Testing Hypotheses in an I(2) Model with Applications to the Persistent Long Swings in the Dmk/$ Rate

Johansen, Søren; Juselius, Katarina; Frydman, Roman; Goldberg, Michael

Citation for published version (APA):
Johansen, S., Juselius, K., Frydman, R., & Goldberg, M. (2007). *Testing Hypotheses in an I(2) Model with Applications to the Persistent Long Swings in the Dmk/$ Rate*. Department of Economics, University of Copenhagen.

Discussion Papers
Department of Economics
University of Copenhagen

No. 07-34

Testing Hypotheses in an I(2) Model with Applications to
the Persistent Long Swings in the Dmk/$ Rate

Søren Johansen, Katarina Juselius,
Roman Frydman, Michael Goldberg

Testing hypotheses in an I(2) model with applications to the persistent long swings in the Dmk/$ rate

Søren Johansen[*] Economics Department, University of Copenhagen
Katarina Juselius, Economics Department, University of Copenhagen
Roman Frydman, Economics Department, New York University
Michael Goldberg, Economics Department, University of New Hampshire

December 18, 2007

Abstract

This paper discusses a number of likelihood ratio tests on long-run relations and common trends in the $I(2)$ model and provide new results on the test of overidentifying restrictions on $\beta' x_t$ and the asymptotic variance for the stochastic trends parameters, $\alpha_{\perp 1}$. How to specify deterministic components in the I(2) model is discussed at some length. Model specification and tests are illustrated with an empirical analysis of long and persistent swings in the foreign exchange market between Germany and USA. The data analyzed consist of nominal exchange rates, relative prices, US inflation rate, two long-term interest rates and two short-term interest rates over the 1975-1999 period. One important aim of the paper is to demonstrate that by structuring the data with the help of the $I(2)$ model one can achieve 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 puzzle, Forward premium puzzle, Cointegrated VAR. Likelihood inference.

1   Introduction

This paper discusses a number of likelihood ratio tests on long-run relations and common trends in the $I(2)$ model and provides new results on the test of overidentifying

restrictions on $\beta' x_t$ and the asymptotic variance for the stochastic trends parameters, $\alpha_{\perp 1}$. How to specify deterministic components in the I(2) model is discussed at some length. Model specification and tests are illustrated with an empirical analysis of long and persistent swings in the foreign exchange market between Germany and USA. The data analyzed consist of the nominal exchange rate, relative goods prices, the U.S. inflation rate, and German and U.S. long-term and short-term interest rates over the 1975-1999 period during which the \$/Dmk rate was floating.

It is clear from the past three decades of floating currencies that exchange rates have a tendency to undergo persistent swings away from purchasing power parity (PPP) for extended periods of time only to be followed by time periods in which exchange rates move persistently towards this benchmark. International macro economists have long puzzled over the long-swings behavior of floating exchange rates. They have uncovered much evidence that, although departures from PPP are ultimately bounded, the rate at which they damp out is much too slow to be consistent with the standard sticky-price monetary model of Dornbusch (1976) or its New Open Economy Macroeconomics formulations. Slow adjustment can be rationalized with the aid of an equilibrium model (Stockman, 1980, among others), but the volatility of real exchange rates (that is, the volatility of departures from PPP) is much too large to be consistent with these models.[1] The inability of exchange rate theory to explain both the high volatility and high persistence of real exchange rates is called the PPP puzzle.

As Dornbusch himself had recognized, his influential monetary model of the exchange rate is grossly inconsistent with the long-swings behavior of exchange rates.[2] Frydman and Goldberg (2007) (hereafter FG) show, however, that once the Rational Expectations Hypothesis (REH) is replaced with an Imperfect Knowledge Economics (IKE) representation of forecasting behavior, the traditional monetary model is able to generate exchange rate swings away from PPP.[3] Such behavior can occur in the model even if goods prices are assumed to be fully flexible and market participants are assumed to form their forecasts solely on the basis of macroeconomic fundamentals. In Frydman, Goldberg and Juselius (2007) (hereafter FGJ), the authors show that this IKE model provides a resolution to the PPP puzzle: with exchange rate swings, goods prices can adjust quickly to equilibrium levels while departures from PPP can damp out very slowly. The empirical analysis provided in FGJ and the one in the present paper are both based on estimating a cointegrated I(2) model of the exchange rate, relative goods prices, and short- and long-run interest rates using the methodology outlined here. To provide support for its resolution of the PPP puzzle, however, FGJ focuses on only a subset of the results from the I(2) model. By contrast, in the present

---

[1] For reviews of the PPP literature, see Rogoff (1996), Taylor and Taylor (2004), Taylor and Sarno (2003), and Mark (2003).

[2] For example, see Dornbusch and Frankel (1983).

[3] In contrast to the REH and behavioral models of aggregate outcomes, IKE only partially prespecifies how individual forecasting strategies might change over time. By only partially prespecifying change, IKE models are able to recognize the importance of imperfect knowledge without presuming that individuals are irrational. For an extensive comparsion of IKE and extant approaches to modeling individual forecasting behvaior see part I and chapter 6 in Frydman and Goldberg (2007).

paper, we examine more broadly whether the estimated I(2) model is consistent with the data.

Our aim is to demonstrate that by structuring the data with the help of the cointegrated $I(2)$ model, one can achieve a better understanding of the empirical regularities underlying the persistent swings in nominal exchange rates that are typical in periods of floating currencies. We shall argue that the cointegrated $I(2)$ model is well designed to study empirical problems characterized by different levels of persistent behavior, as it allows us to study highly persistent $I(2)$, persistent $I(1)$, and transitory $I(0)$ behavior in one model. By structuring the data in this way, we are able to present a number of 'sophisticated stylized facts' that any theory model should replicate in order to claim empirical relevance.

The econometric theory of this paper builds on Johansen (1992, 1995, 1997, 2006), Kongsted, Rahbek and Jørgensen (1999), Paruolo (2000), Paruolo (2002), Nielsen and Rahbek (2007). There are several studies in the literature that have found nominal variables, such as exchange rates, goods prices, and money supplies, to be well approximated as I(2).[4] Despite this evidence, however, there is a resistance among economists to consider economic data to be $I(2)$. This resistance can be traced to the fact that the popular REH macro models imply time series that are at most integrated of order 1. Consequently, macro economists make use of the I(1) framework, often without testing whether this framework provides an appropriate structuring of the data.

The common practice of ignoring trends in data that exhibit two roots near the unit circle may lead economists to draw erroneous inferences from their "statistical" analyses. Instead of forcing such data into an I(1) framework, it would be more useful to construct economic models that are consistent with I(2) behavior. Indeed, FGJ show that, under plausible assumptions, the IKE model of swings in Frydman and Goldberg (2007) implies near I(2) behavior for exchange rates, relative goods prices, and interest rate spreads. Thus, the finding in FGJ and in the present study that the 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 organization of the paper is as follows: Section 2 gives the theoretical background for the $I(2)$ analysis with particular attention to the role of deterministic components in the model. Section 3 discusses the Maximum Likelihood parametrization of the $I(2)$ model and section 4 shows how to test structural hypotheses in that model. Section 5 provides an ocular inspection of the data, which suggests that the nominal exchange rate, goods prices, and interest rate spreads should be modeled as I(2) variables. Section 6 estimates an unrestricted VAR model with particular attention to problem of specifying its deterministic component. Section 7 discusses the choice of rank indices in the $I(2)$ model. Section 8 reports a number of test results based on non-identifying hypotheses as a general description of the properties of the data. Finally, Section 9 reports an overidentified long-run structure and describes the dynamic adjustment of the international transmission mechanisms between Germany and the USA from the

---

[4]See Johansen (1992), Juselius (1994), Kongsted (2003, 2005), Kongsted and Nielsen (2004), and Bacchiocchi and Fanelli (2005).

mid seventies until the beginning of the EMU. Section 10 reports the estimates of the long-run common trends and discusses how they have pushed the variables of this system. Section 11 concludes.

2 Theoretical background for $I(2)$ analysis

For simplicity, the discussion of the various components in the $I(2)$ model will be based on the VAR(3) model formulated in acceleration rates, changes and levels:

$$
\begin{aligned}
\Delta^2 x_t &= \Gamma_1 \Delta^2 x_{t-1} + \Gamma \Delta x_{t-1} + \Pi x_{t-1} \\
&\quad + \Phi_s D_{s,t} + \Phi_p D_{p,t} + \Phi_{tr} D_{tr,t} + \mu_0 + \mu_1 t + \mu_2 t_{91:1} + \varepsilon_t,
\end{aligned} \tag{1}
$$

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, $D_{s,t}$ is a step dummy (...0,0,1,1,...), $D_{p,t}$ is a permanent impulse dummy (...0,0,1,0,0...), $D_{tr,t}$ is a transitory impulse dummy (...0,0,1,-1,0,0...), and all parameters are unrestricted.

Similar to the $I(1)$ model, we define the concentrated $I(2)$ model:

$$
R_{0,t} = \Gamma R_{1,t} + \Pi R_{2,t} + \varepsilon_t \tag{2}
$$

where $R_{0,t}$, $R_{1,t}$, and $R_{2,t}$ are defined by:

$$
\Delta^2 x_t = \hat{B}_1 \Delta^2 x_{t-1} + \hat{B}_2 D_{s,t} + \hat{B}_3 D_{p,t} + \hat{B}_4 D_{tr,t} + R_{0,t}, \tag{3}
$$

$$
\Delta \tilde{x}_{t-1} = \hat{B}_5 \Delta^2 x_{t-1} + \hat{B}_6 D_{s,t} + \hat{B}_7 D_{p,t} + \hat{B}_8 D_{tr,t} + R_{1,t}, \tag{4}
$$

and

$$
\tilde{x}_{t-1} = \hat{B}_9 \Delta^2 x_{t-1} + \hat{B}_{10} D_{s,t} + \hat{B}_{11} D_{p,t} + \hat{B}_{12} D_{tr,t} + R_{2,t}. \tag{5}
$$

where $\tilde{x}_t = [x_t', t, t_{91:1}]$ indicates that x_t has been augmented with a trend, and a broken trend $t_{91:1}$. Note that we need to define three types of 'residuals' in the $I(2)$ model rather than two in the $I(1)$ model. Similar to the $I(1)$ model all estimation and test procedures are based on (2).

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

$$
\Pi = \alpha \beta' \text{, where } \alpha, \beta \text{ are } p \times r \tag{6}
$$

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

$$
\alpha_\perp' \Gamma \beta_\perp = \xi \eta', \text{ where } \xi, \eta \text{ are } (p-r) \times s_1. \tag{7}
$$

Thus, the Γ matrix is no longer unrestricted in the $I(2)$ model. The first reduced rank condition (6) is associated with the variables in levels and the second (7) with the

variables in differences. The intuition is that the differenced process also contains unit roots when data are $I(2)$.

There is, however, an important difference between the first and the second condition. The former is formulated as a reduced rank condition directly on Π, whereas the latter is on a transformed Γ. Below we shall show that this is the basic reason why the ML estimation procedure needs a different parameterization than the one in (1) .

The intuition behind (7) can be seen by pre-multiplying (1) with $\alpha'_\perp$. This makes the levels component $\alpha\beta' x_{t-1}$ disappear and reduces the model to a $(p-r)$-dimensional system of equations in first- and second order differences. In this system the hypothesis of reduced rank of the matrix $\alpha'_\perp \Gamma \beta_\perp$ is tested in the usual way. Thus, the second condition is similar to the first except that the second reduced rank is formulated on the $p-r$ common driving trends, rather than on the p variables.

