

Identifying Causal Structures of Cointegrated Vector Autoregressions with an Application to the G7 Interest Rates.

Marco R. Barassi
Imperial College Management School

A thesis submitted for the degree of Doctor of Philosophy in Economics
31300 Words



A mio padre, mia madre e Ariadna

Acknowledgements

I wish to thank all the people who have made this work possible. Absolutely necessary has been the support I have received from my supervisor Professor Stephen G. Hall. This research is due to his constant motivation and insightful advice without which this thesis would never have been completed.

I am grateful to many people in the Management School, Imperial College for their help. In particular, I would like to mention Professor Keith Cuthbertson and Dr Dirk Nitzsche for their comments and suggestions. My indebtedness extends to all the staff of the Management School, especially to Julie Paranics, Deborah Salmon, and Craig Perrott. I would also like to thank all the colleagues of the research office of the Management School, especially Eugenia Garganas, Ali Al-Eyd, and Efstratios Rappos.

I wish to express my gratitude to all the people of the Centre for Monetary and Financial Economics, at South Bank University. In particular, the director of the CEMFE, Professor Guglielmo Maria Caporale for his interest in this research and his constant motivation, and the director of research Professor Philip Arestis for all our discussions. I would like to mention my colleagues Andrea Cipollini and Kostas Mouratidis for their comments on this research.

Some special thanks are due to my friends Nicola Spagnolo and Fabio Spagnolo, with whom this long “academic adventure” started on a day of September 1995. I am also grateful to Walter Lo Faro, Alberto Cirio and all my Italian friends who helped to create a homely atmosphere in London. A special thank goes to my friend Gherardo Catanzaro for being present and supportive in all the crucial moments as only brothers can be.

And of course I wish to express all my gratitude and love to my family. To my parents for their love and financial support, and to my sisters Marina and Patrizia and their families for their encouragement.

Finally a very special thank to Ariadna for her constant encouragement, support and invaluable love.

Abstract

In this thesis we investigate the identification of the long-run causal structure that links non-stationary variables in cointegrated systems with particular reference to the case in which this structure breaks down. We review the relevant literature in chapter 1. In chapter 2 we introduce the Extended Davidson Methodology, and we use it to examine the structural linkages that exist between the G7 interest rates as well as the causal structure that gives rise to cointegration. In chapter 3, we examine, using Monte Carlo methods, several alternative approaches to detecting structural changes in the causality structure of cointegrated VARs. We adopt recursive and rolling FIML-based techniques as misspecification tests for the detection of instabilities in the adjustment coefficients and of the breakpoints. To overcome the relative inefficiency of these two techniques, we propose an OLS-based sequential test. We derive its distribution and the critical values for different sample sizes and discuss its statistical properties. We will show that the test has a non-standard distribution and has reasonable power for smaller samples. Also, it is graphically efficient in showing the breakpoint. As a benchmark, we also test the performance of a time-varying version of the Kalman Filter on the same DGPs. In chapter 4 we applied the proposed sequential technique as well as the Kalman Filter to the causal linkages of the G7 short-term rates with interesting results. In Chapter 5 we summarise the results obtained and draw some final conclusions.

Index

CHAPTER 1. ON IDENTIFICATION OF THE LONG-RUN STRUCTURES OF
COINTEGRATED VARS 11

1.1 Introduction 11

1.2 Identification of cointegrating relations in cointegrated VARs 13

1.3 Irreducible Cointegrating Vectors and Identification of Structural Relations: The
Davidson Methodology 21

1.4 Weak-Exogeneity and Granger-Causality..... 26

CHAPTER 2 IDENTIFICATION OF STRUCTURAL COINTEGRATING VECTORS: AN
APPLICATION OF THE EXTENDED DAVIDSON METHODOLOGY TO OF THE G7
INTEREST RATES 36

2.1 Introduction 36

2.2 Analysing Interest Rate Linkages 38

2.3 The data-set and the methodology 39

2.4 Empirical results..... 45

2.4.a Short-term rates: the empirical evidence..... 45

2.4.b Long term rates: the empirical evidence 48

2.5 Summary and Conclusions..... 51

CHAPTER 3. TESTING FOR CHANGES IN THE LONG-RUN CAUSAL STRUCTURE OF COINTEGRATED VECTOR AUTOREGRESSIONS..... 63

3.1	Introduction	63
3.2	The Monte Carlo experiment	69
3.3	A sequential test for structural change	71
3.4	Kalman Filtering	73
3.5	Results	75
3.6	Conclusions	80

CHAPTER 4. TESTING FOR STRUCTURAL BREAKS IN THE CAUSAL LINKAGES BETWEEN THE G7 SHORT-TERM INTEREST RATES 89

4.1.	Introduction	89
4.2	Testing for changes in the long-run causality of Cointegrated Vector Error Correction Models.....	92
4.3	Empirical Results	95
4.3.a	Sequential test results	98
4.3.b	Kalman Filter results	99
	<i>USA</i>	100
	<i>UK</i>	100
	<i>The Eurozone</i>	101
	<i>Canada and Japan</i>	102

4.4 Conclusions 103

CHAPTER 5. CONCLUSIONS..... 126

APPENDIX A- GAUSS CODE FOR KALMAN FILTER ESTIMATION 130

REFERENCES.....141

List of Tables

Table 2.1. The relationship of cointegrating errors between structural and solved vectors.....54

Table 2.2. Unit root test on G7 short-term rates..... 54

Table 2.3. Unit root test on G7 long-term rates 54

Table 2.4. G7 short term rates cointegration test 55

Table 2.5 Pairwise cointegration tests for G7 short-term rates 56

Table 2.6. Short-term rates, ranking of IC vectors..... 57

Table 2.7 Cross tabulation of Standard deviations for the G7 short rates..... 58

Table 2.8. G7 long-term rates cointegration test..... 58

Table 2.9 long-run rates pairwise cointegration tests..... 59

Table 2.10. Cointegration test between the inflation rates of Italy and Germany.....60

Table 2.11. ADF-test on inflation rates of Italy and Germany 60

Table 2.12 ADF-test on Cointegrating vectors indicated from rank test in table 4.10 60

Table 2.13 Ranking of irreducible cointegrating vectors..... 61

Table 2.14 Ranking of IC vectors per country..... 62

Table 3.1.Full Sample FIML Estimates (1000 replications)..... 82

Table 3.2 t-values of recursive estimates of α_2 83

Table 3.3 t-values of rolling estimates of α_2 84

Table 3.4 Sequential test – Statistical properties based on 2500 replications..... 84

Table 4.1. Canada: OLS-based estimates for the ECMs 105

Table 4.2. France: OLS-based estimates for the ECMs 106

Table 4.3. Germany: OLS-based estimates for the ECMs 107

Table 4.4. Italy: OLS-based estimates for the ECMs..... 108

Table 4.5. Japan: OLS-based estimates for the ECMs 109

Table 4.6. UK: OLS-based estimates for the ECMs 110

Table 4.7. USA: OLS-based estimates for the ECMs 111

Table 4.8. Results of sequential test for structural change in causality for Canada..... 112

Table 4.9. Results of sequential test for structural change in causality for France 112

Table 4.10. Results of sequential test for structural change in causality for Germany 112

Table 4.11. Results of sequential test for structural change in causality for Italy 112

Table 4.12. Results of sequential test for structural change in causality for Japan..... 113

Table 4.13. Results of sequential test for structural change in causality for UK..... 113

Table 4.14. Results of sequential test for structural change in causality for USA..... 113

List of Figures

Figure 3.1 Recursive eigenvalues for DGP2 of size 100 (average of 1000 replications).....85

Figure 3.2 Recursive estimates of α_2 for DGP2 of size 100 (average of 1000 replications).....85

Figure 3.3 Rolling estimates of α_2 for DGP2 of size 100 with a rolling window of 30 (average of 1000 replications).....86

Figure 3.4 Rolling estimates of α_2 for DGP2 of size 100 with rolling window of 40 (average of 1000 replications).....86

Figure 3.5 Rolling estimates of α_2 for DGP2 of size 100 with rolling window of 50 (average of 1000 replications).....87

Figure 3.6 Coefficient of sequential dummy for DGP2 of size 100 (average of 2500 replications).....87

Figure 3.7 One step ahead Kalman Filter predictions of α_2 for DGP2 of size 100 (average of 1000 replications).....88

Figure 3.8 Smoothed kalman Filter estimates of α_2 for DGP2 of size 100 (average of 1000 replications).....88

Figure 4.1 Test statistic for a break on the French adjustment coefficient towards the long-run relationship France-Canada.....114

Figure 4.2 Test statistic for a break on the French adjustment coefficient towards the long-run relationship France-Germany.....114

Figure 4.3 Test statistic for a break on the French adjustment coefficient towards the long-run relationship France-Japan.....115

Figure 4.4 Test statistic for a break on the Italian adjustment coefficient towards the long-run relationship Italy-France.....115

Figure 4.5 Test statistic for a break on the Italian adjustment coefficient towards the long-run relationship Italy-Germany.....116

Figure 4.6 Test statistic for a break on the Italian adjustment coefficient towards the long-run relationship Italy-Japan.....116

Figure 4.7 Test statistic for a break on the Japanese adjustment coefficient towards the long-run relationship Japan-Canada.....117

Figure 4.8 Test statistic for a break on the Japanese adjustment coefficient towards the long-run relationship Japan-Usa.....117

Figure 4.9 Test statistic for a break on the UK adjustment coefficient towards the long-run relationship UK-Germany.....118

Figure 4.10. Time path of adjustment coefficients in error correction equations for Canada119

Figure 4.11. Time path of adjustment coefficients in error correction equations for France.120

Figure 4.12. Time path of adjustment coefficients in error correction equations for Germany
.....121

Figure 4.13. Time path of adjustment coefficients in error correction equations for Italy...122

Figure 4.14. Time path of adjustment coefficients in error correction equations for Japan..123

Figure 4.15. Time path of speed of adjustment coefficients in error correction equations for the UK.....124

Figure 4.16. Time path of adjustment coefficients in error correction equations for USA...125

Chapter 1

On identification of the long-run structures of cointegrated VARs

1.1 Introduction

In econometrics the identification problem is related to the question of whether it is possible to identify economic structures of some interest within a statistical model. This problem within the cointegration framework assumes very different attributes with respect to the usual ones of vector autoregressive models (VARs) involving only stationary variables (for an exhaustive treatment of VAR models see Hamilton 1995, Hendry 1995). To see this, let us consider a p -dimensional VAR,

$$(1.1.1) \quad X_t = \sum_{i=1}^p \Pi_i X_{t-i} + \gamma D_t + \varepsilon_t$$

where X_t is a $nx1$ vector of variables for fixed values of $X_{k+1}, \dots, X_0$ and $\varepsilon_t \sim iid N_p(0, \Omega)$. The deterministic term D_t can contain a constant and linear term, seasonal or intervention dummies or other non-stochastic regressors. Since each of the variables in the system is either $I(0)$ or $I(1)$, the determinant polynomial $|\Pi(z)|$ contains at most n unit roots. When there are $r < p$ unit roots, then the variables in X are cointegrated in the sense of Granger (1983) and Engle and Granger (1987), βX_t being the cointegrating relations.

Following standard practice we reparameterise the VAR as a structural Vector Error Correction Model as (see Davidson and Hall (1991) and Urbain (1992)),

$$(1.1.2) \quad A_0 \Delta X_t = \Pi^* X_{t-1} + \sum_{i=1}^{p-1} \Psi_i \Delta X_{t-i} + u_t$$

where Π^* has rank r meaning that there are r linear combination of the variables in X that are stationary. This rank may be imposed by defining $\Pi^* = \alpha^* \beta^{*'}$, where α^* and β^* are pxr

matrices containing respectively the structurally identified loading weights and cointegrating vectors. Notice that the estimates of α and β as obtained from the maximum likelihood procedure are not uniquely identified. In fact, given any choice of α and β and any $(r \times r)$ non-singular matrix P , the choice αP and $\beta(P')^{-1}$ will give the same matrix Π determining the same probability distribution for the variables. This is equivalent to say that the data can determine only the space spanned by the cointegrating vectors β and the adjustment space spanned by α . Generally speaking the only functions of α and β that can be estimated are those which can be expressed as functions of the matrix Π .

The structural VECM will be usually estimated as an unrestricted version of the reduced form that is

$$(1.1.3) \quad \Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{p-1} K_i \Delta X_{t-i} + \eta_t,$$

where $K_i = A_0^{-1} \Psi_i$, $\eta_t = A_0^{-1} u_t$, and $\Pi = A_0^{-1} \Pi^*$. We see that there are two different and separate components to the identification problem. The first one concerns the identification of the contemporaneous coefficients in A_0 whilst the second entails the identification of the long run structure of the model, namely the coefficients in β . In general restrictions imposed on the latter cannot tell us anything about the identification of A_0 as this comes from the dynamics of the model. Similarly the dynamic part of a model is of no help in the identification of the long run structure because given that $\Pi = A_0^{-1} \alpha^* \beta^{*'}$, even if we knew A_0 we could not identify β^* without additional restrictions on β .

To summarise, the complete exact (over) identification of the system will involve a combination of four types of restrictions

1. Restrictions on the cointegrating rank of Π ,
2. Restrictions on the cointegrating vectors β ,

3. Restrictions on the matrix of adjustment coefficients α to identify exogeneity or long run causality, and
4. Restrictions on the dynamics K_i .

For unique identification, at least $p(p-1)$ restrictions must be imposed on the short run structural form parameters, while the long run part of the system, being the same in both structural and reduced form, may be identified in either form.

The purpose of this thesis is to investigate some issues related to the analysis of cointegrated time series. Specifically:

- (a) the identification of the structural cointegrating relations (this relates to point 2);
- (b) the identification of the long-run causal structure that links the variables with particular reference to the case in which this structure breaks down (see point 3).

In this introductory chapter we review the literature on identification of the cointegrating relations and the causal structure that generates these relations. The next section contains a review of the most popular contributions (Johansen 1995a Boswijk 1995 and Pesaran and Shin 1994), in terms of the identification of the structural relations. Section 1.3 presents a revolutionary method due to Davidson (1998) which relies on the concept of an Irreducible Cointegrating vector. Section 1.4 will recall some concepts such as weak exogeneity and Granger causality. We will see how these concepts are related and we review the testing procedures available.

1.2 Identification of cointegrating relations in cointegrated VARs

Following Johansen (1995a and 1995b), in order to be able to recognise or identify uniquely the individual relations in $\Pi = \alpha\beta'$ we formulate linear restrictions R_i as full rank $p \times r_i$ matrices and assume that the coefficients in the i -th equation, β_i satisfy

$$(1.2.1) \quad R_i' \beta_i = 0, \quad i = 1, \dots, r$$

We can also use a different formulation of the problem by defining $H_i = R_i^\perp$, that is, we define a $px(p-r_i)=pxs_i$ matrix of full rank such that $H_i' R_i = 0$. In this case (1.2.1) is replaced by $\beta \theta_i = H_i \phi_i$ for some s_i -dimensional vector ϕ_i . If H_i contains zeros and ones only, then H_i can be considered the selection matrix that indicates which elements of β_i vary unrestrictedly.

A direct parameterisation is

$$(1.2.2) \quad \beta_i = (H_1 \phi_1, \dots, H_r \phi_r)$$

The set of restrictions define a linear statistical model by the parameter space

$$(1.2.3) \quad \aleph = \{\beta_{pxr}, \Omega \mid R_i' \beta_i = 0, i = 1, \dots, r\}.$$

The classical result on identification (see Sargan, 1988) is given by the rank condition and is stated by Johansen (1995) with the following theorem.

Theorem 1. A necessary and sufficient condition that a parameter value (θ, Ω) is identified is that

$$(1.2.4) \quad \text{rank}(R_i' \beta) = r - 1, \quad i = 1, \dots, r.$$

The meaning of this is that, when applying the restriction of one equation to the other $r-1$ equations we get a matrix of rank $r-1$. This result implies that it is not possible by taking linear combinations of, for instance, $\beta_2, \dots, \beta_r$ to construct a vector which is restricted in the same way as β_1 . Hence β_1 will be, among all linear combinations of $\beta_1, \dots, \beta_r$ the only one that lies in the $sp(H_1)$, the space spanned by the columns in H_1 , and the only one that satisfies the restrictions R_1 .

Together with $\aleph$ we consider another model $\Im$ that comprises the set of identified parameters, that is

$$(1.2.5) \quad \Im = \{\theta_{pxr}, \Omega \mid R_i' \theta_i = 0, \text{rank}(R_i' \theta) = r - 1, i = 1, \dots, r\}.$$

Note that while $\mathcal{N}$ is a linear model, $\mathcal{S}$ is non-linear for it contains the rank condition that is a nonlinear restriction. We can see that $\mathcal{S} \subset \mathcal{N}$, that is $\mathcal{S}$ is an open dense subset of $\mathcal{N}$ unless the latter is empty (Johansen 1995a).

The implication of this is that we cannot use the likelihood function to distinguish between the two models $\mathcal{N}$ and $\mathcal{S}$, due to the fact that $\mathcal{S} \neq \emptyset$ is dense in $\mathcal{N}$, and the maximum likelihood estimator from $\mathcal{N}$ satisfies the rank conditions and therefore corresponds to a point in $\mathcal{S}$ with probability equal to one. For a given set of restrictions either all parameters in $\mathcal{N}$ defined by the restrictions are unidentified, or “almost all” are identified. Note that a given set of linear restrictions defines the statistical model $\mathcal{N}$. Thus the properties of the model are given by the properties of the restrictions. If we want to stay in the linear world, we need to define a set of identifying restrictions that give rise to a linear model in which the identified parameters are dense.

Johansen (1995a) states that, *The restrictions $R_1, \dots, R_r$ and the statistical linear model $\mathcal{N}$ are called identifying (generically identifying) if there exists a parameter value θ such that the rank condition (1.2.5) is satisfied, or equivalently such that $\mathcal{S}$ is non-empty.*

What we need is a simple algebraic condition such that before a linear model is estimated, a condition can be checked. This would give us an immediate criterion for determining whether a model of the class of $\mathcal{N}$ is identifying since one only needs to check the rank condition for a parameter value imposed randomly but always on a “equation by equation” basis. This implies that only orthonormalising restrictions are imposed, and that this procedure does not involve cross-equation restrictions and restrictions on the covariance matrix for the criterion does not allow for these cases.

Let us consider a cointegrated vector autoregressive model in structural error correction form,

$$(1.2.6) \quad \Delta Y_t = \alpha \beta' Y_{t-1} + \Gamma \Delta Y_{t-1} + \mu + \varepsilon_t$$

where the ε_t are iid $N(0, \Omega)$. Suppose that this model is defined by some restrictions $R_1, \dots, R_r$.

A parameter point has a representation $\theta_i = (H_1\phi_1, \dots, H_r\phi_r)$. For it to be identified it must satisfy the rank condition

$$a) \quad \text{rank}(R_i'H_1\phi_1, \dots, R_i'H_r\phi_r) = r - 1, \quad i = 1, \dots, r$$

and in this case we have $r-1$ column vectors $\{R_i'H_j\phi_j, j = 1, \dots, r, j \neq i\}$ that are linearly independent. This result implies a condition that does not involve the parameters:

$$b) \quad \text{rank}(R_i'H_{i_1}, \dots, R_i'H_{i_k}) \geq k,$$

for any sets of k indices $(i_1, \dots, i_k)$ not containing i . Johansen (1995a) formulates this condition as a theorem:

Theorem 2. The linear statistical model Σ defined by the restrictions $R_1, \dots, R_r$ is identifying if and only if for $K=1, \dots, r-1$ and any set of indices $1 \leq i_1 < \dots < i_K \leq r$ not containing I , it holds that $\text{rank}(R_i'H_{i_1}, \dots, R_i'H_{i_K}) \geq K$. If this holds for a particular value of I , then the restrictions are identifying for that equation.

If we apply the above condition to all possible vectors given by $H_1, \dots, H_r$, we would be able to check whether a linear statistical model contains in general identified parameters. This provide us with a general criterion for identification which does not depend on the unknown parameters but which has the limitation of being applicable only on an equation by equation basis. Also, it supports only orthonormalising restrictions on the matrix of cointegrating coefficients without allowing for cross equation restrictions or restrictions on the covariance matrix.

Another line of research is based on error correction models that are conditional upon a set of variables (see Boswijk 1992, Kremers et al. 1992). In this framework, conditional models are logically related to economic theory, which often specify a number of exogenous variables in addition to the endogenous ones. According to this approach the linkages between a model

and the economic theory underlying it, is better observable in error correction models in structural form. Boswijk's approach (1995) constitute a synthesis between the two multivariate analytical techniques due to Phillips and Hansen (1990), Phillips (1991) Saikkonen (1991) on one side, and Johansen (1991,1995a) on the other. The starting point is a general cointegrated VAR in ECM form,

$$(1.2.7) \quad \Delta X_t = \Pi X_{t-1} + \sum_{j=1}^{p-1} A_j \Delta X_{t-j} + \varepsilon_t, \quad t = 1, \dots, T.$$

with $\{\varepsilon_t\}$ that is assumed to be an independent normal sequence $(0, \Omega)$. As before Π has deficient rank $0 < r < n$, and can be factorised as $\Pi = \alpha\beta'$, where α and β are $n \times r$ matrices of full column rank, so that (1.2.7) can be rewritten as

$$(1.2.8) \quad \Delta X_t = \alpha\beta' X_{t-1} + \sum_{j=1}^{p-1} A_j \Delta X_{t-j} + \varepsilon_t, \quad t = 1, \dots, T.$$

Let us partition the n -dimensional X_t as $X_t = (Y_t', Z_t')'$, where the k -dimensional Y_t contains the endogenous variables and the d -dimensional Z_t consists of exogenously determined variables with $(k + d = n)$. Having partitioned A_j , $j = 1, \dots, p-1$, α and Ω conformably with X_t and rearranging terms we can write the model in terms of the conditional and the marginal models, that is

$$(1.2.9) \quad \begin{aligned} \Delta Y_t &= \Lambda_0 \Delta Z_t + \alpha_{1.2} \beta' X_{t-1} + \sum_{j=1}^{p-1} A_{1.2,j} \Delta X_{t-j} + \varepsilon_{1.2t} & t = 1, \dots, T \\ \Delta Z_t &= \alpha_2 \beta' X_{t-1} + \sum_{j=1}^{p-1} A_{2j} \Delta X_{t-j} + \varepsilon_{2t} & t = 1, \dots, T \end{aligned}$$

where $\Lambda_0 = \Omega_{12} \Omega_{22}^{-1}$, $\alpha_{1.2} = \alpha_1 - \Lambda_0 \alpha_2$, $A_{1.2,j} = A_{1j} - \Lambda_0 A_{2j}$ $j = 1, \dots, T$, also note that

$\varepsilon_{1.2t} = \varepsilon_{1t} - \Lambda_0 \varepsilon_{2t}$. The necessary and sufficient condition for Z_t to be weakly exogenous for α and β is that $\alpha_2 = 0$ and all the relevant information on the long run parameters can be obtained from the analysis of the conditional or partial model only. The main difference between Boswijk's and Johansen's approaches lies in the fact that the former consider over-

identified cointegrating vectors, therefore using ECM in structural form like a reparameterisation of the first equation in (1.2.9), that is obtained pre-multiplying it by some non-singular $k \times k$ matrix Ψ_0

$$(1.2.10) \quad \Psi_0 \Delta Y_t = \delta \beta' X_{t-1} + \Gamma_0 \Delta Z_t + \sum_{j=1}^{p-1} (\Psi_j \Delta Y_{t-j} + \Gamma_j \Delta Z_{t-j}) + \eta_t$$

with $\delta = \Psi_0 \alpha_{1,2}$, $\Gamma_0 = \Psi_0 \Lambda_0$, $[\Psi_j : \Gamma_j] = \Psi_0 A_{1,2,j}$, $j = 1, \dots, p-1$, and the sequence $\{\eta_t\} = \Psi_0 \varepsilon_{1,2,t}$, that is a Gaussian innovation process with covariance matrix $\Sigma = \Psi_0 \Omega_{1,2} \Psi_0'$.

Although β is already identified, we need further restrictions to identify the remaining parameters in (1.2.10). We can first normalise the system for $\Psi_{0,ii}$ $i = 1, \dots, k$. Also other restrictions may be suggested by economic theory¹. In Boswijk (1995), the starting point is the following assumption,

Assumption 1. In (1.2.10) it is the case that i) $k=r$ and ii) $\text{rank}(\delta \beta'_y) = r$, where

$\beta' = [\beta'_y : \beta'_z]$ with β_y an $r \times r$ matrix.

Then, if we know the cointegrating rank, and it equals the number of endogenous variables in the conditional model, and this is stable², we can obtain the cointegrating rank using the Johansen (1988) ML procedure applied to the VAR in error correction form. Conditional on this estimate one can choose r endogenous variables, associated with the r cointegrating vectors, and test for stability as in Boswijk (1994). Note that condition ii) imposes a limit to the choices of the matrix H_i in $\beta = H_i \phi_i$. Under these conditions we can restrict δ to be a

¹ Boswijk (1995) suggests starting with a just identified model which can be given a straightforward economic interpretation, and restrict this model according to some statistical model selection criteria. Two approaches are suggested:

- The first class of identified structures is obtained by setting $\Psi_0 = I_k$ so to obtain a model in structural form identical to the first equation in (1.2.10), and the over-identifying restrictions are restrictions imposed only on the conditional model,

- The second class is obtained by restricting Ψ_0 to be a triangular matrix and Σ to be diagonal, therefore obtaining what is commonly known as a block recursive system.

The main feature of this second class of structures is that it imposes an ordering in the contemporaneous effects or a causal chain (see Johansen and Juselius 1994 for an example).

diagonal matrix so that the restrictions on the i -th column of β are also imposed on the i -th structural error correction equation. If we let θ_i denote the i -th vector of $\beta\delta'$ we have

$$(1.2.11) \quad R_i\theta_i = R_i\beta_i\delta_i = 0.$$

Because of this, the usual rank condition $\text{rank}(R_i\beta) = r - 1$ for $i=1, \dots, r$ is sufficient for identification of the i -th equation of the conditional part of the system in (1.2.10), if assumption 1 also holds. The main result is that there are not other linear combinations of the remaining $k-1$ equations that satisfy (1.2.11). Obviously a large class of models is excluded by assumption 1 and by imposing diagonality restrictions. The reason for such a structure is that a diagonal error correction matrix provides a simple interpretation of the individual equations from the structural system. If the number of endogenous variables of interest is larger than the cointegrating rank (that is assumption 1 does not hold), then the method proposed is not applicable. If instead δ is found not to be diagonal, the model can always be transformed in a system with a diagonal error correction matrix, but in this case it will not represent a genuine structural form.

The limit of Johansen (1995) and Boswijk (1995) lies in that they allow only for linear restrictions on one vector at a time, therefore not allowing for non-linear restrictions or for restrictions across different cointegrating regressions. Pesaran and Shin (1994, updated 1999) deal with the identification problem in such cases deriving rank and order conditions for identification of cointegrating vectors. Their approach emphasises the use of economic theory and does not require an *a priori* partition of the system variables as in Boswijk (1995), Phillips (1991) or Saikkonen (1993).

² Stability in this sense has implies the condition that the characteristic equation of the ECM does not have any unit roots (see also Boswijk 1993 and 1994).

Starting from a Cointegrated VAR in Structural Error correction Form such as the one in (1.1.2), they show that if the matrix $\Pi = \alpha\beta'$ has reduced rank then it is subject to $(m-r)^2$ non linear restrictions and it is determined uniquely in terms of the $(2mr-r^2)$ underlying parameters. Pesaran and Shin (1994) therefore, consider the problem of identification when $\vartheta = \text{vec}(\beta)$ is subject to general k non-linear restrictions as in the case:

$$(1.2.12) \quad h(\vartheta) = 0,$$

where $\vartheta \in \Theta$, $\Theta \subset \mathbb{R}^{mr}$, $h(\vartheta) = (h_1(\vartheta), h_2(\vartheta), \dots, h_k(\vartheta))'$, and $h_i(\vartheta)$, $i = 1, \dots, k$, is a known continuous differentiable scalar function of ϑ . Letting $\vartheta_0 = f(\phi_0)$ be the true value of ϑ , and assuming that $H(\vartheta) = \partial h(\vartheta) / \partial \vartheta'$ has full rank $k \leq mr$, the analysis can now be conducted recalling that $\Pi_0 = \alpha_0 \beta_0' = \alpha_0 Q'^{-1} Q' \beta_0' = \alpha \beta'$ for any non-singular matrix Q . Vectorize $\beta = \beta_0 Q$ so to obtain

$$(1.2.13) \quad \vartheta = (I_r \otimes \beta_0) \text{vec}(Q).$$

Consider the mean value expansion of $h(\vartheta)$ around ϑ_0 as,

$$(1.2.14) \quad h(\vartheta) = h(\vartheta_0 + H(\tilde{\vartheta})(\vartheta - \vartheta_0)),$$

where as before the (i,j) element of $H(\tilde{\vartheta})$ is evaluated at $(\tilde{\vartheta}_i, \tilde{\vartheta}_j)$ and $\tilde{\vartheta}_i$ is a convex combination of ϑ_{0i} and ϑ_i . Under the restrictions in (1.2.12) we have

$$(1.2.15) \quad H(\tilde{\vartheta})\vartheta = g(\tilde{\vartheta}, \vartheta_0),$$

with $g(\tilde{\vartheta}, \vartheta_0) = H(\tilde{\vartheta})\vartheta - h(\vartheta_0) = 0$. And substituting for ϑ in (1.2.15) from (1.2.13) we get:

$$(1.2.16) \quad H(\tilde{\vartheta})(I_r \otimes \beta_0) \text{vec}(Q) = g(\tilde{\vartheta}, \vartheta_0)$$

and a unique solution for $\text{vec}(Q)$ exists if and only if the following rank condition holds,

$$(1.2.17) \quad \text{rank}\{H(\tilde{\vartheta})(I_r \otimes \beta_0)\} = r^2, \quad \forall \tilde{\vartheta}_i, \tilde{\vartheta}_j \in \mathfrak{R}_{\vartheta}.$$

Notice that the rank condition given above holds as a sufficient condition complementing the order condition given by $k \geq r^2$.

1.3 Irreducible Cointegrating Vectors and Identification of Structural Relations: The Davidson Methodology

A revolutionary approach has been recently proposed by Davidson (1998a), who introduces the concept of an *irreducible cointegrating* (IC) relation, one from which no variable can be omitted without loss of the cointegration property. The focus is on the identification of long-run structural relations, where the word structure is used to refer to relations that are consistent with theory-based restrictions and therefore have a clear economic interpretation.

Davidson (1998a) provides a structural interpretation of single cointegrating regressions a la Engle and Granger (1987). The advantage of the procedure he suggests is that, under certain circumstances, when the model is over-identified, it enables the researcher to obtain information about the underlying structure directly from the data. To illustrate the methodology, consider a cointegrated VAR(p), as analysed by Johansen (1988, 1995b):

$$(1.3.1) \quad A(L)x_t = \alpha\beta'x_t + A^*(L)\Delta x_t = \varepsilon_t \quad (px1),$$

where $x_t \sim I(1)$, L is the lag operator, $A(L) = \alpha\beta' + A^*(L)(1-L)$ such that $A(1) = \alpha\beta'$, and α and β are pxk matrices, the loading weights matrix and the matrix of cointegrating vectors

respectively³. When $k < p$ the system incorporates a set of long run relationships of the form $\beta'x_t = s_t$, where $s_t = (\alpha'\alpha)^{-1} \alpha' (\varepsilon_t - A^*(L)\Delta x_t) \sim I(0)$.

In this model there are k linearly independent cointegrating vectors, the columns of β . Recall that without restrictions on β we can always scale the matrix of the cointegrating relations by post-multiplying it by any non-singular $k \times k$ matrix C , to get $C\beta'x_t = Cs_t$, that is observationally equivalent to $\beta'x_t = s_t$ with loading matrix αC^{-1} . The identification problem within the procedures discussed so far is tackled by estimating a collection of linearly or non linearly restricted vectors spanning the same space as β that are identified by the rank condition. Davidson (1998a) proposes to follow a method that allows the researcher to use the data to identify the structural relations in the case of over-identified systems. We recall its main points.

Theorem 1 (Davidson, 1994). If a column of β (say β_l) is identified by the rank condition, the OLS regression which includes just the variables having unrestricted non-zero coefficients in β_l is consistent for β_l .

The issue raised by this theorem is that within a non-stationary world if another variable is added to a cointegrating regression, its coefficient might not necessarily converge to zero as we would expect in the case of an irrelevant variable within a regression involving stationary variables. In the case of cointegration the regression coefficients would generally converge to some other element of the cointegrating space. The main result of this is that, if a collection of $I(1)$ variables is found to be cointegrated, it does not necessary follow that the estimated vectors can be interpreted as structural. In this framework it is useful to recall the definition of irreducible cointegrating vector introduced by Davidson's (1998a), that is,

³ We have assumed for simplicity the absence of any deterministic terms in this representation of the system under analysis. The modifications necessary to relax these assumptions are straightforward and would not alter the substance of the results obtained using a simpler model.

Definition 1. *A set of $I(1)$ variables will be called irreducibly cointegrated (IC) if they are cointegrated, but dropping any of the variables leaves a set that is not cointegrated.*

IC relations have the following important property.

Theorem 2. *An IC vector is unique, up to the choice of normalisation.*

This is proved by the following argument. Assume that there exists for the IC variables a set of cointegrating vectors of rank at least two. We have seen that any linear combination of these vectors would lie in an observationally equivalent cointegrating space. If this is true, we can always generate a combination having a zero element by choosing the weights appropriately. This would allow us to drop the variable in question without losing cointegration, but this contradicts the definition of IC itself.

