



Working Paper

Interpretation of Cointegration Coefficients: A Paradox, a Solution and Empirical Evidence a paradox, a solution and empirical evidence

Author(s):

Müller, Christian

Publication Date:

2003-12

Permanent Link:

<https://doi.org/10.3929/ethz-a-004642626> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

K O F

Konjunkturforschungsstelle
Swiss Institute for
Business Cycle Research

Arbeitspapiere/ Working Papers

Christian Müller

Interpretation of Cointegration
Coefficients: A Paradox, a Solution
and Empirical Evidence

No. 82, December 2003

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Interpretation of Cointegration Coefficients A Paradox, a Solution and Empirical Evidence

Christian Müller

Konjunkturforschungsstelle (KOF) an der
Eidgenössischen Technischen Hochschule Zürich

CH-8092 Zürich, Switzerland

Tel.: ++41-(0)1-632 46 24

Fax: ++41-(0)1-632 12 18

Email: cmueller@kof.gess.ethz.ch

This version: December 5, 2003

Abstract

The concept of cointegration (see e.g., Engle and Granger, 1987; Johansen, 1988) has extensively been used to model equilibrium relationships (see e.g., Johansen and Juselius, 1990; Ericsson, 1998). The links between economic and econometric concepts are now well understood and they have become part of the standard tools of empirical analysis. At the same time, however, the dynamics of the off-equilibrium situation have been met with relatively little interest on part of economic interpretations. This paper derives a paradox in which the econometric analysis is more likely to reveal the true causal links within an economic model the less valid this model actually is. A testing procedure is proposed and the results are illustrated using U.S., Japanese, German and Swiss data.

JEL classification: C32, D40, E31, E32

Keywords: cointegration, equilibrium adjustment, forecasting, rational expectations

*I do thank Erdal Atukeren for many helpful comments. The usual disclaimer applies.

1 Introduction

The concept of cointegration (see e.g., Engle and Granger, 1987; Johansen, 1988) has been extensively used to model economic equilibrium relationships (see e.g., Johansen and Juselius, 1990; Johansen, 1995; Hubrich, 2001). The links between economic and econometric concepts in modelling equilibria are now well understood and are part of the standard tools of empirical analysis. Loosely speaking, economic equilibrium relationships have their counterparts in cointegration relationships whose existence can be tested and whose parameters can be estimated. The other side of the coin is the necessary adjustment back to the equilibrium once it has temporarily been distorted. Technically, this adjustment mechanism has also been analysed quite extensively. For instance, Ericsson, Hendry and Mizon (1998) and Ericsson (1992) look at the implications for inference in cointegrated systems in the presence, or rather absence of equilibrium adjustment in one direction or another.

However, there is hardly an agreement on the economic interpretation of a significant reaction to deviations from equilibrium. A common approach is to view those variables which do not show significant equilibrium correction behaviour as the long-run forcing, or weakly exogenous variables. This sometimes has led to the conclusions that the remaining variables are the dependent or endogenous variables with respect to the long-run equilibrium relationship. For example, Juselius and MacDonald (2003) investigating interest rate parities between Japan and the USA find that the US bond rate is not weakly exogenous with respect to the Japanese-US equilibrium relationship. This leads them to the conclusion that this

”seems to be against the simple expectation’s hypothesis, which predicts that short-term interest rates drive long-term rates.” (p. 13).

The purpose of this paper is to provide a framework for economically and econometrically consistent interpretation of equilibrium adjustment coefficients. It will turn out that the reasoning crucially depends on the economic priors underlying the econometric analysis. In particular, when rational expectations are part of the hypothesis to be tested, then a paradox situation may arise. In such a situation the correct econometric and economic

inference would label disjunct sets of variables as exogenous or endogenous, respectively.

The remainder of the paper is organised as follows. After a brief description of the problem, the framework of the analysis will be sketched. Empirical examples in section 4 illustrate the main issues and section 5 discusses the results. An informal test is proposed to cope with the issue.

2 Economic Hypotheses and Their Econometric Counterparts

Very often, economic hypotheses pose the existence of equilibrium relationships between observable variables. A very convenient and also very popular way of thinking about such relationships is the concept of cointegration. For example, let y_t and z_t represent $(n_1 \times T)$ and $(n_2 \times T)$ vectors of time series which are all integrated of order $d, d = 1, 2, \dots$. Also, let x_t be a linear combination of y_t and z_t given by $x_t = \beta'(y_t', z_t)'$ (β is a $((n_1 + n_2) \times r)$ matrix, $0 < r < n_1 + n_2$) such that x_t is a $(r \times T)$ time series or vector of time series which is integrated of order $d - j, j = 1, 2, \dots$. Then, the r linear combinations x_t are called cointegration relations and the matrix β is the cointegration matrix.

Not surprisingly, the number of possible cointegration relationships is large considering the large number of economic time series that exist. In order to find meaningful β , usually economic theory is employed. For instance, collecting interest rates, gross domestic product and the monetary aggregate M3 into a vector, one should expect to find one such β which represents a demand for money relationship (see, e.g., Ericsson, 1998). Further typical examples are the Fisher hypothesis, the expectation hypothesis about the term structure (EHTS), the purchasing power parity, and the uncovered interest rate parity (UIP) to name but a few. In general, all kinds of market equilibria relationships are often considered candidates for cointegration relations.

However, establishing the existence of an equilibrium is only one part of the story. Since all cointegration relations are defined by their long-run properties, short-term deviations from the equilibrium occur frequently. Therefore, the other part is the necessary

correction towards the equilibrium once it has been shocked. In general, for each single linear combination within $\beta'(y'_t, z'_t)'$ at least one variable in y_t or z_t must change when an off-equilibrium situation is observed. This reaction ensures that the equilibrium will eventually be re-established. At a first glance, it appears natural to regard those variables which do not adjust as the exogenous or predetermined variables with respect to the long run equilibrium. In the following, this standard situation will be generalised to a case in which expectations rather than realisations of variables are considered.

