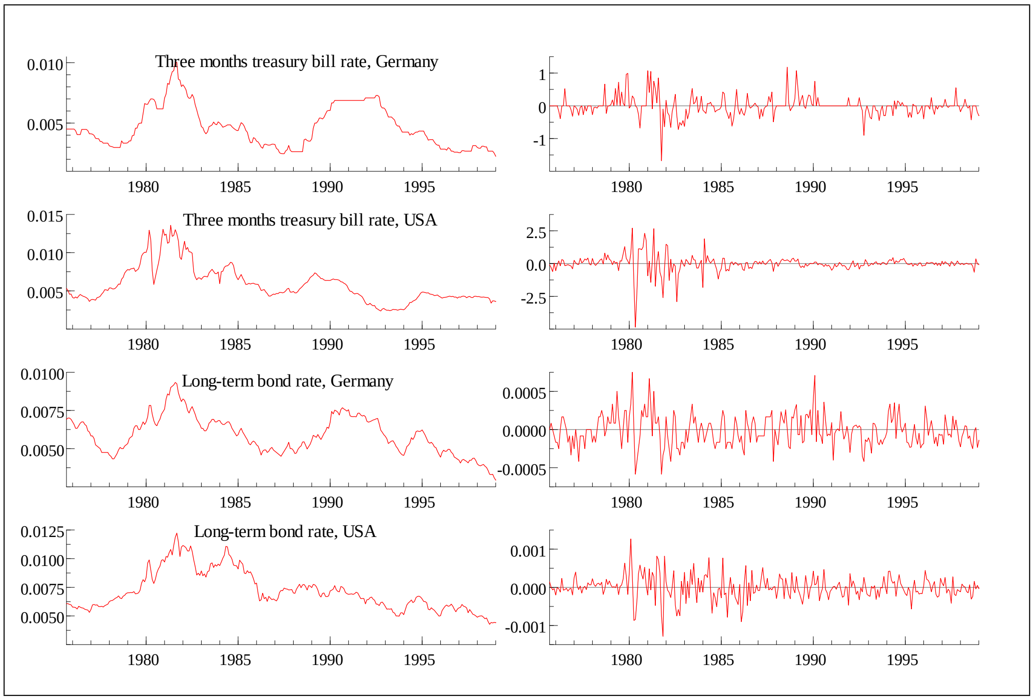

Figure 2: The graphs of German and US short-term and long-term interest rates in levels (left hand side) and differences (right hand side).

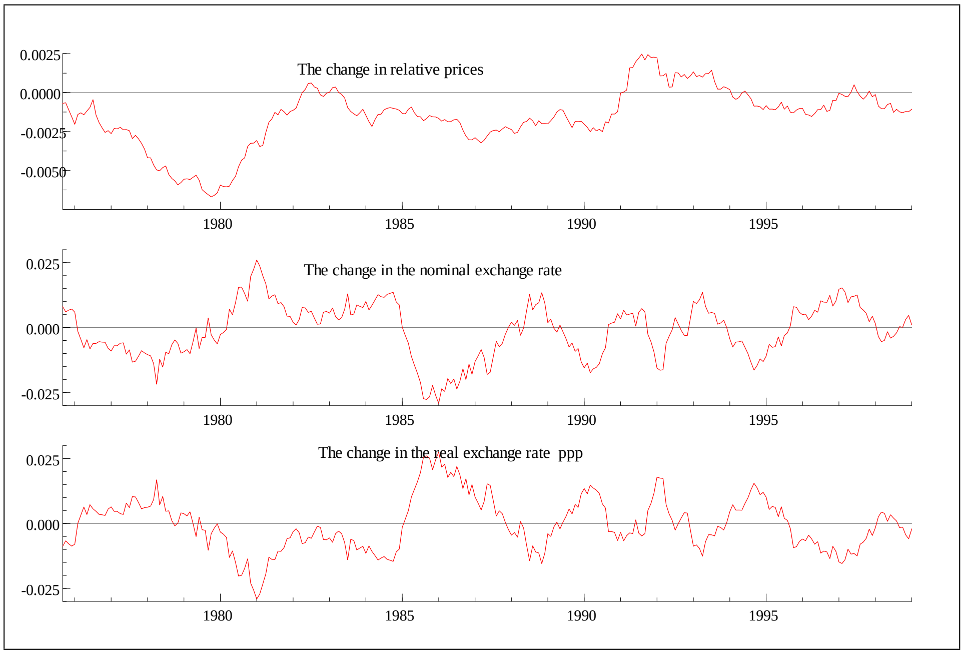

Figure 3: The graph of Δpp_t (upper panel), of $\Delta s_{12,t}$ (middle panel), and Δppp_t (lower panel. All three series are smoothed by a 12 months moving average filter.

