

Strictly preliminary! Please do not quote!

Comments welcomed.

A Note on the Interpretation of Error Correction Coefficients

Christian Müller

Swiss Federal Institute of Technology Zürich

Swiss Institute for Business Cycle Research

CH-8092 Zürich, Switzerland

Tel.: +41.1.6324624

Fax: +41.1.6321218

Email: cmueller@kof.gess.ethz.ch

First version: December 5, 2003

This version: September 23, 2004

Abstract

This paper provides empirical evidence for standard economic equilibrium relationships and shows that the estimated disequilibrium adjustment mechanisms appear running counter to intuition. For example, the German interest rate seems to adjust to Swiss rates but not vice versa implying a driving role of the Swiss with respect to the German rates. It is argued that this phenomenon is due to the well-known difficulties of unveiling causal structures by regression. However, under certain circumstances a simple regression sequence can produce valuable information about the true causalities. The results are illustrated using U.S., Japanese, German and Swiss data.

JEL classification: C51, E37, E47, C32

Keywords: policy analysis, forecasting, rational expectations, error correction

*I do thank Erdal Atukeren and Jürgen Wolters 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, or, more generally speaking, error correction 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. 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.

The reactions to deviations from equilibrium have also been interpreted as evidence for causality or non-causality of variables within a system. Applying Hosoya's (1991) strength of causality measure Granger and Lin (1995) show that for nearly integrated systems lack of adjustment to equilibrium of one variable can be considered evidence for long run Granger causality of that variable for the other one in a bivariate system. Looking at vector autoregressive (VAR) models for policy analysis and advice Hendry and Mizon (2004) point out that potential policy instruments must not feature error correction behaviour if they are supposed to have a lasting impact. In other words, the forecast impulse response to a shock in the policy variable that is part of a cointegration relationship and that does not adjust to disequilibrium situations will not be zero in the long run. Otherwise it would be and hence it could not be considered a useful policy tool. Needless to say, it is important to know which variable may be used for policy making and which not.

Pesaran, Shin and Smith (2000) argue that knowledge about the directions of error correction yields useful information about large scale VAR modelling which would otherwise be haunted by the curse of dimensionality.

The relationship between error correction and causality has also become popular in applied research. Among others Juselius (1996), Lütkepohl and Wolters (2003) and Juselius and MacDonald (2004) use it to qualify certain variables as causal for other variables based on the characteristics of the equilibrium correction mechanism.

This paper focuses on long run causality. It develops an example of an economic model which is fairly general in nature and which can easily be estimated with standard econometric tools. Statements are made about the relationship between the assumed economic causality and the coefficients of the related econometric model. It will turn out that the reasoning crucially depends on the economic priors underlying the econometric analysis. In particular, when expectations are part of the hypothesis to be tested the coefficients of interest are likely to be estimated with a bias from which 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. In the light of the above mentioned references this can be considered rather bad news since the interpretation of the econometric results would be severely misguided.

The remainder of the paper is organised as follows. After a brief description of the problem, empirical examples in section 3 illustrate the main issues and section 4 discusses the results. An informal test is proposed to cope with the issue while a rough formal statement of the problem is sketched in the appendix.

2 An economic model and the econometric approach to it

2.1 The economic example model

We consider a variable or set of variables y which is a function of another variable or set of variables z . The economic rational might tell that variations in z will change y while the opposite does not hold. Calling the realisations of y and z through time y_t and z_t respectively, a relationship of this kind could be cast in the following form.

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 ($n_2 \geq n_1$) as

$$y_t = f(\mathcal{E}_t(z_{t+s})), \quad s > 0$$

$$\mathcal{E}_t(z_{t+s}|z_{t+s-j}) \neq \mathcal{E}_t(z_{t+s}), \quad \text{for some } j \geq s. \quad (1)$$

The expression $f(\cdot)$ denotes some function and the expectations at time t about a value of a variable x_t at time $t + s$ is denoted $\mathcal{E}_t(x_{t+s})$.

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

2.2 The econometric model

The next step is to produce an estimable version of (1) in order to check if the observations match the implications of the economic considerations. Of course, the econometric model must encompass the features of the economic model. One way to proceed could be to consider stochastic processes of the form

$$y_t = \Lambda \mathcal{E}_t(z_{t+s}) + \eta_{1,t}, \quad s > 0$$

$$z_t = \Phi z_{t-1} + \eta_{2,t}$$

$$\eta_{i,t} = A_i(L) \varepsilon_{i,t}, \quad i = 1, 2$$

$$\mathcal{E}_t(z_{t+s}) = z_{t+s} + \iota_{t+s}. \quad (2)$$

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 strictly outside the unit circle and the innovations $\varepsilon_{i,t}$, $i = 1, 2$ are independent multivariate white noise. The $(n_2 \times n_1)$ matrix Λ' has full column rank n_1 and ι_t is a stationary variable.

We should add the following comments on (2). First, the linear relationship between y_t and z_t shall be considered an approximation of the true but probably unknown underlying function $f(\cdot)$. The error processes $\eta_{i,t}$, $i = 1, 2$ are approximations of the (true) structure of the data generating process of y_t and z_t which is not captured by the assumed functional form. This interpretation can be justified by the fact that autoregressive processes

in general represent useful linear approximations to a wide range of (possibly also nonlinear) functions. Under rational expectations one could consider the additional assumption $\mathcal{E}(l_{t+s}, z_{t+s-j}) = 0, \forall s > 0, j > 0$, which is not needed though for the results below.