3 A General Representation

Assume there exists a $(n_1 \times 1)$ vector of endogenous variables, y_t , which depends on a $(n_2 \times 1)$ vector of exogenous variables, z_t ($n_2 \geq n_1$) in the following way:

$$\begin{aligned} y_t &= \Lambda \mathcal{E}_t(z_{t+s}) + \eta_{1,t} \\ \Delta z_t &= \eta_{2,t} \\ \eta_{i,t} &= A_i(L)\varepsilon_{i,t}, \quad i = 1, 2 \end{aligned} \tag{1}$$

As usual, L is the lag-operator with $x_t L^i = x_{t-i}$, and $\Delta = 1 - L$ denotes the first difference operator. The terms $A_i(L) = I_{n_i} - A_{i,1}L - A_{i,2}L^2 - \dots - A_{i,p+1}L^{p+1}$, $i = 1, 2$ denote polynomials in the lag operator of length $p + 1$ at most. Their roots are lying strictly outside the unit circle and the innovations $\varepsilon_{i,t}$, $i = 1, 2$ are independent multivariate white noise. The expectations at time t about a value of a variable x_t at time $t + s$ is denoted $\mathcal{E}_t(x_{t+s})$. The $(n_2 \times n_1)$ matrix Λ' has full column rank n_1 . Of course, (1) represents a cointegrated system and n_1 is the cointegration rank.

In the following we shall call the model (1) the economic model. It shows a dependence of y_t on z_t but not vice versa. The objective of the econometric exercise is to unveil this relationship.

To do so, the model has to be transformed in an estimable version. Therefore, the expectation in the equation for y_t is replaced by

$$\mathcal{E}_t(z_{t+s}) = z_{t+s} + \nu_t$$

$$= \Delta z_{t+s} + \Delta z_{t+s-1} + \cdots + \Delta z_{t+1} + z_t + \iota_t$$

to yield

$$y_t = \Lambda (\Delta z_{t+s} + \Delta z_{t+s-1} + \cdots + \Delta z_{t+1} + z_t + \iota_t) + \eta_{1,t},$$

and ι_t is the expectation error. Furthermore, the first difference of y_t given by

$$\begin{aligned} \Delta y_t &= \Lambda \Delta z_{t+s} + \Lambda (\iota_t - \iota_{t-1}) + \eta_{1,t} - \eta_{1,t-1} \\ &= \Lambda \eta_{2,t+s} + \Lambda (\iota_t - \iota_{t-1}) + \eta_{1,t} - \eta_{1,t-1}. \end{aligned} \quad (2)$$

Now, define the $(n \times n)$ matrices Γ_i with $n = n_1 + n_2$ and $s \leq i \leq p$, the $(n \times 1)$ variables vector $Y_t = (y_t', z_t')$, and the $(n \times 1)$ vector of innovations $\eta_t = (\eta_{1,t}', \eta_{2,t}', \eta_{2,t}')'$.¹ The model can then be written as

$$Y_t = \sum_{i=-1}^{-s} \Gamma_i \Delta Y_{t-i} + B Y_{t-1} + \eta_t \quad (3)$$

All matrices Γ_i , $i < 0$ are of the structure $\Gamma_i = \begin{bmatrix} 0_{n_1} & \Lambda \\ 0_{n_2 \times n_1} & 0_{n_2} \end{bmatrix}$ and $B = \begin{bmatrix} 0_{n_1 \times n_2} & I_{n_1} \\ 0_{n_2 \times n_1} & I_{n_2} \end{bmatrix}$. The representation in (3) can easily be transformed into the usual error correction form by defining $\Gamma_0 = I_n$ and $\Pi = \alpha \beta' = I_n - B$.

$$\Gamma_0 \Delta Y_t = \Pi Y_{t-1} + \sum_{i=-1}^{-s} \Gamma_i \Delta Y_{t-i} + \eta_t \quad (4)$$

The $(n \times n_1)$ matrices α and β' have full column rank. The error correction term, $\beta' Y_{t-1}$, can also be expressed as

$$\beta' Y_{t-1} = \beta' \begin{bmatrix} \Lambda \left(\sum_{i=1}^{s-1} \Delta z_{t+i} + \Delta z_t + \iota_{t-1} \right) + \Lambda z_{t-1} + \eta_{1,t-1} \\ z_{t-1} \end{bmatrix}.$$

Thus, the matrix β' is easily identified because there exists exactly one combination which ensures stationarity of $\beta' Y_{t-1}$. Define the sub-matrices β_1 , α_1 and β_2 , α_2 of dimension $n_1 \times n_1$ and $n_2 \times n_1$, respectively. Then, stationarity is obtained for $\beta_1' = I_{n_1}$ and $\beta_2' = -\Lambda$. As a consequence, $\beta' Y_{t-1} = (\beta_1', \beta_2') Y_{t-1}$ can be written as

$$\beta' Y_{t-1} = \Lambda \left(\sum_{i=1}^{s-1} \Delta z_{t+i} + \Delta z_t + \iota_{t-1} \right) + \eta_{1,t-1}$$

¹The model is written in reduced form at once.