Theorem 3 (Davidson, 1994). *If and only if a structural cointegrating relation is identified by the rank condition, it is irreducible.*

This tells us that at least some IC vectors are structural. When the cointegrating rank of the system is k , an IC relation can contain at most $p-k+1$ variables. There are between k and $(p-k+1)$ of these vectors in total, the actual number depending on the degrees of over-identification of the relations of the system. This is to say that in addition to up to k identified structural relations, which, by theorem 3, are among the IC vectors, there might also be a number of solved vectors that can be defined as follows:

Definition 2. *A solved vector is a linear combination of structural vectors from which one or more common variables are eliminated by choice of offsetting weights such that the included variables are not a superset of any of the component relations.*

A solved vectors lies in the cointegrating space by construction. It may also be irreducible provided that it is a function of identified structural vectors. It is worth highlighting that

solved relations are comparable to the reduced form equations of the conventional simultaneous equations models as they are solved from the structure.⁴

At this point, a reasonable question would be: why does one need to test for irreducible cointegrating relations? And having done this, how does one distinguish between structural and solved cointegrating vectors?

The answers are contained in the following argument. It is common practice to build a presumed cointegrating regression in the light of some economic theory, the theory being considered to receive support if the hypothesis of non-cointegration is rejected. However, economic theory might suggest including some variable which is in fact not really involved in that cointegrating relation but which, interacting with the other variables, might display a coefficient which does not converge to zero. This could well provide us with a stationary relation that could indeed be a wrong one, for that, as learned from *theorem 3*, a cointegrating relation that contains redundant elements is not of interest. The theory could be wrong, in which case this is just an arbitrary element of the cointegrating space. If the theory is correct, the relation is revealed to be underidentified. The estimate is inconsistent and it represents a hybrid of different structural equations. Irreducibility is an important diagnostic property of a cointegrating regression, and testing for it allows us to determine what are the redundant variables in the system removing any unwanted effects. Once an IC relation is found, interest focuses on the problem of distinguishing between structural and solved forms. Of course, the theoretical model might answer this question for us, but this would then simply be using the theory to identify the model, and so in the absence of overidentifying restrictions we could learn nothing about the validity of the theory itself. Davidson (1998a) argues that important clues may also be provided by the data alone and we extend the argument later. To show this

⁴ Note that in standard systems of simultaneous equations the reduced forms are defined with respect to a particular normalisation which is based on the distinction between endogenous and exogenous variables, which is not relevant in the cointegrating framework.

he uses the example discussed below. Prior to this, it is necessary to mention some more of his results, which will also be useful to understand our extension.

Lemma 1. Provided β is restricted only by zero and normalisation restrictions, a solved IC relation contains at least as many variables as each of the identified structural relations from which it derives.

In general, therefore, the fewer variables an IC relation contains, and the fewer it shares with other IC relations, the better the chance that it is structural and not a solved form. In the extreme cases, we can actually draw definite conclusions, as the following pair of results show.

Theorem 4. If an IC relation contains strictly fewer variables than all those others having variables in common with it then, subject to the condition of Lemma 1, it is an overidentified structural relation.

Theorem 5. If an IC relation contains a variable, which appears in no other IC relation, it is structural.

Thus, it is possible, in the context of simultaneous cointegrating relations, to discover structural economic relationships directly from a data analysis, without the use of any theory. To understand this assume a system that consists of four $I(1)$ variables, x , y , z and w . Suppose we had tested for cointegrating rank and had found a rank of two. The second step would involve testing for cointegration on pairs of the variables. If the pairs (x, y) and (z, w) are found to be cointegrated (but not the pairs (x, z) or (y, w)), these two cointegrating relations, necessarily irreducible of course, are also necessarily structural (Davidson 1998a). Neither can have arisen as a result of solving out some more fundamental relationships. This is a case of maximal over-identification and is the framework within which Davidson's (1998a) methodology performs at its best. This result is achieved by Davidson using an algorithm implemented in Gauss called Minimal. The essence of it can be summarised as:

start with a pair of variables, test these for cointegration, and then add to the set one variable at a time until a cointegrated subset is found. This procedure is then repeated in every possible way. When the routine has terminated, every subset of the variables has been tested for cointegration unless a subset of it has been previously found cointegrated. At the end of the procedure we should obtain a list of cointegrated subsets which do not have any cointegrated subsets. Furthermore, one can rank the cointegrating vectors according to the value of the Wald statistic for the vector itself, so as to establish which IC relations are most supported by the data. All the testing can be performed indifferently using the Engle-Granger OLS method or the Likelihood-based technique due to Johansen (1988). In chapter 2 we borrow Davidson's main idea and extend it to analyse the structural relationships between the G7 interest rates as well as the underlying causal structure that links them. In particular we use an alternative version of the Davidson Methodology (1998a), which involves testing for irreducibility using the Johansen (1988) rank test on all the possible pairs of $I(1)$ series and subsequently we introduce the ranking of the irreducible cointegrating vectors according to the criterion of lowest variance in order to distinguish the structural from the solved forms. Likelihood ratio tests on the adjustment coefficients towards the IC relations will provide us with the causal structure that generates the cointegrating relations.

1.4 Weak-Exogeneity and Granger-Causality

The basis for this discussion is provided by the analysis of joint and conditional densities and sequential factorisation (see Hendry 1995 for a detailed account). Let

$$(1.4.1) \quad D_x(y_t, z_t / X_{t-1}, \theta)$$

be the sequential density at time t of the random vector $x_t = (y_t : z_t)'$ conditional on $X_{t-1} = (X_0, x_1, \dots, x_{t-1})$, where $\theta = (\theta_1, \dots, \theta_n)' \in \Theta$ which is a subset of $\mathcal{R}^n$.

Generally speaking, z_t is endogenous in the framework of the joint density function, but if z_t is weakly exogenous it is possible to factorise the joint density such that knowledge of how the process z_t is determined is not necessary in order to investigate the properties of the process y_t .

Let us allow for the existence of many one-one transformations from the original n parameters $\theta = (\theta_1, \dots, \theta_n)' \in \Theta$ to any new set of parameters $\phi \in \Phi$, and also let $\phi = (\phi_1', \phi_2')$.

We can then factorise the joint density function as:

$$(1.4.2) \quad D_x(y_t, z_t | X_{t-1}, \theta) = D_{y|z}(y_t | z_t, X_{t-1}, \phi_1) D_z(z_t | X_{t-1}, \phi_2).$$

Let the joint density under analysis involve a subset ψ of the parameters θ , where ψ is a vector of parameters of interest. The first requirement for a variable z_t to be regarded as weakly exogenous for a set of parameters of interest ψ is that the marginal process for z_t should add no useful information about ψ , that is we must be able to learn about ψ from ϕ_1 alone. The second condition we need to justify taking z_t as given is that ϕ_1 should not depend on ϕ_2 . If this were the case we could learn indirectly about ψ from ϕ_2 .

We can then say that z_t is weakly-exogenous for ψ if and only if

- ψ is function of ϕ_1 and does not depend on ϕ_2 ;
- ϕ_1 and ϕ_2 are variation free.

In a famous paper, Granger (1969) shows that given two multivariate processes $\{x\}$ and $\{y\}$, and the information on them contained in their past behaviour X_t and Y_t , $\{y\}$ causes $\{x\}$ at time t if the past of $\{y\}$ provides additional information for the forecast of x_t with respect to considering the past of $\{x\}$ alone. Indeed stating that a variable y has no role in the prediction of another variable x is tantamount to saying that the lagged values of y do not enter the equation for x , i.e. there is no feedback from y to x . Let us explain this point formally.

Let us define two multivariate processes $\{x_t\}$ and $\{y_t\}$ and associate with each process the information contained in the past behaviour of the process itself, so that we can write,

$$\begin{aligned}
(1.4.3) \quad & \bar{x}_t = (x_t, x_{t-1}, \dots) = (x_{t-i}, i \geq 0), \\
& \bar{y}_t = (y_t, y_{t-1}, \dots) = (y_{t-i}, i \geq 0), \\
& (\bar{x}_t, \bar{y}_t) = (x_{t-i}, i \geq 0, y_{t-i}, i \geq 0).
\end{aligned}$$

The information contained in (1.1) can be used for forecasting the future values of the variables. The best linear forecast of a vector x based on the information I is denoted as $E(x / I)$. The corresponding forecast error is $\varepsilon(x / I) = x - E(x / I)$. The associated mean square error can be computed from the estimated variance-covariance matrix as $\text{var}(\varepsilon(x / I))$.

Granger (1969) has suggested that

- y causes x (in mean) at time t if and only if: $E(x_t | \bar{x}_{t-1}, \bar{y}_{t-1}) \neq E(x_t | \bar{x}_{t-1})$
- y causes x instantaneously (in mean) at time t if and only if: $E(x_t | \bar{x}_{t-1}, \bar{y}_t) \neq E(x_t | \bar{x}_{t-1}, \bar{y}_{t-1})$

As a result of the properties of linear regression models, the variable forecast based on more information is necessarily the best one. Then we can always write

$$(1.4.4) \quad \text{var}(\varepsilon(x_t | \bar{x}_{t-1}, \bar{y}_{t-1})) \leq \text{var}(\varepsilon(x_t | \bar{x}_{t-1})),$$

which says that the mean square error of the forecast based on more information is always smaller or equal to the one of the forecast based on limited information. We can state this formally saying that y does not cause x at time t if and only if

$$(1.4.5) \quad \text{var}(\varepsilon(x_t | \bar{x}_{t-1}, \bar{y}_{t-1})) = \text{var}(\varepsilon(x_t | \bar{x}_{t-1})),$$

and y does not cause x instantaneously at time t if and only if

$$(1.4.6) \quad \text{var}(\varepsilon(x_t | \bar{x}_{t-1}, \bar{y}_t)) = \text{var}(\varepsilon(x_t | \bar{x}_{t-1}, \bar{y}_{t-1})),$$

that is y causes x at time t if the past of y provides additional information for the forecast of x_t with respect to considering the past of x alone. In simple terms the fact that a variable y has no role in the prediction of another variable x can be intended as the fact that lagged values of y do not enter the equation that specifies x , that is there is absence of feedback from y to x .

In the early works, the kind of approach used to test for non-causality was based on the study of bivariate models to which were applied *F-type* tests for instantaneous, unidirectional or bi-directional causality. More recently during the 90s the application of Granger non-causality tests regained popularity with the introduction of cointegration analysis as this renders the topic more fascinating. Itself the concept of cointegration, entails at least one direction of Granger-causality between the variables that gives rise to the long run relationship. Moreover, in cointegrated vector error correction models we can explicitly distinguish two sources of causation, the first through the lagged differences, the other through the error correction term (short and long run causality respectively).

Granger (1988) pointed out that the usual modelling techniques based on some form of ARMA models that does not incorporate the effect of any error correction term, would be misspecified because it would ignore some of the effect from one variable to another is ignored. This implies that when the series are cointegrated causality research based on classical procedures misses some of the forecastability and reaches incorrect conclusions about causality in mean within the variables of the model. Precisely, it is not appropriate to apply standard causality tests to non-stationary processes, because their distribution is non-standard and involves nuisance parameters (Toda and Phillips, 1993).

In the context of bivariate cointegration, Lutkepohl and Reimers (1992) present a version of the Wald test for linear restrictions on a VAR in levels using the maximum likelihood approach proposed by Johansen (1988) and Johansen and Juselius (1990). Toda and Phillips (1993) and Phillips (1995), extend the test to the general case of *p-dimensional* Fully Modified VARs also deriving the distribution of the test statistic. Toda and Yamamoto (1995) propose a simple Wald test for non-causality based on an augmented VAR approach for possibly integrated and/or cointegrated systems which is asymptotically distributed as standard χ^2 . Always in the context of VARs in levels Dolado and Lutkepohl (1996) propose a

method that yields standard χ^2 asymptotics for Wald test performed on the coefficients of cointegrated VARs with I(1) variables if at least one coefficient matrix is unrestricted under the null. Standard asymptotics also apply in the case of all matrices being restricted if one extra lag is added and we concentrate on the original parameter set.

In the context of VARs in error correction form Mosconi and Giannini (1992) have suggested imposing the constraint of cointegration under both the null and the alternative hypotheses when testing for Granger non-causality, so to achieve a gain in efficiency. The testing procedure proposed is a likelihood ratio test that, as they show, is distributed asymptotically χ^2 . The computation of this test statistic is somewhat more cumbersome than the Wald versions of non-causality test, but the limited Monte Carlo evidence available shows that it has better statistical properties (see Mosconi and Giannini, 1992 and Zapata and Rambaldi, 1997).

They consider a the following *p*-dimensional VAR:⁵

$$(1.4.7) \quad \Phi(L)Z_t = \begin{bmatrix} \Phi_{11}(L) & \Phi_{12}(L) \\ \Phi_{21}(L) & \Phi_{22}(L) \end{bmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix} = e_t \quad t, 1, \dots, T$$

where $\Phi(0) = I_p$, $e_1, \dots, e_T$ are iid $N(0, \Omega)$, and the maximum lag in $\Phi(L)$ is k . We assume that some of the roots of $|\Phi(L)|$ are equal to 1 while the others lie outside the unit circle in the complex plane. Let also y_t and x_t be vectors of variables of dimension p_1 and p_2 respectively such that $p_1 + p_2 = p$.

In order to check whether the variables in Z_t are cointegrated and y_t does not cause x_t we have to test whether $\Phi(L)$ is upper block triangular and $\Phi = \Phi(1)$ is non-zero and has reduced rank.

Reparameterise the model in an Error Correction Form as follows:

$$(1.4.8) \quad \Delta Z_t = \Gamma_1 + \dots + \Gamma_{k-1} \Delta Z_{t-k+1} + \Phi Z_{t-k} + e_t$$

or more compactly as

$$(1.4.9) \quad Z_0 = \Gamma Z_1 + \Pi Z_k + E,$$

where $\Gamma = (\Gamma_1, \dots, \Gamma_{k-1})$, $Z_0 = \Delta Z_t$, $Z_1 = (\Delta Z'_{t-1}, \dots, \Delta Z'_{t-k+1})'$, $Z = (Z_1, \dots, Z_{t-k})$.

In this framework (Mosconi and Giannini 1992), y does not Granger-cause x if the hypothesis

$$(1.4.10) \quad H_0: U'IV=0, \quad U'\Pi U\perp=0, \quad \text{and } \Pi=\alpha\beta'$$

holds, where:

$$(1.4.11) \quad U = \begin{bmatrix} 0 \\ Ip_2 \end{bmatrix} \quad U\perp = \begin{bmatrix} Ip_1 \\ 0 \end{bmatrix} \quad V = I_{k-1} \otimes U\perp,$$

and U is $(p \times p_2)$, $U\perp$ is $(p \times p_1)$, V is $(p(k-1) \times p_1(k-1))$. The alternative hypothesis is that

$$(1.4.12) \quad H_1: \Pi \neq \alpha\beta'.$$

The likelihood ratio test for the hypothesis of Granger non-causality conditional on cointegration H_0 on H_1 is given by:

$$(1.4.13) \quad LR = -2 \ln \{ L_{\max}[H_0(r)] / L_{\max}[H_1(r)] \}.$$

Mosconi and Giannini (1992) show that this is distributed χ^2 with degrees of freedom that will vary depending on the value of k in the initial system⁶.

⁵ We are omitting deterministic terms to keep the example as simple as possible, but these could be included without complications.

⁶ Mosconi and Giannini (1992) state that the hypothesis $U'\Pi U\perp=0$ holds if and only if

$$\Pi = \alpha\beta' = \begin{bmatrix} \alpha_{11} & (\alpha_{11}\beta'_{21} + \alpha_{12}\beta'_{22}) \\ 0 & \alpha_{22}\beta'_{22} \end{bmatrix} = \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix}.$$

This partition of $\Pi=\alpha\beta'$ is obtained by noting that $U'\Pi U\perp=0$ implies that

$$U'\Pi U\perp = U'\alpha\beta'U\perp = U'U\perp\alpha_{11}\beta'_1U\perp + U'\alpha_2\beta'_{22}U'U\perp = 0, \quad \text{and } \alpha_2\beta'_1 = 0 \text{ implies that}$$

$$\alpha = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ 0 & \alpha_{22} \end{bmatrix} = (U\perp\alpha_{11}|\alpha_2) \text{ and } \beta = \begin{bmatrix} \beta_{11} & 0 \\ \beta_{21} & \beta_{22} \end{bmatrix} = (\beta_1|U\beta_{22}),$$

where α_{11} is $p_1 \times r_1$, α_2 is $p \times r_2$, β_1 is $p \times r_1$, β_{22} is $p_2 \times r_2$, and $r_1 + r_2 = r$. Notice that, in this case, the rank of Π_{11} is equal to r_1 , and the rank of Π_{22} is r_2 . If we use the partition above, the model to be estimated under $H_0(r)$ or more precisely $H_0(r, r_1, r_2)$ written in compact form will be:

$$Z_0 = \Gamma Z_1 + U\perp\alpha_{11}\beta'_1Z_k + \alpha_2\beta'_{22}U'Z_k + E$$

the likelihood ratio for this hypothesis against the alternative of reduced rank of Π only, will be distributed χ^2 $(pr_1r_1 - p_2r_2 - r_1r_2)$ when $k=1$, and $(pr_1r_1 - p_2r_2 - r_1r_2 + p_1p_2(k-1))$ when $k>1$.

If we do not impose any restrictions on the Γ s, that is if we are only interested on long run causality, we can estimate the parameters α and β using the usual reduced rank regression technique as in Johansen (1988),

$$(1.4.14) \quad R_0 = \alpha\beta' R_k + e,$$

or more precisely

$$(1.4.15) \quad R_{0t} = U\perp\alpha_{11}\beta_1'R_{kt} + \alpha_2\beta_{22}'U'R_{kt} + e,$$

where R_0 are the residuals from the regression of Z_0 on Z_1 , and R_k are the residuals from the regression of Z_k on Z_1 . Finally, it is important mentioning that if we assume that in a bivariate as well as in a multivariate system the cointegrating rank is $r=1$, we can rewrite the model as

$$(1.4.16) \quad Z_0 = \Gamma Z_1 + U\perp\alpha_{11}\beta_1'Z_k + E$$

and the restrictions are only applied to the matrix α , and more specifically the presence of zeros in the rows of this matrix determine the causal structure within the model.

All these testing procedures also matter in the context of testing for weak exogeneity. We will see that, under the hypothesis of cointegration, long-run (weak) non-causality is necessary and sufficient for weak exogeneity. Let us rewrite our system in VECM form as

$$(1.4.17) \quad \Delta X_t = \Gamma_1 \Delta X_{t-1} + \dots + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi X_{t-k} + \varepsilon_t$$

where the parameters are defined as before. Assume for simplicity the absence of deterministic terms. The matrix $\Pi = \alpha\beta'$ contains information on the long-run relationships among the series in the model, with β containing the cointegrating relations and α representing the speed of adjustment to equilibrium. Also, we know that if there are $r \leq (n-1)$ cointegrating vectors in β , this implies that the last $n-r$ columns of α are zero. To test how many $r \leq (n-1)$ cointegrating vectors exist in β is equivalent to testing how many columns of α are zero.

Focusing our attention on the non-zero columns of α , let the process X_t be decomposed into

$$(1.4.18) \quad X_t = (X_{1t}', X_{2t}')' \text{ and } \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \Gamma_j = \begin{bmatrix} \Gamma_{1j} \\ \Gamma_{2j} \end{bmatrix}.$$

We can now rewrite the equations of the model as

$$(1.4.19) \quad \begin{aligned} \Delta X_{1t} &= \alpha_{1t}' \beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_{1i} \Delta X_{t-i} + \varepsilon_{1t} \\ \Delta X_{2t} &= \alpha_{2t}' \beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_{2i} \Delta X_{t-i} + \varepsilon_{2t} \end{aligned}$$

where ε_{ij} are iid $N(0, \Omega)$, and $\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$.

Let us now consider the conditional model for ΔX_1 given the past and ΔX_2 , i.e.

$$(1.4.20) \quad \Delta X_{1t} = \omega \Delta X_{2t} + (\alpha_1 - \omega \alpha_2)' \beta' X_{t-1} + \sum_{i=1}^{k-1} \tilde{\Gamma}_{1i} \Delta X_{t-i} + \tilde{\varepsilon}_{1t}$$

where $\omega = \Omega_{12} \Omega_{22}^{-1}$, $\tilde{\Gamma}_{1i} = \Gamma_{1i} - \omega \Gamma_{2i}$, and $\tilde{\varepsilon}_{1t} = \varepsilon_{1t} + \omega \varepsilon_{2t}$ with variance equal to

$\Omega_{11.2} = \Omega_{11} - \Omega_{12} \Omega_{22}^{-1} \Omega_{21}$. We will now make the following statement:

The presence of all zeros in the i -th row of the matrix α_{ij} , $j = 1, \dots, r$ indicates that the cointegrating vectors in β do not enter the equation determining ΔX_{it} . This implies that there is no loss of information from not modelling the determinants of ΔX_{it} , which can therefore enter only the right-hand side of the system since there is no feedback from the other variables in the system.

This implies long-run non-causality as well as weak exogeneity as mentioned before. We can formalise it as follows, in the context of a system in error correction form:

$$(1.4.21) \quad \begin{aligned} \Delta X_{1t} &= \omega \Delta X_{2t} + (\alpha_1 + \omega \alpha_2)' \beta' X_{t-1} + \sum_{i=1}^{k-1} \tilde{\Gamma}_{1i} \Delta X_{t-i} + \tilde{\varepsilon}_{1t} \\ \Delta X_{2t} &= \alpha_{2t}' \beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_{2i} \Delta X_{t-i} + \varepsilon_{2t} \end{aligned}$$

If $\alpha_2 = 0$ then X_{1t} is not causing X_{2t} , and X_{2t} is weakly exogenous for the parameters of interest (β, α_1) and the maximum likelihood estimator of β and α_1 can be inferred from the

conditional model alone. This can be seen rewriting the system under the hypothesis $\alpha_2=0$, that is:

$$(1.4.22) \quad \begin{aligned} \Delta X_{1t} &= \omega \Delta X_{2t} + (\alpha_1 + \omega \alpha_2) \beta' X_{t-1} + \sum_{i=1}^{k-1} \tilde{\Gamma}_{1i} \Delta X_{t-i} + \tilde{\varepsilon}_{1t} \\ \Delta X_{2t} &= \sum_{i=1}^{k-1} \Gamma_{2i} \Delta X_{t-i} + \varepsilon_{2t} \end{aligned}$$

In fact, we can see that there is no trace of α_1 and β in the marginal model and therefore *there* is no trace of X_{1t} , which is therefore not causing X_{2t} , namely, there is no feedback from the former to the latter. Also, the condition that requires the parameters of the marginal to be unrelated to the parameters of the conditional model is fulfilled as a property of multivariate Gaussian distributions that do not have joint restrictions. Note that if $\alpha_2=0$ then the $sp((0, I)')$ is contained in $sp(\alpha L)$, which means that $\sum_{i=1}^T \varepsilon_{2t}$ is a common trend in the sense that the errors in the equations for X_{2t} cumulate in the system giving rise to non-stationarity. X_2 will still be cointegrated with X_1 of course, as implied by the first equation in (1.4.22). The key point to note here is that as long as Π_{11} is of full rank then $\sum_{i=1}^T \varepsilon_{1t}$ will not be a common stochastic trend of the system and hence there will be no long-run link from X_1 to X_2 . A related proof was presented in Hall and Wickens (1993).

In chapter 3 we will examine the case when the causal structure as determined by the various techniques available breaks down. In particular, using Monte Carlo methods we examine the performance of the Johansen (1988) FIML based procedure in the event of a structural change in the causality between the variables. Furthermore we will assess the performance of recursive and rolling FIML based techniques in the detection of such structural change. We then propose a sequential test for the detection of an endogenous break in the causal structure, deriving the empirical distribution and the critical values under the null of no-break. We will also assess its performance and discuss its statistical properties. We will also compare its performance to a time varying parameter version of the Kalman Filter. In chapter 4 we apply

both the sequential test and the Kalman Filter to the G7 interest rates with interesting results.

In chapter 5 we will draw some final conclusions.

Chapter 2

Identification of Structural Cointegrating Vectors: An application of the Extended Davidson Methodology to the G7 Interest Rates

2.1 Introduction

The uncovered interest parity condition (or the open arbitrage condition) suggests that in general international interest rate differentials should equal the (expected) change in exchange rates. Almost all empirical studies have found that the G7 exchange rates are at most $I(1)$ series. If we then make the reasonable assumption that any risk premium, which may exist, in the relationship is stationary (or that there exists a stationary relation between the G7 risk premia), the implication of these theories is that interest rates should be cointegrated on a bilateral basis. In itself therefore cointegration between interest rates is neither surprising nor particularly informative. However if these interest rates are cointegrated then there must exist a causal structure, which gives rise to cointegration and is of great policy interest.

Much of the empirical evidence on interest rate linkages is based on causality test statistics, even though interest rates are typically $I(1)$ and hence the tests do not follow standard distributions. So the inference is often invalid. Recent work using an appropriate testing procedure put forward by Toda and Yamamoto (1995) shows that in fact, at least in the case of long rates, interest rate movements are determined mainly by domestic policy objectives or country-related factors (see Caporale and Williams, 1998a, 1998c). International linkages in short-term rates appear to be much stronger, although the direction of causality is not always

consistent with common priors about the functioning of the international financial system and of the ERM (see Caporale and Williams, 1998b, 1998c).

In this chapter we examine interest rate linkages in the G7 as structural relations, using the Extended Davidson's Methodology (EDM), which involves testing for irreducibility of cointegrating relations and their ranking according to the criterion of minimum variance. The interesting feature of this method is that it may allow us to learn about the structural relationships linking cointegrated series from the data only, without imposing any arbitrary identifying conditions. Our analysis will also include Likelihood Ratio (LR) tests for exogeneity restrictions on the adjustment coefficients towards irreducible cointegrating (IC) relations to shed some light on the causal structure as resulting from the impact of the creation of an integrated capital market in Europe. By suppressing exchange rate risk within the area and by fostering harmonisation measures, EMU will have an impact on asset prices and monetary and fiscal policy, which in turn will affect investment, real activity, capital flows and hence global interest rate linkages (see Portes and Rey, 1998).

It is worth highlighting that, even in a system like the ERM which aims to produce policy co-ordination it has been possible for monetary authorities to disengage their long-term interest rate policy from developments elsewhere and pursue an independent policy agenda over long periods (possibly implying lack of cointegration between some of the series). Such an option should remain available for non-participating countries, like the UK, after the establishment of the Euro. Therefore the UK authorities will not necessarily find their freedom of action greatly constrained by what is happening in the Euro zone. Within the Euro zone the policies of the European Central Bank (ECB) will not necessarily be as stable or credible as those adopted so far by the German authorities, since smaller countries will also have an influence on monetary policy (see Begg et al, 1998). If in fact Germany has not been able to impose its interest rate policy on the other ERM countries, and if this becomes true of fiscal policy as

well (notwithstanding the Growth and Stability Pact), long-term rates might rise (rather than decline) in the EU following its inception.

The layout of the chapter is as follows. Section 2 analyses interest rate linkages. The data-set and the methodology are described in Section 3. The empirical results are presented in Section 4. Some concluding remarks are offered in Section 5.

2.2 Analysing Interest Rate Linkages

In broad terms one can identify two views on how interest rates may be linked. If they are treated as analogous to other asset prices, then their movements are naturally interpreted as being determined by financial flows in fluid, profit-seeking capital markets. Alternatively, they can be viewed as policy instruments, so that their time paths may be determined by a policy objective such as an exchange rate parity or an inflation target. Interest rate linkages have therefore often been analysed in the context of a specific policy framework such as the Exchange Rate Mechanism (ERM). For instance, numerous studies have attempted to test the so called “German Leadership Hypothesis” (GLH), according to which Germany acts as the dominant player within the system, and monetary authorities in other ERM countries are unable to deviate from the course set by the Bundesbank (see, e.g., Fratianni and Von Hagen, 1990).

Taking this view, co-movement in interest rates arises because of policy convergence. Early studies had concluded that there is no cointegration between German rates and other EMS rates (see Karfakis and Moschos, 1990), and that there is stronger evidence of cointegration between US rates and EMS rates (see Katsimbris and Miller, 1993). Subsequent papers reported convergence in European rates after 1986 (see Caporale et al, 1996). Similar conclusions were reached by Hall et al (1992) using time-varying techniques. In a global context, Caporale and Williams (1998b) found a marked difference between linkages in long-

term rates (10-year bond yields) and in short-term rates (3-month Treasury bills) in the G7 economies. Whilst there is little evidence that the former have been linked to one another over the last two decades, for the latter the evidence of co-movement is more compelling. Furthermore, the causality structure is not consistent with the standard characterisation of the ERM as an asymmetric system in which Germany was the dominant player - it suggests instead that there was German accommodation of French monetary policy within the ERM. This result could be interpreted in the context of the “size effects” identified in recent theoretical research, according to which larger, more stable countries can achieve policy objectives more successfully via accommodation rather than by compulsion (see Martin, 1997). The system was actually more flexible than normally recognised, as there were various “escape clauses” built into it (for instance, the options of exchange rate realignments, wider fluctuation bands, and capital controls). In this study our objective is to identify the fundamental relationships linking interest rates among the G7, and to analyse their causal structure in order to test for hypotheses such as the GLH.

2.3 The data-set and the methodology

The sample under investigation covers the period between 1977:1-1998:3. The source for the data is International Financial Statistics of the IMF. The short-term interest rates we employed are the 3-month Treasury bills, the only exception being Japan, for which only the discount rate was available for the whole period under investigation. The long-term interest rates used are the 10-year government bond rates for all the G7.

Our analysis is performed using the Extended Davidson Methodology, which relies on the concept of *irreducible cointegrating* (IC) relation (Davidson 1998), one from which no variable can be omitted without loss of the cointegration property. This means that, for the purpose of identifying the structure, cointegrating vectors with redundant variables are not

useful. Not all the IC vectors, though, are structural. Some of them are *solved vectors*, namely linear combinations of structural vectors. Therefore one should first perform cointegration tests in order to eliminate all non-cointegrated sets and cointegrated supersets, and then concentrate on the cointegrated sets, which yield IC relations. The original Davidson's (1998b) methodology tries to avoid just identifying restrictions from economic theory by searching for subsets of cointegrated variables (due to zero restrictions) which cannot be further reduced without losing the cointegration property. This irreducible set is found in a bottom-up approach by adding (and dropping) variables until cointegration is satisfied, after a certain rank has been chosen using a Johansen (1988) rank test. In practical terms, one is analysing all possible cointegrated relations and testing exclusion restrictions by means of suitably constructed Wald tests, which can be shown to follow standard distributions (see Davidson, 1998b). Furthermore, one can rank the cointegrating vectors according to the value of the Wald statistic for the vector itself, so as to establish which IC relations are most supported by the data and are therefore structural.

The EDM differs from Davidson's original one in two elements. The first difference consists in that after testing for cointegration for the whole group of series, we perform a rank test in each step of the bottom up strategy without necessarily starting from the cointegrating rank indicated by the rank test on the complete set of variables. The advantage of this amounts to the possibility to extend the IC methodology to exactly identified systems. These might have been suggested by economic theory, and therefore our technique allows us to test for the validity of the theory itself (in the case under investigation). The second difference consists in the fact that we use the criterion of the minimum variance rather than the value of the Wald test statistic to rank the IC vectors and distinguish between structural and solved vectors.

To introduce our extension of Davidson's procedure, we will analyse the case of bivariate cointegration as this gives rise to the largest number of IC vectors and solved cointegrating

relationships for any number of variables. In addition, this is also the most relevant case for our later application of the methodology to the G7 interest rates.

Consider an N-dimensional cointegrating system as analysed by Johansen (1988,1991),

$$(2.3.1) \quad \Pi(L)X_t = \alpha\beta'X_t + \Pi^*(L)\Delta X_t = u_t,$$

where L is the lag operator, $\Pi(L) = \alpha\beta' + \Pi^*(L)(1 - L)$, such that $\Pi(1) = \alpha\beta'$, where α and β are of dimension $N \times R$ and are respectively the loading matrix and the matrix of cointegrating vectors. When $R < N$, it can be shown that $X_t \sim I(1)$ and the system incorporates a set of long run equilibrium relations of the form $\beta'X_t = u_t$.

In general, in the case of bivariate cointegration between each pair of variables in a set of N variables there will be R structural IC vectors where R is N-1, and there will exist K irreducible vectors, which are simply combinations of the R structural ones where K is $((R-1)^2 + (R-1))/2$. Now, if we designate the first R cointegrating residuals as the structural ones, so that for $u_1 \dots u_R \sim NI(0, \sigma^2_1 \dots \sigma^2_R)$, then clearly the solved cointegrating residuals will be combinations of these⁷. Because for any IC vector the variance may vary with the normalisation, we need two starting assumptions for our ranking criterion to be operative.

This first assumption involves the possibility of normalising the IC vectors using the normalisation, which yields a minimum variance. What we mean is that, in formal terms, given two cointegrated series x_{1t} and x_{2t} we choose to normalise their cointegrating relation as

$$(2.3.2) \quad x_{1t} = \delta x_{2t} + u_{1t}, \quad u_{1t} \sim (\mu_1, \sigma_1^2)$$

rather than as

$$(2.3.3) \quad x_{2t} = \rho x_{1t} + u_{2t} \quad u_{2t} \sim (\mu_2, \sigma_2^2),$$

if $\sigma_1^2 \leq \sigma_2^2$.

⁷ Yet we will see that the set of K solved residuals need not all display a variance that is strictly greater than all of the R structural residuals.

A second assumption is that, the residuals from the structural cointegrating relations are normally independent distributed with full rank diagonal covariance matrix Σ^8 (Sims 1980 formulates a similar assumption on error terms of the unobservable structural VAR⁹, see also Hendry 1995 pp.784 and 807). Given these two starting assumptions, we formulate the following statement.