The question of what combination of economic model and choice of variables does best to explain y will usually be the focus of the data analysis. The answer is very often interpreted as an indicator of what economic model is the best representation of real economic relations. We use the following definitions:

1. **Validity** of the economic model:

The economic model is true if $\alpha_1 \neq 0_{n_1 \times n_1} \wedge \alpha_2 = 0_{n_2 \times n_1}$. Otherwise the model will be called wrong.

2. **Quality** of the economic model:

The economic model is poor in content if $\Sigma_\iota \Sigma_z^{-1}$ or $\Sigma_{\eta_1} \Sigma_y^{-1}$ are 'large'. Otherwise the model will be called good.

While the definition of validity of the model may be seen easy to reconcile the concept of a poor model may not so. It is basically given to make discussion below more handy. Its meaning is that we would generally like to have an economic model which ceteris paribus can attribute more of the variation in y to observable variables than to pure noise. For example, if too little of the variance in y is explained it would mean that our understanding of how the world is structured is not increased by large once we know how y is related to z . Or, in other words, even if we knew what impact z has on y , learning about z would not help us much to learn about y since the effect would be so small that it could likewise be ignored.

As an example one might consider the notion by Wang and Jones (2003). They look at the economic model suggesting that today's forward rate is a good predictor of tomorrow's spot rate. They then assert that the total variance in the spot rate is far too large compared to the variation in the forward rate to derive meaningful empirical results. Thus, the underlying efficient market hypothesis might be an economic model which correctly reflects reality, but according to our terminology it is likewise a poor model. We call the

model also poor when the forecast error is very large compared to the forecasted variable.

We now briefly check if the causal features of the economic model (1) are appropriately mapped onto the econometric model. To do so we chose the Granger causality concept and rewrite (2) in the VAR format. For simplicity we first chose $s = 1$ and generalise the result later on. Define

$$A_0 = \begin{bmatrix} I_{n_1} & -\Lambda & 0_{n_1 \times n_2} \\ 0_{n_2 \times n_1} & I_{n_2} & -\Phi \\ 0_{n_2 \times n_2} & 0_{n_2 \times n_2} & I_{n_2} \end{bmatrix} \quad A_1 = \begin{bmatrix} 0_{n_1 \times n_1} & 0_{n_1 \times n_2} & 0_{n_1 \times n_2} \\ 0_{n_2 \times n_1} & 0_{n_2 \times n_2} & 0_{n_2 \times n_2} \\ 0_{n_2 \times n_2} & 0_{n_2 \times n_2} & \Phi \end{bmatrix}$$

$$\eta_t = \begin{bmatrix} \eta_{1,t} + \Lambda t_{t+1} \\ \eta_{2,t+1} \\ \eta_{2,t} \end{bmatrix} \quad Y_t = \begin{bmatrix} y_t \\ z_{t+1} \\ z_t \end{bmatrix}$$

and write (2) as

$$A_0 Y_t = A_1 Y_{t-1} + \eta_t, \quad (3)$$

where $0_{l_1 \times l_2}$ denotes a $(l_1 \times l_2)$ matrix of zeros and I_l a l -dimensional identity matrix.

Pre-multiplying (3) by the inverse of A_0 obtains

$$\begin{bmatrix} y_t \\ z_{t+1} \\ z_t \end{bmatrix} = \begin{bmatrix} 0_{n_1 \times n_1} & 0_{n_1 \times n_2} & \Lambda \Phi^2 \\ 0_{n_2 \times n_1} & 0_{n_2 \times n_2} & \Phi^2 \\ 0_{n_2 \times n_2} & 0_{n_2 \times n_2} & \Phi \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_t \\ z_{t-1} \end{bmatrix} + A_0^{-1} \eta_t,$$

$$= A_1^* Y_{t-1} + \eta_t^* \quad (4)$$

In the terminology of Granger (1969), the triangularity of the coefficient matrix A_1^* indicates that z_{t+1} and z_t are Granger causal for y_t but not vice versa (see Lütkepohl (1993), pp. 37-39). Thus, if an analyst would know A_1^* she could correctly identify the causality features of the economic model. Hence, the Granger causality concept is a suitable approach. In the next step, (4) is transformed into the error correction form by subtracting Y_{t-1} from both sides to arrive at

$$\Delta Y_t = \begin{bmatrix} -I_{n_1} & 0_{n_1 \times n_2} & \Lambda \Phi^2 \\ 0_{n_2 \times n_1} & -I_{n_2} & \Phi^2 \\ 0_{n_2 \times n_2} & 0_{n_2 \times n_2} & -(I_{n_2} - \Phi) \end{bmatrix} Y_{t-1} + \eta_t^*$$

$$= \Pi Y_{t-1} + \eta_t^*. \quad (5)$$

The matrix Π can be decomposed into two full column rank matrices α and β with

$$\alpha = \begin{bmatrix} -I_{n_1} & 0_{n_1 \times n_2} & 0_{n_1 \times n_2} \\ 0_{n_2 \times n_1} & -I_{n_2} & 0_{n_2 \times n_2} \\ 0_{n_2 \times n_2} & 0_{n_2 \times n_2} & -I_{n_2} \end{bmatrix} \quad \beta' = \begin{bmatrix} I_{n_1} & 0_{n_1 \times n_2} & -\Lambda^* \\ 0_{n_2 \times n_1} & I_{n_2} & -\Phi^2 \\ 0_{n_2 \times n_2} & 0_{n_2 \times n_2} & \Psi \end{bmatrix}$$

where $\Psi = I_{n_2} - \Phi$, $\Lambda^* = \Lambda\Phi^2$, and $\Pi = \alpha\beta'$.¹ The Granger non-causality of y_t for z_t is thus equivalent to saying that the last bottom two elements in the first column of α are zero. In other words, z_t is the long-run causal variable and y_t is caused by z_t .