Returning to (4), we can now write

$$\begin{aligned}\Delta Y_t &= \alpha \left[\Lambda \left(\sum_{i=1}^{s-1} \Delta z_{t+i} + \Delta z_t + \iota_{t-1} \right) + \eta_{1,t-1} \right] + \sum_{i=-1}^{-s} \Gamma_i \Delta Y_{t-i} + \eta_t \\ &= \alpha [\Lambda (\Delta z_t + \iota_{t-1}) + \eta_{1,t-1}] + \sum_{i=-1}^{-s} \Gamma_i^* \Delta Y_{t-i} + \eta_t\end{aligned}\quad (5)$$

where $\Gamma_i^* := \begin{cases} \Gamma_i, & \text{for } i = -s \\ (0_{n \times n_1}, \alpha \Lambda) + \Gamma_i, & \text{for } -s < i < 0 \end{cases}$. One further transformation is needed before the properties of standard estimates for α are available. Observe that $\Delta y_t = \Lambda \Delta z_{t+s} + \Lambda(\iota_t - \iota_{t-1})$, $\Delta z_t = \eta_{2,t}$, define the Variance-Covariance matrices $\Sigma_{\tilde{\eta}_2} = \eta_{2,t+s} \eta'_{2,t}$, $\Sigma_{\eta_i} = \eta_{i,t} \eta'_{i,t}$, $i = 1, 2$, and $\Sigma_\iota = \iota_t \iota'_t$. It is also assumed that all elements of η_t are independent of all elements of ι_t at all leads and lags.

The OLS estimates of α are then obtained as²

$$\begin{aligned}X &= [\Lambda \eta_{2,t} + \Lambda \iota_{t-1} + \eta_{1,t-1}] \\ Y &= \begin{bmatrix} \Delta y_t \\ \Delta z_t \end{bmatrix} \\ &= \begin{bmatrix} \Lambda \eta_{2,t+s} + \Lambda(\iota_t - \iota_{t-1}) + \eta_{1,t} - \eta_{1,t-1} \\ \eta_{2,t} \end{bmatrix} \\ \hat{\alpha} &= Y X' (X X')^{-1} \\ &= \begin{bmatrix} \Lambda \Sigma_{\tilde{\eta}_2} \Lambda' - \Lambda \Sigma_\iota \Lambda' - \Sigma_{\eta_1} \\ \Sigma_{\eta_2} \Lambda' \end{bmatrix} [\Lambda \Sigma_{\eta_2} \Lambda' + \Lambda \Sigma_\iota \Lambda' + \Sigma_{\eta_1}]^{-1}\end{aligned}\quad (6)$$

The most interesting implication of (6) certainly is the fact that in contrast to (1), one will now identify a dependence of z_t on y_t . This is because the lower n_2 rows of $\hat{\alpha}$ will generally be nonzero. At the same time, however, it cannot be ruled out that the first n_1 rows turn out to be zero or close to zero. In fact, if $\Lambda \Sigma_{\tilde{\eta}_2} \Lambda'$ and $(\Lambda \Sigma_\iota \Lambda' + \Sigma_{\eta_1})$ are of the same order of magnitude, this appears rather likely. If so, one would conclude, that y_t is weakly exogenous with respect to β and hence, the long-run forcing variable.

Needless to say, the true relationship is just the opposite. It can also be obtained, though at a very high price. The last n_2 rows of $\hat{\alpha}$ are going to be zero for large Σ_{η_1} and $\Lambda \Sigma_\iota \Lambda'$. At the same time a significant value for the first n_1 rows would result, its limiting value being the identity matrix. However, this constellation occurs only if the

²The error terms are assumed to follow stationary processes. This could be accounted for by including additional lagged left hand side variables as regressors. For simplicity, we disregard all explanatory variables except for the expression in parentheses in (5). This simplification does not affect the principal outcomes.

variance of the forecast error is very large, and if the relative importance of the innovations in y_t outweighs those in z_t by far. A large forecast error variance implies unreliable forecasts while a large variance in $\varepsilon_{1,t}$ compared to $\varepsilon_{2,t}$ means that the information about z_t would not very useful anyway for explaining y_t . Taken together, in economic terms this is equivalent to having a poor model.

3.1 A Paradox

In any case, the correct results according to the economic model would only be obtained for the special case of very large expectation errors and large innovations in the DGP for y_t in relation to $\varepsilon_{2,t}$. However, if this holds, then expectations are based on poor grounds, the variance of $\varepsilon_{2,t}$ is relatively small, and the economic model would be less relevant. Likewise, if expectations are very good, and $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are equally important, the model would be rich in content. The standard econometric exercise however, would easily lead to the opposite conclusion. It would in particular do so if the true data generating model as of (1) could be recovered.

This, of course, is a paradox: The poorer in content the economic model the more likely it will be revealed by the econometric analysis.

4 Empirical Examples

The Fisher relation (Fisher, 1930), the expectation hypothesis about the term structure of interest rates and the uncovered interest rate parity condition are popular examples of economic hypotheses with rational expectations, among others. In all these cases the objective is to explain the levels of (typically) the long-term interest rate.

4.1 The Fisher Relation

The starting point is the notion that rational individuals focus on the real return on investments, that is, after accounting for inflation. Therefore, the (long-term) nominal interest rate (i_t^l) needed to convince people to lend money is the sum of the desired real

return (real interest rate, r_t) plus the expected inflation ($E_t(\pi_{t+s})$):

$$i_t^l = r_t + E_t(\pi_{t+s}) \tag{7}$$

The difficulty that arises, is, of course, that expected inflation is not observable. That's why it is usually approximated by the current inflation rate, which would be the best linear forecast if inflation followed a random walk.

Therefore, the Fisher hypothesis can be cast in the framework of section 2 with i_t^l being the endogenous variable and π_t playing the role of z_t .

4.2 The Expectation hypothesis

The main hypothesis of the EHTS is that short-term bond yields (i_t^s) are lower than long-term bond yields because of lower risk of default or inflation because it appears easier to calculate the nearer than the more distant future. However, for our purposes another aspect is more relevant.