Using (7) it is possible to decompose $\alpha_\perp$ and $\beta_\perp$ into the $I(1)$ and $I(2)$ directions $\alpha_\perp = [\alpha_{\perp 1}, \alpha_{\perp 2}]$ and $\beta_\perp = [\beta_{\perp 1}, \beta_{\perp 2}]$. The matrices $\alpha_{\perp 1}$ and $\beta_{\perp 1}$ of dimension $p \times s_1$, and $\alpha_{\perp 2}$ and $\beta_{\perp 2}$ of $p \times s_2$ are defined by $\alpha_{\perp 1} = \overline{\alpha}_\perp \xi$, $\beta_{\perp,1} = \overline{\beta}_\perp \eta$, $\alpha_{\perp 2} = \alpha_\perp \xi_\perp$ and $\beta_{\perp 2} = \beta_\perp \eta_\perp$, where $\xi_\perp, \eta_\perp$ are the orthogonal complements of ξ and η, respectively[5] and $\overline{\alpha} = \alpha_\perp (\alpha'_\perp \alpha_\perp)^{-1}$ denotes a shorthand notation used all through the chapter.

The moving average representation of the $I(2)$ model was derived in Johansen (1992). The baseline VAR model (1) contains a constant, a trend and several dummy variables that will have to be restricted in certain ways to avoid undesirable effects. Without such restrictions the MA model can be given in its completely unrestricted form:

$$
\begin{aligned}
x_t &= C_2 \sum_{j=1}^{t} \sum_{i=1}^{j} (\varepsilon_i + \mu_0 + \mu_1 i + \Phi_s D_{s,i} + \Phi_p D_{p,i} + \Phi_{tr} D_{tr,i}) \\
&\quad + C_1 \sum_{j=1}^{t} (\varepsilon_j + \mu_0 + \mu_1 j + \Phi_s D_{s,j} + \Phi_p D_{p,j} + \Phi_{tr} D_{tr,j}) \\
&\quad + C^*(L)(\varepsilon_t + \mu_0 + \mu_1 t + \Phi_s D_{s,t} + \Phi_p D_{p,t} + \Phi_{tr} D_{tr,t}) + A + Bt
\end{aligned}
\tag{8}
$$

where A and B are functions of the initial values $x_0, x_{-1}, ..., x_{-k+1}$, and the coefficient matrices satisfy:

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

$$
\begin{aligned}
\beta' C_1 &= -\overline{\alpha}' \Gamma C_2, \quad \beta'_{\perp 1} C_1 = -\overline{\alpha}'_{\perp 1}(I - \Psi C_2), \\
\Psi &= \Gamma \overline{\beta} \overline{\alpha}' \Gamma + \Gamma_1
\end{aligned}
\tag{9}
$$

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

$$C_2 = \tilde{\beta}_{\perp 2} \alpha'_{\perp 2}. \tag{10}$$

---

[5]Note that the matrices $\alpha_{\perp 1}, \alpha_{\perp 2}, \beta_{\perp 1}$, and $\beta_{\perp 2}$ are called $\alpha_1, \alpha_2, \beta_1$ and β_2 in the many papers on $I(2)$ by Johansen. The reason why we deviate here from the simpler notation is that we need to distinguish between different β and α vectors in the empirical analysis and, hence, use the latter notation for this purpose.

It is now easy to see that the C_2 matrix has a similar reduced rank representation as C_1 in the $I(1)$ model, so that it is straightforward to interpret $\alpha'_{\perp 2} \sum \sum \varepsilon_i$ as a measure of the s_2 second order stochastic trends which load into the variables x_t with the weights $\tilde{\beta}_{\perp 2}$.

From (9) we note that the C_1 matrix in the $I(2)$ model cannot be given a simple decomposition as it depends on both the C_2 matrix and the other model parameters in a complex way. Johansen (2005) derived an analytical expression for C_1, essentially showing that:

$$C_1 = \omega_0 \alpha' + \omega_1 \alpha'_{\perp 1} + \omega_2 \alpha'_{\perp 2} \tag{11}$$

where ω_i are complicated functions of the parameters of the model (not to be reproduced here).

3 The ML procedure

The full ML procedure derived in Johansen (1997) exploits the fact that the $I(2)$ model contains $r + s_1$ cointegration relations of which the r relations, $\beta' x_t \sim I(1)$, can become stationary by polynomial cointegration, $\beta' x_t + \psi \Delta x_t \sim I(0)$, and the s_1 relations, $\beta'_{\perp 1} x_t \sim I(1)$, can become stationary by differencing, $\beta'_{\perp 1} \Delta x_t \sim I(0)$. Thus, $\tau = (\beta, \beta_{\perp 1})$ define the $r + s_1 = p - s_2$ directions in which the process is cointegrated from $I(2)$ to $I(1)$. This means that the space spanned by $\tau = (\beta, \beta_{\perp 1})$ can be determined by solving just one reduced rank regression, after which the vector space can be separated into β and $\beta_{\perp 1}$. However, this necessitates a re-parametrization of the $I(2)$ model. The following parametrization (here extended with the deterministic components discussed above) was suggested by Johansen (1997):

$$\Delta^2 x_t = \alpha(\rho' \tilde{\tau}' \tilde{x}_{t-1} + \tilde{\psi}' \Delta \tilde{x}_{t-1}) + \omega' \tilde{\tau}' \Delta \tilde{x}_{t-1} + \Phi_p D_{p,t} + \Phi_{tr} D_{tr,t} + \varepsilon_t,$$
$$\varepsilon_t \sim i.i.d.\ N_p(0, \Omega),\ t = 1, ..., T \tag{12}$$

where ρ is a $(r + s_1) \times r$ matrix which picks out the r cointegration vectors $\beta' x_t$ (so that $\rho' \tau' = \beta'$), $\psi' = -(\alpha' \Omega^{-1} \alpha)^{-1} \alpha' \Omega^{-1} \Gamma$, $\omega' = \Omega \alpha_\perp (\alpha'_\perp \Omega \alpha_\perp)^{-1} (\alpha'_\perp \Gamma \bar{\beta}, \xi)$, $\rho' \tilde{\tau}' = [\beta', \rho_0, \rho_{01}]$, $\tilde{\psi}' = [\psi', \gamma_0, \gamma_{01}]$, $\tilde{x}'_t = [x'_t, t, t_{91:1}]$ and $\Delta \tilde{x}'_t = [\Delta x'_t, 1, Ds91 : 1]$, where $t_{91.1}$ is a linear trend starting in 1991:1and $Ds91.1$ is a step dummy starting in 1991:1.

The relations $\tilde{\tau}' \Delta \tilde{x}_t$ define the $p - s_2$ stationary relations between the growth rates, of which r correspond $\beta' \Delta x_t + \rho_0 + \rho_{01} Ds91.1$ and s_1 to $\beta'_{\perp 1} \Delta x_t + \tilde{\gamma}_0 + \tilde{\gamma}_{01} Ds91 : 1$. In some cases they might be given an interpretation as medium run steady-state relations. Based on an iterative estimation algorithm, α and β are estimated subject to the reduced rank restriction(s) (7) on the Γ matrix. This is the reason why the estimates of α and β based on the ML procedure can differ to some degree from the estimates based on the $I(1)$ model.

The $FIML$ estimates of $\tau = (\beta, \beta_{\perp 1})$ are obtained using an iterative procedure which at each step delivers the solution of just one reduced rank problem. In this case the eigenvectors are the estimates of the $CI(2, 1)$ relations, $\tau' x_t$, among the variables x_t, i.e. they give a decomposition of the vector x_t into the $r + s_1$ directions $\tau = (\beta, \beta_{\perp 1})$

in which the process is $I(1)$ and the s_2 directions $\tau_\perp = \beta_{\perp 2}$ in which it is $I(2)$. The ML parameterization avoids the problem of quadratic trends by restricting the constant term, the linear trend, and the step dummy to the various cointegration relations.

The matrix ψ in (12) does not make a distinction between stationary and nonstationary components in Δx_t. For example, when x_t contains variables which are $I(2)$ as well as $I(1)$, then some of the differenced variables picked up by ψ will be $I(0)$. As the latter do not contain any stochastic $I(1)$ trends, they are by definition excludable from the polynomially cointegrated relations. The idea behind the parametrization in Paruolo and Rahbek (1999) was to express the polynomially cointegrated relations exclusively in terms of the $I(1)$ differenced variables by noticing that

$$\psi' \Delta x_{t-1} = \psi'(\bar{\tau}\tau' + \bar{\tau}_\perp \tau'_\perp)\Delta x_{t-1}$$

so that (12) can be reformulated as:

$$
\begin{aligned}
\underbrace{\Delta^2 x_t}_{I(0)} &= \alpha \left\{ [\beta', \rho_{01}, \rho_0] \right\} \begin{bmatrix} x_{t-1} \\ t_{91:1} \\ t \end{bmatrix} + [\delta', \gamma_{01}, \gamma_0] \begin{bmatrix} \Delta x_{t-1} \\ D_s 91:1_{t-1} \\ const \end{bmatrix} \\
&+ \zeta \begin{bmatrix} \beta', \rho_0, \rho_{01} \\ \beta'_{\perp 1}, \tilde{\gamma}_0, \tilde{\gamma}_{01} \end{bmatrix} \begin{bmatrix} \Delta x_{t-1} \\ D_s 91:1_{t-1} \\ const \end{bmatrix} + \Phi_p D_{p,t} + \Phi_{tr} D_{tr,t} + \varepsilon_t, \\
&\varepsilon_t \sim N_p(0, \Omega), \, t = 1, ..., T
\end{aligned}
\tag{13}
$$

where $\delta' = \psi'\bar{\tau}_\perp \tau'_\perp{}^6$ and $\zeta = \alpha\tilde{\psi}'\bar{\tilde{\tau}} + \omega'$. Note that this parameterization defines the $I(2)$ model directly in terms of stationary components.

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

We discuss in this section hypotheses on the parameters α, $\alpha_{\perp 1}$, $\alpha_{\perp 2}$, β, and τ, in he maximum likelihood parametrization (12) written as

$$\Delta^2 x_t = \alpha(\rho'\tau'x_{t-1} + \psi'\bar{\tau}_\perp \tau'_\perp \Delta x_{t-1}) + (\alpha\psi'\bar{\tau} + \omega')\tau'\Delta x_{t-1} + \varepsilon_t,$$

ignoring deterministic terms. The general theory for likelihood ratio tests for such hypotheses is given in Boswijk (2000) and Johansen (2006) and we here discuss the interpretation of the hypotheses and apply the result that likelihood ratio statistics are generally asymptotically distributed as χ^2, except in a few cases, which we describe in some more detail. In all cases the likelihood ratio tests are calculated by maximizing the likelihood function, L, with the parameters restricted by the hypothesis and without the restriction. The test statistic is $-2\log(\max_{restricted} L / \max_{unrestricted} L)$. For each case we give the degrees of freedom for the asymptotic χ^2 distribution.

---

[6]Note that the definition of δ differs from the one used in Johansen (1997, 2006).

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

$$\alpha = \begin{pmatrix} \alpha_1 \\ 0 \end{pmatrix} = \begin{pmatrix} I_{p-m} \\ 0 \end{pmatrix} \alpha_1, \tag{14}$$

or $\alpha_2 = 0$, means that the acceleration $\Delta^2 x_{2t}$ does not react 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

$$\alpha = (e_1, e_{1\perp} \phi). \tag{15}$$

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

4.2 Test on $\alpha_{\perp 1}$ and $\alpha_{\perp 2}$

When testing hypotheses on the adjustment coefficients $\alpha_{\perp 1}$ and $\alpha_{\perp 2}$ it is useful to have expressions for the asymptotic variances, so that t-test and Wald test become feasible without having to estimate the model with the restrictions imposed on $\alpha_{\perp 1}$ and $\alpha_{\perp 2}$. The maximum likelihood procedure determines the superconsistent estimators for the parameters τ, ρ, $\psi' \tau_\perp$, and $\beta = \tau \rho$, which can therefore be treated as known when discussing inference on $\alpha_{\perp 1}$ and $\alpha_{\perp 2}$.