In general the matrix Π is unknown and has to be estimated. The corresponding estimates will then be used to assess the congruence of the observed data with the suggested economic model. Naturally, the quality of the estimates will be essential for the quality of the judgement. In the following, it will be demonstrated that some standard econometric technique to handle models like (2) by (4) or rather, some derivative of (4) turns out inappropriate.

2.3 The estimation approach

The econometric counterpart of (1) can be and usually is set up as an cointegrated system. This is plausible for integrated variables, i.e. $\Phi = I_{n_2}$. If so, y_t and z_t are cointegrated with $\beta' = (I_{n_1} : -\Lambda)$ as the cointegration matrix its rank being n_1 . The corresponding estimation model can be written as

$$\Delta Y_t = \alpha\beta'Y_{t-1} + \sum_{i=1}^d \Gamma_i \Delta Y_{t-i} + \eta_t \quad (6)$$

with $Y_t = (y_t', z_t')$, the $(n \times n_1)$ coefficient matrices α and β , the $(n \times n)$ coefficient matrices Γ_i with $n = n_1 + n_2$ and $p \leq d$, and the $(n \times 1)$ vector of innovations $\eta_t = (\eta_{1,t}' + \eta_{2,t}'\Lambda' + \nu_{t+s}'\Lambda', \eta_{2,t}')'$.² Putting aside the estimation of the Γ_i we focus on the derivation of the estimates for α . As Granger and Lin (1995) have shown the structure of α can be used to make statements about the subset of Y_t that is driving another subset of Y_t in the long run. Letting α_1 correspond to y_t and α_2 to z_t conditions 1 and 2 represent necessary (but not sufficient) conditions for model (1) to be congruent with the actual observations.

Condition 1: $\alpha_2 = 0_{n_2 \times n_1}$

Condition 2: $\alpha_1 \neq 0_{n_1 \times n_1}$

¹Note that for $\Phi = I_{n_2}$ (3) is a cointegrated system and the last columns of α and β would vanish.

²The lagged dependent variables in (6) are meant to account for the autocorrelation structure in $\eta_{i,t}$ which is therefore ignored.

Estimating (6) should thus yield estimates for α which comply with conditions 1 and 2. It is noteworthy that the elements of α_1 should in general be negative in order to obtain a stationary system. Therefore, condition 2 is rather weak. The appendix shows that under pretty reasonable circumstances a clear statement about the estimation outcome cannot be made even if the data is generated in line with (2). In general, the estimates for α_1 and α_2 will both be biased upward. Two main factors drive the outcome.

1. **Good-forecast-bias:** If the expectation about z_{t+s} hardly deviates from the realisation (small variance of ι_{t+s}) the elements of α_2 will be more biased.
2. **No-poor-model-bias:** Instead of having a poor economic model, it might be very rich in content (small variance of $\eta_{1,t}$). Then the estimate of α_1 will be more biased *ceteris paribus*.

The above list implies that the result do not improve with the quality of the supposed economic model. If, for example, expectations approach perfection ($\iota_t \rightarrow 0_{n_2} \forall t$) the estimates do not improve but worsen. The same holds if the economic model explains most of the variation in y_t . As an extreme case, one could even obtain estimates for α where α_1 meets the condition 1 and α_2 meets condition 2.

This of course is rather bad news. Turning the argument around one can expect that the poorer the economic model the more likely the true causal links between y and z will be revealed. The principal problem therefore remains since a logical implication is that the presented and very widely used empirical approach cannot discriminate between having made a mistake when building the model or not. Section 4 proposes a convenient albeit not always feasible procedure to identify and circumvent this pitfall.

3 Empirical Examples

The Fisher relation (Fisher, 1930) 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.

3.1 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 ($\mathcal{E}_t(\pi_{t+s})$):

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

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

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 .

3.2 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.

In all these examples autoregressive processes are commonly used to establish a link between the observable values of the exogenous variable and the unobservable expectations about it.

3.3 Estimation Results

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). 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 (6). 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 (2004) international parity relationship the cointegration rank test supports the hypothetical number of cointegration relations. In the column labelled " β coefficients", it is checked whether the hypothetical cointegration coefficients can be imposed. These coefficients also imply that z_t is an unbiased estimate for $\mathcal{E}_t(z_{t+1})$. Again, this is the case in almost all instances at the 10 percent level of significance. Where this is the case (examples 1-3), the following test for the 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 causal variable significantly reacts to deviations from equilibrium while the theoretically endogenous variable (2nd line) does not.

Table 1: Empirical Evidence

No. Variables ^a	Country	Sample	Cointegration Test		β coefficients		α coefficients		Method ^d		
			H_0^c	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 (2004) 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]	
									$H_0^1 \wedge \alpha_2 = 0$	$\chi^2(6) = 11.74$ [.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

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

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 was 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.

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 desirable. The way to ensure this most efficiently is, of course, to anticipate future movements of the German rate. Consequently, even though the German rate is the long-run driving force with respect to the 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 (2004). They find an "international parity relationship" between Japanese and U.S. short and long-term interest rates. Contrary to what they expected 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 (2004) 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 (2004) cannot be rejected after the rank one is imposed on the system. Likewise in contrast to the estimates of Juselius and MacDonald (2004), the weak exogeneity of the Japanese interest rates cannot be confirmed.³ 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.