It is assumed that investors choose a portfolio that balances returns and risk of the individual assets such that all components yield the same real returns from the perspective of the investor. Without that balance, the investor could earn more by choosing a different basket. As one consequence, the investor should be indifferent between buying short-term bonds consecutively or buying a long-term bond at once. Therefore, if the short-term bonds rate unexpectedly increases, the balance will be distorted to the effect that long-term bonds are sold in order to acquire more short-term bond yields. This induces a rise in the long-term bonds and re-establishes equilibrium.

Naturally, the reaction of the long-term bond yields will depend on the perception of the future path of short-term bond yields. Hence, the framework of section 2 can again be applied with i_t^l being y_t and i_t^s as the independent variable.

4.3 The Uncovered Interest Rate Hypothesis

Taking again the perspective of an investor, the portfolio choice will also be made considering foreign bonds. If the foreign bond rates are determined exogenously (e.g., the U.S.

bonds with respect to the rest of the world), then the choice to buy or sell domestic bonds will depend on what is expected about the future level of the foreign alternative. Again, the role of expectations becomes central and the setting of section 2 applicable.

4.4 Estimation Results

To illustrate the theoretical findings, the following exercise presents results for the USA, Japan, Germany and Switzerland. The standard setup is a reduced rank regression as it has been suggested by Johansen (1988).³ In all cases but one, the choice of variables makes sure that the cointegration rank, as implied by the theory, is exactly one. The test statistic for the cointegration rank test is also provided. The general model for estimation is (4) without leads but with lagged dependent variables as regressors. The lag order p is chosen according to selection criteria. If the suggested lag order is not sufficient to account for residual autocorrelation, further lags are added. Most of the time, this procedure solves the problem. In one instance (example 1 below), the residual autocorrelation cannot be coped with in the multivariate setting. Therefore, single equation methods are also used. With these, a more flexible lag structure can be implemented that also solves the problem of autocorrelation.

Another difficulty with the data is heteroscedasticity and non-normality of residuals which can often be observed when modelling interest rates. Here, no definite answer can be given. It has not always been possible to eliminate *ARCH* effects and excess kurtosis. All results are presented in table 1 except for the residual properties which are, of course, available on request.

In Table 1, the information regarding the model setup is in columns 1-7. In all cases except for the Juselius and MacDonald (2003) international parity relationship, the cointegration rank test supports the hypothetical number of cointegration relations. In the column labelled as " β coefficients", it is checked whether the hypothetical cointegration coefficients can be imposed. Again, this is the case in almost all instances at the 10 percent level of significance. Where this is the case (examples 1-3,5-7), the following test for the

³In this standard modelling approach, no leading variables are included as regressors.

Table 1: Empirical Evidence

No. Variables ^a	Country	Sample	Cointegration Test		β coefficients		α coefficients		Method ^d	
			H_0^0	LR ^b q	H_0^1	LR stat.[prob.]	H_0^2	LR stat. ^c [prob.]		
Fisher Relation										
1	CPI infl. (β_1)	USA	89 : 12 – 03 : 08	$rk = 0$	22.76 [.02]	5 $\beta_1 = -\beta_2 = 1$	$\chi^2(1) = 3.15$ [.07]	$\alpha_1 = 0$	$\chi^2(1) = \mathbf{13.06}$ [.00]	MV
	Bond y. (β_2)	USA	$T = 161$	$rk = 1$	3.28 [.54]			$\alpha_2 = 0$	$\chi^2(1) = 2.70$ [.10]	
				$rk = 0$	74.67 [.00]	2 $\beta_1 = -\beta_2 = 1$	$\chi^2(1) = 13.98$ [.00]	$\alpha_1 = 0$	$\chi^2(1) = \mathbf{64.38}$ [.00]	MV
				$rk = 1$	3.45 [.51]			$\alpha_2 = 0$	$\chi^2(1) = 2.70$ [.10]	
								$\alpha_1 = 0$	$\chi^2(1) = \mathbf{9.5}$ [.00]	SEQ
								$\alpha_2 = 0$	$\chi^2(1) = 1.94$ [.16]	
2	CPI infl. (β_1)	CH	90 : 02 – 03 : 07	$rk = 0$	20.50 [.05]	4 $\beta_1 = -\beta_2 = 1$	$\chi^2(1) = 1.78$ [.18]	$H_0^1 \wedge \alpha_1 = 0$	$\chi^2(2) = \mathbf{15.86}$ [.00]	MV
	Bond y. (β_2)	CH	$T = 158$	$rk = 1$	2.61 [.66]			$H_0^1 \wedge \alpha_2 = 0$	$\chi^2(2) = 3.34$ [.18]	
Uncovered Interest Rate Parity Condition										
3	LIBOR (β_1)	GER	92 : 01 – 03 : 07	$rk = 0$	744.41 [.00]	2 $\beta_1 = -\beta_2 = 1$	$\chi^2(1) = 2.73$ [.10]	$H_0^1 \wedge \alpha_1 = 0$	$\chi^2(2) = \mathbf{13.26}$ [.00]	MV
	LIBOR (β_2)	CH	$T = 139$	$rk = 1$	6.23 [.18]			$H_0^1 \wedge \alpha_2 = 0$	$\chi^2(2) = 2.80$ [.25]	
Juselius and MacDonald's (2003) International Parity Relation										
4	Bond y. (β_1)	USA	83 : 01 – 03 : 07	$rk = 0$	46.50 [.20]	2 $-\beta_1 = -\beta_2$	$\chi^2(3) = 6.66$ [.08]	$\alpha_1 = 0$	$\chi^2(1) = 1.73$ [.19]	MV
	Bond y. (β_2)	JAP	$T = 244$	$rk = 1$	26.86 [.30]	$= \beta_3 = \beta_4$		$\alpha_2 = 0$	$\chi^2(1) = 5.50$ [.02]	
						$= -1$		$H_0^1 \wedge \alpha_2 = 0$	$\chi^2(6) = 11.74$ [.07]	
	LIBOR (β_3)	USA						$\alpha_3 = 0$	$\chi^2(1) = .08$ [.77]	
	LIBOR (β_4)	JAP						$\alpha_4 = 0$	$\chi^2(1) = \mathbf{.86}$ [.35]	
								$H_0^0 \wedge \alpha_4 = 0$	$\chi^2(6) = 18.45$ [.00]	