The minimum variance normalised structural residuals will have a strictly lower variance than any solved residual coming from an IC vector containing the same variable.

This proposition provides us with the immediate rule for distinguish structural from solved irreducible vectors. We can prove it as follows.

Let

$$(2.3.4) \quad \beta' X_t = u_t, \beta (rxn) \text{ and } X_t (nx1)$$

be the structural cointegrating relations with β normalised such that $u_t \sim iN(0, \Omega)$. Here we assume that β has been normalised on the j^{th} element of its columns such that the cointegrating relations have variances $\sigma_{i,j}^2 < \sigma_{i,k}^2$, where σ_i^2 are the elements on the main

⁸ However, given that the solved vectors are linear combinations of the structural ones it is likely that their cointegrating errors will be correlated with one or both the residuals of the solving relations.

⁹ Sims' 1980 argument can be summarised as: The structural model is not directly observable, however a VAR can be estimated as the reduced form of the underlying structural model

$$\begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = A^{-1}C(L) \begin{pmatrix} X_{1t-1} \\ X_{2t-1} \end{pmatrix} + \begin{pmatrix} e_t^{x_1} \\ e_t^{x_2} \end{pmatrix},$$

where e denotes the VAR residual vector, normally independent distributed with full variance-covariance matrix Ω . The relation between the residuals in e and the structural disturbances in u is,

$$A \begin{pmatrix} e_t^{x_1} \\ e_t^{x_2} \end{pmatrix} = B \begin{pmatrix} u_t^{x_1} \\ u_t^{x_2} \end{pmatrix}, \text{ so that } e_t = A^{-1}Bu_t, \text{ and } E(e_t e_t') = A^{-1}BE(u_t u_t')B' A^{-1}. \text{ Thus, } \Omega \text{ contains}$$

$(n^2+n)/2$ different elements, that indicate the maximum number of identifiable parameters in the matrices A and B . In his seminal paper, Sims proposed the following identification strategy based on the Choleski decomposition of matrices:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a_{21} & 1 & 0 & 0 \\ . & & 1 & 0 \\ a_{n1} & a_{nn-1} & 1 \end{bmatrix}, B = \begin{bmatrix} b_{11} & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 \\ 0 & 0 & b_{33} & 0 \\ 0 & 0 & 0 & b_{nn} \end{bmatrix}.$$

diagonal of Ω , and $j \neq k$ indicate different normalisations. Any solved vector will be given by a combination of the structural $\beta' X_t = u_t$ of the form

$$(2.3.5) \quad A\beta' X_t = Au_t,$$

where A is $(1 \times r)$ normalised on the corresponding element j .

It follows that $Au_t \sim N(0, A\Omega A')$ and more explicitly $A\Omega A' = \sum_i a_i^2 \sigma_i^2$. Now recall that as at

least one of the terms of this summation is given by $a_j^2 \sigma_j^2$, and $a_j = 1$, we can state that

$$(2.3.6) \quad A\Omega A' = \sum_i a_i^2 \sigma_i^2 > \sigma_{i,j}^2. \text{ QED.}$$

As an illustration, consider the case when $N=4$, such that $X_t = \{x_{1t}, x_{2t}, x_{3t}, x_{4t}\}$. Assume that, in testing for irreducibility, we have found that cointegration holds between all the possible pairs of series. In this case the system has a rank of $(N-1)=3$ and we have a collection of 6 irreducible cointegrating relations, of which 3 are structural and 3 are just solved relations. Suppose that the minimum variance normalisation delivers us the following irreducible structural relations

$$(2.3.7) \quad x_{1t} = \delta x_{2t} + u_{1t}, \quad u_{1t} \sim N(0, \sigma_1^2)$$

$$(2.3.8) \quad x_{2t} = \gamma x_{3t} + u_{2t}, \quad u_{2t} \sim N(0, \sigma_2^2)$$

$$(2.3.9) \quad x_{3t} = \varphi x_{4t} + u_{3t}, \quad u_{3t} \sim N(\mu_3, \sigma_3^2).$$

The three solved vectors will be linking respectively x_{1t} to x_{3t} , x_{1t} to x_{4t} , and x_{2t} to x_{4t} . They will look like

$$(2.3.10) \quad x_{1t} = \theta x_{3t} + u_{4t} \quad \text{with } \theta = \delta\gamma, \text{ and } u_{4t} = u_{1t} + \delta u_{2t},$$

$$(2.3.11) \quad x_{1t} = \lambda x_{4t} + u_{5t}, \quad \text{with } \lambda = \delta\gamma\varphi, \text{ and } u_{5t} = u_{1t} + \delta u_{2t} + \delta\gamma u_{3t}, \text{ and}$$

$$(2.3.12) \quad x_{2t} = \psi x_{4t} + u_{6t}, \quad \text{with } \psi = \gamma\varphi, \text{ and } u_{6t} = u_{2t} + \gamma u_{3t}.$$

Basic statistics tells us that u_{4t} will be distributed as

$$(2.3.13) \quad u_{4t} \sim N(0, [\sigma_{1t}^2 + \delta^2 \sigma_{2t}^2 + 2\delta \text{cov}(u_{1t}, u_{2t})]).$$

And assuming that u_{1t} and u_{2t} are uncorrelated

$$(2.3.14) \quad u_{4t} \sim N(0, \sigma_{1t}^2 + \delta^2 \sigma_{2t}^2)$$

In the same way

$$(2.3.15) \quad u_{5t} \sim N(0, \sigma_{1t}^2 + \delta^2 \sigma_{2t}^2 + \delta^2 \gamma^2 \sigma_{3t}^2), \text{ and}$$

$$(2.3.16) \quad u_{6t} \sim N(0, \sigma_{2t}^2 + \gamma^2 \sigma_{3t}^2) \text{ respectively.}$$

We can see that the magnitude of the variances of the solved vector depend on the values of the cointegrating parameters δ, γ , and ϕ , and it seems somewhat difficult to distinguish whether for example $\sigma_1^2 < (\sigma_{2t}^2 + \gamma^2 \sigma_{3t}^2)$ without any prior knowledge about γ . This does not mean that we cannot use the criterion of the minimum variances to detect the structural residuals. This is achieved by carrying out a more complex comparison, which involves comparing the variances of the irreducible cointegrating relations not in absolute terms (simply with each other), but comparing the variances of vectors relative to the same variable. Table 2.1 shows a four variable case, such as the example presented above where the structural bivariate relationships are between x_1 and x_2 , x_2 and x_3 and x_3 and x_4 . The variances reported are in fact the ones of $u_1, \dots, u_6$. There we can see that the structural relationship between x_1 and x_2 always has the smallest variance in the row corresponding to vectors normalised for x_1 . In fact we can observe that σ_1^2 must be strictly smaller than both $\sigma_{1t}^2 + \delta^2 \sigma_{2t}^2$, and $\sigma_{1t}^2 + \delta^2 \sigma_{2t}^2 + \delta^2 \gamma^2 \sigma_{3t}^2$ regardless of the values or the signs of the cointegrating parameters δ and γ . The same applies to the structural relationship between x_2 and x_3 as compared to the one linking x_2 and x_4 , and so on.

It is interesting to notice that irreducibility might also be useful to simplify the analysis of causal linkages between variables. This is because by ruling out non cointegrated series we would be working with smaller systems with a gain in efficiency. Therefore, in order to shed

some light on the causal structure of the cointegrating system we perform likelihood ratio tests on the adjustment coefficients towards the irreducibly cointegrated relations.

Below we apply the proposed methodology to the G-7 short-run interest rates with interesting results.

2.4 Empirical results

We begin the analysis by pre-testing for the order of integration of the series using standard Augmented Dickey-Fuller (ADF) tests. The number of lagged differences included in the test is decided on the basis of a criterion advised by Doornik and Hendry (1997), so as to ensure non-autocorrelated residuals on the auxiliary regressions. In each case the tests deliver the expected result that the series are all integrated of order one $I(1)$, and hence follow stochastic trends. The results are shown in tables 2.2 and 2.3 for short- and long-term rates respectively.

Having obtained confirmation that all interest rates are integrated of order one we proceed by running cointegration tests for the complete set of G7 long-term interest rates. For this and subsequent analysis we have used Johansen's (1988, 1991) likelihood based cointegration tests. As suggested in Hall (1991) and Caporale *et al* (1997), in performing the rank tests we have specified the unrestricted VAR to include as many lags as necessary to ensure non-autocorrelation in the residuals, as well as impulse dummies to correct for non-normality or heteroscedasticity of the disturbances. Notice that for clarity of exposition we have preferred to discuss short and long term rates separately.

2.4.a Short-term rates: the empirical evidence

We start by performing cointegration tests on the complete G7 set of interest rates obtaining the results displayed in table 2.4. The test statistics indicate that the cointegrating rank of the

system is four. Of course these tests only allow us to reject the hypothesis that there are less than four cointegrating vectors - they do not necessarily mean that there is not more. So, in order to learn something more about the structure of the linkages among these series, we perform cointegration tests on each pair of series to investigate whether cointegration holds among all of the series of the group. The aim is to establish the number of irreducible cointegrated relations and what series are involved in them. Of course if we find that pairwise cointegration holds between each pair of rates this tells us that the rank of the whole seven variable system is in fact 6. The conflict between the two test procedures is then seen as simply one of the small sample power and size of the tests in different contexts.

The results for pairwise cointegration tests are presented in table 2.5. By looking at the results we get confirmation of what we suspected about the possible true rank of the system. The main result is that cointegration holds among every pair of series and with a unit elasticity in all cases but in two, namely the relationship Japan-Canada and Japan-France. The likely reason is the fact that in the case of Japan we used the discount rate, which appears to behave differently and to change much more slowly compared to other rates, therefore producing different results.

Cointegration is clearly a property of the series and this is confirmed by the existence of nineteen irreducible cointegrating relations with unit coefficients out of twenty-one cointegrating tests. Of course, given such widespread cointegration it is almost certain that in fact all pairwise interest rates cointegrate. For example Italy and Japan clearly cointegrate, as do Italy and Germany, but the test between Germany and Japan is less significant. However, as the first two cointegrate, Germany and Japan must also cointegrate.

The second part of our analysis involves the ranking of the cointegrating vectors according to the criterion of minimum variability and the discussion of exogeneity issues which will help to clarify the relationships among the series (table 2.6). The first two interesting results are

that two of the minimum variance vectors link the US to Canada and Italy to France, both of which can be considered irreducible structural relationships. It then becomes rather hard to interpret the ranking when presented in the form of table 2.6. We therefore collate the information from table 2.6 in the form given in table 2.7.

This table now presents a set of results, which are relatively easy to interpret. We have highlighted in bold the relationships which seem to be structural on a column by column interpretation. Clearly, on a minimum standard deviation criteria the US and Canada is a structural relationship, as is Germany and Japan, France and Italy, UK and Japan, Germany and Japan, and US and Japan.

In the simple ranking of standard deviations given in table 2.6 these all appear in the lowest 7 places; we can exclude the relation between Canada and France, which is in position 6, from being structural as in the French column France and Japan have a lower standard deviation. The LR tests on the alphas reported in Table 2.6 add a little to this understanding. In the Japan Germany relationship, Japan is exogenous, and in the France-Italy relationship, France is exogenous. Perhaps most significantly, the US is found to be exogenous with respect to Japan. This suggests that Italy was largely following French rates and Germany was following Japanese rates, while Japan followed the US. All the other relationships seem to be bi-directional.

Overall we have a picture of the US and Canada being a clear block, then most of the other countries being linked primarily through Japanese rates (these are not of course acting in a dominant way as no exogeneity can be established except with respect to Germany). Italy was clearly following France. The UK responded more to non-European rates than to the other European countries, and, surprisingly, France and Germany seem to be responding more to world rates than to each other.

These results seem to confirm previous studies such as Caporale and Williams (1998b), which rejected the so-called German Leadership Hypothesis (GLH), since Germany does not appear to be driving the European financial markets.

This is suggested by exogeneity of US and Japanese rates in the pairwise cointegration tests with the German one. Also, in the tests with other European countries, Germany does not appear to be exogenous in any but one of the cases. It does not seem that a leader exists in Europe. This conclusion is justified by the fact that we find feedback effects in causality among most of the series involved and in the other cases we find a sort of “*circular*” causality among the series, indicating that there is not a country that represents the leader.

2.4.b Long term rates: the empirical evidence

As in the case of short-term rates we start by performing cointegration tests on the complete set of G7 long-term interest rates obtaining the results displayed in table 2.8.

The results indicate that the cointegrating rank of the system is two. Again these tests only allow us to reject the hypothesis that there are less than two cointegrating vectors. They do not necessarily mean that there are not more. So, in order to investigate further the linkages among these series, we perform cointegration tests on each pair of series. What we want to establish is the number of irreducible cointegrated relations and which are the series involved in them.

The results for pairwise cointegration tests are presented in table 2.9 and they can be summarised in three points:

- The true rank of the system is actually four.
- Direct cointegration holds among every pair of series (and with unit elasticity in all cases) but not between Italy and Germany. Therefore,

- We can rule out the possibility that Italy and Germany are involved in any of the structural relations.

In the light of the findings reported in our earlier study on short-term rates, we would expect convergence of Italian and German rates, and therefore the fact that these rates do not cointegrate with each other and consequently with the others comes as a surprise. Therefore, we investigate the issue further.

It is common to model the expected long-term rates as some weighted average of short-term rates (Expectation Hypothesis)¹⁰ plus a country related risk premium. Recall that we had assumed earlier that either risk premia in the G7 group were stationary, or there existed a stationary relation between them. More formally we can define long-term interest rates as (see Cuthbertson, 1996):

$$2.4.1 \quad (1 + LR)^T = \prod_{t=1}^T (1 + rw_{t+1} + \lambda)$$

where LR are long-term rates, r refers to short-term rates, w are weights attached to each of the rs and λ is the term that we called risk premium, which is made up of country-specific factors. If we make the reasonable assumption that expectations on future inflation enter the risk premium term, having found that cointegration holds between the short rates of Italy and Germany, we can test for cointegration between the inflation rates of Italy and Germany to investigate the cause of non-convergence in long-term rates. The results displayed in table 2.10 indicate full rank of the system and therefore stationarity of the two inflation rates. This result is clearly unacceptable because we had already performed ADF-tests on the two series obtaining results that indicate non-stationarity of the two series as displayed in table 2.11.

We can then test the stationarity of the two cointegrating vectors indicated by the rank test to check whether either of them is indeed stationary. The answer is definitely no. The ADF-tests

¹⁰ For an exhaustive account of the Expectation Hypothesis see Cuthbertson (1996), "Quantitative Financial Economics" chapters 9 and 10, J.Wiley & Sons.

(see Table 2.12) cannot reject the hypothesis of a unit root in both cointegrating vectors indicating that the inflation rates of Italy and Germany are not cointegrated and can actually play a role in explaining non-convergence of Italian and German long-term interest rates.

Moving on to the second part of our investigation, we next rank the irreducible cointegrating vectors indicated from the pairwise tests previously carried out to obtain table 2.13. Some comments are in order. First, notice that cointegration with homogeneous coefficients has been imposed between the non-cointegrated series. Therefore, the standard deviation has been calculated for these cointegrating vectors and they have been ranked accordingly. Second, the LR test for exogeneity restrictions on the vectors involving Italy and Germany have been omitted for the reason that they are not very meaningful, as they are performed on non-cointegrated series under the hypothesis of cointegration.

We find that in general, not surprisingly, non-cointegrated series display a higher variability than cointegrated ones, with the exception of the relations between German rates, and Japanese and UK rates respectively. This is not something to worry about because the ranking in table 2.13 is made in terms of the absolute magnitude of the standard deviations. As for short rates, to make this point clear and to interpret more easily the ranking of the IC vectors it is convenient to use a different format for the results and to display them as in table 2.14.

Given a rank of four, on a minimum standard deviation criterion, US and Canada is a structural relationship and so are UK and Canada and UK and Japan. The fourth involves France and Canada again. If we collate this information with the one on the "exogeneity restrictions" column in table 2.13, we obtain the overall picture. On the whole, we can conclude that the evidence supports the hypothesis that larger and more stable economies can achieve policy objectives more successfully by accommodating rather than driving other countries' policies (see Martin, 1997). This would explain the endogeneity of the US long

rates in the IC relations with all the other countries. The Canadian rate can be identified as the driving force. This is a reasonable result given the strong linkage between Canadian and US markets. Notice that the French rate follows the Canadian one in another of the fundamental relationships. Interestingly, Canadian long rates accommodate the UK ones according to one of the IC relations, and Japan and UK are linked by a relation with a causal feedback.

2.5 Summary and Conclusions

In this chapter we have used the extended Davidson Methodology to examine the structural linkages that exist between the G7 short- and long-term interest rates. Application of this method has shown the importance of testing for irreducibility as a diagnostic. The criterion of minimum variance has allowed us to distinguish between structural and solved IC vectors without any prior theoretical assumptions. Furthermore, we have performed Likelihood Ratio tests on the loading factors to all IC relations in order to gather information on the causal structure that links the rates.

As far as the short rates are concerned, the results can be summarised as follows. First, the rank of the system appears to be 6 when the EDM approach is taken, compared to 4, which is the result obtained performing a rank test *a la* Johansen (1988) on the full system. Second, we have been able to isolate six irreducible structural relations, of which the two most significant ones involve the US and Canada, and Italy and France. Third, exogeneity tests seem to indicate a US world-wide leadership and reject the hypothesis of a German leadership in Europe, therefore confirming the findings of Caporale and Williams (1998b) and of other authors (see, e.g., Katsimbris and Miller, 1993). In brief, US and Canada appear to constitute the fundamental block, UK rates respond more to non-European rates than to other European countries, Italy is clearly following France, and France and Germany respond to world rates

rather than to each other. Last, it is important to mention the role of Japan that acts as the link between European rates and the US one.

Going to the results concerning the long rates, we again need to highlight the usefulness of the EDM approach. We have in fact obtained a rank of four for the system of long-term rates compared to a rank of two as indicated by the rank test on the whole group of series. More crucially we have been able to rule out the Italian and German rates as they appear to be non-cointegrated with the other series.

We have been able to isolate four irreducible structural relations that link the US and Canada, and the latter with the UK, France and Japan respectively. It appears that the driving force is Canada, which is linked to the US, UK and France in three out of the four fundamental relations, and which seems to be a reference point for the US that behave in an accommodating fashion. Notice that long-term rates in Italy and Germany, being non-cointegrated, seem to be determined by factors that are specific to these two countries. To investigate this matter further, we have analysed their expected inflation rates that are likely to affect the respective risk premia. The results obtained provide evidence that non-convergence of Italian and German rates is justifiable on the basis of non-convergence of their risk premia, and therefore interest rates in these two countries are independent from every other country's rates.

These results on the structural relation and the consequential causal structure derived are obtained and discussed on the basis of the belief in constancy of the parameters of both cointegrating regression and adjustment coefficients. This might not always be the case. It is possible, for example, that the lack of cointegration between Italian and German long-term rates with the other ones is due to a structural break in the cointegrating parameters due to differences in long term policies in the two countries with respect to each other and the other G-5. It also possible that the causal structure as derived from EDM technique is not constant

overtime. It would be very interesting to shed some light on both these issues, but we prefer to confine the focus of this study on the identification of causal structures of cointegrated systems. Moreover the results obtained for the short-term rates are somewhat more interesting than the ones obtained for the long term ones. Therefore, in the remainder of this thesis, we are going to concentrate on the analysis of the linkages^{*} between short-term rates, in particular on the changes in their causal structure.

Table 2.1. The relationship of variances of cointegrating errors between structural and solved vectors				
	x_1	x_2	x_3	x_4
x_1	----	σ^2_1	$\sigma^2_{1t} + \delta^2 \sigma^2_{2t}$	$\sigma^2_{1t} + \delta^2 \sigma^2_{2t} + \delta^2 \gamma^2 \sigma^2_{3t}$
x_2	----	----	σ^2_2	$\sigma^2_{2t} + \gamma^2 \sigma^2_{3t}$
x_3	----	----	----	σ^2_3
x_4	----	----	----	----

Table 2.2. Unit root test on G7 short-term rates			
	ADF test value	Critical value 95%	Critical value 99%
Canada	-1.852	-2.89	-3.51
France	-1.383	-2.89	-3.51
Italy	-0.829	-2.89	-3.51
USA	-2.202	-2.89	-3.51
UK	-2.340	-2.89	-3.51
Germany	-2.749	-2.89	-3.51
Japan	-1.465	-2.89	-3.51

Table 2.3. Unit root test on G7 long-term rates			
	ADF test value	Critical value 95%	Critical value 99%
Canada	-0.990	-2.89	-3.51
France	-0.938	-2.89	-3.51
Italy	-1.010	-2.89	-3.51
USA	-1.338	-2.89	-3.51
UK	-0.442	-2.89	-3.51
Germany	-2.073	-2.89	-3.51
Japan	-0.720	-2.89	-3.51

Table 2.4. G7 short term rates cointegration test						
Eigenvalue	Log-likelihood	for rank				
	481.013	0				
0.7557	540.215	1				
0.5184	570.903	2				
0.3961	592.086	3				
0.2975	606.915	4				
0.1550	613.99	5				
0.0877	617.845	6				
0.0370	619.428	7				
Ho:rank=p	$-T \log(1 - \hat{\lambda}_p)$	(T-nm) [°] 95% cv		$-T \sum_{i=p+1}^n \log(1 - \hat{\lambda}_i)$	(T-nm) [°] 95%cv	
P == 0	118.4**	88.8**	45.3	276.8**	207.6**	124.2
P <= 1	61.38**	46.03**	39.4	158.4**	118.8**	94.2
P <= 2	42.37**	31.77	33.5	97.05**	72.79*	68.5
P <= 3	29.66*	22.24	27.1	54.68**	41.01	47.2
P <= 4	14.15	10.61	21	25.03	18.77	29.7
P <= 5	7.71	5.783	14.1	10.87	8.156	15.4
P <= 6	3.164	2.373	3.8	3.164	2.373	3.8
*indicates rejection of the null at the 95% level						
** indicates rejection of the null at the 99% level						
° -(T-nm) is a small sample correction replacing (-T) in the λ -max and λ -trace statistics						

Table 2.5 Pairwise cointegration tests for G7 short-term rates						
	$-T \log(1 - \hat{\lambda}_p)$	$-(T-nm)^\circ$	95% cv	$-T \sum_{i=p+1}^n \log(1 - \hat{\lambda}_i)$	$-(T-nm)^\circ$	95% cv
USA-Canada	14.23*	9.678	14.1	16.58*	11.27	15.4
USA-Japan	22.46**	14.57*	14.1	23.68**	15.36	15.4
USA-UK	14.71*	9.936	14.1	17.31*	11.7	15.4
USA-France	14.48*	11.92	14.1	15.48*	12.74	15.4
USA-Italy	16*	13.16	14.1	19*	15.63*	15.4
USA-Germany	24.45**	20.17**	14.1	24.45**	20.17**	15.4
Canada-France	15.52*	11.8	14.1	17.94*	13.64	15.4
Canada-UK	14.72*	12.15	14.1	17.39*	14.35	15.4
Canada-Italy	21.09**	18.52**	14.1	21.34**	18.74*	15.4
Canada-Japan	24.71**	20.39**	14.1	24.76**	20.43**	15.4
Canada-Germany	17.47*	13.93	14.1	17.47*	13.93	15.4
UK-France	15.95*	12.55	14.1	19.46*	15.31	15.4
UK-Germany	30.99**	23.74**	14.1	31.24**	23.93**	15.4
UK-Italy	17.8*	14.48	14.1	19.71*	16.03	15.4
UK-Japan	19.9**	15.31*	14.1	24.62**	18.94*	15.4
Germany-France	16.31*	14.32*	14.1	18.43*	16.18*	15.4
Germany-Italy	22.21**	18.7*	14.1	25.75**	21.68*	15.4
Germany-Japan	14.8*	13	14.1	17.03*	14.95	15.4
Italy-Japan	27.96**	25.96**	14.1	29.79**	27.67**	15.4
Italy-France	24.09**	21.15*	14.1	30.62**	26.89**	15.4
Japan-France	27.46**	21.9**	14.1	28.91**	23.05**	15.4
* indicates rejection of the null of no cointegration at 95% level						
** indicates rejection of the null of no cointegration at 99% level						
° $-(T-nm)$ is a small sample correction replacing $(-T)$ in the λ -max and λ -trace statistics						

Table 2.6. Short-term rates, ranking of IC vectors

IC vectors	eigenvalue	standard deviation	exogeneity restrictions
Italy-France	0.254565	1.173084	France exogenous
Germany-Japan	0.165175	1.571049	Japan exogenous
USA-Canada	0.172846	1.782502	Feedback
UK-Japan	0.225159	1.883635	Feedback
Japan-France*	0.293614	1.967761	Feedback
Canada-France	0.186971	2.108287	Canada exogenous
USA-Japan	0.261784	2.126364	USA exogenous
Canada-UK	0.168093	2.138164	Feedback
Germany-France	0.18824	2.21158	Feedback
USA-UK	0.180218	2.359356	USA exogenous
USA-France	0.167506	2.405468	France exogenous
UK-Germany	0.307000	2.467542	Feedback
Canada-Japan*	0.265748	2.496532	Feedback
Canada-Italy	0.226752	2.497663	Feedback
Italy-Japan	0.283097	2.540980	Feedback
USA-Germany	0.263376	2.545527	USA exogenous
USA-Italy	0.183315	2.647584	Italy exogenous
UK-France	0.191583	2.674037	UK exogenous
Canada-Germany	0.198368	2.772337	Canada exogenous
Germany-Italy	0.253395	2.820897	Germany exogenous
UK-Italy	0.211309	3.190997	Feedback

* indicates cointegration with non-homogeneous coefficients

Table 2.7 Cross tabulation of Standard deviations for the G7 short rates.							
	USA	Canada	Japan	Germany	France	Italy	UK
USA	-	1.78	2.1	2.5	2.4	2.6	2.3
Canada	1.78	-	2.5	2.7	2.1	2.5	2.1
Japan	2.1	2.5	-	1.57	1.96	2.5	1.88
Germany	2.5	2.7	1.57	-	2.2	2.8	2.4
France	2.4	2.1	1.96	2.2	-	1.17	2.6
Italy	2.6	2.5	2.5	2.8	1.17	-	3.2
UK	2.3	2.1	1.88	2.4	2.6	3.2	-
Structural IC relations are in highlighted in bold							

Table 2.8. G7 long-term rates cointegration test						
Eigenvalue	Log-likelihood	for rank				
	648.865	0				
0.681239	695.169	1				
0.469721	720.860	2				
0.297578	735.166	3				
0.200772	744.242	4				
0.118348	749.344	5				
0.0792845	752.689	6				
0.0312077	753.973	7				
Ho:rank=p	$-T \log(1 - \hat{\lambda}_p)$	(T-nm) [°]	95% cv	$-T \sum_{i=p+1}^n \log(1 - \hat{\lambda}_i)$	(T-nm) [°]	95%cv
P = 0	92.61**	52.59**	45.3	210.2**	119.4	124.2
P ≤ 1	51.38**	29.18	39.4	117.6**	66.79	94.2
P ≤ 2	28.61	16.25	33.5	66.23	37.61	68.5
P ≤ 3	18.15	10.31	27.1	37.61	21.36	47.2
P ≤ 4	10.2	5.794	21	19.46	11.05	29.7
P ≤ 5	6.691	3.8	14.1	9.259	5.258	15.4
P ≤ 6	2.568	1.458	3.8	2.568	1.458	3.8
*indicates rejection of the null at the 95% level						
** indicates rejection of the null at the 99% level						
° -(T-nm) is a small sample correction replacing (-T) in the λ-max and λ-trace statistics						

Table 2.9 long-run rates pairwise cointegration tests							
	Rank	$-T \log(1 - \hat{\lambda}_p)$	$(T-nm)^\circ$	95%cv	$-T \sum_{i=p+1}^n \log(1 - \hat{\lambda}_i)$	$(T-nm)^\circ$	95%cv
USA-Canada	$p = 0$	20.39**	17.91*	14.1	23.08**	20.27**	15.4
USA-Japan	$p = 0$	18.64**	16.82*	14.1	19.66*	17.74*	15.4
USA-UK	$p = 0$	14.43*	12.67	14.1	14.73	12.94	15.4
USA-France	$p = 0$	24.07**	21.75**	14.1	24.34**	22**	15.4
USA-Italy	$p = 0$	12.13	10.66	14.1	12.88	11.31	15.4
USA-Germany	$p = 0$	5.524	4.985	14.1	6.001	5.415	15.4
Canada-France	$p = 0$	36.9**	31.43**	14.1	38.76**	33.01**	15.4
Canada-UK	$p = 0$	16.55*	13.66	14.1	18.56*	15.31	15.4
Canada-Italy	$p = 0$	9.61	8.438	14.1	11.26	9.891	15.4
Canada-Japan	$p = 0$	16.55*	14.51*	14.1	16.86*	14.78	15.4
Canada-Germany	$p = 0$	8.095	7.305	14.1	8.731	7.88	15.4
UK-France	$p = 0$	30.31**	26.32**	14.1	33.9**	29.44**	15.4
UK-Germany	$p = 0$	11.59	7.781	14.1	14.1	11.9	15.4
UK-Italy	$p = 0$	28.23**	22.96**	14.1	42.31**	34.41**	15.4
UK-Italy	$p \leq 1$	14.08**	11.45**	3.8	14.08**	11.45**	3.8
UK-Japan	$p = 0$	16.85*	15.21*	14.1	17.1*	15.43*	15.4
Germany-France	$p = 0$	10.71	9.401	14.1	14.93	13.99	15.4
Germany-Italy	$p = 0$	8.323	7.242	14.1	8.327	7.245	15.4
Germany-Japan	$p = 0$	13.96	12.24	14.1	14.7	12.89	15.4
Italy-Japan	$p = 0$	17.07*	15.84*	14.1	21.88*	20.3**	15.4
Italy-Japan	$p \leq 1$	4.805*	4.458*	3.8	4.805*	4.458*	3.8
Italy-France	$p = 0$	10.64	9.848	14.1	10.91	10.1	15.4
Japan-France	$p = 0$	20.05**	17.57*	14.1	20.1**	17.62*	15.4
* indicates rejection of the null of no cointegration at 95% level							
** indicates rejection of the null of no cointegration at 99% level							
° $-(T-nm)$ is a small sample correction replacing $(-T)$ in the λ -max and λ -trace statistics							

Table 2.10. Cointegration test between the inflation rates of Italy and Germany						
Eigenvalue	Log-likelihood	For rank				
	130.258	0				
0.210912	139.259	1				
0.085497	142.655	2				
Ho:rank=p	$-T \log(1 - \hat{\lambda}_p)$	$(T-nm)^\circ$	95% cv	$-T \sum_{i=p+1}^n \log(1 - \hat{\lambda}_i)$	$(T-nm)^\circ$	95%cv
P == 0	18.0*	14.6*	14.1	24.8**	20.2**	15.4
P <= 1	6.7**	5.5*	3.8	6.7**	5.5*	3.8
*indicates rejection of the null at the 95% level						
** indicates rejection of the null at the 99% level						
° -(T-nm) is a small sample correction replacing (-T) in the λ -max and λ -trace statistics						

Table 2.11. ADF-test on inflation rates of Italy and Germany			
	ADF	Critical value 95%	Critical value 99%
Italy	-0.9596	-2.9	-3.51
Germany	-2.326	-2.9	-3.51

Table 2.12 ADF-test on Cointegrating vectors indicated from rank test in table 4.10			
	ADF	Critical value 95%	Critical value 99%
C1vec1	-1.872	-2.9	-3.51
C1vec2	-2.584	-2.9	-3.51

Table 2.13 Ranking of irreducible cointegrating vectors			
IC vectors	eigenvalue	standard deviation	exogeneity restrictions
USA-Canada	0.220188	0.706543	Canada exogenous
Canada-UK	0.186922	0.879522	UK exogenous
UK-Japan	0.185781	0.947805	Feedback
Canada-France	0.365902	0.970788	Canada exogenous
USA-France	0.251763	1.041766	Feedback
Canada-Japan	0.184848	1.125332	Japan exogenous
USA-UK	0.16136	1.210084	UK exogenous
Germany-Japan**	0.158347	1.282394	-
UK-Germany**	0.139702	1.307491	-
UK-France	0.328857	1.403686	UK exogenous
USA-Japan	0.203316	1.421847	Japan exogenous
Japan-France	0.219263	1.445948	Japan exogenous
Canada-Germany**	0.0620838	1.499267	-
Italy-France**	0.0506678	1.584742	-
USA-Germany**	0.0645576	1.726671	-
Germany-France**	0.10124	2.062937	-
Canada-Italy**	0.110588	2.179342	-
USA-Italy**	0.137558	2.275801	-
Italy-Japan*	0.163975	2.467952	-
UK-Italy*	0.260275	2.480295	-
Germany-Italy**	0.0911157	2.865093	-
* test indicates full rank			
** no cointegration has been found between the series			

Table 2.14 Ranking of IC vectors per country.							
	US	CA	JP	GER	FR	IT	UK
US	-	0.70	1.42	<i>1.72</i>	1.04	<i>2.27</i>	1.21
CA	0.70	-	1.12	<i>2.7</i>	0.97	<i>2.17</i>	0.87
JP	1.42	1.12	-	<i>1.28</i>	1.44	<i>2.46</i>	0.94
GER	<i>1.72</i>	<i>2.7</i>	<i>1.28</i>	-	<i>2.06</i>	<i>2.86</i>	<i>1.3</i>
FR	1.04	0.97	1.44	<i>2.06</i>	-	<i>1.58</i>	1.4
IT	<i>2.27</i>	<i>2.17</i>	<i>2.46</i>	<i>2.86</i>	<i>1.58</i>	-	<i>2.48</i>
UK	1.21	0.87	0.94	<i>1.3</i>	1.4	<i>2.48</i>	-
- bold indicates vectors irreducible and structural - italic for the ones involving Italy and Germany.							