4 Discussion

4.1 Is there a cure?

Having described and illustrated the 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 yet incurs the 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

³In table 5 of Juselius and MacDonald (2004), 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.

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											
5	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]	
6	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]	
7	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_t 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.

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 5 of table 2, the economic and econometric notion of dependence and independence finally coincide.⁴ A cross-check is provided by example 6, where instead of the 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.

4.2 Relevance

The literature has so far paid not too much 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.

Another stream of literature makes use of the fact that y_t needs to be a good *predictor* for *future* z_{t+s} if (1) is the true model. Thus, regressing z_t on y_{t-s} (or, rather on $(z_{t-s} - y_{t-s})$) should yield a significant coefficient and significance would be interpreted being consistent with the economic model. This conclusion, however clouds the fact that according to (1) variations in z_{t-s} would have to affect y_t highlighting that the coefficients of such a model are very likely inefficiently estimated and more or less useless for economic

⁴Note that theoretically, lagging one variable of the system should not alter the cointegration test 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 5 and 6 - the framework of section 2 in principal still holds without the additional complication of non-stationarities.

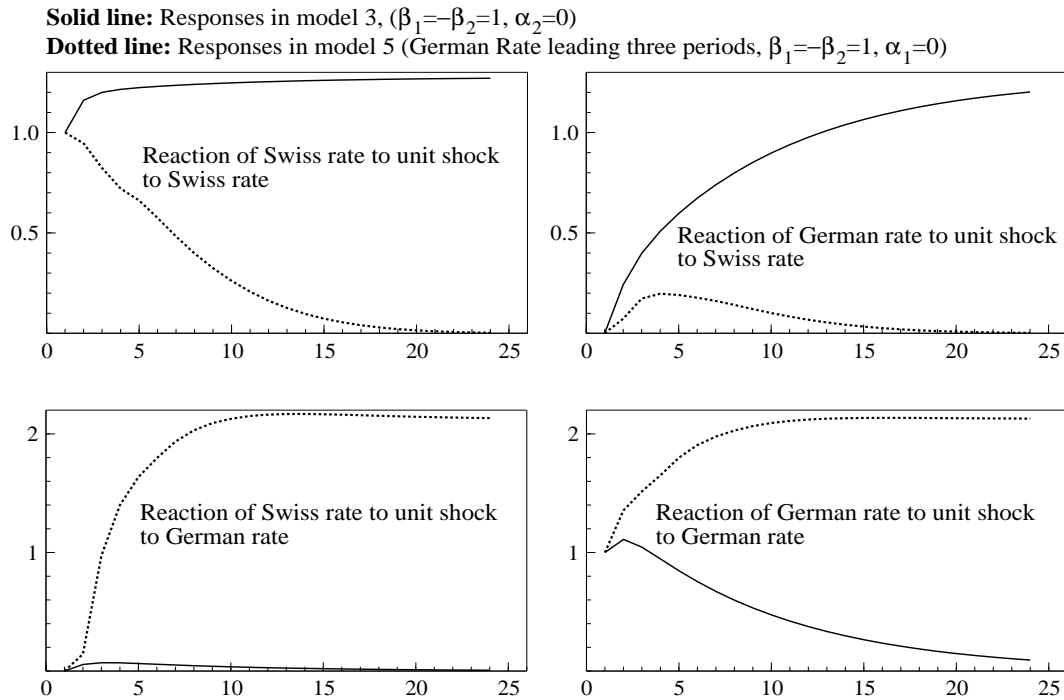


Figure 1: Impulse-responses in systems 3 and 5.

policy analysis (see Ericsson et al., 1998, p. 377). The latter is the ultimate goal of many econometric studies, however.

4.2.1 Forecasting and policy simulation

Following up on the last point, there are also at least two situations where the difference matters. The first is forecasting.⁵ Figure 1 illustrates the effect. Systems 3 and 5 have been subjected to forecast error shocks. This means that one equation is shocked once while no shock is allowed in the other equation at the same time. This resembles a hypothetical attempt of a policy maker who may regard either of the variables as a policy instrument. The corresponding reactions of the 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

⁵Policy simulation and analysis are naturally related concepts, see Hendry and Mizon (2004), for example.

while in model 5 it remains above two for the whole simulation period.⁶ Likewise striking 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 strongly, while this cannot be observed in model 5.

Therefore, if one bases forecasts for the Swiss interest rate assuming a change in the ECB interest rate, for example, on model 3, not only would one obtain results which are at odds with conventional wisdom about the relationship between the German and Swiss economies, but one would also be diverted from the "true" causal links. Considering model 5 instead, solves the puzzle. These opposite reactions are a direct result of Hendry and Mizon's (2004) analysis of instrument-target relationships. They show that weak exogeneity of the instrument variable (z in our case) with respect to Λ is a sufficient condition for a long-run zero response to a shock to the target variable (y in our case). The following argument shows that the choice between model 3 and 5 may not need to be purely arbitrary.

4.2.2 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. For example, for monetary policy analysis it would matter if money is causal for inflation or not. If one would assume that the demand for money were a function of the expected future price level instead of current prices an analysis based on impulse-responses can be severely impaired. In such cases rivalling economic models exist which pose causality in opposite directions and hence, not taking into account the possibility of biased estimates may result in a wrong conclusion. Even worse, since either direction of causality might be possible and since the $\hat{\alpha}$ would in both cases indicate equilibrium adjustment, the chance of ever noticing is very low.