It turns out that $\alpha_{\perp 1}$ and $\alpha_{\perp 2}$ are functions of α and the coefficient matrix $\zeta = \alpha \tilde{\psi}' \bar{\tau} + \omega'$ to $\tau' \Delta X_{t-1}$. The parameters α and ζ are determined by regression of $\Delta^2 x_t$ on the stationary processes $\beta' x_{t-1} + \psi' \bar{\tau}_\perp \tau'_\perp \Delta x_{t-1}$ and $\tau' \Delta x_{t-1}$. The asymptotic variance of $(\hat{\alpha}, \hat{\zeta})$ is therefore given by

$$asVar(\hat{\alpha}, \hat{\zeta}) = \Psi \otimes \Omega, \tag{16}$$

8

where

$$\Psi^{-1} = E \begin{pmatrix} \beta' x_{t-1} + \psi' \bar{\tau}_\perp \tau'_\perp \Delta x_{t-1} \\ \tau' \Delta x_{t-1} \end{pmatrix} \begin{pmatrix} \beta' x_{t-1} + \psi' \bar{\tau}_\perp \tau'_\perp \Delta x_{t-1} \\ \tau' \Delta x_{t-1} \end{pmatrix}'.$$

From the expression $\alpha_{\perp 1} = \bar{\alpha}_\perp \alpha'_\perp \zeta \bar{\rho}_\perp$ and $\alpha_{\perp 2} = \alpha_\perp (\alpha'_\perp \zeta \bar{\rho}_\perp)_\perp$ we can therefore find the asymptotic variances of the estimators for these from those of α and ζ using the $\delta-$ method, see Paruolo (2002).

Lemma 1 *Let the asymptotic variance of $\hat{\alpha}, \hat{\zeta}$ be $\Psi \otimes \Omega$. The asymptotic variance of $\hat{\alpha}_{\perp 2}$ is given by*

$$\alpha'_{\perp 2} \Omega \alpha_{\perp 2} \otimes [(I_p - \theta \zeta')\bar{\alpha}; \theta] \Psi [(I_p - \theta \zeta')\bar{\alpha}; \theta]', \tag{17}$$

where $\theta = \alpha_\perp \alpha'_\perp \zeta \rho_\perp (\rho'_\perp \zeta' \alpha_\perp \alpha'_\perp \zeta \rho_\perp)^{-1} \rho'_\perp$.
  For any vectors u, v the asymptotic variance of $u' \hat{\alpha}_{\perp 1} v$ is given by

$$\begin{aligned}
&(u' \bar{\alpha}(I_r; 0) \Psi (I_r; 0)' \bar{\alpha}' u)(v' \alpha'_{\perp 1} \Omega \alpha_{\perp 1} v) \tag{18} \\
&+ (v' \bar{\rho}'_\perp (\zeta' \bar{\alpha}, -I_p) \Psi (\zeta' \bar{\alpha}, -I_p)' \bar{\rho}_\perp v)(u' \bar{\alpha}_\perp \alpha'_\perp \Omega \bar{\alpha}_\perp \alpha'_\perp u) \\
&+ 2(v' \bar{\rho}'_\perp (\zeta' \bar{\alpha}, -I_p) \Psi (I_r; 0)' \bar{\alpha}' u)(v' \alpha'_{\perp 1} \Omega \bar{\alpha}_\perp \alpha'_\perp u).
\end{aligned}$$

Note that by choosing the unit vectors $u = e_i$ and $v = e_j$ we find the asymptotic variance of the element $(\hat{\alpha}_{\perp 1})_{ij} = e'_i \alpha_{\perp 1} \varepsilon_j$, and for $u = e_i + e_k, v = e_j$ we can then also find the asymptotic covariance between $(\hat{\alpha}_{\perp 1})_{ij}$ and $(\hat{\alpha}_{\perp 1})_{kj}$. The proof is found in Appendix A.

4.3 Tests on β

We consider in Section 9 test for linear restrictions on each β vector

$$\beta = (h_1 + H_1 \psi_1, \ldots, h_r + H_r \psi_r) \tag{19}$$

where h_i is $p \times 1$ and H_i is $p \times m_i$ both known, and ψ_i is an unknown parameter of dimension $m_i \times 1$.

Lemma 2 *Under the assumption that the restrictions (19) are identifying the asymptotic distribution of the likelihood ratio test is $\chi^2(\sum_{i=1}^r (p - r - m_i))$.*

The proof is given in Appendix B. The hypothesis will be applied to simplify the estimated polynomial cointegration relations. The hypothesis does not involve the coefficient δ, because the asymptotic theory for such hypotheses has not been worked out.

4.4 Tests on τ

We consider first the same restriction on all vectors in τ, that is,

$$\tau = H\phi \tag{20}$$

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

The other hypothesis corresponds to (15), that is,

$$\tau = (b, b_\perp \psi) \tag{21}$$

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

The test statistic for the first test is asymptotically distributed with degrees of freedom $(p-m)(r+s_1)$, 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 β, that is, when the hypothesis $\beta = (b, b_\perp \xi)$ is satisfied.

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

The above problem is related to the conditions (36) and (37) in the Appendix which have to be checked for this case. It is shown in Johansen (2006) that for a vector b, inference on the hypothesis $\tau = (b, b_\perp \psi)$ is χ^2 if the dimension of $sp(\beta) \cap sp(b)$ is $max(0, 1 - s_1)$. Thus if $s_1 \geq 1$ then the condition states that $b \notin sp(\beta)$.

5 An ocular inspection of the persistent behavior in the data

An ocular inspection of the data offers a first impression of the time series properties of the nominal variables and illustrates their tendency to undergo long swings. Figure 1, upper panel, shows the graphs of the price differential and nominal exchange rate. There are three features in the upper panel that are important to notice: (1) the downward sloping trend in price differentials which should be considered a stochastic trend (as a deterministic trend would not make sense); (2) the big swings in the nominal exchange rate that evolve around a downward sloping trend which looks very similar to the one behind relative goods prices; and (3) a possible change in the slope of the time trend of goods prices around 1991 (together with a shift in the level) and possibly one around 1980-81. Figure 1, lower panel, shows the mean and range adjusted $ppp = p_1 - p_2 - s_{12}$ and the real bond rate differential, $(b_1 - \Delta p_1) - (b_2 - \Delta p_2)$, where the latter is given as a 12 months moving average. The close co-movements between the two series are quite remarkable and have previously been found among others
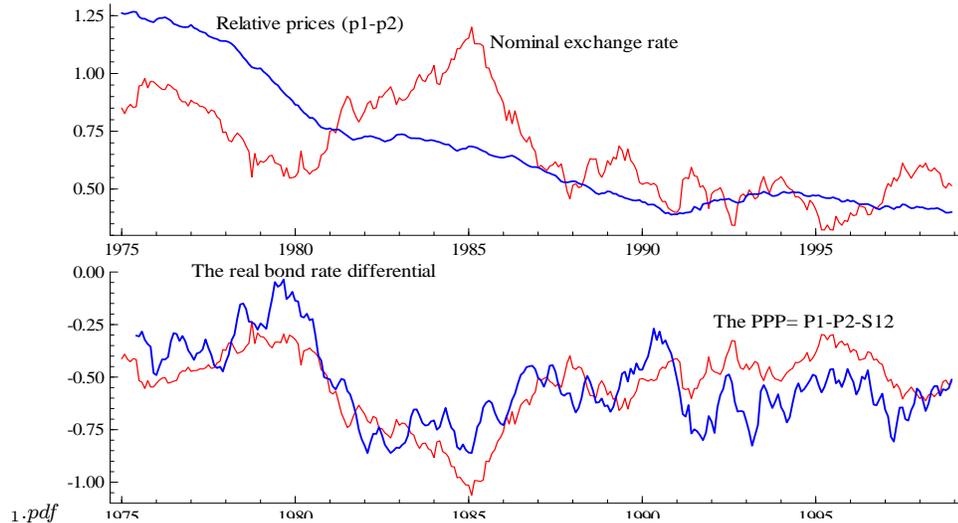
ppp b1$_b$2

Figure 1: The graphs of price differential and the nominal exchange rate in logs (upper panel), and the PPP together with bond rate differential (lower panel).

in Juselius (1995, 2006, Chapter 21) and Juselius and MacDonald (2004, 2006). An important question is whether these persistent swings should be modelled as I(1) or I(2). Since an $I(2)$ variable typically exhibits smooth behavior, which can be difficult to distinguish from an $I(1)$ variable with a linear trend, the graphs of the differenced data rather than the levels are often more informative about such $I(2)$ behavior. Also, inspecting the graph of the differenced process gives a first hint of whether the average growth rate has been approximately zero or not, whether the growth rates have been changing over the sample period, and whether there is significant mean reversion in the differences.

Monthly differences are, however, often noisy and it is useful to apply a moving average filter to the original data to single out the long-run movements from the transitory noise. This has been done in Figure 2, where we have graphed $\Delta(pp)_t$ and $\Delta s_{12,t}$ filtered through a twelve-month moving average. The very persistent trending behavior in $\Delta(pp)_t$ and $\Delta s_{12,t}$ is apparent, suggesting that it may be useful to treat both pp_t and $s_{12,t}$ as $I(2)$. We note that the inflation rate differentials exhibit more persistent behavior than exchange rate changes. In the middle panel, we have graphed the short-term interest spread and in the bottom panel the long-term interest rate spread. In both cases, the long persistent swings suggest that these variables too exhibit I(2) behavior.

Indeed, as we mentioned in the introduction, FGJ show that, under plausible assumptions, the IKE model of swings in Frydman and Goldberg (2007) implies that the exchange rate, relative goods prices, and interest rate spreads display I(2) behavior. This is the case if the macroeconomic fundamentals on which market participants form

The change in nominal exchange rates

Inflation differential (Ge-US)

The short interest rate spread

The long interest rate spread
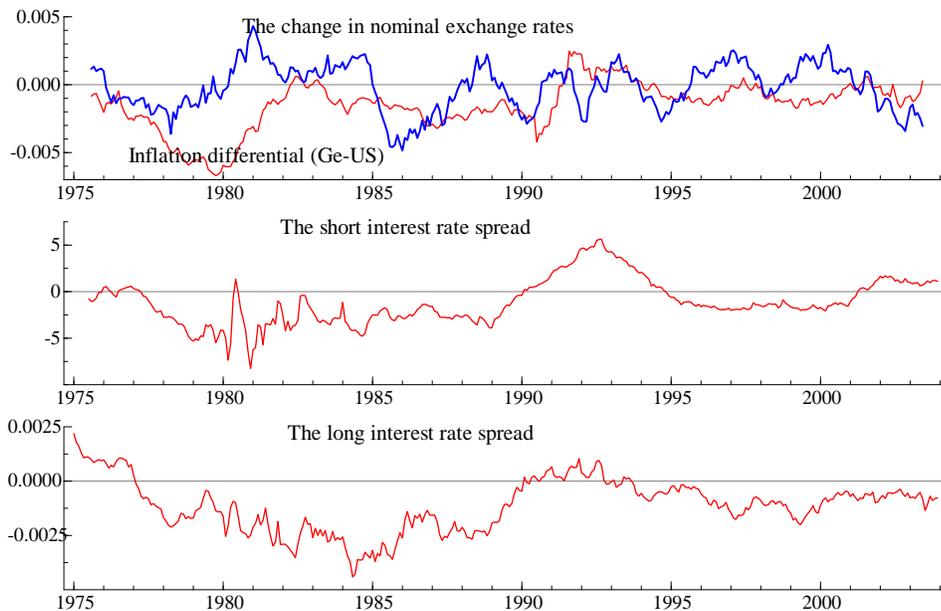
their heterogeneous exchange rate forecasts are assumed to follow random walks with drift. With this assumption, revisions of market participants' forecasting strategies lead to temporally unstable drift terms in the processes driving the endogenous variables of the model. FGJ show that these broken-trends processes can be approximated as $I(2)$.