Chapter 3

Testing for Changes in the Long-Run Causal Structure of Cointegrated Vector Autoregressions

3.1 Introduction

The underlying motivation for this chapter is a simple apparent contradiction, which exists in two major branches of the applied econometric literature. On the one hand we have a very large literature which attests to the pervasive presence of structural breaks in the real world - this is not simply the formal econometric literature (Perron 1989, Hansen 1992, Lumsdaine and Papell 1997, Stock and Watson 1993 among many others) but also includes the general world of economics (Lucas 1988) where we see such evident structural breaks as the formation of the European Monetary Union, the unification of Germany or the restructuring of the financial system which happened throughout the 1980's. Contradicting this, however, we find many examples of stable long-run relationships in the form of co-integrating vectors which appear to exist and remain stable despite the presence of such large changes. The contention of this chapter is that these two stylised facts can be reconciled if we recognise that many of the most important structural changes come in the form of a changing causal structure. So major changes may occur while the basic long-run relationships remain stable. A good example of this is the European Monetary Union. There has been a long established relationship between national price levels and currencies which has been widely tested in the purchasing power parity (PPP) hypothesis. Under floating exchange rates it is usually assumed that the exchange rate will move to equalise prices between any two countries. When a monetary union is formed this obviously no longer happens, as the exchange rate is

irrevocably fixed. However, since PPP does not disappear, what must happen is that the causal structure must change so that the price levels react to each other to remain in line. The long-run relationship does not change but the way the system works in a causal sense changes. In a modern cointegration framework (Johansen 1988, 1991) this is governed by the α matrix. It is structural change in this part of the system which we wish to focus on.

Within the cointegration framework several tests for structural change have been proposed. Perron (1989) first showed the importance of testing for structural change in such a framework, arguing that if there is a break in the deterministic trend, the unit root tests tend to under-reject the null of a unit root. Taking the breakpoint as exogenous, he suggests a modification of the Dickey-Fuller test with three different types of deterministic trend functions. These allow, in turn, for a one-off change in the intercept, a change in slope of the trend, and both of them. The null of a unit root is then tested against the alternative of a broken trend stationary process. After Perron's (1989) seminal paper several testing methods have been developed for endogenising the breakpoint. These procedures comprise recursive (using sub-samples), rolling (using a fixed-size window that moves along the sample), and sequential methods (including switching dummies in the full sample).

Zivot and Andrews (1992) use a sequential unit root test and derive the distribution of the test and tabulate its critical values. Banerjee, Lumsdaine and Stock (1992) use various recursive and sequential tests which endogenise the breakpoint. They consider the recursive maximum and minimum DF test and the difference between them deriving the asymptotic distribution of the recursive and sequential test statistics and tabulating the relative critical values. Lumsdaine and Papell (1997) extend the analysis to the case of multiple breaks with unknown breakpoints.

This set of procedures for unit root tests has been extended to the tests for cointegration. Here we need to distinguish between breaks in the relationships and breaks in the individual

variables. Gregory, Nason and Watt (1996) study the sensitivity of the ADF test for cointegration in the presence of a single permanent break. Using Monte Carlo methods they show that the presence of a break results in under-rejection of the null of no cointegration implying the inappropriateness of constant parameter cointegration analysis in such cases. A simple diagnostic test for structural change is suggested by Hao and Inder (1996) that extends the CUSUM test to the case of non-stationary regressors using the FM-OLS residuals and replacing the error variance with the long-run variance estimate. They derive the asymptotic distribution of the FM-OLS based CUSUM test statistics and tabulate the critical values.

Hansen (1992) derives the asymptotic distribution of a LM test for parameter instability against several alternatives in the context of cointegrated regression models. Quintos and Phillips (1993) develop a test for the null of parameter constancy in cointegrated regressions against the alternative that the coefficients follow a random walk. Gregory and Hansen (1996) propose several tests for the null of no cointegration against the alternative of cointegration in the presence of a possible break in the intercept or the slope coefficients in the cointegrating relation at an unknown point in time.

All the testing procedures considered so far are performed within a FM-OLS environment and are therefore suitable for FM models or single equations. In cointegrated systems a la Johansen (1988,1991) we consider four types of structural changes of which three occur in the long-run structure of the model. To explain this let

$$(3.1.1) \quad \Delta X_t = \Gamma_1 \Delta X_{t-1} + \dots + \Gamma_{k-1} \Delta X_{t-k+1} + \alpha \beta' X_{t-k} + \varepsilon_t$$

be the usual cointegrated model in error correction form where X is a p -dimensional matrix of non-stationary $I(1)$ variables, Γ_i is the matrix of short-run parameters, the error term $\varepsilon_t \sim \text{iid}N(0, \Omega)$, β is the cointegrating matrix and α is the matrix of adjustment coefficients that determines the long-run causal structure, and $\alpha \beta' = \Pi$ has reduced rank $r < p$. It is convenient to rewrite the model as

$$(3.1.2) \quad Z_{0t} = \Theta Z_{1t} + \alpha\beta' Z_{2t} + \varepsilon_t,$$

where $\Theta = (\Gamma_1, \dots, \Gamma_{k-1})$, $Z_{0t} = \Delta X_t$, $Z_{1t} = (\Delta X'_{t-1}, \dots, \Delta X'_{t-k+1})'$, $Z_{2t} = (X_{t-k})$. Using this parameterisation we can see that first of all changes may occur over time in the short-run parameters Θ . This case does not give rise to any complications and can in general be dealt with by redefining Z_1 and Θ . The other three types of changes have different implications for the statistical reliability of the model, and they concern the long-run parameters α and β (suitably normalised) as well as the rank of the matrix $\Pi = \alpha\beta'$.

Quintos (1995) uses the trace statistic to assess the constancy of the cointegrating rank against the hypothesis of a changing rank of $\Pi = \alpha\beta'$. Assuming a known breakpoint she divides the sample into different periods and derives the likelihood ratio test for H_0 and its asymptotic distribution. Hansen and Johansen (1999) suggest graphical procedures to evaluate the constancy of the long-run parameters of the model. They look at the time paths of the recursively estimated eigenvalues and argue that by inspecting them it is possible to identify the breakpoints. Furthermore they show that fluctuation tests like the one by Ploberger, Kramer and Kontrus (1989) and the LM-test by Nyblom (1989) can be applied to test the constancy of long-run parameters in cointegrated VARs. Seo (1998) defines LM tests statistics for structural changes in both the cointegrating vector and the vector of adjustment parameters for both the cases of a known and unknown breakpoint. Using Monte Carlo methods he finds that the tests for structural change of the cointegrating vector have a non-standard distribution that is equal to the one found by Hansen (1992) using the FM technique. The tests on the vector of loading weights are found to have the same asymptotic distribution for models with stationary variables.¹¹

¹¹ Similar results are obtained by Andrews and Ploberger (1994) for the test on the matrix of loading weights but they find a different (even though still non-standard) distribution for the tests for constancy of the parameters in cointegrating relations.

Because changes in the cointegrating rank occur as a result of shifts in the parameters, it is possible to concentrate only on the latter types of breaks. We start by rewriting the model as

$$(3.1.3) \quad Z_{0t} = \Theta Z_{1t} + \alpha(t)\beta(t)' Z_{2t} + \varepsilon_t,$$

thereby incorporating the time dependency of the two parameters. Within this framework there are three possible types of structural change in the model:

$$A) \quad \alpha(t)\beta(t)' = (\alpha\beta_1 I_{1t} + \alpha\beta_2 I_{2t} + \dots + \alpha\beta_k I_{kt})$$

$$B) \quad \alpha(t)\beta(t)' = (\alpha_1\beta I_{1t} + \alpha_2\beta I_{2t} + \dots + \alpha_k\beta I_{kt})$$

$$C) \quad \alpha(t)\beta(t)' = (\alpha_1\beta_1 I_{1t} + \alpha_2\beta_2 I_{2t} + \dots + \alpha_k\beta_k I_{kt})$$

The first case amounts to a shift in the cointegrating coefficients. This may take the form of a shift in the parameter attached to one or more variables and therefore a change in the cointegrating space, but may also take the extreme form of a breakdown of the long-run relationship. Case B constitutes a structural break in the matrix of adjustment coefficients and it implies a change in the long-run causal structure of the model but not a change in the cointegrating space. A relevant example would be an attempt on the part of the monetary authorities to change the structure of the economy using specific policies. Case C involves a contemporaneous shift in both the cointegrating and the adjustment coefficients.

The key feature to note here is that almost all the work cited above has focused on analysing structural change of the form given by model A, rather than model B. Yet we stressed in the opening paragraph that we believe structural change of the form given by B is both more empirically relevant and more interesting from a policy perspective. However, we have very little insight into how such change will affect our estimation techniques, how to detect it or what its small sample implications are.

In this chapter we concentrate on such cases and investigate, using Monte Carlo methods, the performance and the distribution of the FIML estimates for the α s in the presence of a break in the long-run causality structure. We will adopt recursive as well as rolling FIML

techniques for detecting such changes occurring at different points in time and in samples of different size. As in Hansen and Johansen (1999), we will regard the recursive and rolling estimation techniques as misspecification tests for the detection of possible instabilities in the adjustment coefficients and of the breakpoints.

We show that in the case of a break in the long-run causal structure, recursive techniques like the one of Hansen and Johansen (1999) that relies upon the inspection of recursive eigenvalues are not very useful. Inspecting the recursively-estimated coefficients proves useful in showing their instability but not enough to detect the breakpoints in individual replications. Another major limitation of recursive techniques is that they never provide unbiased estimates of the parameters of interest.

Rolling techniques perform relatively better in that they detect the instability of the coefficients, providing also estimates of their true values. However, they are not entirely satisfactory, since they only detect the region where the break is likely to occur. To overcome such a problem we propose two alternative approaches showing that they are more efficient in detecting the breakpoints as well as providing the true value of the parameters of interest.

The first testing procedure consists in incorporating into the system in structural error correction form a sequentially switching dummy for the speed of adjustment parameter, defined so that it becomes “more” statistically significant when it corresponds exactly to the break in the data. We will derive the critical values for the null of no break for different sample sizes and will discuss the statistical properties of the test. We will see how such a technique has a relatively good power for the detection of the breakpoint in the case of a single persistent shift. We will then compare its performance to that of a time-varying version of the Kalman filter for the estimation of the parameters of interest, the latter being a well established estimation methodology for detecting instabilities of parameters and breakpoints.

Note also that the properties of the Kalman filter are such that it can be used in the case of multiple breaks or gradually changing coefficients.

The layout of the chapter is the following: In the next section we will introduce the DGPs generated for the Monte Carlo experiment. Section 3 and 4 present a sequential test for structural breaks in the causality of cointegrated VARs and the Kalman filter, as they will be used in our experiment. The results will be discussed in section 5. A brief summary and final remarks will conclude.

3.2 The Monte Carlo experiment

We have designed the DGPs according to two simple criteria: model dimension and breakpoint of the causal structure. For better management of the experiment and in order to concentrate only on the long-run causal structure of the models, we have assumed absence of breaks in the cointegrating relations as well as their stability. We have generated a basic cointegrated bivariate model in structural form with dynamics that obey a simple autoregressive process of order one (AR(1)), and have examined the performance of the procedures under investigation allowing for the possibility of three different points in time where the causal structure breaks down. The simulated model is of the following form:

$$(3.2.1) \quad A^0 \Delta X_t = \Gamma \Delta X_{t-1} + \alpha \beta' X_{t-1} + e_t$$

that can be rewritten as

$$(3.2.2) \quad \Delta X_t = (A^0)^{-1} \Gamma \Delta X_{t-1} + (A^0)^{-1} \alpha \beta' X_{t-1} + (A^0)^{-1} e_t$$

Here we allow A^0 to switch to A^1 so as to create a break in the causal structure of the model.

In particular we consider the case where

$$(3.2.3) \quad A^0 = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.3 \end{bmatrix}, \quad \alpha = \begin{bmatrix} 0.25 \\ 0 \end{bmatrix}, \quad \beta = \begin{bmatrix} 1 \\ -2 \end{bmatrix};$$

and allow the matrix A^0 to switch to

$$(3.2.4) \quad A^1 = \begin{bmatrix} 1 & 0 \\ 0.5 & 1 \end{bmatrix}$$

implying that the matrix $\Pi = \alpha\beta'$ will switch to Π^* as follows:

$$(3.2.5) \quad \Pi = \begin{bmatrix} -0.25 & 0.50 \\ 0 & 0 \end{bmatrix}, \quad \Pi^* = \begin{bmatrix} -0.25 & 0.50 \\ 0.125 & -0.25 \end{bmatrix}.$$

The reason for considering bivariate processes is that they are relatively easy to work with and are frequently used in empirical research (e.g. Purchasing Power Parity, Uncovered Interest Parity), while higher dimensional processes would create problems of experiment management and interpretation. In each of these processes we have considered the simple case of a unique break consisting of a zero loading weight switching to a non-zero value at some point in time. For the FIML-based analysis and the Kalman Filter we have generated 1000 samples of size $T+50$, always discarding the first 50 observations. The DGPs differ from each other in the points at which the break occurs. These are $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$ of the samples. We will henceforth identify these three basic DGPs as DGP1, DGP2, and DGP3 respectively. For each of the DGPs, four sample sizes were considered, that is $T = 50, 100, 200$, and 400 . To derive the critical values and the statistical properties of the sequential test we have carried out 2500 replications using the three DGPs of size 50, 60, 80, 100, 200, 400. In our FIML analysis we will perform various types of estimation. Specifically, we will first estimate full sample models performing the usual Johansen (1988, 1991) cointegration test, obtaining the estimates of the parameters of interest in the models. The objective is to investigate the statistical properties of the FIML estimates for the loading weights in the presence of a unique and permanent break for different sample sizes and breakpoints. We will then re-estimate the model recursively to investigate whether the FIML technique is able to detect the presence of the breaks via the inspection of the recursive eigenvalues and the recursive estimates. Furthermore, we will estimate the loading weights from the model in error correction form using OLS, where the cointegrating relation is the one obtained

recursively using FIML, and we will investigate their statistical significance in all the recursions for all the sample sizes. Lastly, we will perform a rolling estimation of the cointegrating coefficients and the adjustment parameters, infer the statistical properties of the FIML estimator in small samples, and investigate the possibility of detecting the break in the DGPs using such a technique. In our experiments, the initial sample size for the recursive estimation is 20. For the rolling estimation the sizes of the moving window are 30, 40 and 50 points.

Next, we will introduce and perform an OLS-based sequential test that is based on defining a sequentially switching dummy for the loading weights in order to detect the occurrence of a break and the breakpoint itself. We will also report the results obtained on the same DGPs using a time-varying version of the Kalman filter. The latter will be used as a benchmark to compare the relative performance of the sequential test and the FIML based procedures in the case of a shift in the causation between variables in the model. These two last techniques are described in the following sections.

All the simulations and the estimation have been carried out using codes written in GAUSS.

3.3 A sequential test for structural change

Consider the following cointegrating VAR in error correction form,

$$(3.3.1) \quad \begin{aligned} \Delta Z_{1t} &= \Gamma_{11}\Delta Z_{t-1} + \dots + \Gamma_{1k-1}\Delta Z_{t-k+1} + \alpha_1(\beta' Z_{t-k}) + e_{1t} \\ \Delta Z_{2t} &= \Gamma_{21}\Delta Z_{t-1} + \dots + \Gamma_{2k-1}\Delta Z_{t-k+1} + \alpha_2(\beta' Z_{t-k}) + e_{2t} \end{aligned}$$

where $\beta'Z_{t-k}$ is the stationary linear combination of the non-stationary levels that is assumed to be stable over time, and $\alpha = [\alpha_1, \alpha_2]$ represents the speed at which the variables in Z adjust to disequilibrium, and consider the case when the matrix α switches over time to some α° like,

$$(3.3.2) \quad \alpha = \begin{bmatrix} \alpha_1 \\ 0 \end{bmatrix}, \quad \alpha^\circ = \begin{bmatrix} \alpha_1^\circ \\ \alpha_2^\circ \end{bmatrix}, \quad \alpha_2^\circ \neq 0.$$

Such a switching may be caused by many phenomena. Consider the case of the UK joining the ERM at some point in time. Suppose that we are estimating a simple model of uncovered interest parity between the rates of Germany and UK. Assuming that the two rates are cointegrated before and after the entry of the UK into the ERM, it is reasonable to expect that from the date at which the UK joins the ERM, its rates receive some feedback from the German ones. In this case, in the UK equation, the coefficient of the speed of adjustment towards the equilibrium (in our example α_2) would switch from a zero to a non-zero value.

Here we propose to estimate such a model including a dummy variable that should be multiplied by the error correction term. This dummy variable is then switched on for subsamples of the data. The strategy consists in observing for which sub-periods this switching dummy is “more” statistically significant. Among these sub-periods, the intervals in which the t-value of this dummy variable is larger indicate the presence of a shift in the coefficients to which it is attached for the period covered. Formally, the estimated model will be,

$$(3.3.3) \quad \begin{aligned} \Delta Z_{1t} &= \Gamma_{11} \Delta Z_{t-1} + \dots + \Gamma_{1k-1} \Delta Z_{t-k+1} + \alpha_1 (\beta' Z_{t-k}) + \alpha_1^\circ D_i (\beta' Z_{t-k}) + e_{1t} \\ \Delta Z_{2t} &= \Gamma_{21} \Delta Z_{t-1} + \dots + \Gamma_{2k-1} \Delta Z_{t-k+1} + \alpha_2 (\beta' Z_{t-k}) + \alpha_2^\circ D_i (\beta' Z_{t-k}) + e_{2t} \end{aligned}$$

with $D_i = (0_0, \dots, 0_{i-1}, 1_i, \dots, 1_T)$, where T is the sample size.

The test statistics for the null hypothesis of no-break against the alternative of a single persistent one is given by

$$(3.3.4) \quad \tau = t - \max = \operatorname{argmax} \{t_{\alpha_i} = 0(n)\}, n = 2, T - 2.$$

This value will be compared against the 95% and 90% critical values that we have derived for the different sample sizes considered.

One limit of this technique consists in the fact that it can only be used to detect a single persistent shift in the coefficient to which it is attached. By contrast, where the suspect is of

multiple breaks, and when the breakpoints are in general unknown, it could prove more feasible to use a time-varying technique such as the Kalman filter presented below.

3.4 Kalman Filtering

The Kalman (1960, 1963) Filter technique is adopted to estimate linear models with time-varying coefficients¹². This class of models consists of two equations: the transition equation, describing the evolution of the state variables, and the measurement equation, describing how the observed data are generated from the state variables. This approach is extremely useful for investigating the issue of parameter constancy, because it is an updating method producing estimates for each time period based on the observations available up to the current period. It is important to realise that recursive OLS estimation (or moving window OLS estimation) is not a suitable technique to use here. Recursive estimation is essentially a test of structural stability. We can set up as the null the hypothesis that the parameters are constant and see if that can be rejected through recursive estimation. But as the underlying assumption of OLS is always that the parameters are constant, recursive estimation does not provide a consistent estimate of a time-varying parameter.

Let the Kalman Filter measurement equation be:

$$(3.4.1) \quad y_t = x_t' \beta + \varepsilon_t \quad \varepsilon_t \sim N(0, H_t)$$

and the transition equation be:

$$(3.4.2) \quad \beta_t = T\beta_{t-1} + \eta_t \quad \eta_t \sim N(0, Q_t)$$

with the initial conditions given by:

$$(3.4.3) \quad \beta_0 \sim N(\beta_0, \sigma^2 P_0)$$

When $T=I$ and $Q_t=0$, the model is reduced to the standard normal OLS regression model. The matrices T , H and Q are assumed to be known, and the problem is obtaining estimates of β_t

using information I_t available up to time t . The process of evaluating the conditional expectation of β_t given I_t is known as filtering. The evaluation of β_t given I_s , with $s > t$, is instead referred to as smoothing, whereas the estimation of β_t with $s < t$ is called prediction. Kalman (1960) derived the basic results to obtain filtered and smoothed estimates of β_t recursively. The prediction equation is given by:

$$(3.4.4) \quad \hat{\beta}_{t/t-1} = T\hat{\beta}_{t-1}$$

and the covariance matrix is defined as:

$$(3.4.5) \quad P_{t/t-1} = TP_{t-1}T' + Q_t$$

Finally, the updating formulae are given by:

$$(3.4.6) \quad \hat{\beta}_t = \hat{\beta}_{t/t-1} + P_{t/t-1}x(y_t - x'\hat{\beta}_{t/t-1})/(x'P_{t/t-1}x + H_t)$$

and

$$(3.4.7) \quad P_t = P_{t/t-1} - P_{t/t-1}x'xP_{t/t-1}' / (x'P_{t/t-1}x + H_t)$$

As the estimates are updated recursively each period, Kalman filtering can be viewed as belonging to the class of Bayesian estimators. Before starting the estimation process, one has to specify the vector of prior coefficients β_t and the matrix Q_t . By estimating the long-run relationship in this way one obtains a vector containing the evolving state coefficients which show whether the relative importance of the factors driving the dependent variable has changed over time.

In our case we start from a model in error correction form such as:

$$(3.4.8) \quad \Delta Z_t = \Gamma_1 \Delta Z_{t-1} + \dots + \Gamma_{k-1} \Delta Z_{t-k+1} + \alpha \beta' Z_{t-k} + e_t,$$

where as usual e_t are assumed to be iid $N(0, \Sigma)$, and taking the Γ_i and β as non time-varying we estimate the matrix of adjustment coefficients α , with the Kalman filter, under the assumption that this matrix follows a random walk process such that

¹² For a more detailed account of Kalman filtering, see Cuthbertson *et al* (1992).

$$(3.4.9) \quad \alpha_t = \alpha_{t-1} + v_t, \quad v_t \text{ is } N(0, \sigma^2 I)$$

Using this technique we will obtain the time path of the adjustment coefficients for all the DGPs considered. This should reveal the changes in the values of the adjustment coefficients as well as the unknown breakpoints.

3.5 Results

We start with the FIML estimation using the whole sample. The first result concerns the estimates of the cointegrating coefficients denoted as β_1 and β_2 respectively in table 3.1. These coefficients (standardised with respect to β_1) are always estimated accurately by the FIML-based procedure in all the DGPs and for all the sample sizes. This indicates that, in general, the presence of a break in the causal structure of the model does not seem to affect the cointegration test or the estimation of the matrix of cointegrating coefficients. As expected, the estimates of α_2 are instead biased being a weighted average of the two values that α_2 takes before and after the break. These are never too close to 0.125, that is the value that α_2 takes after the break, and in some cases (mostly in DGP3) they reach numerical values close to zero.

When it comes to accounting for breaks¹³ in the long-run causality structure of a system, the Johansen procedure fails to do so for breaks in the three DGPs in the case of smaller samples (50 points). In the case of the samples of size 100 it fails to account for the breaks in DGP2 and DGP3. In the case of DGP1 the hypothesis of a coefficient that is significantly different from zero would not be rejected. For the samples of larger dimensions (200 and 400) the results are different depending on the point in time at which the break occurs. If the break occurs at $\frac{1}{4}$ of the sample as in DGP1, this is detected and the hypothesis of a zero coefficient

¹³By failing to account for breaks in the full sample estimation, we mean that the coefficient estimates are distributed such that it would not be possible to reject the hypothesis of a zero coefficient.

on α_2 would be easily rejected¹⁴. When the break occurs in the middle of the sample (DGP2) the results obtained are different for the two sample sizes of 200 and 400.

With a sample size of 200 we would not be able to reject the hypothesis of a zero coefficient on α_2 at the 95% level, while the same hypothesis would be rejected for a sample of 400, where we would find a coefficient on α_2 that is significantly different from zero. In the case of DGP3 the hypothesis of a non-zero coefficient would be rejected for both the samples of 200 and 400. All these results are illustrated in table 3.1.

The recursive FIML estimation is not expected to provide different results in terms of the estimates. Here, following Hansen and Johansen (1999), we use it for the purpose of detecting instability of the coefficients and possible breakpoints. If we examine figures 3.1 and 3.2 in the appendix for DGP2 of size 100 it is clear that the inspection of the recursive eigenvalues does not provide any information on the occurrence of a break in the long-run causal structure of the model. In fact, there is no change in the pattern of the eigenvalues in the graph showing the average of 1000 replications at the points where the break occurs. The recursively estimated matrix of loading weights appears to be relatively more informative in terms of graphic analysis.

The plot of the recursive estimates of the loading weights shows their instability over time. In the graph of the time path of the recursive loading weights for DGP2 as an average of 1000 replications and a sample of size 100, we observe a clear tendency in the value to increase after the breakpoint that is identifiable by a slight kink. At first sight, we could be tempted to conclude that the recursive estimates provide excellent results in terms of the detection of the break as well as the breakpoint. However, these results need to be interpreted with some care. Indeed, the results obtained in single replications (not displayed), confine the recursive analysis to being useful only in the detection of instability in the parameter estimates. As far

¹⁴ Notice that the rejection of the hypothesis of a zero in favour of a non-zero coefficient would not account for

as the detection of the break point is concerned, the recursive technique cannot be relied upon because the breakpoint kink is not so clearly observable in single replications.

In any case, the statistical ability of the recursive procedure in detecting the breaks is cast in doubt by the fact that when the coefficients become significantly different from zero in statistical terms they do so with a considerable time lag. We have obtained these results estimating recursively the loading weights as the coefficients attached to the error correction term (as obtained with recursive FIML) in a VECM using OLS. The OLS estimation and the inspection of the *t-values* for the α_2 coefficient provide us with results that are different depending on the sample size and the breakpoint. They indicate (see table 3.2) that when α_2 becomes significant, it does so with a delay that goes from 60 to 100 points depending on sample size and breakpoint.

More in detail, we see that in samples of size 50 α_2 is only significant at a 90% level in DGP1. For the samples of size 100, we see that it is significant only in DGP1 and it starts to be so only when $\frac{3}{4}$ of the sample points are included. For other sample sizes the story is much the same, with α_2 that takes values significantly different from zero after 100 points have been added from the recursive procedure. Notice that the α_2 coefficients are never significant for DGP3 regardless of the sample size. This implies that if we had a sample of 400 points with the last 100 being different from zero, and we analysed it using recursive methods we would not be able to detect the presence of a break at all. This result highlights a critical limit of the FIML estimation even in its recursive version.

To summarise, while the *t-values* on α_2 could be used in some cases to confirm the presence of a break, they would not be reliable for the precise detection of the breakpoint or its magnitude. At any rate, recursive estimation is not very useful to account for a break in samples that are smaller than 200 and for all sample sizes when the break occurs at $\frac{3}{4}$ of the

the fact that the first $N/4$ values are actually zero.

series. Furthermore, it is important to highlight that the graphs of the recursive coefficients are obtained as averages of 1000 replications, and of course they are able to show the breakpoints relatively clearly while these are not easily detectable in individual replications.

The rolling estimation provides us with results that are a little more efficient. Indeed (see graphs 3.3 and 3.4), if used graphically, the rolling procedure shows the occurrence of a break in the data. Unfortunately, the estimates of the alphas are not statistically significant in the case of the size of a rolling window of 30. This relatively small window and the size of the break in terms of the value to which the coefficient α_2 switches to, do not allow to obtain statistically useful results. The implication is that in individual repetitions we are not able to draw statistically sustainable conclusions from the inspection of the rolling coefficients and relative t-values (table 3.3).

When we use a rolling window of 40 points we get better results in the sense that we observe the presence of a break in the data and we are able to obtain coefficients that are significantly different from zero at the 90% level. In order to get statistically significant results at a conventional 95% level we should use a rolling window of at least 50 points. This implies that in order for the break to be detectable, it has to be of course larger than 50 points, and that therefore, we would need relatively large samples. Furthermore, the larger size of the rolling window complicates the graphic analysis in terms of the detection of the break points as can be gathered from graphs 3.3, 3.4 and 3.5. In empirical work, apart from financial series, we are rarely in possession of very large samples; given also the need for a relatively large minimum size of the break, the reliability of the technique in statistical terms is limited. Nevertheless, it seems that graphically it performs better than the recursive procedure in detecting the interval (even so this is sometimes a wide one) within which the break occurs and the value of the coefficient after the break.

The sequential test represents without doubt a better alternative to the techniques examined so far. It performs better with respect to the accuracy with which it detects both the point and the magnitude (in terms of the value that α_2 takes) of the structural break in the long-run causal structure of the model. In general, the proposed test displays reasonable statistical properties. These are reported in table 3.4.

Specifically, the sequential test has a reasonably good power at the 90% level, even for a sample size of 80, when the break occurs in the middle of the sample. The power is always relatively good for samples of size 100 regardless of the location of the breakpoint, and the test can be considered very reliable for samples of 200 points. For very large samples its precision is extreme. The loss of power in smaller samples can be justified on the grounds of its non-standard distribution and the relatively high critical values for samples of 50, 60 points (that are basically the same as the ones for 80). If we use the test graphically, we can see (graph 3.6) that the coefficient attached to the switching dummy displays a kink that allows us to detect the point at which the break occurs quite accurately. Also, notice that at the same point the coefficient attached to it takes the value of 0.125 showing a certain precision in providing the magnitude of the break. Moreover, as far as statistical significance is concerned, the coefficient attached to the dummy at the breakpoint would results always highly significant at the 95% level in single replications.

The relatively good performance of the sequential test is only limited by the possible occurrence of multiple breaks. In this case the significance of the switching dummy is affected and the test is not very reliable. The Kalman Filter may provide the solution to this kind of problems. In theory, the fact that we obtain the time path of the adjustment coefficients in a consistent way should enable detection of the break quite accurately as far the timing and the magnitude of the break are concerned. A further advantage of using such a

technique would be the possibility of detecting multiple breaks, as the Kalman Filter is sensitive to all possible shifts that may occur because of its continuous updating.

From the results obtained we can conclude that, the Kalman filter may be graphically useful to detect the presence of breaks and their magnitude, and that it is also very efficient. We can see this from the reported graphs (graphs 3.7 and 3.8) that refer to coefficients estimated for samples of size 100 generated according to DGP2. Here the break points and their magnitude are graphically observable, with relatively narrow margins of error, as shown by the confidence interval bands. These highlight the fact that the coefficients estimated with such a technique will always be significantly different from zero after the break has occurred, suggesting a reasonable efficiency of the Kalman. This result applies to all the DGPs and sample sizes.

3.6 Conclusions

In this chapter we have examined, using Monte Carlo methods, several alternative approaches to detecting structural changes in the causality structure of cointegrated VARs. As in Hansen and Johansen (1999), we have adopted recursive techniques as misspecification tests for the detection of instabilities in the adjustment coefficients and of the breakpoints. The results show that the recursive eigenvalues are not able to highlight a break in the causal structure of the models, whilst the recursive coefficients only show non-constancy of the parameters of interest.

We have then adopted a rolling FIML-based technique for the estimation of the model. The main result is that this technique shows more convincingly the instability of the parameters of interest and the true value of after the breakpoint. However, rather than a breakpoint it only detect a region where the break is likely to occur.

To overcome the drawbacks of these techniques, we proposed an OLS-based sequential test which consists in incorporating a sequentially switching dummy for the speed of adjustment parameter such that it becomes more statistically significant when it corresponds to the break in the data. For this test we derived the critical values for different sample sizes and studied its statistical features. We showed that the test has a non-standard distribution and has reasonable power at the 90% level even for smaller samples (e.g. 80, 100). It is interesting to notice that the test is relatively efficient graphically in that it shows a kink at the breakpoint. We also show that, at that point, the numerical value of the coefficient of the sequential dummy corresponds to the true value of the parameter after the break has occurred. The only limit to this technique is provided by the occurrence of multiple breaks. A possible solution to these cases is the use of the Kalman filter. We therefore have tested the performance of a time-varying version of the Kalman filter on the same DGPs, showing the graphic usefulness of such a technique as far as the breakpoint and the magnitude of the break are concerned. Regardless of the location of the breakpoint, this well established technique proves in fact its usefulness in both detecting the breakpoints and providing the true value of the parameters after the break has occurred.