In some situations, however, a not too difficult way exists to detect a bias. To see this, consider again the model of section 2. If it was possible to replace the approximation of

⁶Note 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.

the expected value by the expectation itself, then the standard situation as of, e.g., Engle and Granger (1987), Hendry and Mizon (2004), Ericsson et al. (1998) arises. In terms of the stylised situation of section 2, this results in estimates for α_1 and α_2 ($\hat{\alpha}_1$ and $\hat{\alpha}_2$, respectively) which are in accordance with the conditions 1 and 2.

The crucial point is that now $\hat{\alpha}_2$ will be zero if forecasts are nearly perfect. We now also obtain $\text{plim}(\hat{\alpha}_1) = \alpha_1$.⁷ 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). 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 5). 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, but some "overall" future price change. That difficulty of course limits the potential for obtaining useful test results. The estimation result of models 3 versus 5 may represent examples where the two-step procedure proved useful, however.

Another obstacle to detecting the bias lies with the fact that the estimates for α are biased indeed, but not necessarily to the extent that a paradoxical situation becomes obvious. Thus, in most applications it will go unnoticed if diagnostic analysis points to significant equilibrium adjustment and this is considered sufficient.

⁷The expression plim denote the probability limit. Of 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.

5 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. This paper has shown that this result also holds for cointegrated relationships where the direction of adjustment towards the equilibrium is used to identify the long run dependent and independent variables. Moreover, a paradox may arise in which the true links are more likely 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.
- Fisher, I. (1930). *The Theory of Interest*, Macmillan, New York.
- Granger, C. W. J. and Lin, J. (1995). Causality in the long run, *Econometric Theory* **11**(1): 530 – 536.

- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods, *Econometrica* **37**(3): 424 – 438.
- Hendry, D. F. and Mizon, G. E. (2004). The role of exogeneity in economic policy analysis, *Journal of Business and Economic Statistics* **forthc.**(-9): 999 – 999.
- Hosoya, Y. (1991). The decomposition and measurement of the interdependence between second-order stationary processes, *Probability Theory and Related Fields* **88**: 429 – 444.
- 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. (2004). International Parity Relationships Between the USA and Japan, *Japan and the World Economy* **16**: 17 – 34.
- Juselius, K. (1996). An Empirical Analysis of the Changing Role of the German Bundesbank after 1983, *Oxford Bulletin of Economics and Statistics* **58**: 791 – 819.
- Lütkepohl, H. and Wolters, J. (2003). Transmission of German Monetary Policy in the Pre-Euro Period, *Macroeconomic Dynamics* **7**(5): 711 – 733.
- Lütkepohl, H. (1993). *Introduction to Multiple Time Series Analysis*, 2nd edn, Springer-Verlag, Berlin.

Pesaran, H. H., Shin, Y. and Smith, R. J. (2000). Structural Analysis of Vector Error Correction Models with Exogenous I(1) Variables, *Journal of Econometrics* **97**: 293 – 343.

Wang, P. and Jones, T. (2003). The Impossibility of Meaningful Efficient Market Parameters in Testing for the Spot-Forward Relationship in Foreign Exchange Markets, *Economics Letters* **81**: 81 – 87.

A From the economic to the econometric approach

A.1 Step 1 - the standard regression

The model will first be re-stated and then the estimation of α will be discussed.

$$\Delta \begin{bmatrix} y_t \\ z_{t+1} \\ z_t \end{bmatrix} = \begin{bmatrix} -I_{n_1} & 0_{n_1 \times n_2} & 0_{n_1 \times n_2} \\ 0_{n_2 \times n_1} & -I_{n_2} & 0_{n_2 \times n_2} \\ 0_{n_2 \times n_2} & 0_{n_2 \times n_2} & -I_{n_2} \end{bmatrix} \begin{bmatrix} I_{n_1} & 0_{n_1 \times n_2} & -\Lambda^* \\ 0_{n_2 \times n_1} & I_{n_2} & -\Phi^2 \\ 0_{n_2 \times n_1} & 0_{n_2 \times n_2} & \Psi \end{bmatrix} \times \begin{bmatrix} y_{t-1} \\ z_t \\ z_{t-1} \end{bmatrix} + \eta_t^*$$

In the standard estimation approach the second line is usually disregarded and inference is only made with respect to the first and last line of α as well as the first line of β' . For simplicity it is assumed that the matrix β can be super consistently estimated (e.g. in cointegrated systems), or that the economic prior regarding Λ is so strong that it need not be estimated at all. Then, the central causality analysis is with respect to α alone.

Let's write

$$\Delta \begin{bmatrix} y_t \\ z_{t+1} \\ z_t \end{bmatrix} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{3,2} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{3,2} \\ \alpha_{3,1} & \alpha_{2,3} & \alpha_{3,3} \end{bmatrix} \begin{bmatrix} I_{n_1} & 0_{n_1 \times n_2} & -\Lambda^* \\ 0_{n_2 \times n_1} & I_{n_2} & -\Phi^2 \\ 0_{n_2 \times n_1} & 0_{n_2 \times n_2} & \Psi \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_t \\ z_{t-1} \end{bmatrix} + \eta_t^*.$$