6 Specification of deterministic terms in the empirical model

A proper specification of deterministic terms in the $I(2)$ model is mandatory for the model to yield statistically good estimates. From (8) we know that linear trends in the data can originate from several sources: (1) $E(\Delta x_{i,t}) \neq 0$ implying that some of the variables have exhibited significant linear growth over the sample period, (2) $E(\beta' x_t) = \rho_0 t \neq 0$ implying that some of the cointegration relations are trend-stationary, (3) from initial conditions. Even though one would expect a linear trend to be present in nominal prices (reflecting the fact that average inflation rates have been nonzero in most economies) it is less obvious that one should expect a linear trend in relative prices and nominal exchange rates. Since deterministic trends in $p_{1,t}$ and $p_{2,t}$ are likely to cancel in $p_{1,t} - p_{2,t}$, one should generally expect the deterministic components to be different in a model for $p_{1,t}$, $p_{2,t}$, and $s_{12,t}$ as compared to a model for $pp_t = p_{1,t} - p_{2,t}$ and $s_{12,t}$. The graph in Figure 1 shows a downward sloping trend in relative prices

over the sample period and the question is whether this trend should be treated as stochastic or deterministic. From an economic point of view, a deterministic trend in relative prices would not be plausible. From a statistical point of view it might, however, work as a local approximation. We shall include a linear trend in the VAR and then test whether it can be excluded from the model.

In the present data, the re-unification of Germany is a very significant event which is likely a priori to have strongly affected the German prices, but not the US. There are several possibilities, for example: (1) an additive effect on price levels measured by a step dummy outside the VAR dynamics at the time when the two economies were merged, (2) an additional innovational effect measured by a step dummy inside the VAR dynamics, (3) a change in the nominal growth rates which corresponds to a broken linear trend in the data.[7]

The additive step dummy effect on German prices has been removed prior to the empirical analysis using a procedure in Nielsen (2004), the remaining two effects will be properly tested within the model. There are several possibilities for how these effects will influence the specification of deterministic components:

1. p_1 and p_2 are $I(2)$ with a level shift and linear, but no broken trends, i.e. $\{p_1 - b_0 D_s 91 - b_1 t\} \sim I(2)$ and $\{p_2 - b_2 t\} \sim I(2)$, and, provided that p_1 and p_2 share the same stochastic $I(2)$ trend, $\{p_1 - p_2 - b_0 D_s 91 - (b_1 - b_2)t\} \sim I(1)$, i.e. the price differential adjusted for a linear deterministic trend is $I(1)$. If, in addition, $b_1 = b_2$ then $\{p_1 - p_2 - b_0 D_s 91\} \sim I(1)$.

2. p_1 and p_2 are $I(2)$ with a level shift and broken linear trends, i.e. $\{p_1 - b_0 D_s 91 - b_{11} t - b_{12} t_{91:1}\} \sim I(2)$ and $\{p_2 - b_{21} t\} \sim I(2)$, and, provided that p_1 and p_2 share the same stochastic $I(2)$ trend, $\{p_1 - p_2 - b_0 D_s 91 - (b_{11} - b_{21})t - b_{12} t_{91:1}\} \sim I(1)$, i.e. the price differential adjusted for a broken linear trend is $I(1)$. If, in addition, the linear trend is identical in two prices then

   $$\{p_1 - p_2 - b_0 D_s 91 - b_{12} t_{91:1}\} \sim I(1).$$

3. p_1 and p_2 are $I(2)$ with a level shift and linear (or broken linear) trends but the stochastic $I(2)$ trends do not cancel in $p_1 - p_2$. In this case, $p_1 - p_2 \sim I(2)$ and $s_{12,t}$ would also need to be $I(2)$ in order for the ppp_t to be $I(1)$.

We shall allow for a dynamic step dummy and a broken linear trend in the cointegration relations as well as in the data to be able to test the hypothesis whether the latter is significant (assuming that the step dummy has to be there). The subsequent results suggest that case 3 works best with our information set, but they also show that the $I(2)$ trend in pp_t and $s_{12,t}$ are not necessarily identical (which is also what the graphs showed) implying two stochastic I(2) trends. How can this make sense? The highly persistent, downward sloping, trend in price differentials (see Figure 1, upper panel) looks as a near $I(2)$ trend and the long swings in nominal exchange rates can be

---

[7]For a detailed description of the role of deterministic trends in the I(2) model, see Juselius, 2006, Chapter 16.

considered another highly persistent trend. It may not be exactly $I(2)$, but persistent enough for the trace test not to reject it as a double unit root.[8]

Unfortunately, the assumption of two stochastic I(2) trends seems to introduce an inconsistency in the analysis. If pp_t and $s_{12,t}$ are the only near I(2) variables then they will be long-run excludable from τ given two common stochastic I(2) trends. Thus, to maintain the assumption of two I(2) trends, there must be other variables in the system which exhibit approximately the same persistency profile. Figure 1, lower panel, clearly suggests that the real interest rate spread is a likely candidate because it is closely co-moving with the ppp, the long swings of which are associated with nominal exchange rate.

The graphs of the data in the Appendix C show a number of outlier observations. Most of them belong to the short-term interest rates in the period 1980-1982 which coincide with the period of monetary targeting. This was a very volatile period which does not seem representative for the rest of the sample. The hypothesis that the parameters of the VAR model are unchanged in this period was tested in Hansen and Johansen (1999) and clearly rejected. Because of this we have excluded the observations from 1980:2-1982:3 from the model analysis.

The properties of 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). Therefore, among the remaining outliers only those being extra-ordinarily large and those producing skewed residuals have been corrected for[9]. 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 (...,0,1,-1,0,..), are impulse dummies (...,0,1,0,...).

Thus, provided that we are willing to consider broken linear trends in the variables, but no quadratic or cubic trends, we need to restrict the trend, t, and the broken linear trend, $t_{91:1}$, to exclusively enter the $\beta' x_{t-1}$ relations, and the constant and the shift dummy $D_s 91 : 1_t$ to exclusively enter the $\delta' \Delta x_{t-1}$ and $\tau' \Delta x_{t-1}$ relations, whereas the permanent blip dummy, can enter the VAR model unrestrictedly. See the specification in (13).

Given this specification, Table 2 shows that the model passes most of the specification tests, though there are still some problems with the normality of the short-term interest rates and nominal exchange rates due to excess kurtosis and with residual ARCH for the two bond rates. However, the cointegrated VAR results are reasonably

---

[8] The nonlinear adjustment process in Bec and Rahbek (2004) might even provide a better approximation, but testing for this possibility is outside the scope of this paper.

[9] 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 statistical interpretation of the results is less reliable.

Table 1: Estimated outlier coefficients

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

robust to moderate ARCH and excess kurtosis.

7 Determining the two reduced rank indices

The number of stationary polynomial cointegrating relations, r, and the number of $I(1)$ trends, s_1, among the common stochastic trends, $p - r$, are determined by the ML trace test procedure in Johansen (1997). Table 3 reports the tests of the joint hypothesis (r, s_1, s_2) for all values of r, s_1 and s_2. The test procedure starts with the most restricted model $(r = 0, s_1 = 0, s_2 = 5)$ in the upper left hand corner, continues to the end of the first row $(r = 0, s_1 = 5, s_2 = 0)$, and proceeds similarly row-wise from left to right until the first acceptance. Based on the tests, the first acceptance is at $(r = 2, s_1 = 4, s_2 = 1)$, whereas the next acceptance is at $(r = 3, s_1 = 2, s_2 = 2)$, which is at a much higher p-values.

Since our model has a broken linear trend restricted to be in the cointegration relations, and a shift dummy restricted to the differences, the asymptotic trace test distribution provided by CATS should be shifted to the right, i.e., the test is likely to be somewhat undersized. The trace tests suggest the possibility of either $r = 2$ or 3. Thus, it is useful to perform a sensitivity check before the final choice of r, s_1, and s_2.

The characteristic roots assuming no I(2) trends show that the choice of $(r = 2, s_1 = 5)$ leaves a large unrestricted root (0.90) in the model, whereas $(r = 3, s_1 = 4)$ leaves two (0.93 and 0.90). Both cases seem to suggest a total of approximately six (near) unit roots in the model, consistent with both $(r = 2, s_1 = 3, s_2 = 2)$ and $(r = 3, s_1 = 2, s_2 = 2)$. Thus, the final decision seems to be between two (three) polynomial cointegration relations $(\beta' x_t + \delta \Delta x_t)$ and three (two) medium-run relations in differences $(\beta'_{\perp 1} \Delta x_t)$. 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

15

Table 2: Misspecification tests

Multivariate tests (p-values in brackets)						
Residual autocorrelation LM(1)		$\chi^2(49)$	$=$	67.19 (0.04)		
Residual autocorrelation LM(2)		$\chi^2(49)$	$=$	11.56 (1.00)		
Test for normality		$\chi^2(14)$	$=$	87.24 (0.00)		
Test for ARCH LM(1)		$\chi^2(784)$	$=$	944.20 (0.00)		
Univariate tests:						
	$\Delta^2 pp$	$\Delta^2 s_{12}$	$\Delta^2 \Delta p$	$\Delta^2 b_1$	$\Delta^2 b_2$	$\Delta^2 s_1$	$\Delta^2 s_2$
ARCH	1.08 [0.58]	1.65 [0.44]	4.74 [0.09]	7.57 [0.02]	5.68 [0.06]	1.92 [0.38]	0.44 [0.80]
Skew.	-0.01	0.18	0.10	0.22	-0.00	0.35	-0.02
Kurt.	2.78	4.19	3.66	3.51	3.83	4.41	5.11
Norm.	0.21 [0.90]	14.19 [0.00]	5.91 [0.05]	4.46 [0.11]	8.38 [0.02]	16.69 [0.00]	35.31 [0.00]

relation is indeed stationary. A graphical inspection of Figure 3 in Section 9 confirms that the first three β relations look very stationary. Even though $s_2 = 1$ would be easier to discuss, $s_2 = 2$ seems empirically more correct and we shall continue with the case $(r = 3, s_1 = 2, s_2 = 2)$.

Altogether, the evidence of highly persistent behavior in the data seems compelling.

8   Testing non-identifying hypotheses

The $\hat{\Pi}$ matrix in Appendix C shows that the estimated coefficients in the row describing the nominal exchange rate and the US long-term bond rate are essentially all insignificant, suggesting that there might be no long-run levels feed-back on these two variables. This hypothesis, described in Section 4.1, was individually accepted with $\chi^2(3) = 5.24\,[0.15]$ for nominal exchange rate and $\chi^2(3) = 1.27\,[0.74]$ for the US bond rate, as well as jointly accepted based on $\chi^2(6) = 6.578[0.362]$. As FGJ points out, this is exactly what one would expect to find given the temporal instability of market participants' forecasting strategies and the limited information set employed in this study. Another hypothesis of interest is the unit vector in α, also described in Section 4.1, implying that a variable is purely adjusting, i.e., the opposite of the no long-run feed-back hypothesis. We found that this hypothesis was accepted for the US inflation rate based on $\chi^2(3) = 2.41\,[0.66]$. Thus, nominal exchange rates are pushing and goods prices are adjusting, which is inconsistent with REH models of the exchange rate. This result is, however, completely consistent with the FG model of swings. See FGJ for a detailed discussion.

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

Table 3: The trace 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 [0.00]	**94.27** [0.19]	93.63 [0.02]
4	3				265.40 [0.00]	134.94 [0.00]	**69.81** [0.72]	39.63 [0.99]	40.67 [0.82]
3	4					64.60 [0.77]	32.60 [1.00]	15.58 [1.00]	22.18 [0.90]
2	5						13.04 [1.00]	2.03 [1.00]	8.06 [0.98]
1	6							-0.91 [NA]	0.17 [1.00]

test the following four hypotheses:

1. $\mathcal{H}_1 : R_1'\tau = 0$, 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 τ vectors. If accepted 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. $\mathcal{H}_2 : R_2'\tau = 0$, 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 τ. The hypothesis was borderline accepted based $\chi^2(5) = 9.99\,[0.08]$. 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. $\mathcal{H}_3 : R_3'\tau = 0$, where $R_3' = [0,0,0,0,0,0,0,0,1]$, i.e. we test whether the trend is long-run excludable from τ. 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. $\mathcal{H}_4 : R_4'\tau = 0$, 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 such trends were only borderline rejected, which supports our prior assumption that a deterministic trend should only be considered a local approximation.