^a CPI denotes consumer price index, Bond y. is short for government bond yield, LIBOR is the interest rate for short term credits (3-months) at the London interbank market, and Money is the interest rate on one-month interbank credits. More details can be found in Table 3.

^c Likelihood ratio test for the cointegration rank test (Johansen, 1995, Tab. 15.2). ^d One degree of freedom if no restriction on β -vector imposed, 2 degrees of freedom if H_0^1 is also imposed (no. 2,3,5-6), 6 degrees of freedom (no. 4): H_0^1 and $\alpha_1 = \alpha_3 = 0$ additionally imposed

^e MV abbreviates multivariate model, SEQ single equation model.

Table 1: Empirical Evidence (cont.)

No. Variables ^a	Country	Sample	Cointegration Test		β coefficients		α coefficients		Method
			H_0^0	LR stat.[prob.] q	H_0^1	LR stat.[prob.]	H_0^2	LR stat.[prob.]	
Expectation Hypothesis about the Terms Structure									
5	LIBOR (β_1)	USA	83 : 01 – 03 : 07	$rk = 0$	21.56 [.00]	$2 \beta_1 = -\beta_2 = 1\chi^2(1) = 1.63 [.20]$	$H_0^1 \wedge \alpha_1 = 0$	$\chi^2(2) = \mathbf{8.75} [.01]$	MV
	Bond y. (β_2)	USA	$T = 244$	$rk = 1$	2.27 [.72]		$H_0^1 \wedge \alpha_2 = 0$	$\chi^2(2) = 1.79 [.40]$	
6	Call Mon. (β_1)	CH	89 : 03 – 03 : 07	$rk = 0$	40.13 [.00]	$2 \beta_1 = -\beta_2 = 1\chi^2(1) = .29 [.59]$	$H_0^1 \wedge \alpha_1 = 0$	$\chi^2(2) = \mathbf{21.90} [.00]$	MV
	LIBOR (β_2)	CH	$T = 173$	$rk = 1$	1.72 [.83]		$H_0^1 \wedge \alpha_2 = 0$	$\chi^2(2) = \mathbf{8.01} [.02]$	
7	LIBOR (β_1)	GER	73 : 08 – 03 : 07	$rk = 0$	47.23 [.00]	$4 \beta_1 = -\beta_2 = 1\chi^2(1) = .38 [.54]$	$H_0^1 \wedge \alpha_1 = 0$	$\chi^2(2) = \mathbf{26.36} [.00]$	MV
	Bond y. (β_2)	GER	$T = 360$	$rk = 1$	7.29 [.11]		$H_0^1 \wedge \alpha_2 = 0$	$\chi^2(2) = .96 [.62]$	

^a CPI denotes consumer price index, Bond y. is short for government bond yield, LIBOR is the interest rate for short term credits (3-months) at the London interbank market, and Money is the interest rate on one-month interbank credits. The data is described in detail in Table 3.

restrictions on the adjustment coefficients (α) is performed including the restrictions on the β coefficients. In addition, in example 4, the test result is reported for the significance of the adjustment coefficients in the equations for the Japanese interest rates when weak exogeneity is imposed on the U.S. interest rates.

Each of the first lines in the examples 1-3, 5-7 should, according to the outlines above, feature a rejection of the null hypothesis that the respective adjustment coefficient is zero. It should be born in mind that this variable is always supposed to represent the independent variable for the long-run relationship in economic terms. As expected, the estimation results seem to produce the opposite conclusion, namely that the presumed independent variable significantly reacts to deviations from equilibrium while the theoretically dependent variable (2nd line) does not.

For example, in case 1, where the Fisher parity is tested for U.S. data, the hypothesis is that inflation expectations rule the nominal interest rates. In the econometric model, the expectations are replaced by current inflation which is viewed as a predictor of unobservable inflation expectations. Obviously and independent of the specific model, the null hypothesis that inflation does not adjust to deviations from the long-run equilibrium is strongly rejected. At the same time, however, it is found that interest rates do not adjust significantly. While the latter statement has found to be true at the 10 percent level only, the situation is much clearer in the Swiss case (example 2). Here, the hypothesis that interest rates do not adjust cannot be rejected at the 18 percent level.⁴

Example 3 is concerned with the interest rate parity between Germany and Switzerland. From the Swiss perspective, Germany is a large economy whose bond rates appear exogenous with respect to the Swiss rates. Therefore, the Swiss National Bank would be forced to keep an eye on the German rate if too strong a revaluation of the Swiss Franc versus the Euro (or Deutschmark) is considered not desired. The way to ensure this most efficiently is, of course, to anticipate the future movements of the German rate. Consequently, even though the German rate is the long-run driving force with respect to the

⁴The weak exogeneity property is not always invariant to the inclusion of leads of the l.h.s. variable.

The conclusions are, however, not affected.

Swiss rate, according to the model, the adjustment coefficients should seemingly imply the opposite. This is actually the case. The hypothesis that German rates do not adjust to Swiss rates is very strongly rejected while the hypothesis that Swiss rates do not adjust to German rates passes the test.