Table 3.1.Full Sample FIML Estimates (1000 replications)

	<i>DGP1</i>		<i>DGP2</i>		<i>DGP3</i>	
	<i>Mean</i>	<i>Std. Dev.</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Mean</i>	<i>Std. Dev.</i>
sample size 50						
$\alpha1$	-0.278	0.191	-0.266	0.056	-0.294	0.211
$\alpha2$	0.098	0.110	0.063	0.062	0.0191	0.117
$\beta1$	1	#	1	#	1	#
$\beta2$	-1.967	0.311	-2.008	0.141	-1.957	0.377
sample size 100						
$\alpha1$	-0.255	0.037	-0.254	0.038	-0.255	0.038
$\alpha2$	0.089	0.043	0.053	0.041	0.025	0.037
$\beta1$	1	#	1	#	1	#
$\beta2$	-2.009	0.091	-2.006	0.073	-2.003	0.062
sample size 200						
$\alpha1$	-0.250	0.025	-0.249	0.025	-0.250	0.025
$\alpha2$	0.085	1.029	0.052	0.028	0.024	0.026
$\beta1$	1	#	1	#	1	#
$\beta2$	-2.001	0.043	-2.003	0.036	-2.002	0.031
sample size 400						
$\alpha1$	-0.249	0.018	-0.248	0.018	-0.249	0.018
$\alpha2$	0.084	0.020	0.051	0.020	0.024	0.018
$\beta1$	1	#	1	#	1	#
$\beta2$	-2.001	0.024	-2.001	0.19	-2.000	0.017

Table 3.2 t-values of recursive estimates of α_2							
DGP1	T/4	(3/8)T	T/2	(5/8)T	(3/4)T	(7/8)T	T
ss-50	#	0.849	1.108	1.302	1.493	1.632	1.730*
ss-100	0.266	0.828	1.203	1.548	1.826*	2.080**	2.346**
ss-200	0.193	0.957	1.522	2.010**	2.435**	2.822**	3.178**
ss-400	0.177	1.253	2.083**	2.785**	3.395**	3.953**	4.463**
DGP2	T/4	(3/8)T	T/2	(5/8)T	(3/4)T	(7/8)T	T
ss-50	#	0.382	0.459	0.687	0.908	1.072	0.247
ss-100	0.266	0.236	0.195	0.615	0.937	1.225	1.519
ss-200	0.193	0.155	0.150	0.749	1.232	1.667*	2.059**
ss-400	0.177	0.152	0.126	0.967	1.667*	2.288**	2.853**
DGP3	T/4	(3/8)T	T/2	(5/8)T	(3/4)T	(7/8)T	T
ss-50	#	0.382	0.332	0.282	0.377	0.553	0.739
ss-100	0.266	0.236	0.195	0.156	0.128	0.496	0.811
ss-200	0.193	0.155	0.150	0.137	0.106	0.628	1.062
ss-400	0.177	0.152	0.126	0.111	0.097	0.812	1.428
* denotes a coefficient α_2 significantly different from zero at 90% level							
** denotes a coefficient α_2 significantly different from zero at 95% level							
T = sample size (ss)							
The bold indicates the breakpoints in the DGPs							

Table 3.3 t-values of rolling estimates of α_2							
Rolling Window	T/4	(3/8)T	T/2	(5/8)T	(3/4)T	(7/8)T	T
30	0.134	0.421	0.827	1.207	1.529	1.499	1.505
40	0.358	0.582	0.911	1.292	1.645*	1.848*	1.862*
50	0.500	0.748	0.946	1.265	1.510	1.850*	2.130**
$T = 100$							
Breakpoint is at T/2 (in bold)							
* denotes a coefficient significantly different from zero at the 90% level							
** denotes a coefficient significantly different from zero at the 95% level							

Table 3.4 Sequential test – Statistical properties based on 2500 replications						
<i>Critical values (2500 replications)</i>						
Probability	t=50	T=60	T=80	T=100	T=200	T=400
0.95	2.93	2.93	2.93	2.94	2.98	3.02
0.90	2.65	2.61	2.61	2.64	2.69	2.71
<i>Power (break at 1/2) as percentage of rejections of the null</i>						
Probability	T=50	T=60	T=80	T=100	T=200	T=400
0.95	0.31	0.35	0.45	0.56	0.88	1.00
0.90	0.42	0.49	0.61	0.69	0.98	1.00
<i>Power (break at 3/4) as percentage of rejections of the null</i>						
Probability	T=50	T=60	T=80	T=100	T=200	T=400
0.95	0.23	0.26	0.34	0.44	0.75	0.98
0.90	0.34	0.40	0.50	0.58	0.86	0.99
<i>Power (break at 1/4) as percentage of rejections of the null</i>						
Probability	T=50	T=60	T=80	T=100	T=200	T=400
0.95	0.21	0.26	0.35	0.44	0.74	0.97
0.90	0.32	0.38	0.47	0.55	0.83	0.98

Figure 3.1 Recursive eigenvalues for DGP2 of size 100 (average of 1000 replications)

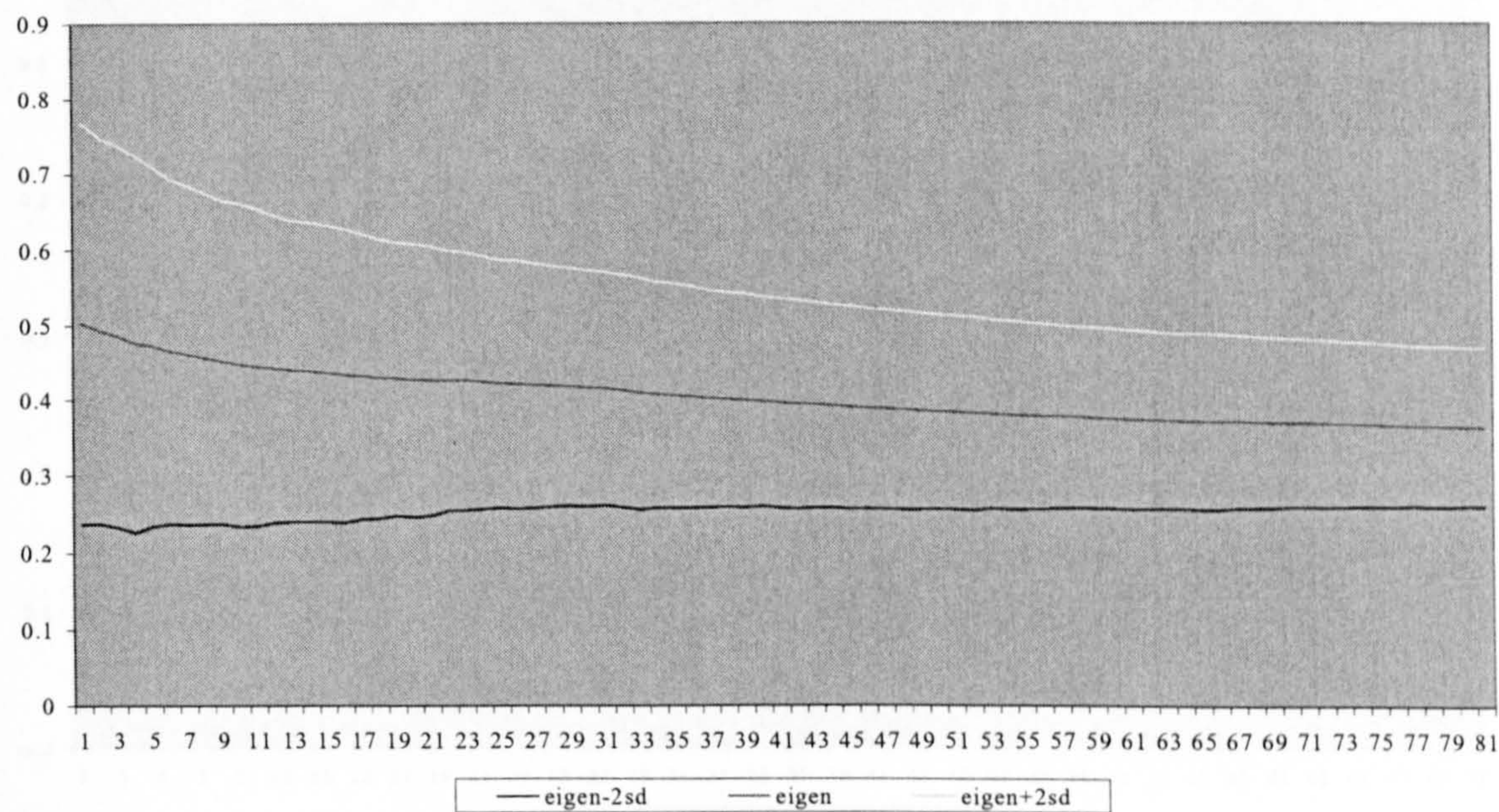


Figure 3.2 Recursive estimates of alpha2 for DGP2 of size 100 (average of 1000 replications)

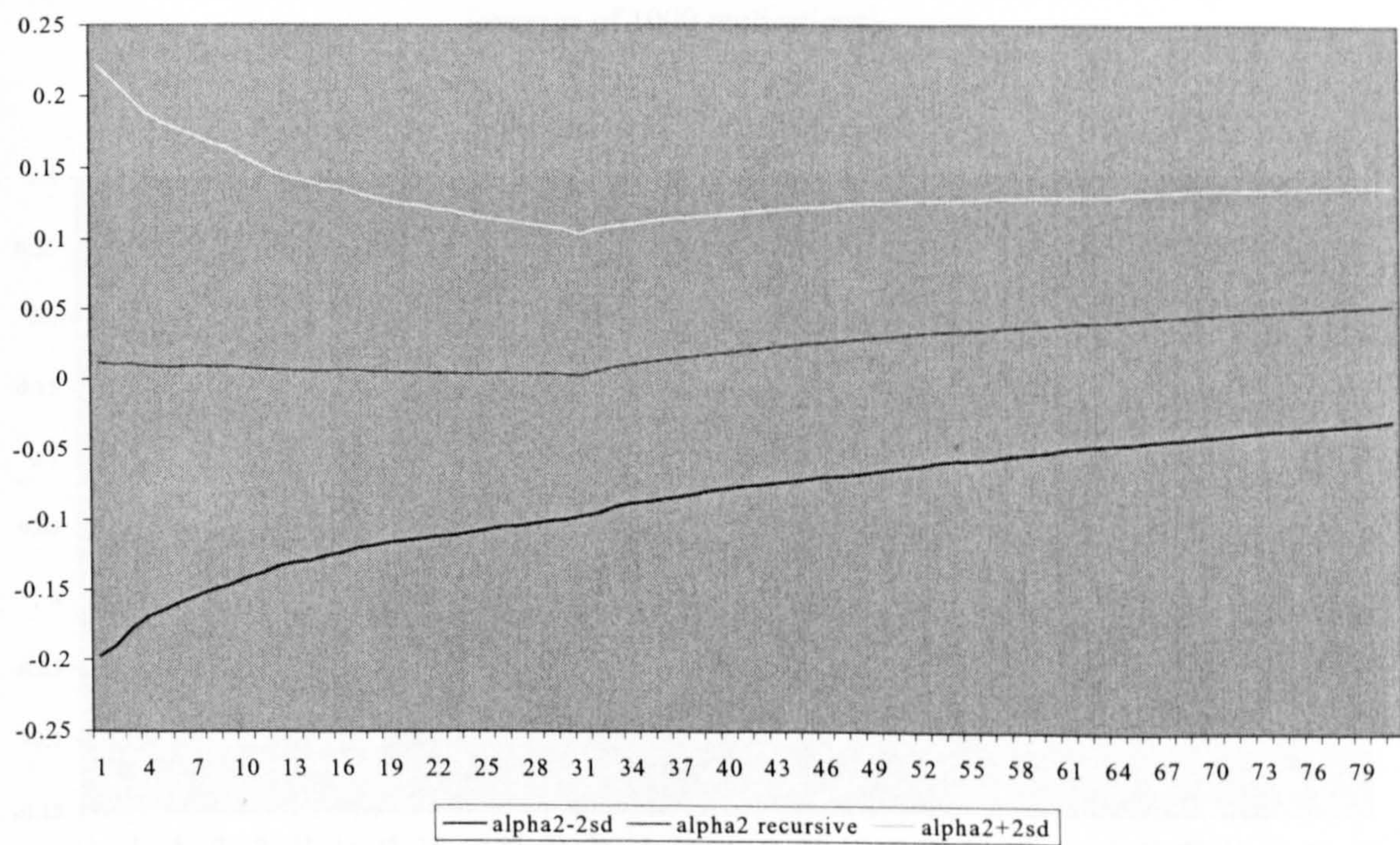


Figure 3.3 Rolling estimates of α_2 for DGP2 of size 100 with a rolling window of 30
(average of 1000 replications)

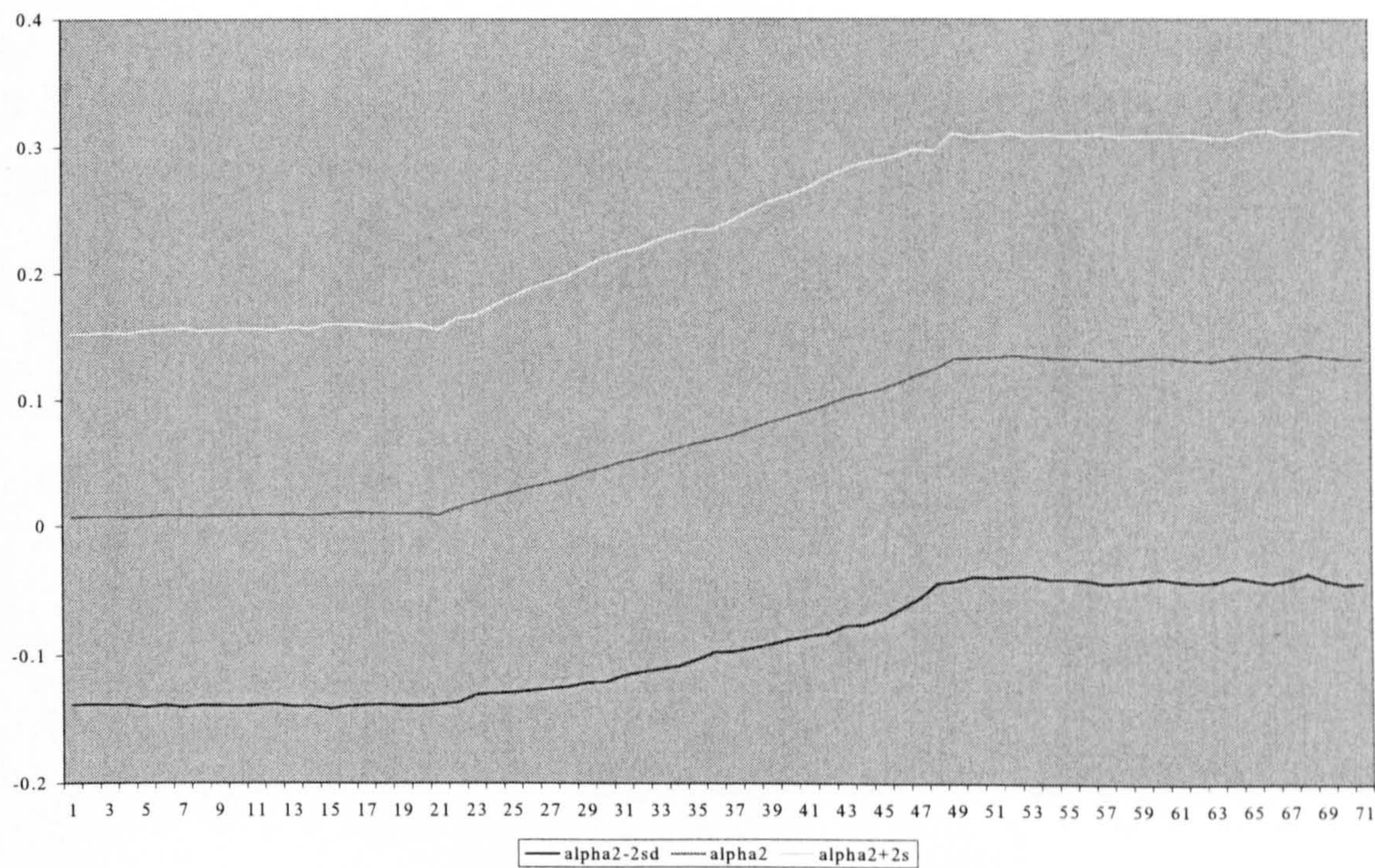


Figure 3.4 Rolling estimates of α_2 for DGP2 of size 100 with rolling window of 40
(average of 1000 replications)

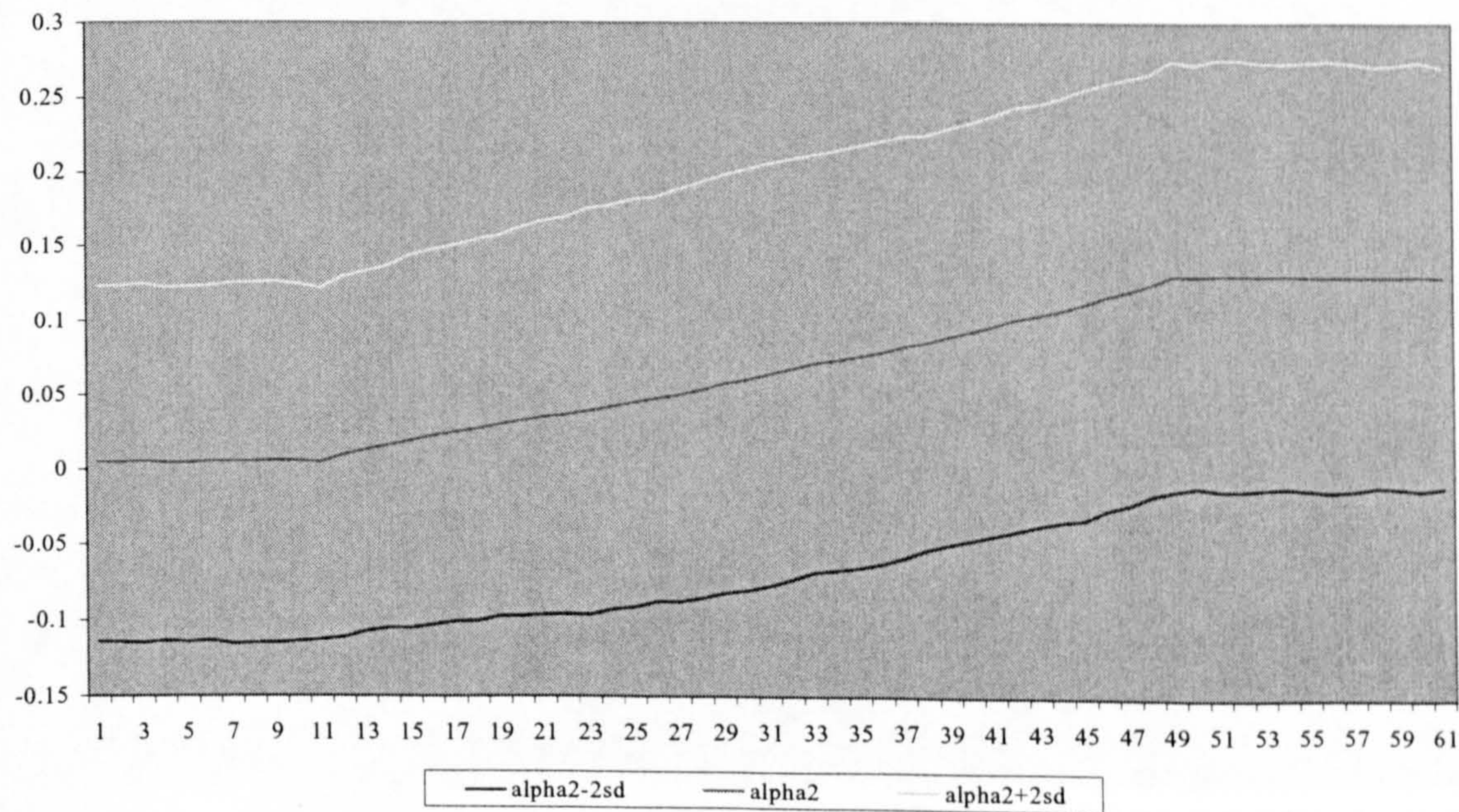


Figure 3.5 Rolling estimates of α_2 for DGP2 of size 100 with rolling window of 50
(average of 1000 replications)

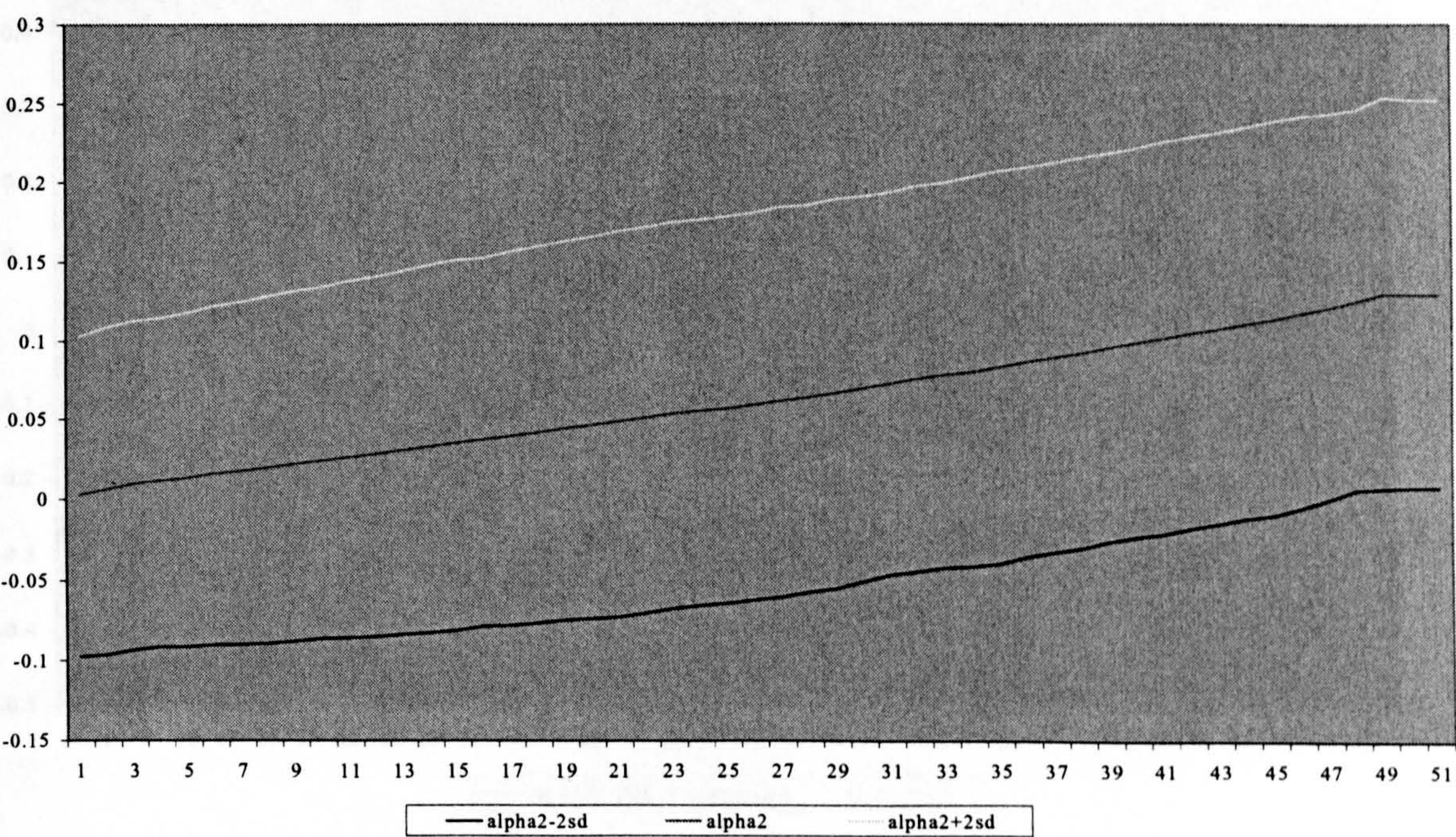


Figure 3.6 Coefficient of sequential dummy for DGP2 of size 100 (average of 2500 replications)

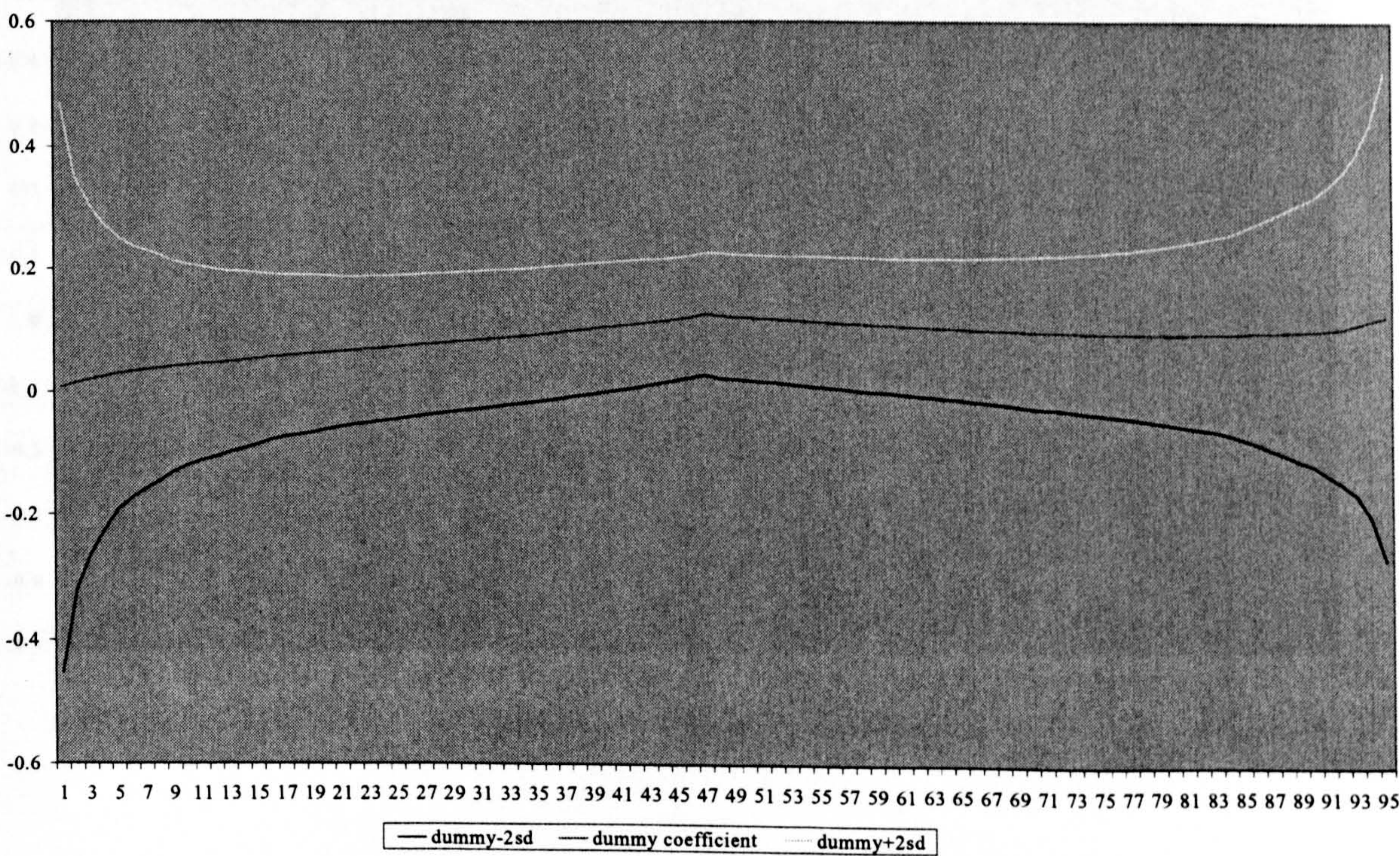


Figure 3.7 One step ahead Kalman Filter predictions of α_2 for DGP2 of size 100
(average of 1000 replications)

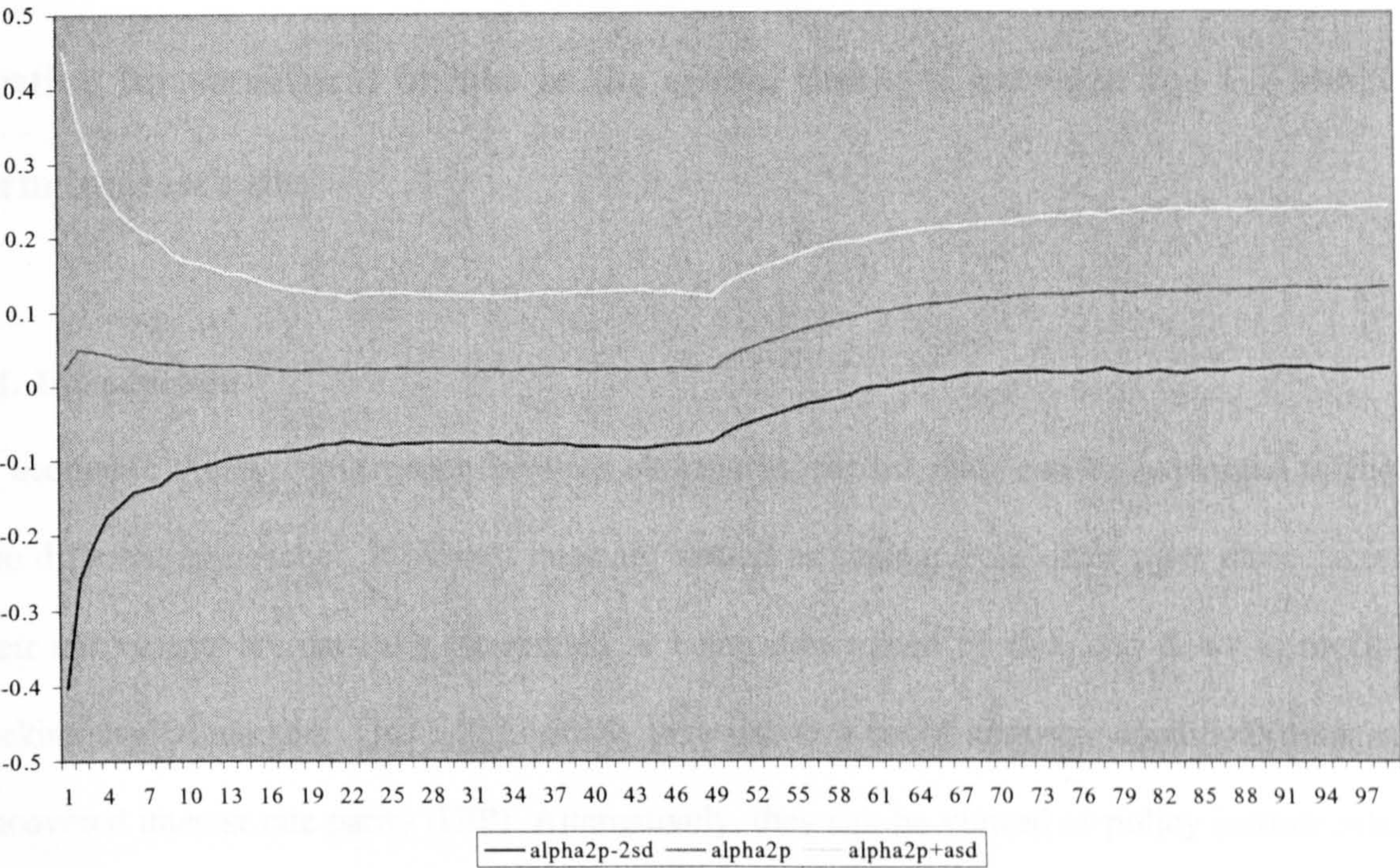
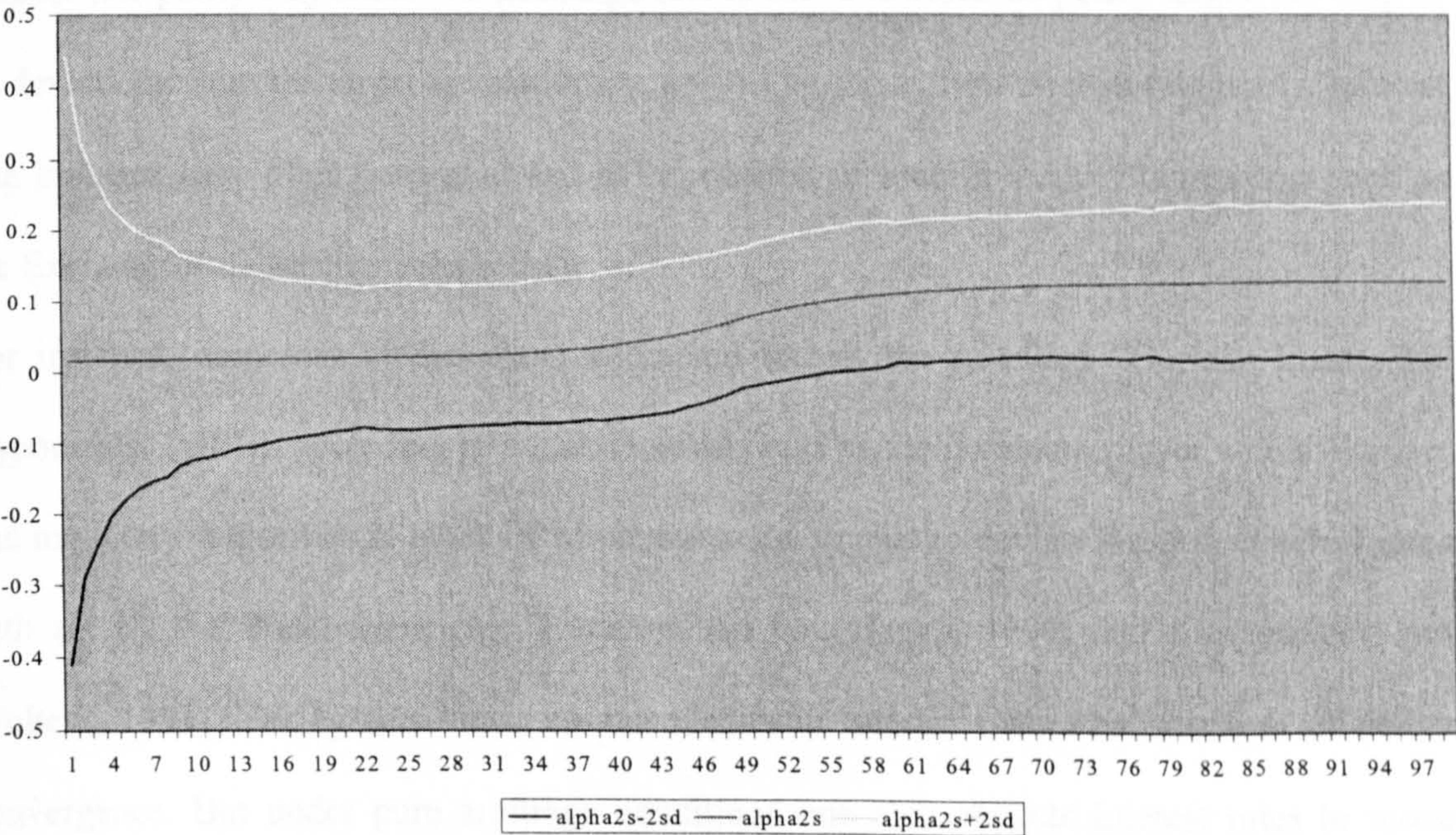


Figure 3.8 Smoothed kalman Filter estimates of α_2 for DGP2 of size 100
(average of 1000 replications)



Chapter 4

Testing for structural breaks in the causal linkages between the G7 short term interest rates

4.1. Introduction

In economic theory convergence between short-term interest rates can be explained taking two different approaches. If interest rates are treated as analogous to other asset prices, then their movements are naturally interpreted as being determined by financial flows in profit-seeking capital markets. This will normally give rise to a set of arbitrage conditions such as uncovered interest rate parity (UIP). Alternatively, they can be viewed as policy instruments, with their time paths being determined by a policy objective such as an exchange rate or an inflation target. These two approaches are not necessarily inconsistent, since deviations from interest rate parity may cause the exchange rate to move towards its policy target. As long as its deviations from the target are stationary, so will be those from interest rate parity. Interest rate linkages have often been analysed in the context of specific policy frameworks such as the Exchange Rate Mechanism (ERM).