Finally, we shall test five hypotheses formulated as a known vector b in τ. If accepted, they imply the variable in question is at most I(1). If, in addition, the variable in question is not a vector in β, then it is $I(1)$. None of the variables tested below can be considered a vector in β, hence, the tests are tests of I(1).

1. $\mathcal{H}_5 : \tau = (b_1, b_{1\perp}\varphi)$ where $b_1 = [1, 0, 0, 0, 0, 0, 0, 0, 0]$, i.e. we test whether relative prices is a unit vector in τ which, if accepted, would imply that pp_t is $I(1)$. The test strongly rejected based on $\chi^2(4) = 55.56\,[0.00]$.

2. $\mathcal{H}_5 : \tau = (b_2, b_{2\perp}\varphi)$ where $b_2 = [0, 1, 0, 0, 0, 0, 0, 0, 0]$, i.e. we test whether nominal exchange rate is a unit vector in τ which, if accepted, would imply that $s_{12,t}$ is (at most) $I(1)$. The test is rejected based on $\chi^2(4) = 9.76\,[0.04]$.

3. $\mathcal{H}_6 : \tau = (b_3, b_{3\perp}\varphi)$ where $b_3 = [1, -1, 0, 0, 0, 0, 0, 0, 0]$, i.e., we test whether the real exchange rate is a unit vector in τ which, if accepted, would imply that ppp_t is $I(1)$. The test is accepted based on $\chi^2(4) = 4.90\,[0.30]$. Thus, real exchange rates can be approximately considered an $I(1)$ process.

4. $H_7 : \quad \tau = (b_3, b_{3\perp}\varphi)$ where $b_4 = [0, 0, 0, 1, -1, 0, 0, 0, 0]$, i.e., we test whether the bond rate spread is a unit vector in τ which, if accepted, would imply that $b_{1,t} - b_{2,t}$ is $I(1)$. The test is accepted based on $\chi^2(4) = 3.65\,[0.46]$. Thus, bond rate differential can be approximately considered an $I(1)$ process.

5. $H_7 : \quad \tau = (b_3, b_{3\perp}\varphi)$ where $b_5 = [0, 0, 0, 0, 0, 1, -1, 0, 0]$, i.e., we test whether the short spread is a unit vector in τ which, if accepted, would imply that $s_{1,t} - s_{2,t}$ is $I(1)$. The test is only borderline accepted based on $\chi^2(4) = 8.43\,[0.08]$. Thus, short spread can be approximately considered an $I(1)$ process, but barely so.

These tests provide an approximate description of the properties of the data and should not be confused with testing structural hypotheses, which is the topic of the next section. For example, the result that ppp_t, $b_{1,t} - b_{2,t}$, and $s_{1,t} - s_{2,t}$ are $I(1)$ seems to be at odds with the previous assumption that the long swings are I(2). The explanation for this apparent inconsistency is that the I(2) approximation of the long swings trend is a borderline case. This is because it consists of a unit root together with a large root of roughly 0.86. Depending on whether size or power is considered more important, one can interpret 0.86 as a unit root or argue that it is small enough to be different from one. But regardless of whether one interprets ppp and the interest rate spreads as I(1) or I(2), the results are inconsistent with the REH monetary model, which implies that ppp and interest rate spreads are I(0). By contrast, the monetary model with IKE implies that these variables are both near I(2). Thus, beyond the inability to formally reject the unit-root hypothesis, the FG model justifies the I(2) interpretation of the results.

Table 4: An identified long-run structure in β

	pp	s_{12}	Δp_2	b_1	b_2	s_1	s_2	$t_{91.1}$	t
	The structure : $\beta = (H_1\varphi_1, \ldots, H_r\varphi_r)$, $\chi^2(10) = 9.19[0.51]$
β'_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]	**0.00** [5.60]
β'_2	0.00 [NA]	**0.00** [6.32]	0.09 [10.25]	0.00 [NA]	**1.00** [NA]	0.00 [NA]	**−1.00** [NA]	**0.00** [4.36]	0.00 [NA]
β'_3	**−0.01** [−5.51]	0.00 [NA]	**1.00** [NA]	0.00 [NA]	0.00 [NA]	**−0.74** [−16.48]	0.00 [NA]	**−0.00** [−4.83]	0.00 [NA]
δ_1	**−0.92**	**0.15**	0.03	0.03	0.04	0.04	0.04	0.00	−0.01
δ_2	**−0.51**	**−0.17**	0.02	0.02	0.02	0.03	0.02	0.00	−0.00
δ_3	**1.31**	**−0.07**	−0.04	−0.04	−0.05	−0.06	−0.05	0.00	0.00
$\beta'_{\perp 1,1}$	1.00	0.35	11.71	−1.62	−1.62	15.80	−0.61	−0.00	0.00
$\beta'_{\perp 1,2}$	2.86	1.00	33.51	−4.67	−4.69	45.23	−1.80	−0.01	0.01

9 Estimating the $I(2)$ model subject to identifying restrictions on the long-run structure

The decomposition $\tau = (\beta, \beta_{\perp 1})$ defines three stationary polynomially cointegrating relations, $\beta'_i x_t + \delta'_i \Delta x_t$, $i = 1, 2, 3$ and five stationary cointegration relations between the differenced variables, $\tau' \Delta x_t$. The difference between $\beta' x_t$ and $\beta'_{\perp 1} x_t$ is that the latter can only become stationary by differencing, whereas the former can become stationary by polynomial cointegration. How to impose and test over-identifying restrictions on β was discussed in Section 4, but not on $\beta_{\perp 1}$ or τ as their asymptotic distributions are not yet worked out.

9.1 The estimated long-run structure

To obtain standard errors of the estimated β coefficients we need to impose identifying restrictions on each of the polynomially cointegrating relations reported above. The asymptotic distribution of an identified $\hat{\beta}$ is given in Johansen (1997).

When interpreting the β relations below we shall only include the first two elements of $\delta' \Delta x_t$, corresponding to the inflation rate differentials and the depreciation/appreciation rate, as they are likely to be more relevant than the other variables.

The first relation is approximately describing the relationship between long-term interest rate spreads and ppp which have been found in many other VAR models of similar data:

$$\beta'_1 x_t + \delta'_1 \Delta x_t = \{(b_1 - 0.92\Delta p_1) - (b_2 - 0.92\Delta p_2) + 0.15\Delta s_{12} - 0.01 ppp + trend\}. \quad (22)$$

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:
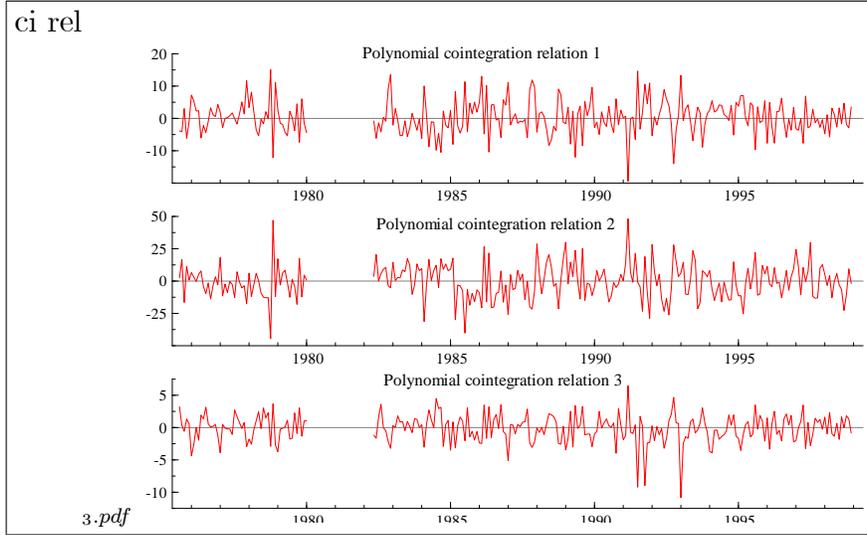
19

Figure 2: 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.

$$\beta_2' x_t + \delta_2' \Delta x_t = \{(b_2 - s_2) + 0.60\Delta p_2 - 0.51\Delta p_1 - 0.17\Delta s_{12} + 0.00 s_{12} + 0.00 t_{91.1}\} \quad (23)$$

The third relation, essentially a relation for German inflation rate, is similar to the relation found in Juselius and MacDonald (2006) and describes the latter as (almost) homogeneously related to US inflation rate, German short-term interest rate, and the change in the Dmk/\$ rate:

$$\beta_3' x_t + \delta_3 \Delta x_t = \{1.31\Delta p_1 - 0.31\Delta p_2 - 0.74 s_1 - 0.07\Delta s_{12} + 0.00 pp - 0.00 t_{91.1}\} \quad (24)$$

All three relations contain a tiny, but significant, trend effect which is more difficult to interpret. The most likely explanation is the usual one 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 (22) might account for some perceived productivity differential between the two economies. In (23) 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 (24) 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 exhibited a long-run adjustment to trend-adjusted relative prices. Figure 2 shows that the three polynomially cointegrating relations are very stationary.

The $\widehat{\beta}_{\perp 1,1}' x_t$ relations are $CI(2,1)$ cointegrating relations which only become stationary by differencing. Thus, $\widehat{\beta}_{\perp 1}' \Delta x_t$ could be interpretable as partially specified

20

medium-run steady-state relations. As we are not yet able to impose and test overidentifying restrictions on the estimated vectors, interpreting the unrestricted estimates does not make much sense and will not be done here.

9.2 The dynamics of the short-run adjustment

Table 5 reports the estimates of the short-run adjustment coefficients associated with the polynomially cointegrating relations, $\alpha_i = 1, 2, 3$, and the coefficients $\zeta_i = 1, ..., 5$ associated with the changes in the five equilibrium errors, $\tau_i' \Delta x_t$. The number of estimated coefficients is large, making it difficult to summarize the main results in a simple way. We shall not make detailed comments on the results, but instead give a cursory description of the basic adjustment mechanisms in this system.

The estimated α_1 shows that all variables, except the German long-term bond rate, react very significantly on the equilibrium error from the IKE relation, pointing to its importance for the international transmission mechanisms. The estimated α_2 shows that relative prices, nominal exchange rates, and US inflation rate and the two German interest rates (whereas not the US rates) react significantly to a deviation between expected inflation and its determinants. The estimated α_3 is consistent with the interpretation of the third cointegration relation as a relation for German inflation rate, as it is essentially prices which are reacting on an equilibrium error, albeit the short-term interest rates show some small effects.

Given that the nominal exchange rate was found to exhibit no long-run feed-back effects in Section 8, it is somewhat surprising that there are two significant α coefficients in the exchange rate equation. However, the test for a zero row in α in Section 8 was for β relations between the levels of variables, whereas the estimated α coefficients in 9.2 correspond to polynomial cointegration relations containing variables in levels and differences. Furthermore, the combined relation $(\alpha_{1,2}\beta_{1,t}' + \alpha_{2,2}\beta_{2,t}')$ suggests that the two significant cointegration relations almost neutralize each other, nonetheless with some small but significant evidence of the first relation being important for the nominal exchange rates. Thus, the result suggests that the *ppp* may act as an anchor for exchange rates, even though we do not expect its relationship to be completely stable over time.

One of the important questions in international macro is why prices and exchange rates adjust so sluggishly to the *ppp*. The answer provided by FGJ is that, with imperfect knowledge, equilibrium in the goods markets is no longer characterized by PPP, but by a cointegrating relationship between *ppp*, $b_1 - b_2$, and $\Delta p_1 - \Delta p_2$. FGJ show that relative goods prices adjust to this equilibrium relation extremely fast, which can be seen from the estimates in Table 5 which are also reported in FGJ. The estimated adjustment coefficient $\hat{\alpha}_{11} = 0.39$ shows that relative goods prices (and US inflation rate with $\hat{\alpha}_{13} = -0.59$) adjust very fast to the first cointegration relation. By contrast, the adjustment of relative prices to *ppp* is very slow -(0.39x0.01). Thus, *ppp* acts as an anchor for prices, but the chain is very long indeed.