The fourth exercise is concerned with the problem raised by Juselius and MacDonald (2003). They find a relationship which they call "international parity relationship" between Japanese and U.S. short and long-term interest rates. They are surprised by the fact that the Japanese interest rates appear weakly exogenous while the U.S. rates show significant adjustment. At a first glance, this seems to fit in the framework of section 2, where similar arguments as in the Germany-Switzerland case could be applied. The table 1 reports an attempt to reproduce the respective results.

In contrast to Juselius and MacDonald (2003) a shorter sample period is chosen in order to avoid some of the modelling hassle. All dummy variables they used and which are suitable for the shorter sample are also included. While the cointegration test does not suggest the existence of a stationary relationship between the variables under consideration, at the ten percent level of significance it is found that the restrictions used by Juselius and MacDonald (2003) cannot be rejected after the rank one is imposed on the system. Likewise in contrast to the estimates of Juselius and MacDonald (2003), the weak exogeneity of the Japanese interest rates cannot be confirmed.⁵ This, however, may well be owed to the smaller set of endogenous variables in the current setup. If the model was correct and the weakly exogenous variables were the true dependent variables with respect to the long-run, then one would have to conclude that the Japanese rates are driving the U.S. rates. Thus, the same surprise would emerge, though due to the opposite argument. A caveat against this line of reasoning is, of course, given by the fact that in the reduced sample a cointegration relationship has not found support.

The remaining examples provide evidence for the EHTS for some sets of interest rates in the USA, Germany and Switzerland. As before, the interest rates of shorter maturity

⁵In table 5 of Juselius and MacDonald (2003), both long-term rates are exogenous with respect to the identified international parity relation but not the short-term rates. Weak exogeneity of the U.S. rates is rejected in the model without identified long-run relationships.

are supposed to be driving those with longer maturity. The standard analysis leads to rejecting the hypothesis that the rate with the longer maturity is weakly exogenous. On the other hand, the hypothesis that the shorter rate reacts to the longer rate cannot be rejected.

5 Discussion

It has been realised before that sequences of events do not automatically imply a causal ordering. The model of section 2 can be viewed as yet another example thereof. However, the notion of the paradox goes beyond that conventional wisdom. The following discussion will touch upon two further aspects.

5.1 Is there a cure?

Having described and illustrated the interpretation problem, a natural question is of course whether there is a cure for it. The most desirable remedy would be an estimation setup where the economic model itself can be tested directly. In the standard situation, an indirect approach is used because the key element, the expectation about z_t , is not observable. Replacing it by an unbiased estimator helps to circumvent the measurement problem but incurs the interpretation paradox. This point can be illustrated by the following additional regression, where the UIP between German and Swiss interest rates is used again. This time however, the unobservable expected German rate is approximated by a very good predictor, which is its own future realisation. Table 2 has the details.

Obviously, the three months ahead realisation of the German 3-months interest rate is a good guess about the German 3-months interest rates three months ahead. A shock to this expectation (now) significantly affects the Swiss interest rate while no effect can be measured in the opposite direction. Thus, the paradox is solved "econometrically". In regression 8 of table 2, the economic and econometric notion of dependence and independence finally coincide.⁶ A cross-check is provided by example 9, where instead of the

⁶Note that theoretically, lagging one variable of the system should not alter the cointegration test

Table 2: Empirical Evidence - 2nd Step

No. Variables ^a	Country	Sample	Cointegration Test		β coefficients		α coefficients		Method ^d	
			H_0^0	LR ^b q	H_0^1	LR stat.[prob.]	H_0^2	LR stat. ^c [prob.]		
Uncovered Interest Rate Parity Condition										
8	LIBOR _{t+3} (β_1)	GER	92 : 01 – 03 : 04	$rk = 0$	25.24 [.00]	3 $\beta_1 = -\beta_2 = 1$	$\chi^2(1) = .63$ [.43]	$H_0^1 \wedge \alpha_1 = 0$	$\chi^2(2) = .72$ [.70]	MV
	LIBOR (β_2)	CH	$T = 136$	$rk = 1$	11.42 [.02]			$H_0^1 \wedge \alpha_2 = 0$	$\chi^2(2) = \mathbf{9.39}$ [.01]	
9	LIBOR (β_1)	GER	92 : 01 – 03 : 04	$rk = 0$	60.727 [.00]	3 $\beta_1 = -\beta_2 = 1$	$\chi^2(1) = 2.56$ [.11]	$H_0^1 \wedge \alpha_1 = 0$	$\chi^2(2) = \mathbf{39.97}$ [.00]	MV
	LIBOR _{t+3} (β_2)	CH	$T = 136$	$rk = 1$	15.0 [.00]			$H_0^1 \wedge \alpha_2 = 0$	$\chi^2(2) = 2.58$ [.27]	
11	Money _{t+3} (β_1)	CH	89 : 05 – 03 : 04	$rk = 0$	58.089 [.00]	4 $\beta_1 = -\beta_2 = 1$	$\chi^2(1) = 3.82$ [.05]	$\alpha_1 = 0$	$\chi^2(1) = 1.70$ [.19]	MV
	LIBOR (β_2)	CH	$T = 168$	$rk = 1$	2.64 [.66]			$\alpha_2 = 0$	$\chi^2(1) = \mathbf{49.58}$ [.00]	

^a CPI denotes consumer price index, Bond y. is short for government bond yield, LIBOR is the interest rate for short term credits (3-months) at the London interbank market, and Money is the interest rate on one-month interbank credits. More details can be found in Table 3.