For instance, numerous studies have attempted to test the so-called “German Leadership Hypothesis” (GLH), according to which Germany acts as the dominant player within Europe, and monetary authorities in other ERM countries are unable to deviate from the interest rates path set by the Bundesbank (see Fratianni and von Hagen, 1990, and Kirchgassner and Wolters, 1993). Taking this view, co-movements in interest rates arise because of policy convergence. But under pure arbitrage conditions one also expects interest rates to move

together in the long run. So the question naturally arises: how is the system affected by a policy regime, and how will it change if there is a regime shift?

In general, one can think of changes in structure as changes either in the long-run relationships themselves (the cointegrating vectors) or in causality links (the loading factors).

To be more precise, consider a Vector Equilibrium Correction Model (VEqCM), where the long-run matrix has been decomposed into a vector of loading weights and one of cointegrating relationships. A change in the structure of the system could occur through changes either in the former or in the latter.

Previous empirical studies, such as Caporale et al (1996), reported convergence in European rates after 1986. Artis and Zhang (1998), using rolling window cointegration techniques¹⁵, found that there is widespread cointegration between both US and German short rates and those on other ERM currencies up to 1995, after which the US influence on world-wide rates vanishes. In the context of the ERM, with its target zones, there might be regime shifts owing to the policies pursued by central banks. Specifically, the stochastic properties of interest rates (volatility, level and speed of adjustment) are likely to be different in periods when the currency has to be defended from speculative attacks, compared to periods when the exchange rate is credible. Because of the UIP relation, switches in the process governing exchange rates are translated into switches in the process followed by interest rates. Such regime shifts tend to be more frequent and not to be as long-lived as changes in monetary policy regimes in the US, say. Dahlquist and Gray (2000) show that a Markov-switching model characterises adequately the behaviour of a number of EMS short rates.

In general, it would be problematic to specify the source of structural change in a model allowing for both types of changes, as such a model would typically not be identified. In the case of interest rates, since almost any theory suggests long-run co-movement, it is

reasonable to assume that the cointegrating vectors are constant but the direction of causality changes. For instance, UK and German interest rates may have been set primarily in line with US rates prior to entering the ERM, but once full membership is achieved the UK rates may be driven by the German monetary stance which is still set with a view to the US. In both cases the three interest rates will move in line and be cointegrated on a pairwise basis, but the direction of causality will be different between the two policy regimes.

In this chapter we test for structural changes on the causal linkages between the G-7 short term interest rates using two of the techniques outlined in the chapter 3. Namely, first we will use the sequential dummy technique, and second we will use a time varying parameter version of the Kalman Filter. It is worth highlighting how the sequential test suffers from two potential pitfalls. The first concerns the fact that standard distributions cannot be relied upon to test for the significance of the t-statistic. The second is related to the fact that the sequential technique shows reasonable power only in detecting single breaks, and cannot be considered reliable when there is the suspect of multiple breaks.

To overcome the first of these problems, we derived, using Monte Carlo methods, the distribution and the critical values for the null of no breaks against the alternative of a single persistent one for different sample sizes. The second issue is dealt with by estimating time-varying parameter models for the loading weights. This has the advantage that one does not have to impose a priori restrictions on the number of breaks and the point in time at which they occur. Instead, the relationships are allowed to evolve freely, and the revealed timing of the structural breaks can be very informative about the effects of policy changes (see, e.g., Haldane and Hall, 1991, who analyse sterling's relationship with the US dollar and the Deutschemark).

¹⁵ Note that the results from this estimation method are highly sensitive to the selection of the window width and the magnitude of the break. (see chapter 3 above).

This chapter is organised as follows. The next section (section 4.2) outlines the testing procedures. Section 4.3 discusses the empirical results and their economic implications. A summary (section 4.4) concludes.

4.2 Testing for changes in the long-run causality of Cointegrated Vector Error Correction Models

Consider the following cointegrating VAR in error correction form,

$$4.2.1 \quad \begin{aligned} \Delta Z_{1t} &= \Gamma_{11}\Delta Z_{t-1} + \dots + \Gamma_{1k-1}\Delta Z_{t-k+1} + \alpha_1(\beta'Z_{t-k}) + e_{1t} \\ \Delta Z_{2t} &= \Gamma_{21}\Delta Z_{t-1} + \dots + \Gamma_{2k-1}\Delta Z_{t-k+1} + \alpha_2(\beta'Z_{t-k}) + e_{2t} \end{aligned}$$

where $\beta'Z_{t-k}$ is the stationary linear combination of the non-stationary levels that is assumed to be stable over time, and $\alpha = [\alpha_1, \alpha_2]$ represents the speed at which the variables in Z adjust to disequilibrium, and consider the case when any of the elements α_i of the matrix α switches over time to some α° . This can happen in two basic ways:

$$\begin{aligned} \text{A} \quad \alpha &= \begin{bmatrix} \alpha_1 \\ 0 \end{bmatrix}, & \alpha^\circ &= \begin{bmatrix} \alpha_1^\circ \\ \alpha_2^\circ \end{bmatrix}, & \alpha_2^\circ &\neq 0, \\ \text{B} \quad \alpha &= \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, & \alpha^\circ &= \begin{bmatrix} \alpha_1^\circ \\ 0 \end{bmatrix}, & \alpha_2 &\neq 0, \end{aligned}$$

In case A, Z_2 is exogenous and is not caused in the long run by Z_1 before the regime shift, but receives feedback from Z_1 afterwards. In case B we can see that the two states are just reversed in terms of time.

In this particular set-up, we assume that once the coefficient switches to its new value it does not go back to the original state. In order to detect the possible occurrence of a single permanent break we estimate such a model including a dummy variable that should be multiplied by the error correction term. This dummy variable is then switched on for sub-samples of the data. The strategy consists in observing for which sub-periods this switching

dummy is statistically significant. The intervals where the t-value of this dummy variable is larger indicate the presence of a shift in the coefficients to which it is attached for the period covered. In formal terms the estimated model will be:

$$4.2.2 \quad \begin{aligned} \Delta Z_{1t} &= \Gamma_{11}\Delta Z_{t-1} + \dots + \Gamma_{1k-1}\Delta Z_{t-k+1} + \alpha_1(\beta' Z_{t-k}) + \alpha_1^* D_i(\beta' Z_{t-k}) + e_{1t} \\ \Delta Z_{2t} &= \Gamma_{21}\Delta Z_{t-1} + \dots + \Gamma_{2k-1}\Delta Z_{t-k+1} + \alpha_2(\beta' Z_{t-k}) + \alpha_2^* D_i(\beta' Z_{t-k}) + e_{2t} \end{aligned}$$

with $D_i = (0_0, \dots, 0_{i-1}, 1_i, \dots, 1_T)$, where T is the sample size.

The test statistics for the null of no-break against the alternative of a single persistent break is given by

$$4.2.3 \quad \tau = \left| \arg \max \left\{ t_{\alpha_i} = 0(n) \right\} \right|, n = 2, T-2. ^{16}$$

Previously we have showed how the obtained test statistic follows a non-standard distribution, therefore, to make the testing procedure operative, we have derived the 90 and 95 per cent critical values for different sample sizes using Monte Carlo methods. These are displayed on table 3.4 in chapter 3.

However, the sequential testing procedure has reasonable power in detecting single permanent breaks only. The presence of multiple breaks may produce results that may not be unequivocally interpreted, especially if the multiple breaks have opposite direction. The restrictive conditions within which our testing procedure works may not allow us to draw strong conclusions by simply comparing the value of the test statistic to the critical values. We will therefore complement it presenting the time varying estimates for the adjustment coefficients of error correction equations as obtained by means of the Kalman filter.

The Kalman (1960, 1963) filter technique is adopted to estimate linear regression models with time-varying coefficients¹⁷. We have seen that this class of models consists of two

¹⁶ We consider the absolute value of the test statistic because its sign differs depending on the type (either A or B) of break that occurs.

equations: the transition equation, describing the evolution of the state variables, and the measurement equation, describing how the observed data are generated from the state variables. This approach is extremely useful for investigating the issue of parameter constancy, because it is an updating method producing estimates for each time period based on the observations available up to the current period current period. It is important to realise that recursive OLS estimation (or moving window OLS estimation) is not a suitable technique to use here. Recursive estimation is essentially a test of structural stability. We can set up a null hypothesis that the parameters are constant and see if that can be rejected through recursive estimation. But as the underlying assumption of OLS is always that the parameters are constant, recursive estimation does not provide a consistent estimate of a time-varying parameter.

In our case we start from a model in error correction form such that the Kalman Filter measurement equation will be:

$$4.2.4 \quad \Delta X_t = \Gamma_1 \Delta X_{t-1} + \dots + \Gamma_{k-1} \Delta X_{t-k+1} + \alpha \beta' X_{t-k} + \varepsilon_t,$$

where ε_t are assumed to be *iid* $N(0, H)$, and taking the Γ_i and β as non time-varying we estimate the matrix of adjustment coefficients α , so that the transition equation will be:

$$4.2.5 \quad \alpha_t = T \alpha_{t-1} + \nu_t, \quad \nu_t \text{ is } N(0, Q)$$

with the initial conditions given by:

$$4.2.6 \quad \alpha_0 \sim N(\alpha_0, \sigma^2 P_0)$$

Here we assume that the coefficients in α follow a random walk process such that $T=I$.

¹⁷ For an exhaustive exposition of Kalman filtering, see Harvey (1991) "Forecasting, Structural Time Series Models and the Kalman Filter", *Cambridge University Press*, Cambridge UK.

In particular we apply this procedure to the bivariate systems linking the G-7 short interest rates as irreducible relations¹⁸ in order to investigate the possibility of breaks in the causal structure of these linkages or reversals in the direction of causality.

Our expectations are to find that exogeneity of US interest rates with respect to European ones holds, implying the absence of breaks in the adjustment coefficients of the error correction equations for the first difference of US rate. We also expect to find evidence of a break in the adjustment coefficients for the UK rates in equations involving the other European rates after the collapse of the ERM. Other testable hypotheses are the presence of breaks in the influence that Germany may exercise on the European rates (as the Germany leadership Hypothesis would require), and the stability of the causal structure within the G-7 system as a whole.

4.3 Empirical Results

For our analysis we used IMF quarterly data on the three-month Treasury bill rates covering the period between 1980:Q3 and 1998:Q3. Notice that for efficiency reasons we have estimated single equations in error correction form rather than bivariate vector error correction models. This is legitimate as earlier on we found that there is bivariate cointegration between all the rates with unit cointegrating vectors. Once the cointegrating vectors are determined then single equation estimation becomes FIML.

More specifically, we analysed causal linkages between the G-7 short-term interest rates by applying our modification of a methodology due to Davidson (1998) relying on the concept of irreducible cointegrating vectors (IC). Evidence was found that cointegration is a property of the G-7 short rates, and that it is important to test for irreducibility as a diagnostic.

¹⁸ The concept of irreducible cointegrating vectors is due to Davidson (1998) and refers to cointegrating subsets of variables, which do not have any cointegrating subsets. The Davidson methodology (see Davidson 1998) decomposes hybrid cointegrating relations in irreducible ones allowing (in the case of over-identified systems)

Ranking the IC relations according to the criterion of minimum variance has allowed us to distinguish between structural and solved IC vectors without any prior theoretical assumptions¹⁹.

Recalling our results, we found that the system of the G-7 interest rates has a rank of six implying six structural irreducible cointegrating regressions which we inferred by ranking the IC relations according to the criterion of the minimum variance. The two most significant relations involve the US and Canada, and Italy and France. The other four relations involve Germany and Japan, and Japan with USA, France and UK.

Here we have estimated the error correction equations for each country, and bearing in mind the structural nature of some IC vectors we have inferred the causal structure that links the G-7 interest rates and tested for its stability. Specifically, six models were estimated for each country; in each equation we allow for the possibility that a particular country adjusts to one of the other G-7, and therefore for the possibility that each country is being influenced by any of the other six. So if we found that for country A all six adjustment parameters are zero or statistically insignificant, this would indicate that this country was not influenced by any of the other countries over the period considered.

The first observation to make is that most of the first differences of the G-7 interest rates seem to follow simple autoregressive processes of order one (AR(1)), apart from the US rates that require including some dummy variables to offset an ARCH effect corresponding to the monetary base targeting pursued by the Fed during the early 1980s. The complete results for all the countries are presented on Tables 4.1-4.7, which report the coefficient estimates and both OLS and Newey-West corrected standard errors. Notice that for ease of interpretation the sign of the adjustment coefficients is always presented as negative (even in the cases in

the identification of the structural relations without imposing theory based restrictions on the cointegrating vectors.

which this is positive due to the ordering of variables in the original bivariate VARs). At any rate, none of the series exhibit explosive behaviour.

The causal structure is obtained from the OLS estimation and is presented in Tables 4.1-4.7. The first important result is the exogeneity of US rates with respect to all the others (Table 4.7). This is particularly evident in the error correction equations including the adjustment coefficients towards the structural long-run equilibrium relations with Canadian and Japanese rates, suggesting a US worldwide leadership. Interestingly, Japan shares four out of the six IC relations with the US, France, Germany and the UK. It seems that Japanese rates (Table 4.4) represent the *trait d'union* between American and European rates, being exogenous with respect to all other rates apart from the US and UK ones.

Another important result is the weak leadership of Germany within the European Monetary Union (EMU). The exogeneity of German rates with respect to the Italian and French ones (Table 4.3) is clear, as they do not share any structural relations with the European ones and receive feedback from Japan (in a structural IC relation) and all the remaining rates included the UK. As for the UK rates, they seem to receive feedback from US and Canadian rates but not from the EU countries. This may be due to the presence of a break in the causal relationships with the latter rates after the third/fourth quarter of 1992, and it is one of the hypotheses we are interested in testing. Essentially, the US and Canada appear to constitute the fundamental block, UK rates respond to non-European rather than to other European rates, Italy is probably following France, and France and Germany respond to world rates rather than to each other. Lastly, Japan acts as the link between US and European rates.

These results, however, are all predicated on the assumption that the causal structure is constant over the sample period. A casual consideration of the structural changes which have

¹⁹ A solved vector is defined as a linear combination of structural vectors from which one or more common variables are removed by choosing offsetting weights such that the included variables are not a superset of any of the constituent relations.

occurred in the monetary policy structure of the world over the last 20 years makes this an unlikely assumption. Consequently, we will now apply the techniques outlined above to investigate possible changes in the adjustment coefficients in the single-equation error-correction models corresponding to the irreducible cointegrating relations for all the G-7 short-term interest rates.

4.3.a Sequential test results

The sequential test results are displayed in tables 4.8-4.14, and can be summarised as follows. First, we do not find evidence of structural changes in causality in the equations for the US, Canada and Germany. The US and Canada appear to have constituted a fundamental block throughout, with the US economy being a constant point of reference for all the other countries. Second, the weak German leadership within the Eurozone is confirmed, as German rates do not appear to be influenced by those of the EMU countries, but are influenced by all the other rates, including the UK ones.

Table 4.9 shows that there is some evidence of breaks in the equations for the French rate. These concern the feedback received from Canadian, German and Japanese rates (see figures 4.1, 4.2 and 4.3 respectively). The breaks seem to occur immediately after the *collapse of the* ERM, between the fourth quarter of 1992 and the first two quarters of 1993. It seems that in that particular period the feedback the French rate receives becomes stronger than in the past. The same conclusions can be drawn for the Italian short-term rate, that is closely linked to the French one in one of the structural relations, and from which it receives a stronger feedback from the third quarter of 1984 (figure 4.4). It appears, in fact, that German and Japanese rates start exerting a strong influence on it at the same time as on the French rate (end of 1992-early 1993), as shown in figures 4.5 and 4.6 respectively.

As highlighted in table 4.12, Japanese rates start following more closely the US and Canadian ones soon after the high volatility period due to the monetary base targeting *pursued by the* Fed, more precisely in the first quarter of 1983 (figures 4.7 and 4.8). This result, given the linkages between the US and Canadian rates with the Japanese ones in two of the structural IC relations, again indicates that these two countries constitute the fundamental block of the G-7 group, the US being a world-wide leader. Lastly, one should notice that the UK rate stops receiving feedback from the German one exactly after the collapse of the ERM in the last quarter of 1992, as shown in figure 4.9.

4.3.b Kalman Filter results

Below we discuss the time paths of the adjustment coefficients in the single-equation error-correction models corresponding to the irreducible cointegrating relations. IN particular we estimate the single equations by OLS, and having imposed the OLS coefficients as the fixed parameters of the observation equation we then re-estimate the same equations with the Kalman filter, assuming that the coefficient of the error correction term follows a random walk.

Again we estimate six models for each country. In each equation we allow for the possibility that that particular country adjusts to one of the other G7 countries, therefore standing for the possibility that each country is being influenced by any of the other six at any point in time. So if we found that for country A all six adjustment parameters were zero for the whole period, it would tell us that this country was not influenced by any of the other countries over this period. If we found that the adjustment coefficient involving country B became significant half way through the period, then this would indicate a shift in policy regime such that country A started to follow country B from that period onwards. The short-run dynamic parameters of the models are obviously the ones displayed on tables 4.1-4.7.

USA

The empirical results seem to support the existence of US leadership. The speed of adjustment coefficients converge towards zero in almost all cases. Exogeneity of US interest rates with respect to the Italian and the German ones is clearly observable (figure 4.16). The linkage with the Canadian rate seems to be the most significant one, although even this coefficient is below 0.1 and is falling. The same (but with even smaller coefficients) can be said about the speed of adjustment to disequilibrium in the linkage with the French and Japanese rates. The situation is slightly different in the error correction equation containing the cointegrating relation that links US to UK rates. It seems that there was some (decreasing) feedback from UK to US rates until the collapse of the ERM in the last quarter of 1992. Afterwards, US rates appear to be exogenous and causality runs from US to UK rates only. Overall we can conclude that the evidence from the time-varying estimation supports the idea that the US economy is a constant point of reference for all the countries of the G-7 group.

UK

The results for the UK are even more interesting. The main result is that the linkage between UK and other European rates has become weaker after the collapse of the ERM. In particular, apart from the clear exogeneity of UK rate with respect to the Italian one, it can easily be seen that in the 1990s the influence of the German and French rates on the British one has decreased overtime, even reaching zero in the case of the linkage with the French rate (see figure 4.15). It seems that the UK rate still responds to the German one (very weakly), but it is also clear that the influence on the UK of conditions in the non-European G-7 countries is still strong and even growing in the case of Canada. It is worth considering what happens in

the causal relation with the US rate. Here we can observe a break in the causal linkage between 1989 and 1991. Notice that after this period the US influence on UK rates seems to grow again. Overall we can then say that UK rates seem to respond to non-European rates, and that, following the breakdown of the ERM, UK monetary authorities have pursued policies that are completely independent from those implemented in the economies of the Eurozone. This might help explain the differences in economic performance between UK and other European economies during the 1990s.

The Eurozone

As already mentioned, one of the hypotheses of interest is the purported German leadership (GLH) within the Eurozone. In order to test it, we should only consider the three countries in our sample belonging to EMU, therefore investigating the speed of adjustment of German rates towards the long-run equilibria shared with the French and Italian rates. On the basis of the results obtained from this “partial” analysis, we might be tempted to conclude that the empirical evidence supports the GLH. This is because of the exogeneity of German rates with respect to the French and the Italian ones (figure 4.12). However, a closer look at the complete G-7 system shows that such a hypothesis does not have empirical support, for three main reasons.

First, while the only causal link to Germany seems to be from the US there are also strong direct links from the US to France and Italy (figures 4.11 and 4.13 respectively) (in both cases larger in magnitude than the link to Germany). So while the three rates may be moving in line there does not seem to be a strong case for arguing for one of them as the leader. Second, one of the six structural IC vectors links Italy and France, which implies that the relations between these two latter countries and Germany belong to the class of irreducible

solved relations rather than to the structural ones. Third, Germany does not share any of the six structural irreducible relations with another European country.

As for French and Italian rates (again figures 4.11 and 4.13 respectively), there appears to be feedback from all the world rates and between themselves. This provides us with enough evidence to conclude that European rates are actually driven more by US rates, which are exogenous with respect to all the Eurozone rates. It is worth noticing that the time path of the speed of adjustment coefficients of the error correction equations of European rates displays a kink in the fourth quarter of 1992, which coincides with the collapse of the ERM. Also, subsequently it becomes a lot smoother, implying more stability within the Eurozone.

Canada and Japan

We have chosen to discuss the results on Canadian and Japanese rates together as they seem to act as the trait d'union between US and European rates within the G-7 (see figures 4.10 and 4.14 respectively). The main feature of the Japanese rate is its high non-variability. Specifically, it is characterised by step-changes, clearly indicating that, policy decisions rather than market conditions determine Japanese rates. This may help to explain the low variability of the time-varying parameter estimates of the adjustment coefficients in its error correction equations as well as their low (but non-zero) values. We find some weak feedback from more or less all the other rates (with the exception of the German one) to Japanese rates²⁰. As already stated, it appears that Japanese rates, together with the Canadian ones, act as a linkage between European and US rates. In fact the feedback from both Japanese and Canadian rates to US rates is substantially weaker than the one in the opposite direction. As for the Canadian rate, it is linked to the other rates by a two-ways feedback relation, apart

²⁰ Notice that the relation between these rates constitutes one of the six irreducible structural linkages between the G-7 rates.

from its clear exogeneity with respect to the French rate in another of the six irreducible structural cointegrating relations.

4.4 Conclusions

In this chapter we investigated changes in the causal structure linking the G-7 short-term rates by means of an OLS-based sequential test and Kalman filter estimation for the constancy of the adjustment coefficients in error correction equations. The latter constitutes a *benchmark* for the performance of the former as this is only suitable to detect single permanent structural breaks in the causal linkages. Besides the constancy of the G-7 causal structure as a whole, hypotheses of interest were the US world-wide leadership, the possible disengagement of UK monetary policy from that pursued in the Eurozone after the collapse of the ERM, and the German leadership hypothesis (GLH) within the European Union (EU).

As a general observation we can say that the two procedures perform well and do not provide conflicting results. However, it has to be said that some of the results are expectedly more clearly interpretable by examining the Kalman filter estimates. This may be partly due to the limited power of the sequential test in detecting the breaks when the coefficients show instabilities in opposite directions. Furthermore the detection of time of the switch may be affected when the shifts occur gradually rather than abruptly.

To outline our findings- first we did not find any cases of reversal of causality, second we did not find evidence of substantial changes in the feedback that the US, Canadian and German rates receive. Clearly, the US and Canada have consistently constituted the fundamental block of the G-7 group, with the US economy being a constant point of reference. Third, the absence of breaks in the adjustment coefficients in the equations for the German rate therefore confirming the weak leadership of Germany within the Eurozone. Moreover, we

found some evidence of breaks in the equations for the French and Italian rates as it seems that they start following German rates more closely after the collapse of the ERM. Another result suggests that Japan link its rate to the US and Canadian ones soon after the FED was targeting the monetary base. Lastly, the sequential test evidence points to a break in the causal linkages between the UK and the German rates after the third-fourth quarter of 1992 implying the disengagement of UK policies from developments in the EU. This progressive disengagement of UK from the policies pursued in the EU is more evident from the Kalman filter estimates. In the following period, UK rates seem to be linked much more to world rates, as shown by the higher speed of adjustment parameter after 1992 in the equations linking UK and world rates.

Table 4.1. Canada: OLS-based estimates for the ECMs		
Dependent Variable: DCanada(t)		Error correction term
Constant	DCanada(t-1)	Canada-France(t-1)
-0.105	0.212	-0.057
(0.148)	(0.127)	(0.077)
[0.151]	[0.108]	[0.059]
Constant	DCanada(t-1)	Canada-Germany(t-1)
0.094	0.220	-0.046
(0.232)	(0.118)	(0.053)
[0.169]	[0.097]	[0.046]
Constant	DCanada(t-1)	Canada-Italy(t-1)
-0.212	0.247	-0.051
(0.257)	(0.122)	(0.062)
[0.218]	[0.104]	[0.049]
Constant	DCanada(t-1)	Canada-Japan(t-1)
0.662	0.276	-0.130
(0.357)	(0.114)	(0.059)
[0.286]	[0.106]	[0.051]
Constant	DCanada(t-1)	Canada-UK(t-1)
-0.164	0.277	-0.172
(0.142)	(0.112)	(0.068)
[0.152]	[0.970]	[0.062]
Constant	DCanada(t-1)	USA-Canada(t-1)
0.362	0.174	-0.230
(0.212)	(0.110)	(0.078)
[0.249]	[0.097]	[0.098]
OLS Standard errors in parenthesis		
Newey-West standard errors in brackets		

Table 4.2. France: OLS-based estimates for the ECMs		
Dependent Variable: DFrance(t)		Error correction term
Constant	DFrance(t-1)	Canada-France(t-1)
-0.114 (0.097) [0.102]	0.132 (0.110) [0.119]	-0.165 (0.047) [0.078]
Constant	DFrance(t-1)	Germany-France(t-1)
0.282 (0.195) [0.166]	0.292 (0.108) [0.110]	-0.093 (0.049) [0.032]
Constant	DFrance(t-1)	France-Italy(t-1)
-0.906 (0.323) [0.481]	0.263 (0.112) [0.104]	-0.232 (0.089) [0.141]
Constant	DFrance(t-1)	France-Japan(t-1)
0.564 (0.311) [0.243]	0.236 (0.112) [0.110]	-0.119 (0.052) [0.042]
Constant	DFrance(t-1)	France-UK(t-1)
-0.150 (0.103) [0.111]	0.174 (0.112) [0.128]	-0.087 (0.037) [0.036]
Constant	DFrance(t-1)	USA-France(t-1)
0.260 (0.130) [0.174]	0.228 (0.103) [0.105]	-0.146 (0.041) [0.062]
OLS Standard errors in parenthesis		
Newey-West standard errors in brackets		

Table 4.3. Germany: OLS-based estimates for the ECMs		
Dependent Variable: DGermany(t)		Error correction term
Constant	DGermany(t-1)	Canada-Germany(t-1)
-0.140 (0.101 [0.068])	0.393 (0.106 [0.088])	-0.049 (0.023 [0.025])
Constant	DGermany(t-1)	Germany-France(t-1)
-0.001 (0.118 [0.071])	0.448 (0.103 [0.103])	-0.001 (0.029 [0.023])
Constant	DGermany(t-1)	Germany-Italy(t-1)
-0.028 (0.178 [0.112])	0.406 (0.107 [0.102])	-0.001 (0.024 [0.021])
Constant	DGermany(t-1)	Germany-Japan(t-1)
0.204 (0.095 [0.102])	0.449 (0.096 [0.095])	-0.100 (0.036 [0.037])
Constant	DGermany(t-1)	Germany-UK(t-1)
-0.247 (0.116 [0.102])	0.340 (0.106 [0.090])	-0.052 (0.025 [0.022])
Constant	DGermany(t-1)	USA-Germany(t-1)
-0.095 (0.069 [0.065])	0.378 (0.101 [0.094])	-0.063 (0.023 [0.025])
OLS standard errors in parenthesis		
Newey-West standard errors in brackets		

Table 4.4. Italy: OLS-based estimates for the ECMs

Dependent Variable: DItaly(t)		Error correction term
Constant	DItaly(t-1)	Canada-Italy(t-1)
0.374	0.277	-0.127
(0.153)	(0.100)	(0.036)
[0.151]	[0.076]	[0.040]
Constant	DItaly(t-1)	France-Italy(t-1)
0.723	0.376	-0.233
(0.302)	(0.103)	(0.081)
[0.289]	[0.083]	[0.091]
Constant	DItaly(t-1)	Germany-Italy(t-1)
0.312	0.344	-0.059
(0.264)	(0.105)	(0.035)
[0.220]	[0.084]	[0.028]
Constant	DItaly(t-1)	Italy-Japan(t-1)
0.749	0.369	-0.090
(0.352)	(0.101)	(0.037)
[0.346]	[0.080]	[0.039]
Constant	DItaly(t-1)	Italy-UK(t-1)
0.063	0.293	-0.057
(0.135)	(0.108)	(0.030)
[0.113]	[0.090]	[0.031]
Constant	DItaly(t-1)	USA-Italy(t-1)
0.478	0.324	-0.099
(0.217)	(0.103)	(0.035)
[0.277]	[0.071]	[0.046]
OLS standard errors in parenthesis		
Newey-West standard errors in brackets		

Table 4.5. Japan: OLS-based estimates for the ECMs

Dependent Variable: DJapan(t)		Error correction term
Constant	DJapan(t-1)	Canada-Japan(t-1)
-0.295	0.381	-0.049
(0.145)	(0.183)	(0.027)
[0.158]	[0.110]	[0.031]
Constant	DJapan(t-1)	France-Japan(t-1)
-0.229	0.360	-0.031
(0.167)	(0.106)	(0.028)
[0.157]	[0.091]	[0.028]
Constant	DJapan(t-1)	Germany-Japan(t-1)
-0.058	0.423	-0.018
(0.088)	(0.105)	(0.034)
[0.104]	[0.105]	[0.036]
Constant	DJapan(t-1)	Italy-Japan(t-1)
-0.088	0.406	-0.006
(0.213)	(0.105)	(0.023)
[0.144]	[0.110]	[0.013]
Constant	DJapan(t-1)	Japan-UK(t-1)
-0.567	0.324	-0.087
(0.190)	(0.102)	(0.029)
[0.186]	[0.102]	[0.031]
Constant	DJapan(t-1)	USA-Japan(t-1)
-0.022	0.321	-0.043
(0.081)	(0.109)	(0.020)
[0.080]	[0.122]	[0.021]
OLS standard errors in parenthesis		
Newey-West standard errors in brackets		

Table 4.6. UK: OLS-based estimates for the ECMs		
Dependent Variable: DUK(t)		Error correction term
Constant	DUK(t-1)	Canada-UK(t-1)
0.032	0.239	-0.117
(0.110)	(0.106)	(0.051)
[0.121]	[0.114]	[0.051]
Constant	DUK(t-1)	UK-France(t-1)
-0.086	0.177	-0.032
(0.116)	(0.108)	(0.049)
[0.118]	[0.113]	[0.047]
Constant	DUK(t-1)	Germany-UK(t-1)
0.071	0.184	-0.045
(0.209)	(0.119)	(0.045)
[0.204]	[0.114]	[0.047]
Constant	DUK(t-1)	Italy-UK(t-1)
-0.118	0.156	-0.004
(0.155)	(0.120)	(0.036)
[0.147]	[0.110]	[0.038]
Constant	DUK(t-1)	UK-Japan(t-1)
0.447	0.262	-0.080
(0.414)	(0.122)	(0.065)
[0.354]	[0.118]	[0.064]
Constant	DUK(t-1)	USA-UK(t-1)
0.245	0.169	-0.132
(0.160)	(0.111)	(0.045)
[0.145]	[0.117]	[0.041]
OLS Standard errors in parenthesis		
Newey-West standard errors in brackets		

Table 4.7. USA: OLS-based estimates for the ECMs		
Dependent Variable: DUSA(t)*		Error correction term
Constant	DUSA(t-1)	USA-Canada(t-1)
0.040	0.167	-0.048
(0.098)	(0.072)	(0.038)
[0.071]	[0.086]	[0.028]
Constant	DUSA(t-1)	USA-France(t-1)
-0.019	0.251	-0.019
(0.108)	(0.091)	(0.036)
[0.078]	[0.113]	[0.023]
Constant	DUSA(t-1)	USA-Germany(t-1)
-0.054	0.273	-0.002
(0.092)	(0.087)	(0.032)
[0.078]	[0.098]	[0.031]
Constant	DUSA(t-1)	USA-Italy(t-1)
0.056	0.248	-0.020
(0.201)	(0.090)	(0.033)
[0.175]	[0.101]	[0.026]
Constant	DUSA(t-1)	USA-Japan(t-1)
0.261	0.227	-0.090
(0.195)	(0.109)	(0.048)
[0.251]	[0.099]	[0.058]
Constant	DUSA(t-1)	USA-UK(t-1)
-0.162	0.203	-0.043
(0.153)	(0.112)	(0.045)
[0.162]	[0.084]	[0.050]
(*) dummies for 1980:3, 1981:1, 1982:1, 1982:4 are included to account for the period corresponding to the Monetary Base Targeting pursued by the FED during the early 1980s		
OLS Standard errors in parenthesis		
Newey-West standard errors in brackets		