The estimated coefficients of ζ_i may not be highly interesting for the following

Table 5: The short-run dynamic adjustment

$$
\begin{bmatrix}
\Delta^2 pp_t \\
\Delta^2 s_{12,t} \\
\Delta^2 \Delta p_{2,t} \\
\Delta^2 b_{1t} \\
\Delta^2 b_{2,t} \\
\Delta^2 s_{1,t} \\
\Delta^2 s_{2,t}
\end{bmatrix}
=
\begin{bmatrix}
\underset{[3.99]}{\mathbf{0.39}} & \underset{[4.53]}{\mathbf{0.29}} & \underset{[-3.00]}{\mathbf{-0.29}} \\
\underset{[-3.19]}{\mathbf{-4.47}} & \underset{[3.11]}{\mathbf{2.88}} & \underset{[-1.05]}{-1.47} \\
\underset{[-8.62]}{\mathbf{-0.59}} & \underset{[-6.48]}{\mathbf{-0.29}} & \underset{[-6.98]}{\mathbf{-0.48}} \\
\underset{[0.67]}{0.00} & \underset{[-6.57]}{\mathbf{-0.03}} & \underset{[1.44]}{0.01} \\
\underset{[-2.86]}{\mathbf{-0.03}} & \underset{[0.29]}{0.00} & \underset{[0.33]}{0.00} \\
\underset{[4.59]}{\mathbf{0.03}} & \underset{[-13.55]}{\mathbf{-0.06}} & \underset{[2.67]}{\mathbf{0.02}} \\
\underset{[-4.86]}{\mathbf{-0.04}} & \underset{[0.66]}{0.00} & \underset{[-2.39]}{\mathbf{-0.02}}
\end{bmatrix}
\begin{bmatrix}
\beta_1' x_{t-1} + \delta_1' \Delta x_{t-1} \\
\beta_2' x_{t-1} + \delta_2' \Delta x_{t-1} \\
\beta_3' x_{t-1} + \delta_3' \Delta x_{t-1}
\end{bmatrix}
+
$$

$$
+
\begin{bmatrix}
\underset{[1.12]}{0.85} & \underset{[0.48]}{0.24} & \underset{[2.58]}{\mathbf{0.93}} & \underset{[-0.02]}{-0.32} & \underset{[0.01]}{0.08} \\
\underset{[-0.68]}{-7.40} & \underset{[1.57]}{11.35} & \underset{[-0.36]}{-1.88} & \underset{[-1.00]}{-274.41} & \underset{[0.99]}{95.90} \\
\underset{[-1.58]}{-0.84} & \underset{[0.05]}{0.02} & \underset{[-5.52]}{\mathbf{-1.39}} & \underset{[0.25]}{3.38} & \underset{[-0.25]}{-1.17} \\
\underset{[-11.16]}{\mathbf{-0.59}} & \underset{[-6.10]}{\mathbf{-0.21}} & \underset{[2.39]}{\mathbf{0.06}} & \underset{[-9.16]}{\mathbf{-12.26}} & \underset{[9.16]}{\mathbf{4.28}} \\
\underset{[1.73]}{0.14} & \underset{[-8.24]}{\mathbf{-0.45}} & \underset{[0.57]}{0.02} & \underset{[-9.86]}{\mathbf{-20.69}} & \underset{[9.85]}{\mathbf{7.23}} \\
\underset{[2.91]}{\mathbf{0.16}} & \underset{[2.95]}{\mathbf{0.11}} & \underset{[17.62]}{\mathbf{0.46}} & \underset{[5.98]}{\mathbf{8.41}} & \underset{[-6.00]}{\mathbf{-2.95}} \\
\underset{[3.52]}{\mathbf{0.23}} & \underset{[14.71]}{\mathbf{0.64}} & \underset{[-1.76]}{-0.05} & \underset{[-7.30]}{\mathbf{-12.02}} & \underset{[7.29]}{\mathbf{4.20}}
\end{bmatrix}
\begin{bmatrix}
\beta_1' \Delta x_{t-1} \\
\beta_2' \Delta x_{t-1} \\
\beta_3' \Delta x_{t-1} \\
\beta_{\perp 1,1}' \Delta x_{t-1} \\
\beta_{\perp 1,2}' \Delta x_{t-1}
\end{bmatrix}
+
\begin{bmatrix}
\varepsilon_{1,t} \\
\varepsilon_{2,t} \\
\varepsilon_{3,t} \\
\varepsilon_{4,t} \\
\varepsilon_{5,t} \\
\varepsilon_{6,t} \\
\varepsilon_{7,t}
\end{bmatrix}
$$

$$
\text{with } \Omega =
\begin{bmatrix}
1.00 & & & & & & \\
0.12 & 1.00 & & & & & \\
-0.58 & -0.07 & 1.00 & & & & \\
-0.04 & -0.04 & 0.15 & 1.00 & & & \\
0.00 & 0.03 & 0.15 & 0.31 & 1.00 & & \\
-0.04 & -0.03 & 0.12 & 0.35 & -0.07 & 1.00 & \\
0.02 & 0.06 & 0.01 & 0.18 & 0.19 & 0.11 & 1.00
\end{bmatrix}
$$

Table 6: The estimates of the common stochastic trends

	ε_{pp}	$\varepsilon_{s_{12}}$	$\varepsilon_{\Delta p_2}$	ε_{b_1}	ε_{b_2}	ε_{s_1}	ε_{s_2}
The estimates of the first order stochastic trends, $\alpha_{\perp 1}$							
$\alpha'_{\perp 1,1}$	0.02 [0.32]	−0.03 [−0.85]	0.08 [1.01]	−0.14 [−0.53]	**−1.61** [−7.89]	**1.32** [2.45]	−0.17 [−0.79]
$\alpha'_{\perp 1,2}$	−0.01 [−0.48]	−0.00 [−0.01]	0.04 [0.88]	−0.33 [−1.75]	**−1.29** [−8.50]	0.57 [1.44]	−0.14 [−0.88]
The estimates of the second order stochastic trends, $\alpha_{\perp 2}$							
$\alpha'_{\perp 2,1}$	−0.00 [−0.13]	0.01 [1.06]	−0.00 [−0.30]	**1.00** [NA]	**−0.47** [−8.03]	**−0.47** [−7.99]	**−0.00** [NA]
$\alpha'_{\perp 2,2}$	0.01 [0.29]	0.00 [0.24]	−0.03 [−0.78]	**0.00** [NA]	−0.12 [−0.77]	−0.01 [−0.05]	**1.00** [NA]

reason: The estimated model is formulated in second differences, which for the interest rates and U.S. inflation rate means over-differencing. Thus, the highly significant coefficients in the last four rows are likely to compensate for this.

10   The driving forces

Table 6 reports the estimates of the common stochastic trends where $\alpha_{\perp 1}$ and $\alpha_{\perp 2}$ define the first and second order stochastic trends as a linear function of the VAR residuals. The two $\alpha_{\perp 1}$ vectors are determined by the chosen normalization of $\beta_{\perp 1}$, whereas $\alpha_{\perp 2}$ has been normalized and just-identified by the choice of the two zero coefficients.

As discussed in Section 2, the estimates of the second order trends are more straightforward to interpret and we shall mostly focus on them. Based on the estimates in Table 6, the first stochastic I(2) trend, $\alpha'_{\perp 2,1} \sum \sum \hat{\varepsilon}_s$, seems to be generated from the twice cumulated shocks to the bond spread and to the German term spread with almost equal weights (roughly 0.5, 0.5), whereas the second trend, $\alpha'_{\perp 2,2} \sum \sum \hat{\varepsilon}_s$, seems to have been generated from the twice cumulated shocks to the US short term interest rate.

Even though the estimates of the I(1) stochastic trends are less straightforward to interpret, it is quite interesting to note that only the interest rates coefficients are significant. Since, cumulated shocks both to the long-term and short-term interest rates are highly significant, it means that there are not just one stochastic trend driving the term structure, but at least two.[10]

The coefficients to Δp_2 and the pp are completely insignificant as are the coefficients to nominal exchange rates. The former result seems very plausible given the previous finding that prices seem to be purely adjusting (see also Juselius and MacDonald, 2004 and 2006). But, the finding that exchange rate shocks are completely insignificant may seem surprising, given that the nominal exchange rate was found to have no long-run levels feed-back. On the other hand, the fact that it was found to significantly adjust to the polynomial cointegration relations can explain the lack of significant effects in

---

[10]This is consistent with the findings in Johansen and Juselius (2001).

Table 7: The common stochastic trends and their loadings

$$
\begin{bmatrix}
pp_t \\
s_{12,t} \\
\Delta p_{2,t} \\
b_{1t} \\
b_{2,t} \\
s_{1,t} \\
s_{2,t}
\end{bmatrix}
=
\begin{bmatrix}
1.56 & -0.69 \\
2.43 & 3.35 \\
-0.00 & 0.07 \\
-0.01 & 0.09 \\
-0.03 & 0.08 \\
0.00 & 0.16 \\
-0.05 & 0.12
\end{bmatrix}
\begin{bmatrix}
\boldsymbol{\alpha}'_{\perp 2,1} \sum \sum \hat{\boldsymbol{\varepsilon}}_s \\
\boldsymbol{\alpha}'_{\perp 2,2} \sum \sum \hat{\boldsymbol{\varepsilon}}_s
\end{bmatrix}
+
$$

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

Table 6. We also note that this adjustment explains a small part of the variation in the nominal exchange rate.

Finally, Table 7 reports the weights with which the I(2) stochastic trends have affected the variables of the system. We do not report the weights of the I(1) trends as these are complicated functions of the estimated matrices. See Section 2.

We note that the first I(2) trend seems to have affected pp and s_{12} with coefficients of the same sign, but not the same magnitude, whereas the second I(2) trend appears to influence these variables with coefficients of opposite signs. Interpreting the first trend as the long-run downward sloping trend visible in both relative prices and nominal exchange rates suggests that it is associated with twice cumulated shocks to the bond spread and the German term spread. Interpreting the second trend as capturing the long swings movements predominantly in nominal exchange rates suggests that it is predominantly associated with the twice cumulated shocks to the US short- term interest rate. Moreover, this second trend has affected relative prices with a fairly small coefficient that is opposite in sign to the one on nominal exchange rates. This reflects, of course, the tendency of the nominal exchange rate to move away from relative prices for extended periods of time. Finally, it is notable that nominal interest rates seem to have been affected in particular by the second I(2) trend which is consistent with the FG model as discussed in FGJ.

11　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 allowing for I(2) trends is important for the econometric analysis of macroeconomic data 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 estimated four stochastic trends summarizing the different levels of persistence seemed primarily to derive from shocks to the interest rates. Thus, many of the basic assumptions behind standard rational expectations models (such as the PPP and Fisher parity) are incompatible with the empirical evidence. The sluggish adjustment of relative goods prices and nominal exchange rates to the *ppp* is also incompatible with standard rational expectations models. The fact that price inflation was found to be 'purely' adjusting, whereas nominal exchange rates showed no long-run feed-back is also in conflict with the assumptions of such models. In striking contrast, these results accord well with the IKE monetary model in FG.

To conclude, we find the general-to-specific approach of a cointegrated VAR model to be 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 address the question: if the standard theory is too simple, what then? In contrast, the specific-to-general approach is designed to replicate the favorite theory model and is, therefore, likely to have a built-in theory bias (see Juselius and Franchi, 2007).

12　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.

Bec, F. and A. Rahbek (2004), Vector Equilibrium Correction Models with Non-Linear Discontinuous Adjustments. *Econometrics Journal* 7, 628–651.

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.

Dornbusch, R. and J.A. Frankel (1988), The Flexible Exchange Rate System: Experience and Alternatives, in S. Borner, ed., *International Finance and Trade*, London:

Macmillan, reprinted in J.A. Frankel ed., *On Exchange Rates*, Cambridge MA: The MIT Press, chapter 1,1995.

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

Frydman, R., Goldberg, M.. K. Juselius, and S. Johansen (2007), Reexamining the Purchasing Power Parity Puzzle with Imperfect Knowledge Economics, Preprint, University of Copenhagen.

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.