Estimation of the first and third row will be called the first step regression. The estimates for $\alpha_{1,1}$ and $\alpha_{3,1}$ would then be used for causality interpretation. Finding them to correspond to a nonzero ($\alpha_{1,1}$) matrix and to a zero ($\alpha_{3,1}$) matrix respectively would yield the correct interpretation, namely that in the long run causality is running from z to y but not vice versa. Let

$$Y = \begin{bmatrix} \Delta y_1 & \Delta y_2 & \cdots & \Delta y_t \\ \Delta z_1 & \Delta z_2 & \cdots & \Delta z_t \end{bmatrix}, \quad X' = \begin{bmatrix} (y_0 - \Lambda^* z_0)' \\ (y_1 - \Lambda^* z_1)' \\ \vdots \\ (y_{t-1} - \Lambda^* z_{t-1})' \end{bmatrix},$$

$$\alpha = \begin{bmatrix} \alpha_{1,1} \\ \alpha_{3,1} \end{bmatrix}, \quad \bar{\eta}_t = \begin{bmatrix} \eta_{1,t} \\ \eta_{2,t} \end{bmatrix},$$

and obtain the regression model $Y = \alpha X + \bar{\eta}_t$ with ordinary least squares (OLS) estimates $\hat{\alpha} = YX'(XX')^{-1}$. Giving the set of elements \mathbb{L} which solve $\det(I_{n_2} - \Phi L) = 0_{n_2 \times n_2}$ obtains

$$\begin{aligned} \Delta z_t &= \begin{cases} \tilde{\eta}_{2,t} \equiv \eta_{2,t} \in \mathbb{L} \\ \tilde{\eta}_{2,t} \equiv -\Psi(\Phi^{t-1}z_0 + \sum_{i=0}^{t-1} \Phi^i \eta_{2,i}^{t-i}) + \eta_{2,t} \in \mathbb{L}, \end{cases} \\ \Delta y_t &= \Lambda \tilde{\eta}_{2,t+1} + \Lambda \Delta \iota_{t+1} + \eta_{1,t} - \eta_{1,t-1}, \end{aligned}$$

$$y_{t-1} - \Lambda^* z_{t-1} = \eta_{1,t-1} + \Lambda \iota_{t+1} + \eta_{2,t}^\circ$$

$$\eta_{2,t}^\circ = \Lambda \sum_{i=t+1-k}^{t+1} \Phi^{t+1-i} \eta_{2,i}, \quad k = 1 \text{ (in general: } k \text{ finite)}$$

The following definitions help to simplify the representation. Assume that the model is constant over time and that the innovations $\eta_{i,t}$, $i = 1, 2$ as well as the forecast error, ι_t has time invariant first and second moments.⁸ We then let for a covariance stationary stochastic variable $W = (w_1, w_2, \dots, w_t)$

$$\text{plim}(WW'/T) = \Sigma_{w_s} = \frac{1}{T} \sum_{t=1}^T w_t w'_{t+s}, \quad |s| = 0, 1, 2, \dots$$

be the probability limit for the covariance estimator of the covariance between w_t and w_{t+s} .⁹ Furthermore we may note that $\Sigma_{w_{s-j}} = \Sigma_{w_{s+j}}$, $|s|, |j| = 1, 2, \dots$. Similar arguments hold for w_t and another covariance stationary stochastic variable τ_t :

$$\Sigma_{w, \tau_s} = \frac{1}{T} \sum_{t=1}^T w_t \tau'_{t+s}, \quad |s| = 1, 2, \dots$$

The asymptotic OLS biases are given as $a_1 = \text{plim}(\hat{\alpha}_{1,1} - \alpha_{1,1})$, and $a_2 = \text{plim}(\hat{\alpha}_{3,1} - \alpha_{3,1})$, the corresponding formulas being $a_i = \text{plim} \left(\frac{\eta_{i,t} X'}{T} \right) \text{plim} \left(\frac{XX'}{T} \right)^{-1}$, $i = 1, 3$. It then follows that

$$\text{plim}(XX'/T) = \left[\Sigma_{\eta_1} + 2\Sigma_{\eta_1 \iota_2} \Lambda' + \Lambda \Sigma_{\iota} \Lambda' + 2\Sigma_{\eta_2^\circ \iota_1} \Lambda' + \Sigma_{\eta_2^\circ} + 2\Sigma_{\eta_1 \eta_{2,1}^\circ} \right] \quad (8)$$

and

$$\begin{aligned} a_1 &= \left[\begin{array}{l} \Sigma_{\eta_{1,1}} + \Sigma_{\eta_{1 \iota_2}} \Lambda' + \Sigma_{\eta_{1 \eta_{2,2}}} \Lambda' + \Sigma_{\eta_{1 \eta_{2,1}}} \Phi' \Lambda' \\ + \Sigma_{\eta_{1 \iota_1}} \Lambda' + \Lambda \Sigma_{\iota} \Lambda' + \Lambda \Sigma_{\eta_{2 \iota}} \Lambda' + \Lambda \Sigma_{\eta_{2 \iota_1}} \Phi' \Lambda' \\ + \Sigma_{\eta_2^\circ \eta_1} + \Sigma_{\eta_2^\circ \iota_1} \Lambda' + \Sigma_{\eta_2^\circ \eta_{2,1}} \Lambda' + \Sigma_{\eta_2^\circ \eta_2} \Phi' \Lambda' \end{array} \right]' \\ &\quad \times \left[\Sigma_{\eta_1} + 2\Sigma_{\eta_{1 \iota_2}} \Lambda' + \Lambda \Sigma_{\iota} \Lambda' + 2\Sigma_{\eta_2^\circ \iota_1} \Lambda' + \Sigma_{\eta_2^\circ} + 2\Sigma_{\eta_1 \eta_{2,1}^\circ} \right]^{-1}. \end{aligned}$$

⁸ Note that stationarity of the forecast error is part of the necessary conditions for writing (1) as a cointegrated system.