Table 4.8. Results of sequential test for structural change in causality for Canada						
	France	Germany	Italy	Japan	UK	USA
Test-statistic	-1.18	-1.71	-0.85	-2.52	1.27	2.50
** denotes rejection of the null of no break at the 95% level						
* denotes rejection of the null of no break at the 90% level						

Table 4.9. Results of sequential test for structural change in causality for France						
	Canada	Germany	Italy	Japan	UK	USA
Test-statistic	-2.82*	-5.18**	-1.60	-3.52**	-1.72	-2.37
** denotes rejection of the null of no break at the 95% level						
* denotes rejection of the null of no break at the 90% level						

Table 4.10. Results of sequential test for structural change in causality for Germany						
	Canada	France	Italy	Japan	UK	USA
Test-statistic	-2.18	-1.77	1.62	1.95	-1.95	1.13
** denotes rejection of the null of no break at the 95% level						
* denotes rejection of the null of no break at the 90% level						

Table 4.11. Results of sequential test for structural change in causality for Italy						
	Canada	France	Germany	Japan	UK	USA
Test-statistic	-2.24	-2.68*	-3.72**	-2.92*	-2.01	-2.50
** denotes rejection of the null of no break at the 95% level						
* denotes rejection of the null of no break at the 90% level						

Table 4.12. Results of sequential test for structural change in causality for Japan						
	Canada	France	Germany	Italy	UK	USA
Test-statistic	-3.25**	-1.83	2.48	-1.90	-2.41	-2.72*
** denotes rejection of the null of no break at the 95% level						
* denotes rejection of the null of no break at the 90% level						

Table 4.13. Results of sequential test for structural change in causality for UK						
	Canada	France	Germany	Italy	Japan	USA
Test-statistic	1.86	2.39	2.72*	1.01	1.78	2.06
** denotes rejection of the null of no break at the 95% level						
* denotes rejection of the null of no break at the 90% level						

Table 4.14. Results of sequential test for structural change in causality for USA						
	Canada	France	Germany	Italy	Japan	UK
Test-statistic	-1.85	2.40	2.00	2.02	1.52	1.36
** denotes rejection of the null of no break at the 95% level						
* denotes rejection of the null of no break at the 90% level						

Figure 4.1 Test statistic for a break on the French adjustment coefficient towards the long-run relationship France Canada

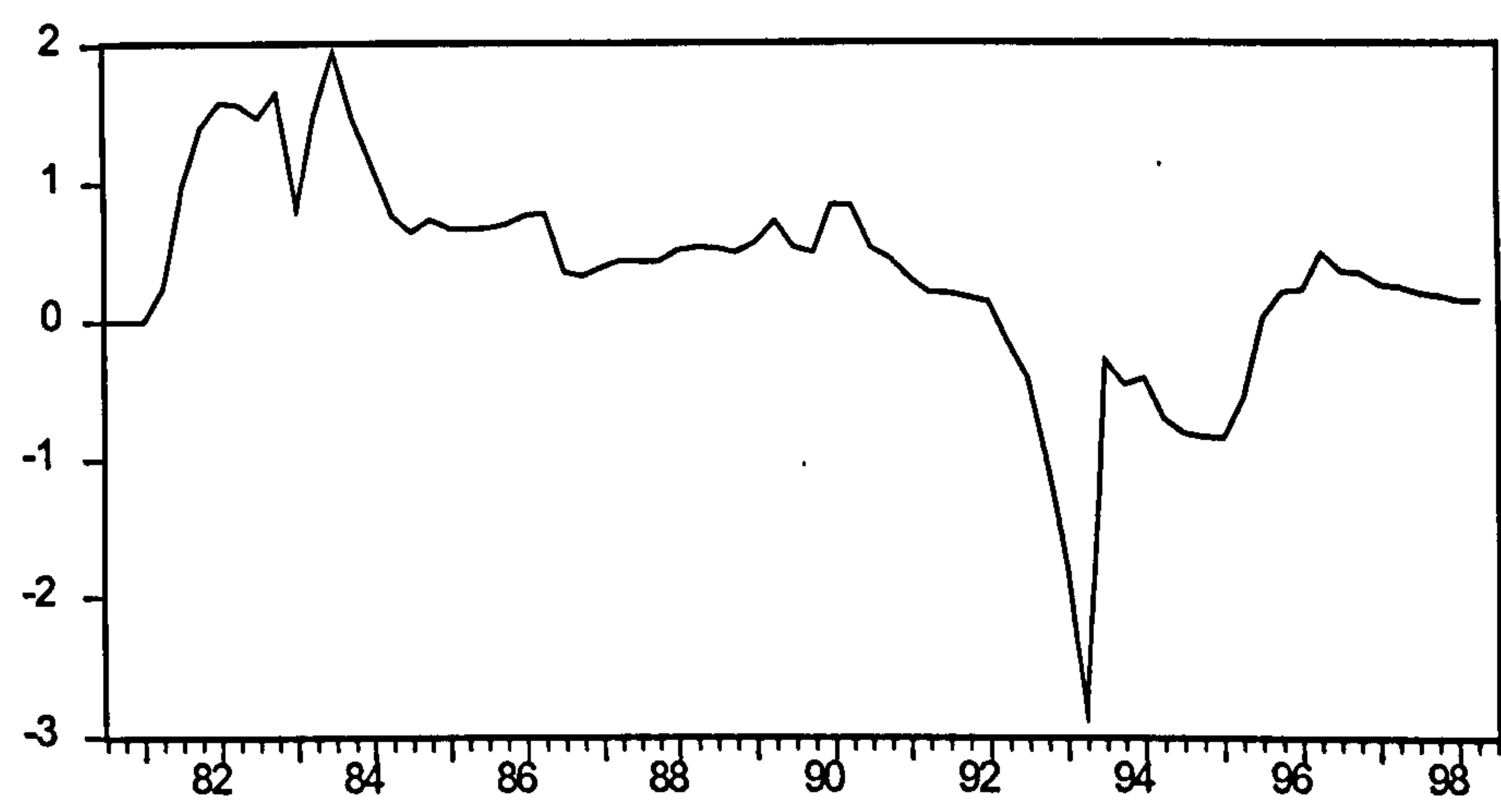


Figure 4.2 Test statistic for a break on the French adjustment coefficient towards the long-run relationship France Germany

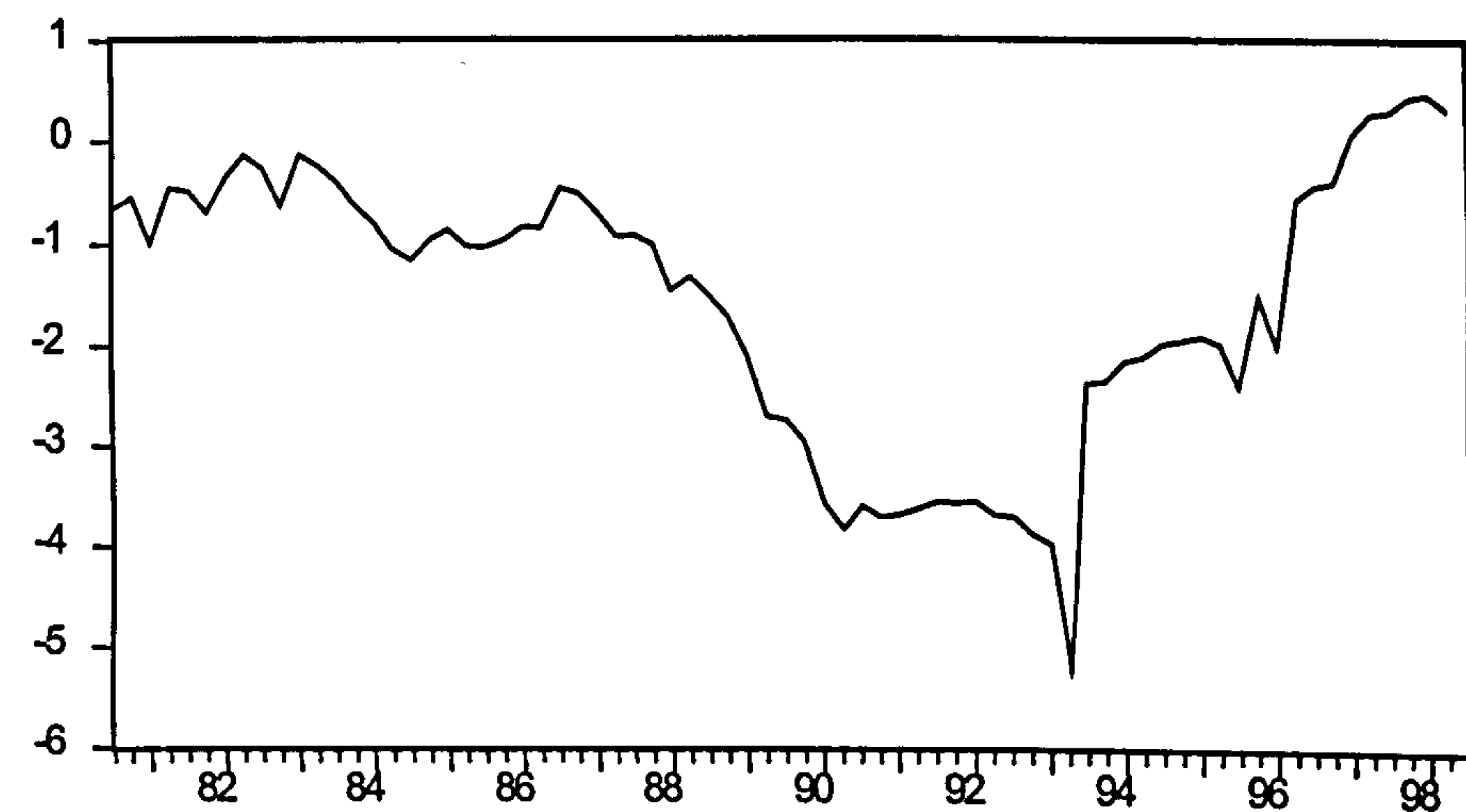


Figure 4.3 Test statistic for a break on the French adjustment coefficient towards the long-run relationship France Japan

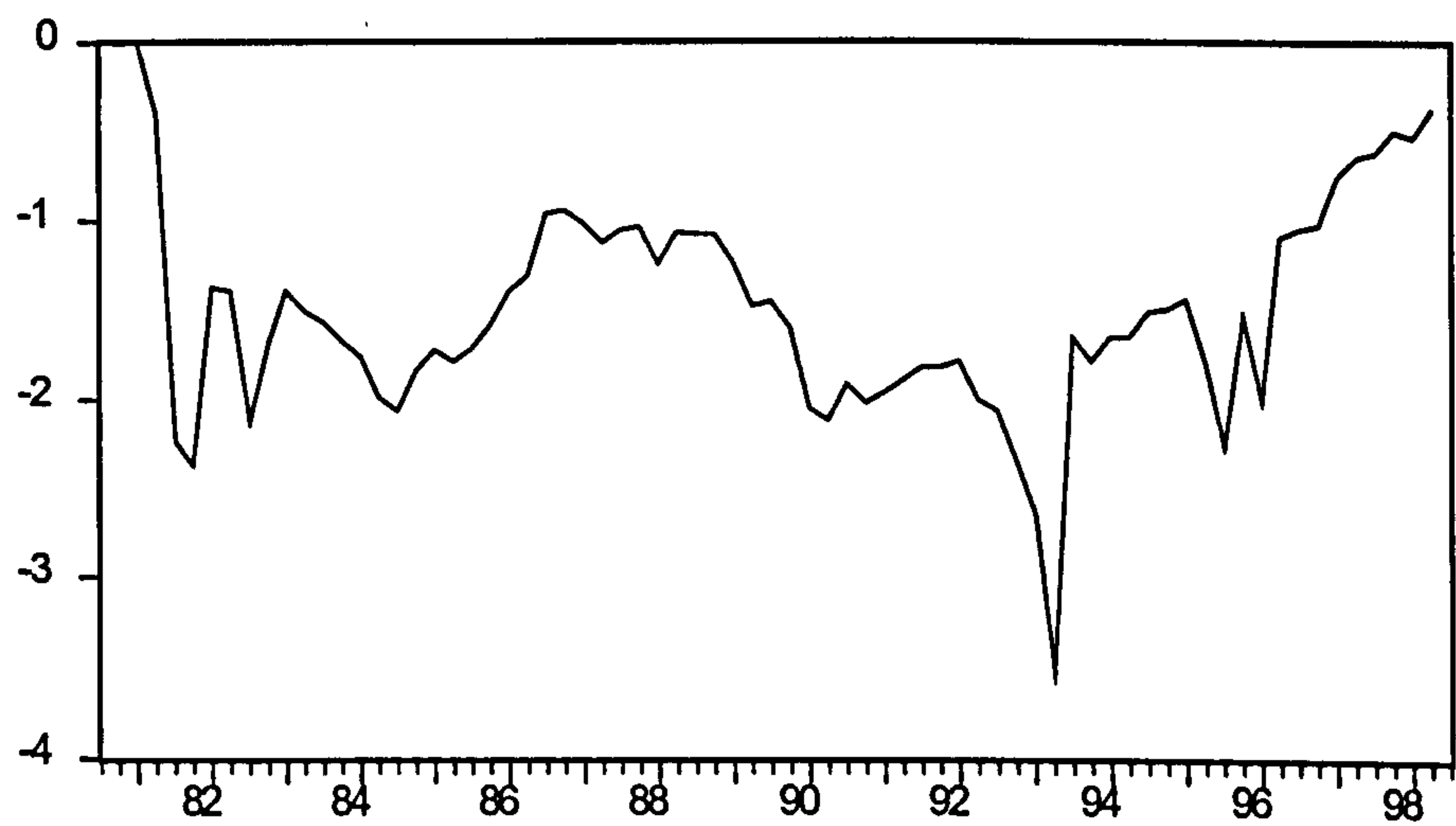


Figure 4.4 Test statistic for a break on the Italian adjustment coefficient towards the long-run relationship Italy-France

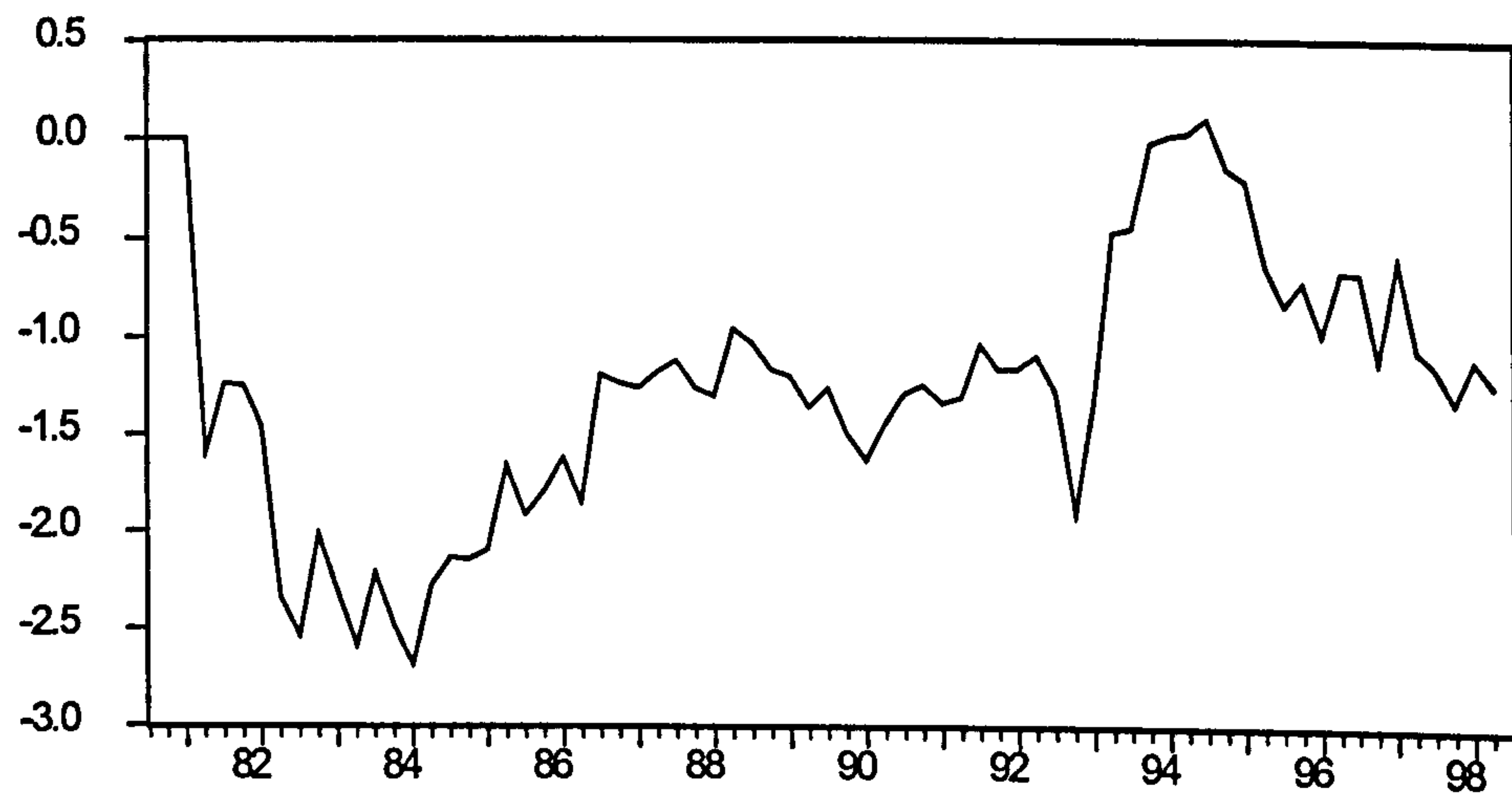


Figure 4.5 Test statistic for a break on the Italian adjustment coefficient towards
the long-run relationship Italy-Germany

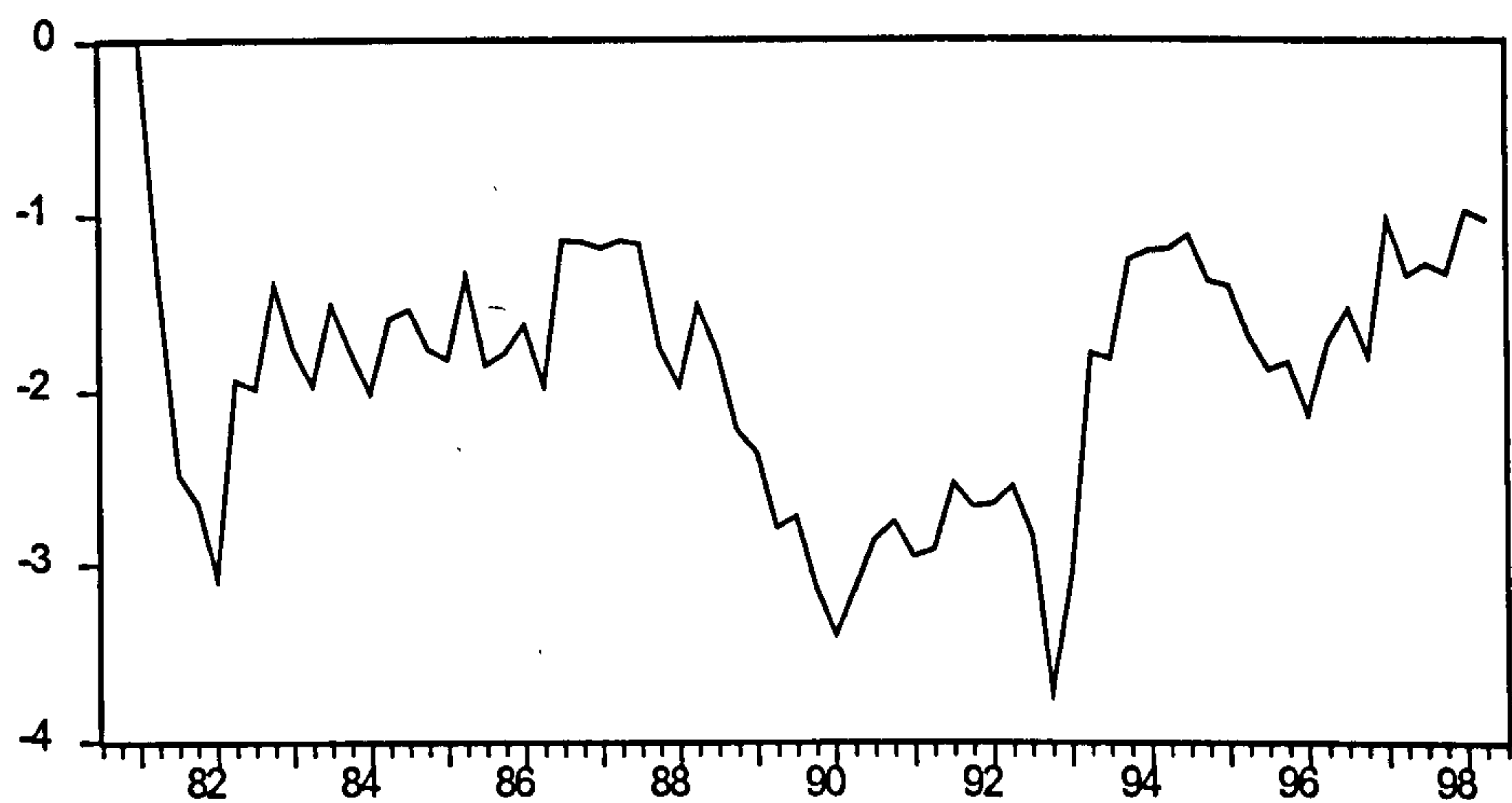


Figure 4.6 Test statistic for a break on the Italian adjustment coefficient towards
the long-run relationship Italy-Japan

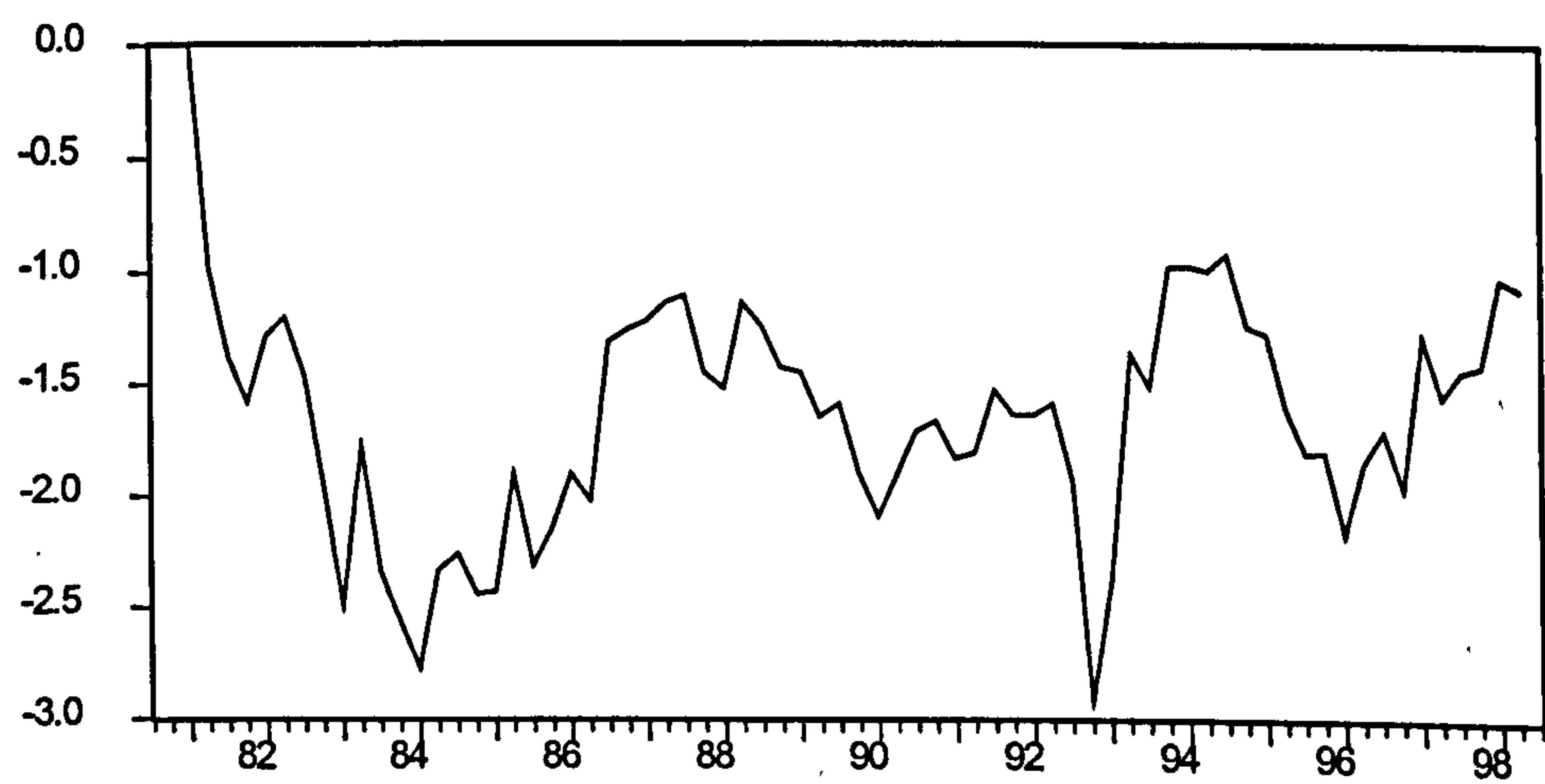


Figure 4.7 Test statistic for a break on the Japanese adjustment coefficient towards the long-run relationship Japan-Canada

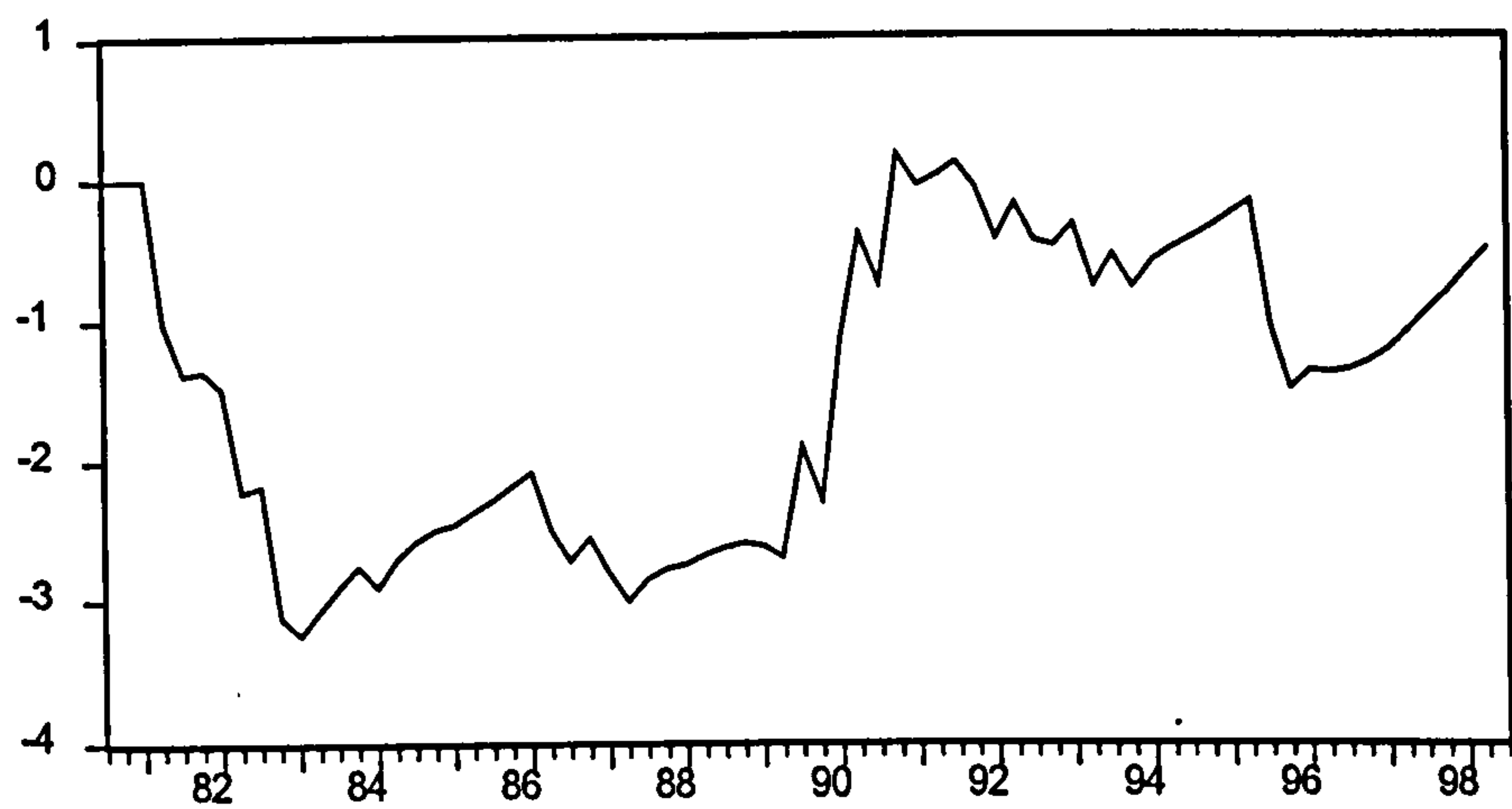


Figure 4.8 Test statistic for a break on the Japanese adjustment coefficient towards the long-run relationship Japan-Usa

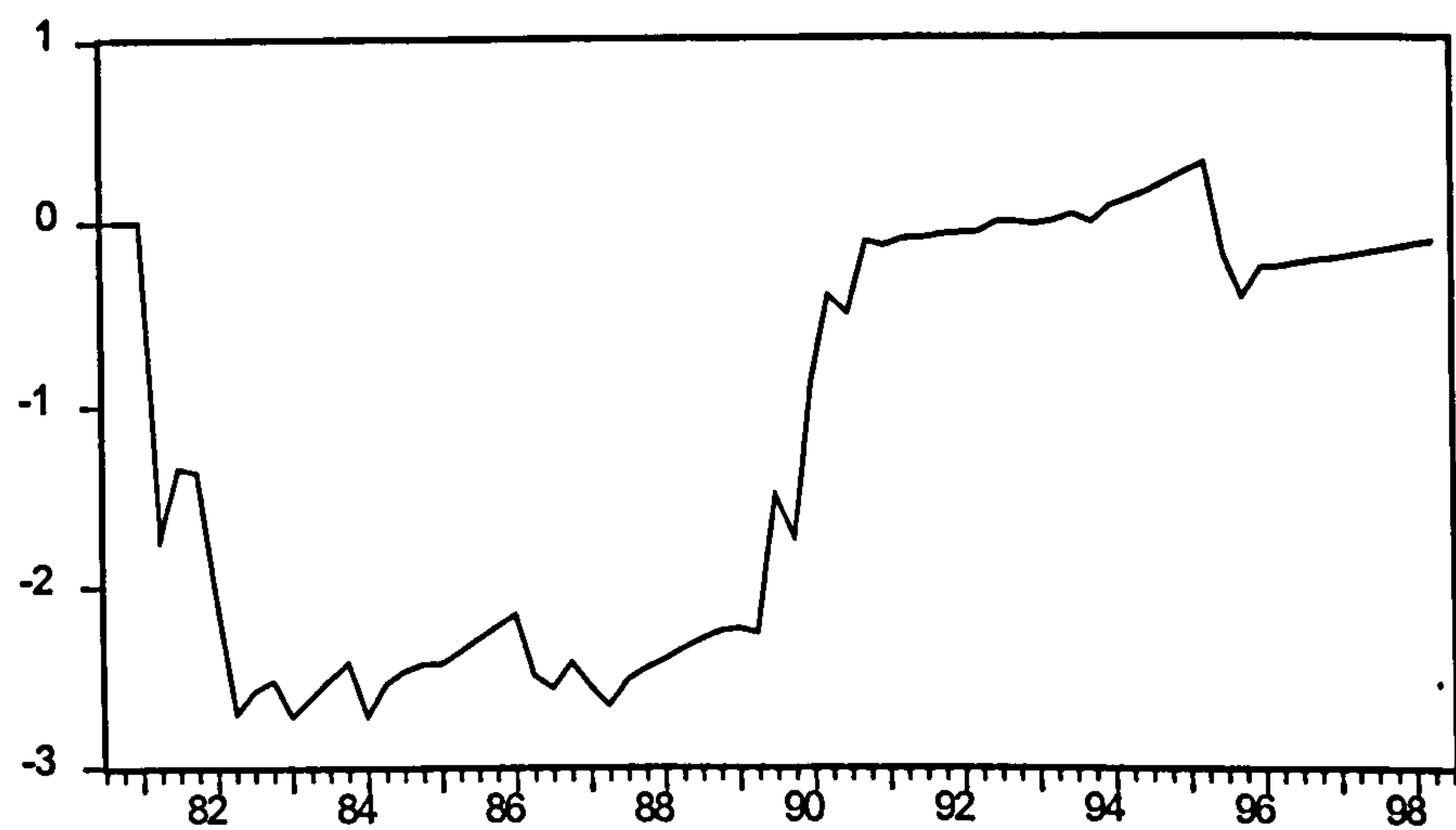


Figure 4.9 Test statistic for a break on the UK adjustment coefficient towards the long-run relationship UK-Germany