Johansen, S. (1992a), An $I(2)$ Cointegration Analysis of the Purchasing Power Parity between Australia and the United States. In C. Hargreaves (Ed.), *Macroeconomic Modelling of the Long Run*, 229–248. Edward Elgar.

Johansen, S. (1992b), 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. (2006), Statistical analysis of hypotheses on the cointegrating relations in the $I(2)$ model. *Journal of Econometrics* 132, 81–115.

Johansen, S and Juselius, K. (2001) Controlling inflation in a cointegrated vector autoregressive model, Discussion paper 03, Institute of Economics, University of Copenhagen, www.econ.ku.dk/okokj/

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*, 13,2.

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

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

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

Juselius, K. and Franchi, M. (2007), Taking a DSGE Model to the Data Meaningfully, *Economics*, 4, 1.

Jørgensen, C. (1998), A simulation study of tests in the cointegrated VAR model. Ph.D. thesis, Department of Economics, University of Copenhagen.

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.

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

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.

Nielsen, B. and A. Rahbek (2000) Similarity issues in cointegration Analysis. *Oxford Bulletin of Economics and Statistics* 62, 5–22.

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.

Rogoff, K. (1996), The purchasing power parity puzzle. *Journal of Economic Literature* 34, 647-68.

13 Appendices

A Proof of Lemma 1

It is convenient to introduce $\xi = \alpha'_\perp \zeta \bar{\rho}_\perp$, so that $\alpha_{\perp 1} = \bar{\alpha}_\perp \xi$ and $\alpha_{\perp 2} = \alpha_\perp \xi_\perp$. Because the estimates of $\rho, \tau, \beta, \delta$ are superconsistent we can find the asymptotic variance of $\hat{\alpha}_{\perp 1}$ and $\hat{\alpha}_{\perp 2}$ by finding derivatives of $\alpha_{\perp 1}$ and $\alpha_{\perp 2}$ with respect to α and ζ, and then transform the asymptotic variance of $(\hat{\alpha}, \hat{\zeta})$ by an application of the $\delta-$ method.

It is convenient to express the matrix expansion

$$(\alpha + u)'_\perp = \alpha'_\perp - \alpha'_\perp u \bar{\alpha}' + O(|u|^2)$$

in terms of differentials

$$(d\alpha'_\perp) = -\alpha'_\perp (d\alpha) \bar{\alpha}'. \tag{25}$$

As an example of the $\delta-$method we can find the asymptotic variance of $\hat{\alpha}_\perp$ from (16) and (25) as

$$asVar(\hat{\alpha}_\perp) = \bar{\alpha}(I_r; 0) \Psi(I_r; 0)' \bar{\alpha}' \otimes \alpha'_\perp \Omega \alpha_\perp.$$

Applying (25) we find derivatives of $\xi = \alpha'_\perp \zeta \bar{\rho}_\perp, \xi_\perp$, and $\bar{\alpha}_\perp$:

$$\begin{aligned}
d\xi &= -\alpha'_\perp (d\alpha) \bar{\alpha}' \zeta \bar{\rho}_\perp + \alpha'_\perp (d\zeta) \bar{\rho}_\perp, \\
d\xi_\perp &= \bar{\xi} \bar{\rho}'_\perp (\zeta' \bar{\alpha}(d\alpha)' - (d\zeta)') \alpha_{\perp 2}, \\
d\bar{\alpha}_\perp &= -\bar{\alpha}(d\alpha)' \bar{\alpha}_\perp.
\end{aligned}$$

Hence from $\alpha_{\perp 2} = \alpha_\perp \xi_\perp$, we get

$$d\alpha_{\perp 2} = (d\alpha_\perp) \xi_\perp + \alpha_\perp (d\xi_\perp) = -(I_p - \alpha_\perp \bar{\xi} \bar{\rho}'_\perp \zeta') \bar{\alpha}; \alpha_\perp \bar{\xi} \bar{\rho}'_\perp) \begin{pmatrix} (d\alpha)' \\ (d\zeta)' \end{pmatrix} \alpha_{\perp 2},$$

where we find

$$\theta = \alpha_\perp \bar{\xi} \bar{\rho}'_\perp = \alpha_\perp \alpha'_\perp \zeta \rho_\perp (\rho'_\perp \zeta' \alpha_\perp \alpha'_\perp \zeta \rho_\perp)^{-1} \rho'_\perp,$$

which gives the expression (17). We next find

$$\begin{aligned}
d\alpha_{\perp 1} &= d(\bar{\alpha}_\perp) \alpha'_\perp \zeta \bar{\rho}_\perp + \bar{\alpha}_\perp d(\alpha'_\perp) \zeta \bar{\rho}_\perp + \bar{\alpha}_\perp \alpha'_\perp (d\zeta) \bar{\rho}_\perp, \tag{26} \\
&= -\bar{\alpha}(d\alpha)' \alpha_{\perp 1} - \bar{\alpha}_\perp \alpha'_\perp (d\alpha) \bar{\alpha}' \zeta \bar{\rho}_\perp + \bar{\alpha}_\perp \alpha'_\perp (d\zeta) \bar{\rho}_\perp.
\end{aligned}$$

In order to find the elements of the asymptotic variance of $\hat{\alpha}_{\perp 1}$ it is enough to have an expression for the asymptotic variance of $v' \hat{\alpha}_{\perp 1} u$ for any vectors v, u. We find from (26)

$$\begin{aligned}
v' d\alpha_{\perp 1} u &= -u' \bar{\alpha}(d\alpha)' \alpha_{\perp 1} v - u' \bar{\alpha}_\perp \alpha'_\perp (d\alpha) \bar{\alpha}' \zeta \bar{\rho}_\perp v + u' \bar{\alpha}_\perp \alpha'_\perp (d\zeta) \bar{\rho}_\perp v \\
&= -v' \alpha'_{\perp 1} (d\alpha) \bar{\alpha}' u + u' \bar{\alpha}_\perp \alpha'_\perp [(d\zeta) - (d\alpha) \bar{\alpha}' \zeta] \bar{\rho}_\perp v \\
&= N_1 + N_2
\end{aligned}$$

Then $Var(v' d\alpha_{\perp 1} u) = Var(N_1) + Var(N_2) + 2Cov(N_1, N_2)$, which is the expression in (18).

B Proof of Lemma 2

We apply the maximum likelihood parameterization (12), especially the parameters ρ, τ and $\beta = \tau\rho$. In order to apply the results in Johansen (2006) we have to normalize τ and β on $\bar{\tau}^0$ and $\bar{\beta}^0$ respectively, that is, so that $\beta'\bar{\beta}^0 = I_r$ and $\tau'\bar{\tau}^0 = I_{r+s}$. We split the parameters τ, ρ into the variation free parameters $\beta = \tau\rho$, $\eta = \tau\rho_\perp$ and ρ, and let $\beta = \beta(\psi)$ be given by (19)

$$\beta(\psi) = (h_1 + H_1\psi_1, \ldots, h_r + H_r\psi_r).$$

B.1 Normalization of parameters

We define the normalized versions of β, τ and the corresponding ρ, as functions of the free parameters ψ, η, ρ, using the decomposition $\tau = \tau\rho\bar{\rho}' + \tau\rho_\perp\bar{\rho}'_\perp = \beta(\psi)\bar{\rho}' + \eta\bar{\rho}'_\perp$:

$$
\begin{aligned}
\tilde{\beta} &= \tilde{\beta}(\psi) = \beta(\psi)(\bar{\beta}^{0\prime}\beta(\psi))^{-1}, \\
\tilde{\tau} &= \tilde{\tau}(\psi, \eta, \rho) = (\beta(\psi)\bar{\rho}' + \eta\bar{\rho}'_\perp)(\bar{\tau}^{0\prime}\beta(\psi)\bar{\rho}' + \bar{\tau}^{0\prime}\eta\bar{\rho}'_\perp)^{-1}, \\
\tilde{\rho} &= \tilde{\rho}(\psi) = \bar{\tau}^{0\prime}\beta(\psi)(\bar{\beta}^{0\prime}\beta(\psi))^{-1}.
\end{aligned}
$$

which satisfies $\bar{\beta}^{0\prime}\tilde{\beta} = I_r$, $\bar{\tau}^{0\prime}\tilde{\tau} = I_{r+s_1}$, and $\tilde{\beta} = \tilde{\tau}\tilde{\rho}$.

The asymptotic theory for hypotheses on the parameters in the $I(2)$ model, Johansen (2006), is developed in terms of the parameters

$$
\begin{aligned}
B_1 &= \bar{\beta}^{0\prime}_{\perp 1}\tilde{\beta}(\psi) = B_1(\phi_{1B}, \phi_2), \\
B_2 &= \bar{\beta}^{0\prime}_{\perp 2}\tilde{\beta}(\psi) = B_2(\phi_{1B}, \phi_2), \\
C &= \bar{\beta}^{0\prime}_{\perp 2}\tilde{\tau}\tilde{\rho}_\perp = \phi_{1C},
\end{aligned}
$$

where $\phi_1 = (\phi_{1B}, \phi_{1C})$ and ϕ_2, are defined below, in such a way that the estimators for B_2 and ϕ_2 are T^2 consistent and the estimators for B_1, C, and $\phi_1 = (\phi_{1B}, \phi_{2B})$ are T consistent. Moreover the asymptotic distribution of $(T\hat{B}_1, T^2\hat{B}_2)$ is mixed Gaussian and so is the asymptotic distribution of $T\hat{C}$. The asymptotic distribution of $T^2\hat{\phi}_2$ and $T\hat{\phi}_1$ are only mixed Gaussian under some further conditions, which we discuss below.

B.2 The parameters ϕ_2, ϕ_{1B}, and ϕ_{1C}

We start defining the parameters ϕ_2 which determine the variation of B_2 through

$$\bar{\beta}^{0\prime}_{\perp 2}\beta(\psi) = \bar{\beta}^{0\prime}_{\perp 2}(h_1 + H_1\psi_1, \ldots, h_r + H_r\psi_r) = \bar{\beta}^{0\prime}_{\perp 2}H_1(\psi_1 - \psi_1^0), \ldots, \bar{\beta}^{0\prime}_{\perp 2}H_r(\psi_r - \psi_r^0)$$

In order to find the effective number of parameters we write $\bar{\beta}^{0\prime}_{\perp 2}H_i = a_i b'_i$, where a_i, $s_1 \times g_i$, and b'_i, $g_i \times m_i$, are of rank $g_i = \text{rank}(\bar{\beta}^{0\prime}_{\perp 2}H_i) \leq \min(m_i, s_2)$. We define the $\sum_{i=1}^r g_i$ parameters

$$\phi_2 = (\phi_{21}, \ldots, \phi_{2r}) = (b'_1(\psi_1 - \psi_1^0), \ldots, b'_r(\psi_r - \psi_r^0)),$$

and the $\sum_{i=1}^{r}(m_i - g_i)$ parameters

$$\phi_{1B} = (\phi_{1B1}, \ldots, \phi_{1Br}) = (b'_{1\perp}(\psi_1 - \psi_1^0), \ldots, b'_{r\perp}(\psi_r - \psi_r^0)),$$

and finally the $s_1 s_2$ parameters

$$\phi_{1C} = \bar{\beta}_{\perp 2}^{0\prime} \eta.$$

Thus the number of parameters is $\sum_{i=1}^{r} m_i + s_1 s_2$. Note that

$$\psi_i - \psi_i^0 = \bar{b}_i b'_i(\psi_i - \psi_i^0) + \bar{b}_{i\perp} b'_{i\perp}(\psi_i - \psi_i^0) = \bar{b}_i \phi_{2i} + \bar{b}_{i\perp} \phi_{1Bi}.$$

The parameters ϕ_{1B}, ϕ_{1C}, and ϕ_2 are varying freely.