⁹ For $s = 0$ the subscript will be omitted.

Both factors on the right hand side involve the term $\Lambda\Sigma_\iota\Lambda'$ which implies that a_1 cannot be assumed to turn out zero. In fact, since $\Lambda\Sigma_\iota\Lambda'$ is positive semi-definite there is a tendency for the $\hat{a}_{1,1}$ to be biased upward. Hence, it will in general not be informative about the true causal links.

Turning to a_3 , we find

$$a_3 = \left[\Sigma_{\eta_1}\tilde{\eta}_{2,1} + \Lambda\Sigma_\iota\tilde{\eta}_{2,1} + \Sigma_{\tilde{\eta}_2\eta_2^\circ} \right]' \\ \times \left[\Sigma_{\eta_1} + 2\Sigma_{\eta_1\iota_2}\Lambda' + \Lambda\Sigma_\iota\Lambda' + 2\Sigma_{\eta_2^\circ\iota_1}\Lambda' + \Sigma_{\eta_2^\circ} + 2\Sigma_{\eta_1\eta_{2,1}^\circ} \right]^{-1}.$$

To interpret z to not depend on y the estimate of $\hat{a}_{3,1}$ should be a matrix of zeros. This again will generally not be the case. The reason is the term $\Sigma_{\tilde{\eta}_2\eta_2^\circ}$ in the first factor on the right hand side which is not going to be zero even if the remaining matrix expressions which involve cross covariances might do. Moreover, it would not even help to find $\iota_t = 0, \forall t$ (perfect forecasts). In that situation the most reasonable effect would be an even larger bias since the 'denominator', $\left(\frac{XX'}{T}\right)$, would be 'smaller' and hence the bias would not be reduced as much as it is due to $\Lambda\Sigma_\iota\Lambda'$ for non-zero forecast errors. On the other hand, if the forecast error had a huge variance in comparison to the innovations in η_i the bias would disappear. This effect gives rise to what has been called 'good forecast bias'. Similarly, for very large Σ_{η_1} implying that the economic model does not make much sense, the bias would also disappear. Therefore, the true causal links between y and z will be obtained only if the underlying economic model is poor as defined in the main text.

A.2 Step 2 - the complementary regression

The standard regression approach appeared to produce unreliable or even totally misleading results with respect to the coefficients of interest. Therefore, a complementary regression has been suggested in section 4. Consider

$$\tilde{A}_0 Y_t = \tilde{A}_1 Y_{t-1} + \tilde{\eta} \\ \begin{bmatrix} I_{n_1} & -\Lambda \\ 0_{n_2 \times n_1} & I_{n_2} \end{bmatrix} \begin{bmatrix} y_t \\ z_{t+1} \end{bmatrix} = \begin{bmatrix} 0_{n_1 \times n_1} & 0_{n_1 \times n_2} \\ 0_{n_2 \times n_1} & \Phi \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_t \end{bmatrix} + \begin{bmatrix} \eta_{1,t} + \Lambda\iota_{t+1} \\ \eta_{2,t+1} \end{bmatrix},$$

Pre-multiplying with the inverse of \tilde{A}_0 and subtracting Y_{t-1} from both sides gives

$$\Delta \begin{bmatrix} y_t \\ z_{t+1} \end{bmatrix} = \begin{bmatrix} -I_{n_1 \times n_1} & \Lambda\Phi \\ 0_{n_2 \times n_1} & -\Psi \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_t \end{bmatrix} + \tilde{A}_0^{-1} \begin{bmatrix} \eta_{1,t} + \Lambda\iota_{t+1} \\ \eta_{2,t+1} \end{bmatrix}$$

which reproduces the triangular structure seen before and which ensures that the economic and Granger causality coincide also in this partial model. The decomposition of the matrix in front of the lagged right hand side variables follows the lines above and the coefficient of interest will be called $\alpha_{1,1}$ and $\alpha_{2,1}$ respectively. The new regressor, \tilde{X} , is now given by

$$\tilde{X}' = \begin{bmatrix} (y_0 - \Lambda\Phi z_1)' \\ (y_1 - \Lambda\Phi z_2)' \\ \vdots \\ (y_{t-1} - \Lambda\Phi z_t)' \end{bmatrix},$$

giving rise to

$$\text{plim} \left(\frac{\tilde{X}\tilde{X}'}{T} \right) = \begin{bmatrix} \Lambda\Sigma_{\tilde{\eta}_2}\Lambda' - 2\Lambda\Sigma_{\tilde{\eta}_2\eta_2}\Lambda' - 2\Lambda\Sigma_{\tilde{\eta}_2\eta_{1,1}} - 2\Lambda\Sigma_{\tilde{\eta}_2\iota}\Lambda' \\ +\Lambda\Sigma_{\eta_2}\Lambda' + 2\Sigma_{\eta_1\eta_{2,1}}\Lambda' + 2\Lambda\Sigma_{\eta_2\iota}\Lambda' \\ +\Sigma_{\eta_1} + 2\Sigma_{\eta_1\iota_1}\Lambda' \\ +\Lambda\Sigma_{\iota}\Lambda' \end{bmatrix}, \quad (9)$$