^c Likelihood ratio test for the cointegration rank test (Johansen, 1995, Tab. 15.2). ^d One degree of freedom if no restriction on β -vector imposed, 2 degrees of freedom if H_0^1 is also imposed (no. 2,3,5-6), 6 degrees of freedom (no. 4): H_0^1 and $\alpha_1 = \alpha_3 = 0$ additionally imposed

^e MV abbreviates multivariate model, SEQ single equation model.

German rate, the Swiss rate is leading three periods, the outcome however, is qualitatively the same as that of model 3.

Unfortunately, there are not always good predictors at hand. For example, when testing the Fisher parity for long-term bonds, it is not clear how the future inflation rates should be weighted in order to produce a good proxy for the inflation in the remaining time to maturity. Similar arguments hold for many other relationships.

5.2 Relevance

The literature has so far paid little attention to the seemingly surprising lack of weak exogeneity of the supposed long-run driving variables. There are, however, also good reasons for that. For example, it is of interest in itself if the spread between nominal interest rates and inflation is stationary or not, because it helps to learn about the Fisher hypothesis. The same holds for the other concepts briefly discussed. This inference can be made without reference to the adjustment characteristics as long as there is adjustment towards equilibrium at all.

5.2.1 Forecasting

However, there are also at least two situations where the difference matters. The first is forecasting. Figure 1 illustrates the effect. Systems 3 and 8 have been subjected to unit shocks. This means that one equation is shocked once while no shock is allowed in the other equation at the same time.⁷ The corresponding reactions of the left hand side variables are then graphed.

Obviously, the responses could hardly be more different. In model 3 the reaction of the Swiss rate to a shock in the German rate (lower left panel) dies out pretty quickly while in model 8 it remains above two for the whole simulation period.⁸ Likewise striking results. In the empirical example it does so. However, this alterations does not matter because in a stationary system - as it is implied by the tests in models number 8 and 9 - the framework of section 2 in principal still holds without the additional complication of non-stationarities.

⁷Further analyses with e.g. orthogonalized impulse-responses do not alter the results quantitatively.

⁸Note that no statement about significance with respect to the distance from zero is made. What matters most, however, is the (principal) difference between the responses in the two models.

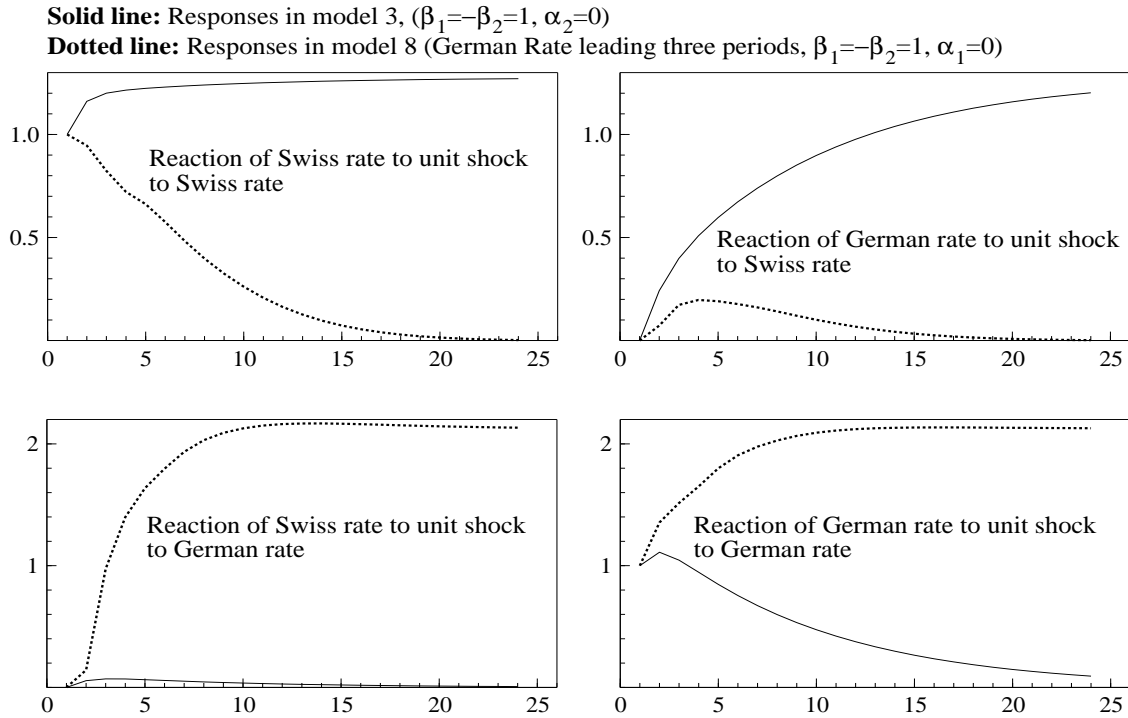


Figure 1: Impulse-responses in systems 3 and 8.

is the reaction of the German rate in model 3 when the Swiss rate is shocked (upper right panel). It appears that the German rate responds rather strongly, while this cannot be observed in model 8.

Therefore, if one bases forecasts for the Swiss rate, for example, on model 3, not only would one obtain results which are at odds with conventional wisdom about the relationship between Germany and Switzerland, but one would also be diverted from the "true" causal links. Considering model 8 instead, solves the puzzle. The following argument shows that the choice between model 3 and 8 may not need to be purely arbitrary.