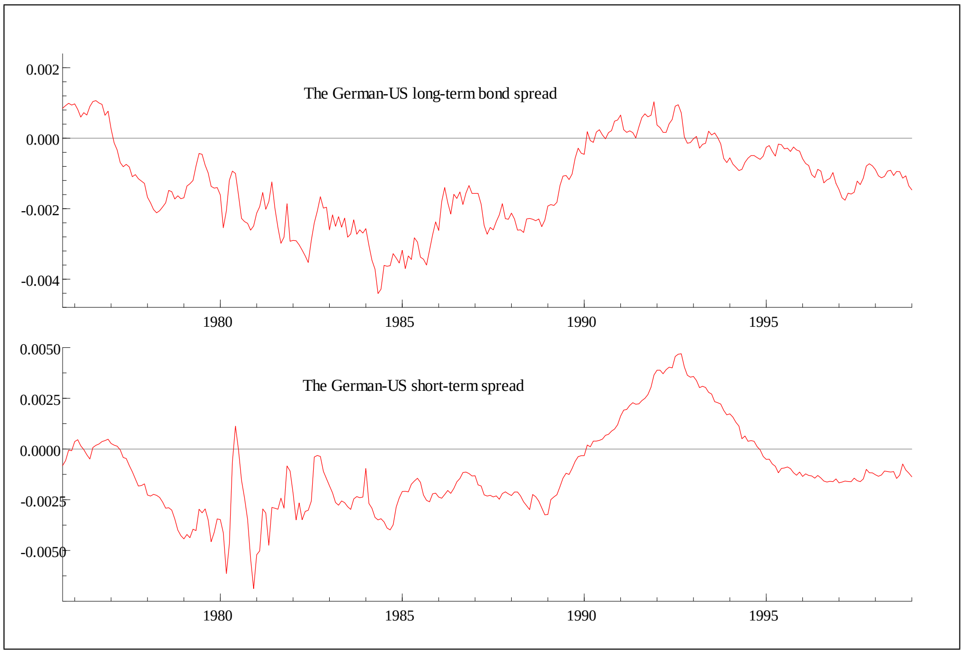

Figure 4: The German-US long-term bond spead (upper panel) and the short-term interest spread (lower panel)

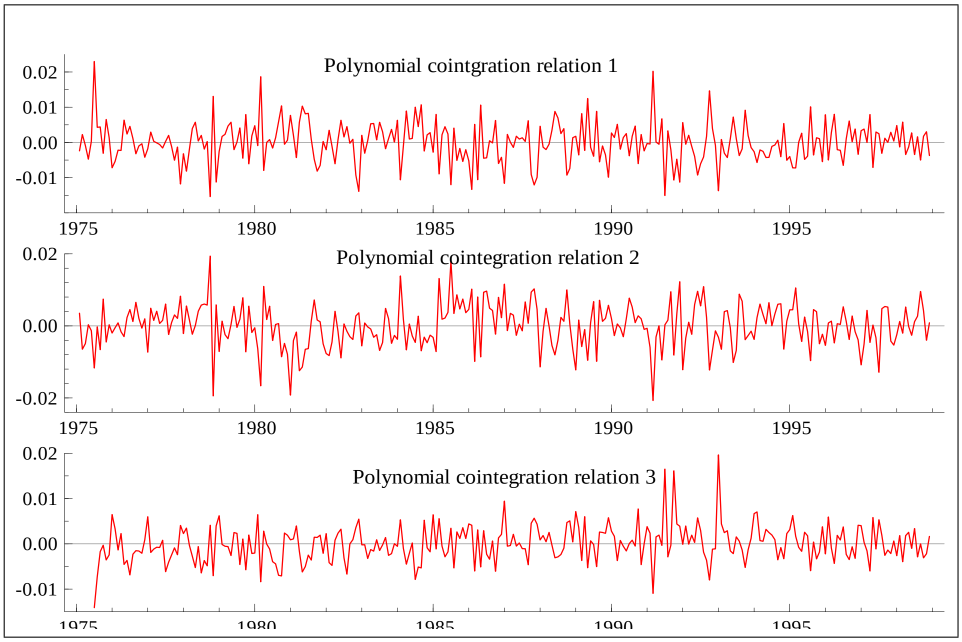

Figure 5: The graphs of the three polynomial cointegration relations. Upper panel describes the IKE relation, the middel panel the inflation expectations relation, and the lower panel the German inflation rate relation.