which can be simplified to yield

$$\text{plim} \left(\tilde{X}\tilde{X}'/T |_{\Phi=I_{n_2}} \right) = [\Sigma_{\eta_{1,0}} + 2\Sigma_{\eta_1\iota_1}\Lambda' + \Lambda\Sigma_{\iota}\Lambda'],$$

in case of $\Phi = I_{n_2}$ which corresponds to the cointegration approach pursued in the application. Furthermore, the biases \tilde{a}_1 , \tilde{a}_2 for the estimates $\alpha_{1,1}$ and $\alpha_{2,1}$ are now

$$\tilde{a}_1 = \begin{bmatrix} -\Lambda\Sigma_{\tilde{\eta}_2\eta_{1,1}}\Lambda' - \Lambda\Sigma_{\tilde{\eta}_2\iota_1}\Lambda' - \Lambda\Sigma_{\tilde{\eta}_2\eta_{2,1}}\Lambda' \\ +\Lambda\Sigma_{\eta_2\eta_{1,1}}\Lambda' + \Lambda\Sigma_{\eta_2\iota_1}\Lambda' + \Lambda\Sigma_{\eta_2\eta_{2,1}}\Lambda' \\ +\Sigma_{\eta_{1,1}} + \Sigma_{\eta_1\iota_2}\Lambda' + \Sigma_{\eta_1\eta_{2,2}}\Lambda' \\ +\Lambda\Sigma_{\iota\eta_1} + \Lambda\Sigma_{\iota_1}\Lambda' + \Lambda\Sigma_{\iota\eta_{2,1}}\Lambda' \end{bmatrix}' \times \text{plim} \left(\frac{\tilde{X}\tilde{X}'}{T} \right)^{-1},$$

$$\tilde{a}_2 = [\Lambda\Sigma_{\tilde{\eta}_2\eta_{2,1}} - \Lambda\Sigma_{\tilde{\eta}_2,1} + \Sigma_{\eta_1\tilde{\eta}_2,2} + \Lambda\Sigma_{\tilde{\eta}_2\iota_1}]' \times \text{plim} \left(\frac{\tilde{X}\tilde{X}'}{T} \right)^{-1}.$$

Again, one might look at the special case for $\Phi = I_{n_2}$. We then have

$$\tilde{a}_1|_{\Phi=I_{n_2}} = \begin{bmatrix} \Sigma_{\eta_{1,1}} + \Sigma_{\eta_1\iota_2}\Lambda' + \Sigma_{\eta_1\eta_{2,2}}\Lambda' \\ +\Lambda\Sigma_{\iota\eta_1} + \Lambda\Sigma_{\iota_1}\Lambda' + \Lambda\Sigma_{\iota\eta_{2,1}}\Lambda' \end{bmatrix}' \times \text{plim} \left(\tilde{X}\tilde{X}'/T |_{\Phi=I_{n_2}} \right)^{-1},$$

$$\tilde{a}_2|_{\Phi=I_{n_2}} = [\Sigma_{\eta_1\tilde{\eta}_2,2} + \Lambda\Sigma_{\tilde{\eta}_2\iota_1}]' \times \text{plim} \left(\tilde{X}\tilde{X}'/T |_{\Phi=I_{n_2}} \right)^{-1}.$$

It is thus straightforward to see that for smaller forecast errors ($\iota_t \approx 0_{n_2 \times 1}$), \tilde{a}_2 will be closer to zero than otherwise. The same holds for small variations in $\eta_{1,t}$. Under the same conditions $\hat{\alpha}_{1,1}$ will approach $\alpha_{1,1}$.¹⁰ Thus, the better the economic model, the higher is the chance that the true causal links will be revealed. However, it is not known a priori, if

¹⁰Independent of Φ the 'numerators' now only contain cross terms which would vanish if zero correlation between $\eta_{1,t}$ and $\eta_{2,r}$ for all r and t is assumed and if the correlation between $\eta_{1,t}$, $\eta_{2,t}$, ι_t and ι_r is zero for all $r \neq t$ (i.e. under rational expectations).

one finds herself or himself in the standard regression or in the complementary regression. That's why the following procedure can be suggested.

1. run a regression as in the standard case
2. lead the set of regressors which do not turn out weakly exogenous to the most likely period for which expectations of these variables may count for the weakly exogenous variables
3. run the complementary regression
4. If the same set of variables turn out weakly exogenous as before they can be considered the driving variables
5. If the set of variables turn out weakly exogenous which have previously been found the endogenous variables, lead the weakly exogenous variables of the first step appropriately and run another regression.
6. If the set of variables turning out weakly exogenous is the same as in step 2, then they can be considered the driving variables

Otherwise no set of variables can be labelled causal. The number 5 of the procedure above could be regarded a third step regression, but in fact it is merely a confirmation of the correct choice and could also be omitted. In Table 2 the results for this third regression are reported in section 6.

A.3 Generalisation for $s \geq 1$

So far, s has been restricted to equal 1. It is easy to see however, that the results can be generalised to any value of s . Consider

$$\tilde{y}_t = \begin{bmatrix} y_t \\ \vdots \\ y_{t-s} \end{bmatrix}, \tilde{z}_t = \begin{bmatrix} z_{t+s-1} \\ \vdots \\ z_{t-1} \end{bmatrix}, \Lambda^+ = I_s \otimes \Lambda, \Phi^+ = I_s \otimes \Phi$$

Then replace Λ by Λ^+ , Φ by Φ^+ , y_t by \tilde{y}_t and z_t by \tilde{z}_t in (1) and the analysis goes through.

B 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