5.2.2 A Two-Stage-Procedure

A second situation where it may pay to account for the paradox is to test for the existence of the paradox itself. To see this, consider again the model of section 2. If it was possible to replace the approximation of the expected value by the expectation itself, then the standard situation as of, e.g., Engle and Granger (1987) arises. In terms of the framework

of section 2, this results in the following estimates for α_1 and α_2 :

$$\hat{\alpha}_1 = (\Lambda \Sigma_{\tilde{\eta}_1} \Lambda') (\Lambda \Sigma_{\tilde{\eta}_1} \Lambda')^{-1} \quad (8)$$

$$\hat{\alpha}_2 = (0_{n_2 \times n_2}) (\Lambda \Sigma_{\tilde{\eta}_1} \Lambda')^{-1} \quad (9)$$

The crucial point is that now $\hat{\alpha}_2$ will be zero because of the fact that the term in the nominator vanishes in accordingly specified models. On the other hand, $\mathcal{E}(\hat{\alpha}_1)$ yields exactly the true value of α_1 .⁹ Thus, the economically sensible result is obtained which implies that z_t drives y_t in the long-run but not vice versa. Therefore, a two-step procedure can be proposed. First, the standard cointegration analysis is performed and the weak exogeneity properties are determined (see models 3, 6). Then, the set of weakly exogenous variables, z_t , is replaced by its best possible s -step ahead forecast (which, e.g., could be z_{t+s}) and the analysis repeated (see models 8, 11). If the results are identical to the ones obtained in the first step, one would be re-assured, that the underlying structural dependence, is as it appears to be from the face values of the estimates. If, however, some variables are now found to belong to the set y_t which in step 1 have been found belonging to z_t , then the true relationship is likely to be of the type sketched in section 2.

Unfortunately, it is not always clear what the best possible forecast is. In the Fisher relationship, expected inflation is certainly not the inflation rate of a specific month in the future, rather than some "overall" future price change. That difficulty of course limits the potential for obtaining useful test results. The estimation result of models 3, 6 versus 8, 11 may represent examples where the two-step procedure proved useful, however.

6 Conclusion

In economic models where expectations about one variable rule the behaviour of another one the standard econometric approach is not very likely to reveal the true causal links if the expectations cannot be directly observed. In this paper we have seen that this result also holds for cointegrated relationships where the direction of adjustment towards the

⁹Of course, the value of α_1 is strictly speaking a function of $A_1(L)$ and $A_2(L)$. For the purpose of demonstration it is relevant to note that it will not be zero.

equilibrium is used to identify the dependent and the independent variables. Moreover, a paradox may arise in which the true links are more likely to be recovered if the underlying economic model is in fact built on poor grounds. Therefore, when it comes to interpreting the adjustment coefficients, one has to be particularly careful.

Various data examples using popular economic hypotheses have illustrated these considerations. The paradox is especially relevant for forecasting and policy simulation. Under some circumstances, however, a simple cure for the paradox exists which also has the potential for testing for the true causal relations.

References

- Engle, R. F. and Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation and Testing, *Econometrica* **55**(2): 251 – 276.
- Ericsson, N. R., Hendry, D. F. and Mizon, G. E. (1998). Exogeneity, Cointegration, and Economic Policy Analysis, *Journal of Business and Economic Statistics* **16**(4): 370 – 387.
- Ericsson, N. R. (1992). Cointegration, Exogeneity, and Policy Analysis: An Overview, *Journal of Policy Modeling* **14**: 251 – 280.
- Ericsson, N. (1998). Empirical Modelling of Money Demand, *Empirical Economics* **23**(3): 483 – 506.
- Fisher, I. (1930). *The Theory of Interest*, Macmillan, New York.
- Hubrich, K. (2001). *Cointegration Analysis in a German Monetary System*, Physica Verlag, Heidelberg.
- Johansen, S. and Juselius, K. (1990). Maximum Likelihood Estimation and Inference on Cointegration - With Applications to the Demand for Money, *Oxford Bulletin of Economics and Statistics* **52**: 169–210.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics* **12**: 231 – 254.

Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, 1st edn, Oxford University Press, Oxford.

Juselius, K. and MacDonald, R. (2003). International Parity Relationships Between the USA and Japan, *Japan and the World Economy* **482**: 1 – 18.

A Data Sources

Table 3: Data Descriptions and Data Sources

Model (Tab. 1)	item / description	code	source
1	CPI infl.: 1200-fold log of 1st difference of Consumer price index, all items less food and energy Base Period: 1982-84=100, seasonally adjusted with X12Arima Bond y.: Rate of interest in money and capital markets, Federal Government securities, Constant maturity Ten-years	CUUR0000SA0L1E	USA, bureau of labor statistics (BLS) Federal Reserve System (FED)
2	CPI infl.: Switzerland, 1200-fold log of 1st difference of Consumer price index, all items Base Period: May 1993=100, seasonally adjusted Bond y.: Switzerland, Rate of interest in Federal Government securities, Constant maturity Ten-years	TS11515102	Switzerland, Federal Bureau of Statistics Swiss National Bank (SNB), Monthly Bulletin (MB) 08/2003, Table E3
3	LIBOR: Germany, Money Market Rate, 3 months LIBOR: Switzerland, Money Market Rate, 3 months	SU0107	Bundesbank, MB 08/2003 SNB, MB 08/2003
4	Bond y.: USA see Model 1 Bond y.: Japan, Government Bond Yield LIBOR: USA Eurodollar deposits, Primary market, three-month maturity LIBOR: Japan, 3-MONTH LIBOR: Offer London	M.15861...ZF... M.15860EA.ZF...	International Monetary Fund (IMF) FED IMF
5	Bond y.: USA see Model 1 LIBOR: USA see Model 4		
6	Call Mon.: Switzerland, money market rate (tomorrow next) LIBOR: Switzerland see Model 3		SNB, MB 08/2003, Table E1
7	LIBOR: Germany, see model 3 Bond y.: Germany, government securities, more than 5 and less than 6 years to maturity	WU0916	Bundesbank, MB 08/2003