B.3 Derivatives of parameter functions

We first investigate the derivatives of B_2 with respect to the parameters ψ, and find

$$\frac{\partial}{\partial \psi} B_2 = \bar{\beta}_{\perp 2}^{0\prime} [\frac{\partial}{\partial \psi} \beta(\psi)] (\bar{\beta}^{0\prime} \beta(\psi))^{-1} + \bar{\beta}_{\perp 2}^{0\prime} \beta(\psi) [\frac{\partial}{\partial \psi} (\bar{\beta}^{0\prime} \beta(\psi))^{-1}]$$

$$\frac{\partial^2}{\partial \psi^2} B_2 = \bar{\beta}_{\perp 2}^{0\prime} [\frac{\partial^2}{\partial \psi^2} \beta(\psi)] (\bar{\beta}^{0\prime} \beta(\psi))^{-1} + \bar{\beta}_{\perp 2}^{0\prime} \beta(\psi) [\frac{\partial^2}{\partial \psi^2} (\bar{\beta}^{0\prime} \beta(\psi))^{-1}]$$
$$+ 2 \bar{\beta}_{\perp 2}^{0\prime} [\frac{\partial}{\partial \psi} \beta(\psi)] [\frac{\partial}{\partial \psi} (\bar{\beta}^{0\prime} \beta(\psi))^{-1}].$$

For $\psi = \psi^0$ we have $\beta(\psi^0) = \beta^0$, so that $\bar{\beta}_{\perp 2}^{0\prime} \beta(\psi_0) = \bar{\beta}_{\perp 2}^{0\prime} \beta^0 = 0$, and $\bar{\beta}^{0\prime} \beta(\psi^0) = I_r$, which means that because $\frac{\partial^2}{\partial \psi^2} \beta(\psi) = 0$, we find

$$\frac{\partial}{\partial \psi} B_2|_{\psi=\psi^0} = \bar{\beta}_{\perp 2}^{0\prime} [H_1(d\psi_1), \ldots, H_r(d\psi_r)] = (a_1 b'_1(d\psi_1), \ldots, a_r b'_r(d\psi_r)), \quad (27)$$

$$\frac{\partial^2}{\partial \psi^2} B_2|_{\psi=\psi^0} = 2 \bar{\beta}_{\perp 2}^{0\prime} [\frac{\partial}{\partial \psi} \beta(\psi)|_{\psi=\psi^0}] [\frac{\partial}{\partial \psi} (\bar{\beta}^{0\prime} \beta(\psi))^{-1}|_{\psi=\psi^0}]. \quad (28)$$

This implies, using $d\psi_i = \bar{b}_i(d\phi_{2i}) + \bar{b}_{i\perp}(d\phi_{1Bi})$, that

$$\frac{\partial}{\partial \phi_2} B_2|_{\psi=\psi^0} = (a_1(d\phi_{21}), \ldots, a_r(d\phi_{2r})), \quad (29)$$

$$\frac{\partial}{\partial \phi_{1B}} B_2|_{\psi=\psi^0} = (a_1 b'_1 \bar{b}_{1\perp}(d\phi_{1B1}), \ldots, a_r b'_r \bar{b}_r(d\phi_{1Br})) = 0. \quad (30)$$

Similarly

$$\bar{\beta}_{\perp 2}^{0\prime} \frac{\partial}{\partial \phi_{1B}} \beta(\psi)|_{\psi=\psi^0} = (a_1 b'_1 \bar{b}_{1\perp}(d\phi_{1B1}), \ldots, a_r b'_r \bar{b}_r(d\phi_{1Br})) = 0$$

so that from (28) we find

$$\frac{\partial^2}{\partial \phi_{1B}^2} B_2|_{\psi=\psi^0} = 0. \tag{31}$$

The derivative of $B_1 = \bar{\beta}_{\perp 1}^{0\prime} \tilde{\beta}(\psi)$ with respect to ϕ_{1B} is

$$\frac{\partial}{\partial \phi_{1B}} B_1|_{\psi=\psi^0} = \bar{\beta}_\perp^{0\prime} H_1 \bar{b}_{1\perp}(d\phi_{1B1}), \ldots, \bar{\beta}_\perp^{0\prime} H_r \bar{b}_{r\perp}(d\phi_{1Br})] = 0, \tag{32}$$

which will be discussed below, and finally we find the derivative of C

$$\frac{\partial}{\partial \eta} C|_{\psi=\psi^0, \eta=\eta^0, \rho=\rho^0} = \bar{\beta}_{\perp 2}^{0\prime}(d\eta) \bar{\rho}_\perp^{0\prime} \rho_\perp^0 = d\phi_{1C}, \tag{33}$$

$$\frac{\partial}{\partial \psi} C|_{\psi=\psi^0, \eta=\eta^0, \rho=\rho^0} = \bar{\beta}_{\perp 2}^{0\prime} \frac{\partial}{\partial \psi} \beta(\psi)|_{\psi=\psi^0} \bar{\rho}_\perp^{0\prime} \rho_\perp^0 = 0, \tag{34}$$

$$\frac{\partial}{\partial \rho} C|_{\psi=\psi^0, \eta=\eta^0, \rho=\rho^0} = \bar{\beta}_{\perp 2}^{0\prime}[\beta^0 \frac{\partial}{\partial \rho} \bar{\rho}'|_{\rho=\rho^0} + \eta^0 \frac{\partial}{\partial \rho} \bar{\rho}'_\perp|_{\rho=\rho^0}]\rho_\perp^0 = 0, \tag{35}$$

because $\bar{\rho}^{0\prime}\rho_\perp^0 = 0$, $\bar{\beta}_{\perp 2}^{0\prime}\beta^0 = 0$, and $\bar{\beta}_{\perp 2}^{0\prime}\eta^0 = 0$.

B.4 Conditions for asymptotic χ^2 distribution

The conditions are expressed in terms of the derivative at the true value, see Johansen (2006, Theorem 5). There are two conditions, the first states that although B_2 may depend on ϕ_1, the dependence is very small close to the true value, and the second implies that we can split the parameter ϕ_1 into ϕ_{1B} which locally determines B_1 and ϕ_{1C} which locally determines C. Thus at the true value it should holds that

$$\frac{\partial B_2}{\partial \phi_1} = 0, \ \frac{\partial^2 B_2}{\partial \phi_1^2} = 0, \tag{36}$$

$$\frac{\partial B_1}{\partial \phi_{1C}} = 0, \ \frac{\partial C}{\partial \phi_{1B}} = 0. \tag{37}$$

We find from (30) and (31) that (36) is satisfied. Because B_1 does not depend on ϕ_{1C}, we have $\frac{\partial B_1}{\partial \phi_{1C}} = 0$. On the other hand C does depend on ϕ_{1B}, but from (34) we have $\frac{\partial C}{\partial \phi_{1B}} = 0$ at the true value so that (37) holds. The consequence of this is that asymptotic inference is χ^2 and we only have to find the degrees of freedom, that is, the difference between the number of identified parameters with and without restrictions.

B.5 Number of parameters

The number of parameters in the unrestricted parameters B_1, B_2, and C are $s_1 r$, $s_2 r$ and $s_1 s_2$ respectively. We next want to show that when β is identified by the linear restrictions, the number of parameters in the model is $s_1 s_2 + \sum_{i=1}^r m_i$, which gives the

degrees of freedom. $\sum_{i=1}^{r}(p-r-m_i)$. From (32), (29), and (33) it follows that at the true value we have

$$
\begin{aligned}
\frac{\partial B_1}{\partial \phi_{1B}}(d\phi_{1B}) &= (\bar{\beta}_{\perp 1}^{0\prime} H_1 \bar{b}_{1\perp}(d\phi_{11B}), \ldots, \bar{\beta}_{\perp 1}^{0\prime} H_r \bar{b}_{r\perp}(d\phi_{1rB})), \\
\frac{\partial B_2}{\partial \phi_2}(d\phi_2) &= (a_1(d\phi_{21}), \ldots, a_r(d\phi_{2r})), \\
\frac{\partial C}{\partial \phi_{1C}}(d\phi_{1C}) &= d\phi_{1C}.
\end{aligned}
$$

We want to determine the rank of these matrices to determine the number of parameters in the restricted model. We now need the result, see Johansen (2007).

Lemma 3 *If $\beta_i = h_i + H_i\psi_i$ is identified, then the $(p-r) \times m_i$ matrix $\beta'_\perp H_i$ has rank m_i.*

It follows that $\bar{\beta}_\perp^{0\prime} H_i$ has rank m_i, and therefor so has the matrix

$$
\bar{\beta}_\perp^{0\prime} H_i(\bar{b}_i, \bar{b}_{i\perp}) = (\bar{\beta}_{\perp 1}^0, \bar{\beta}_{\perp 2}^0)' H_i(\bar{b}_i, \bar{b}_{i\perp}) = \begin{pmatrix} \bar{\beta}_{\perp 1}^0 H_i \bar{b}_i & \bar{\beta}_{\perp 1}^0 H_i \bar{b}_{i\perp} \\ a_i & 0 \end{pmatrix}.
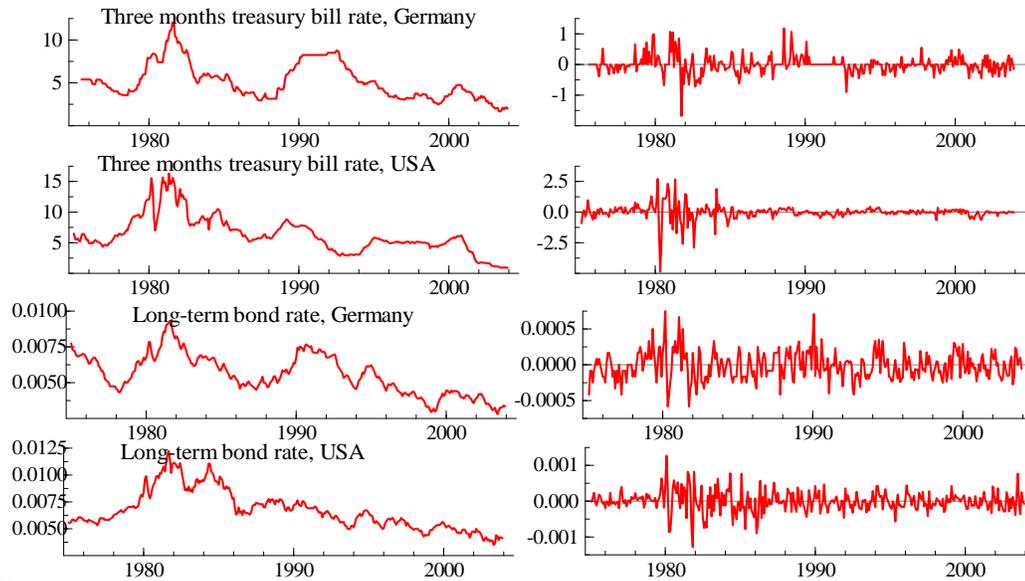$$

The rank of a_i is g_i so the rank of $\bar{\beta}_{\perp 1}^0 H_i \bar{b}_{i\perp}$ is $m_i - g_i$. Therefore

$$
\begin{aligned}
rank(\frac{\partial vec(B_1)}{\partial vec(\phi_{1B})}) &= \sum_{i=1}^{r} \mathrm{rank}(\bar{\beta}_{\perp 1}^0 H_i \bar{b}_{i\perp}) = \sum_{i=1}^{r}(m_i - g_i), \\
rank(\frac{\partial vec(B_2)}{\partial vec(\phi_2)}) &= \sum_{i=1}^{r} \mathrm{rank}(a_i) = \sum_{i=1}^{r} g_i, \\
rank(\frac{\partial vec(C)}{\partial(\phi_{1C})}) &= s_1 s_2.
\end{aligned}
$$

This shows that the number of parameters is $s_1 s_2 + \sum_{i=1}^{r} m_i$, and hence that the test for the identifying restrictions is asymptotically distributed as χ^2 with $\sum_{i=1}^{r}(p-r-m_i)$ degrees of freedom.

C   Data

data

Three months treasury bill rate, Germany

Three months treasury bill rate, USA

Long-term bond rate, Germany

Long-term bond rate, USA

$_4.pdf$
Graphs of the short and the long-term interest rates, levels in the r.h.s. and differences in the l.h.s.