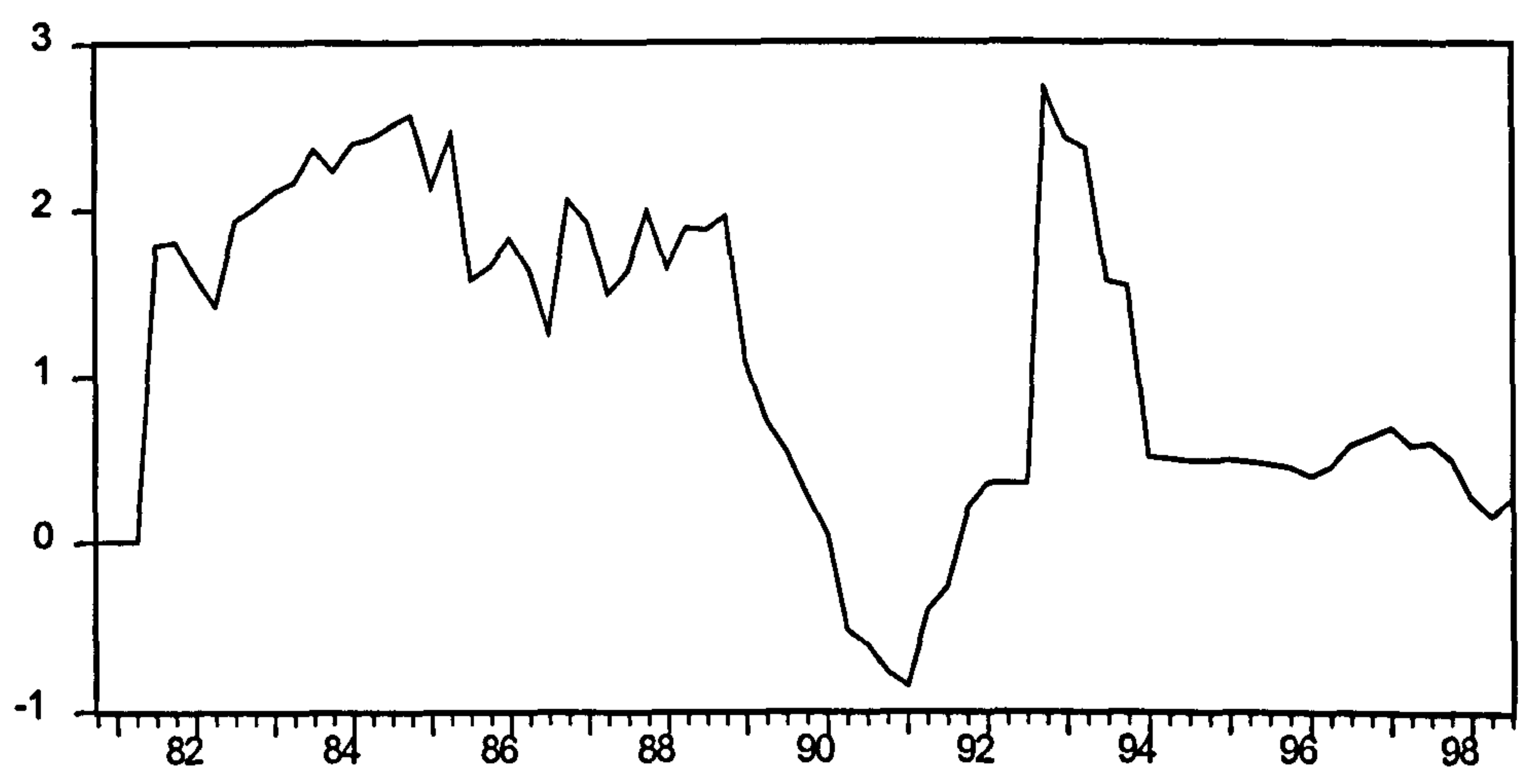


Figure 4.10. Time path of adjustment coefficients in error correction equations for Canada

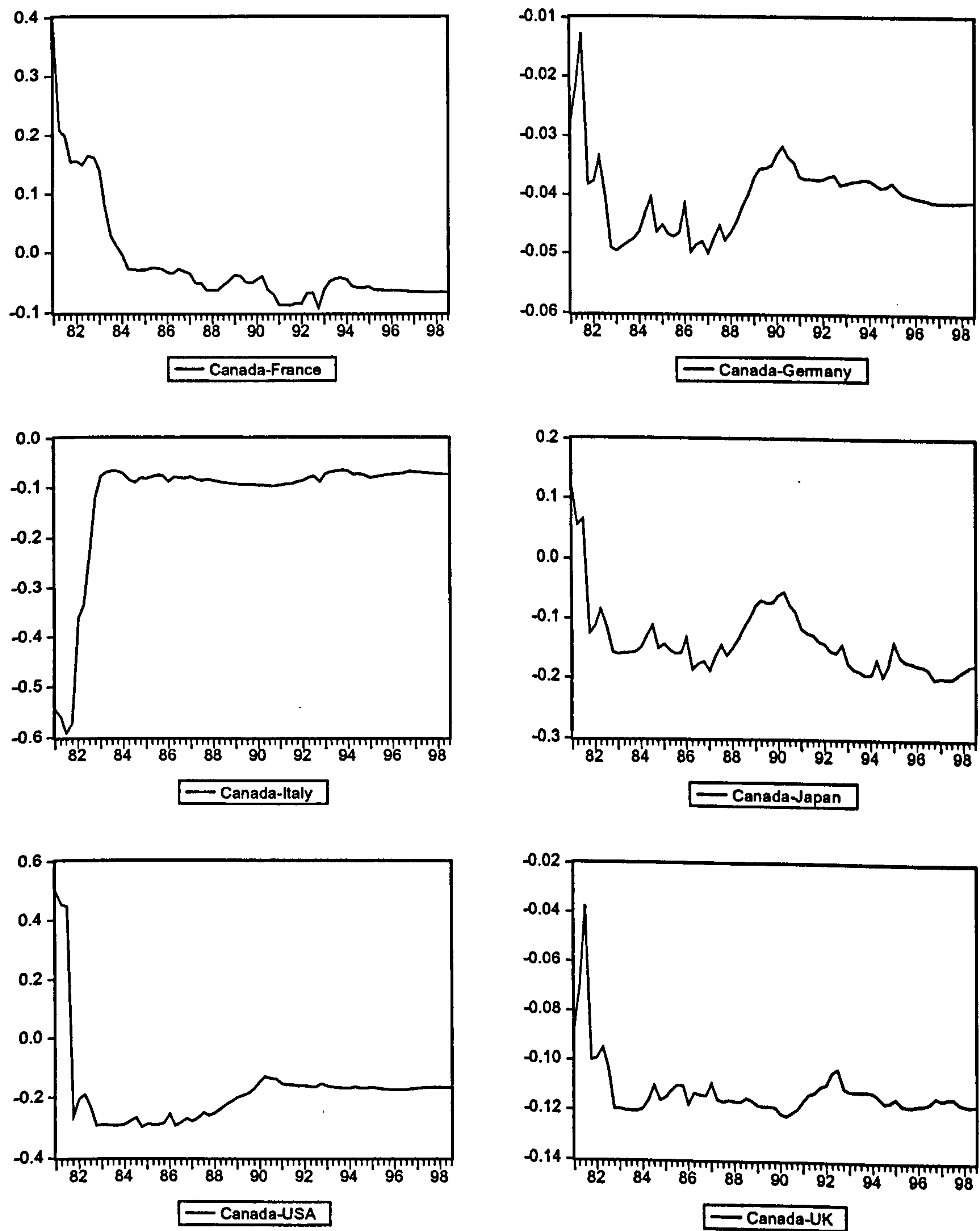


Figure 4.11. Time path of adjustment coefficients in error correction equations for France

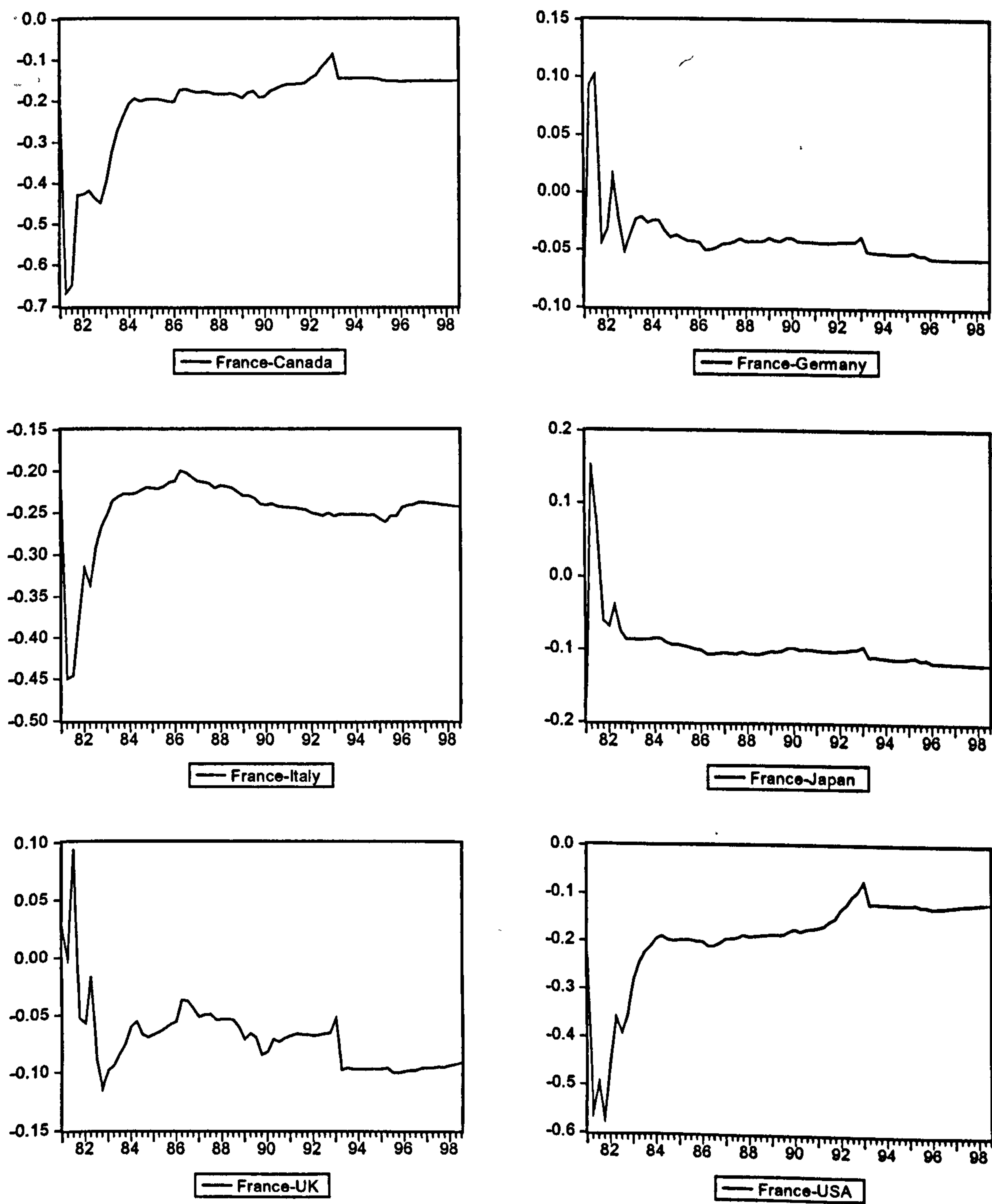


Figure 4.12. Time path of adjustment coefficients in error correction equations for Germany

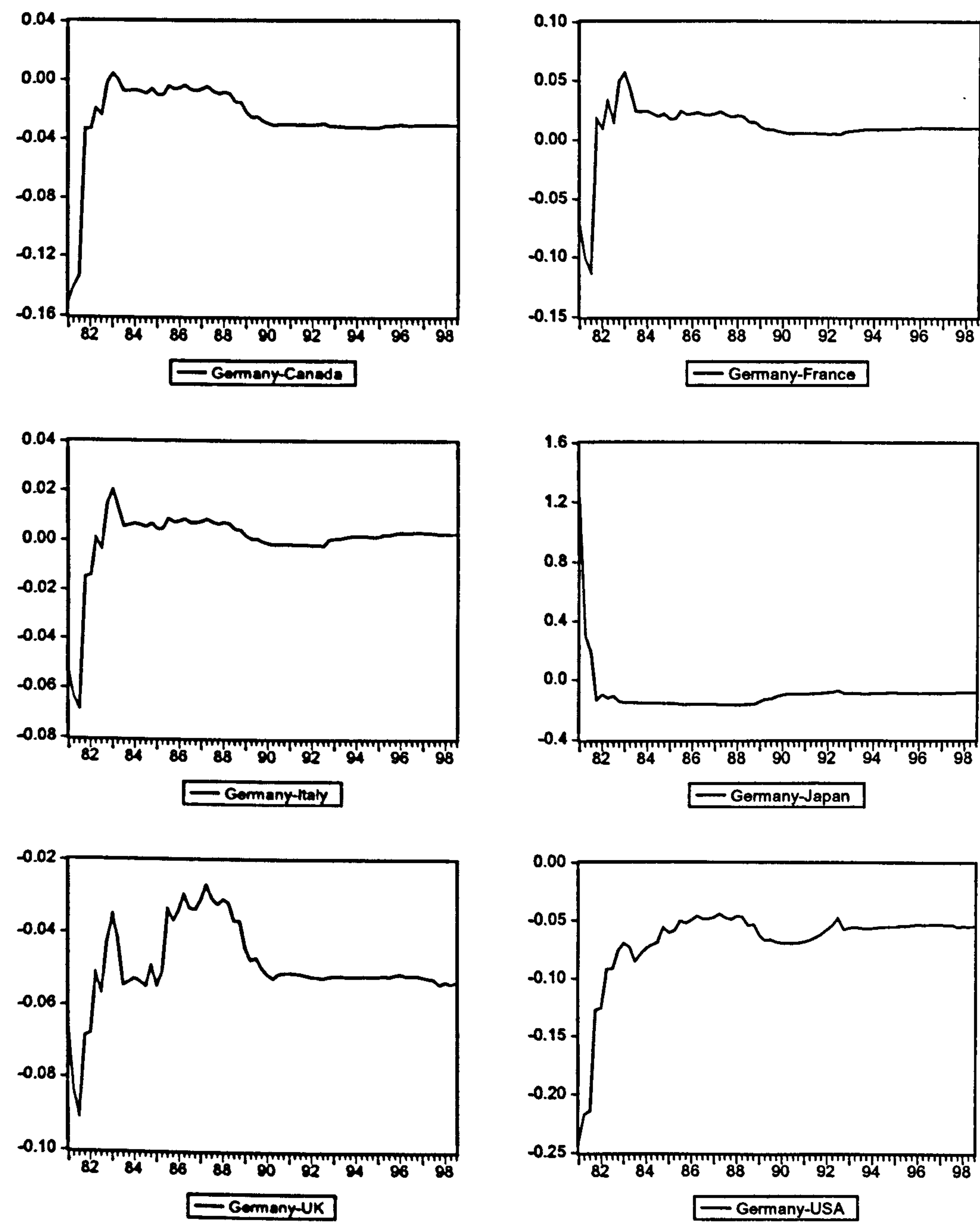


Figure 4.13. Time path of adjustment coefficients in error correction equations for Italy

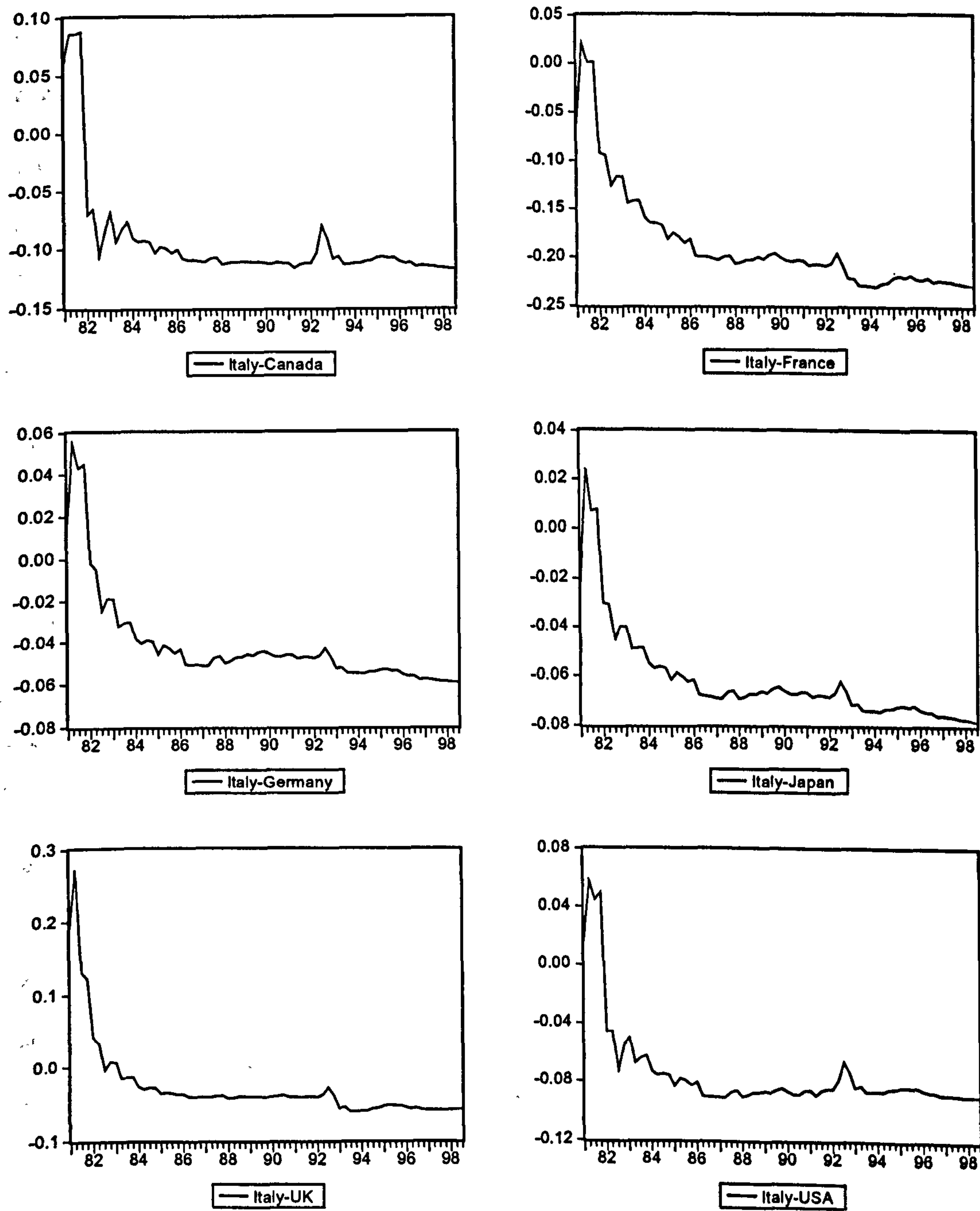


Figure 4.14. Time path of adjustment coefficients in error correction equations for Japan

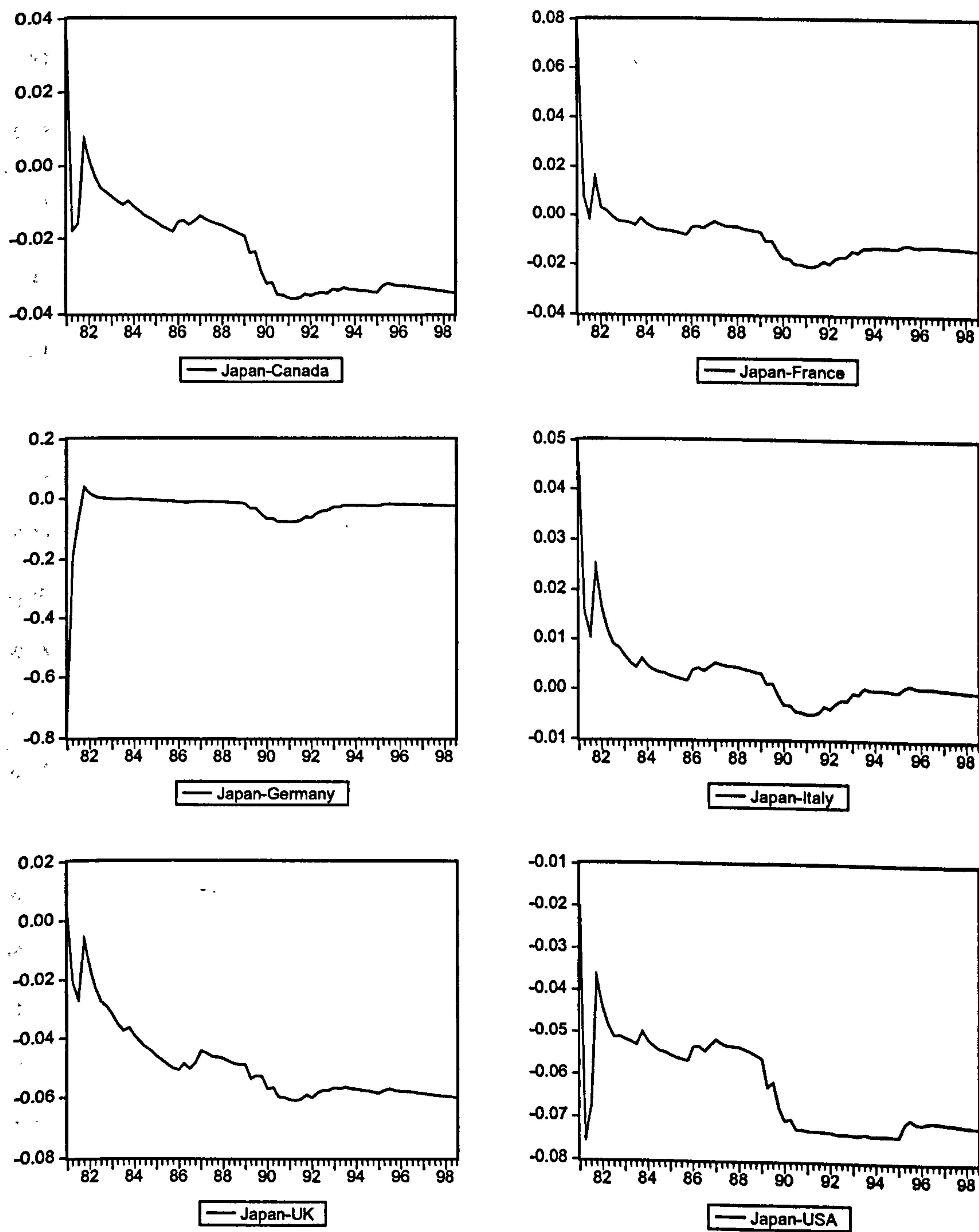
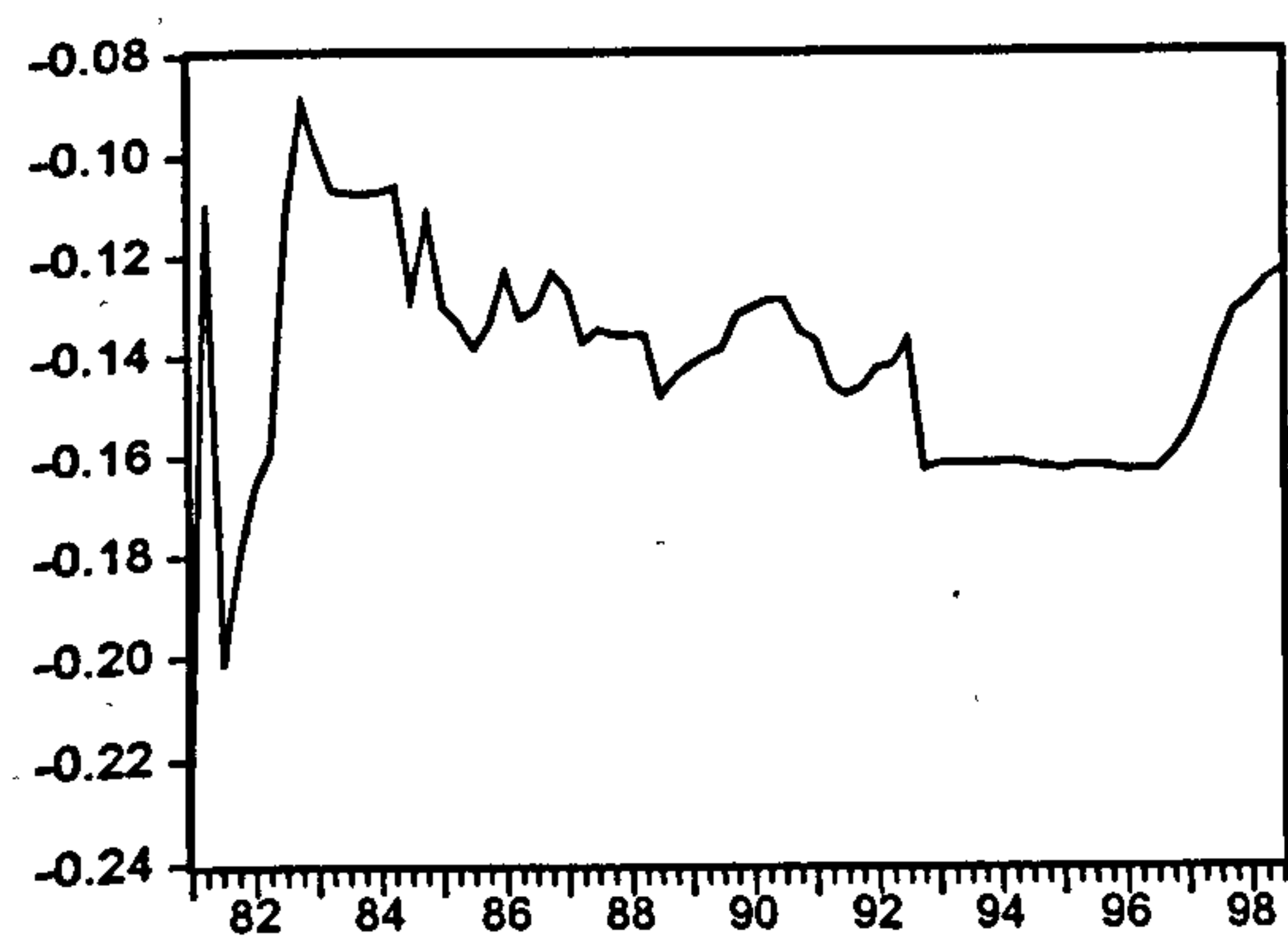
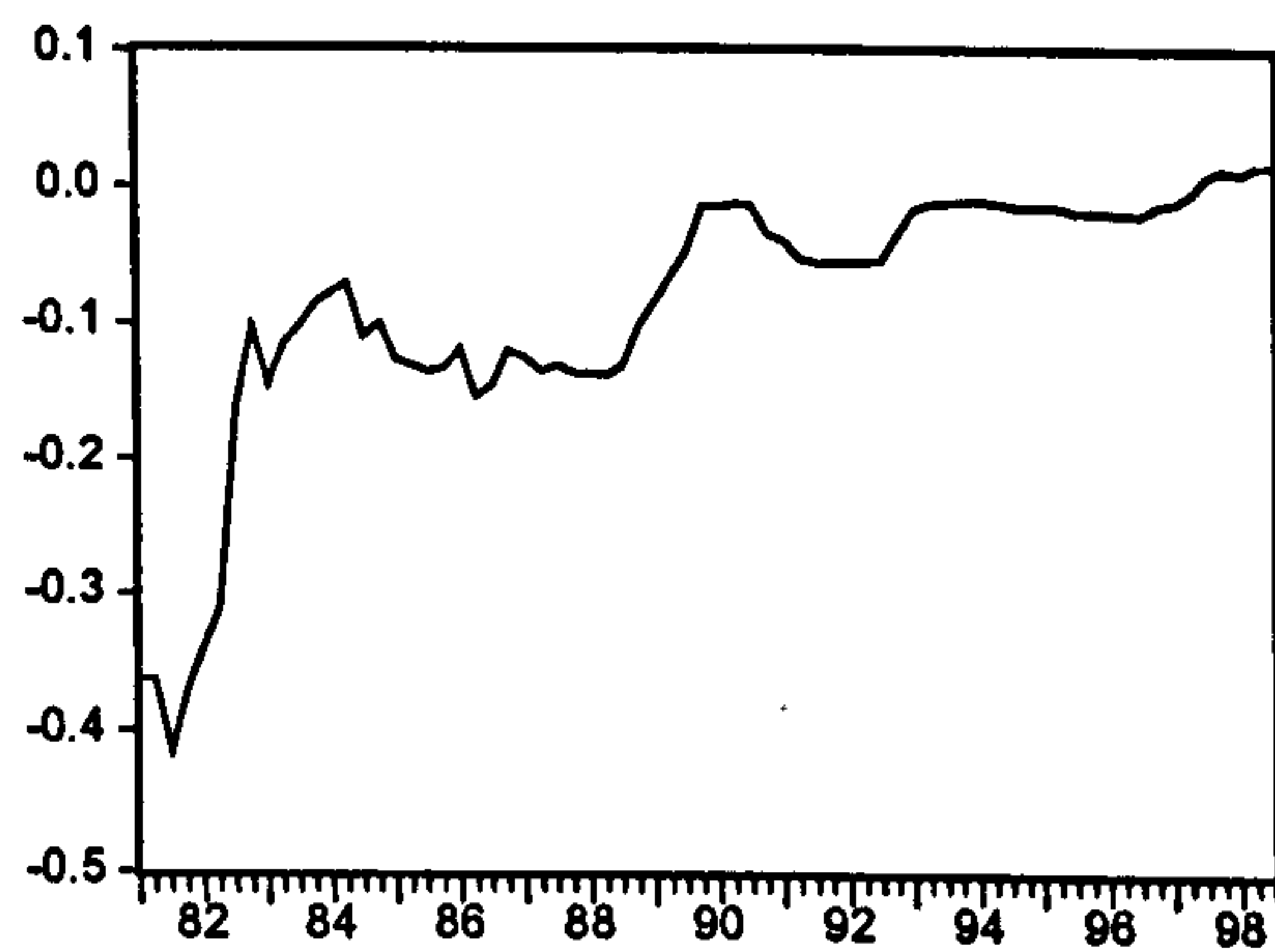


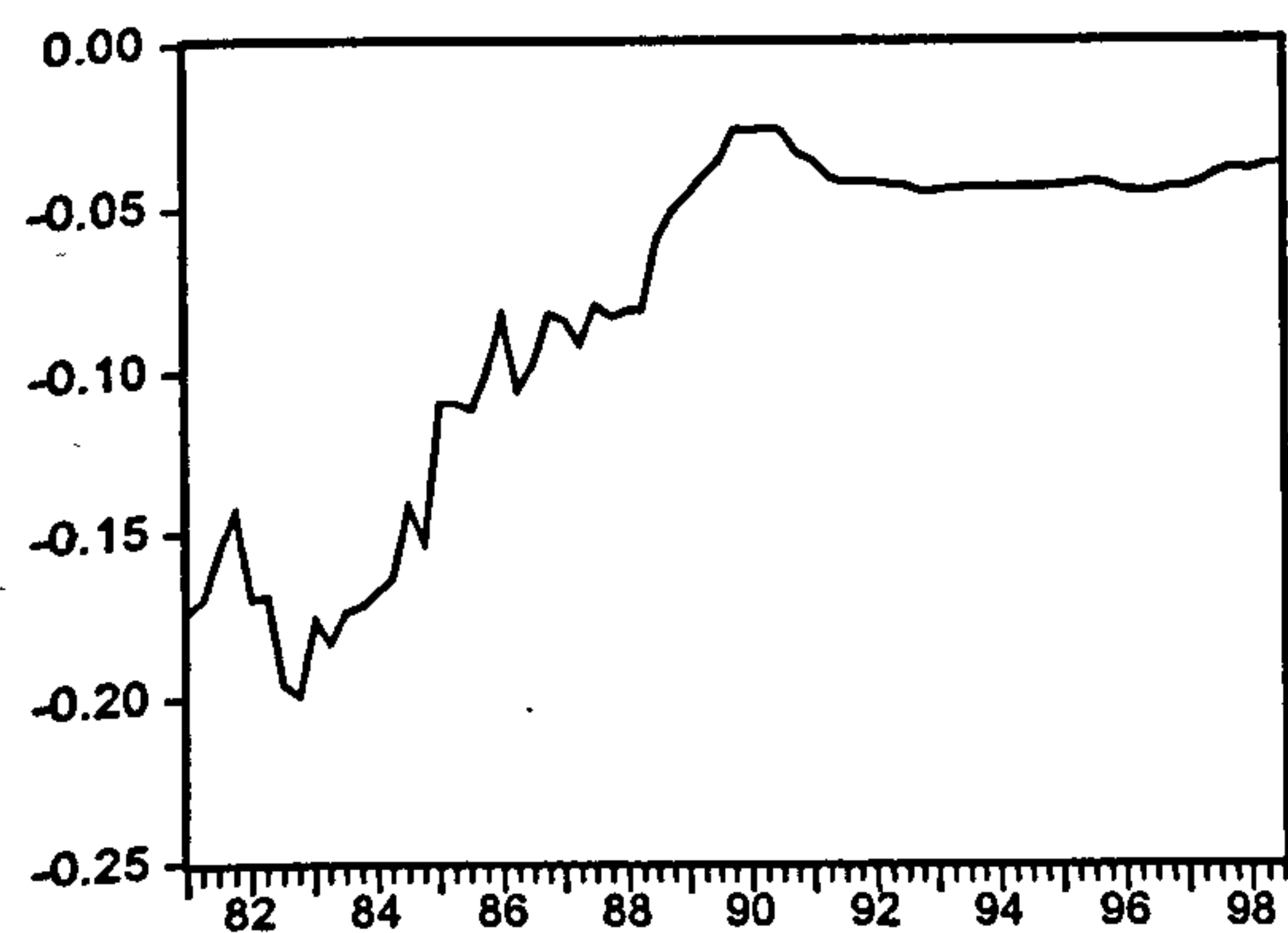
Figure 4.15. Time path of speed of adjustment coefficients in error correction equations for
UK



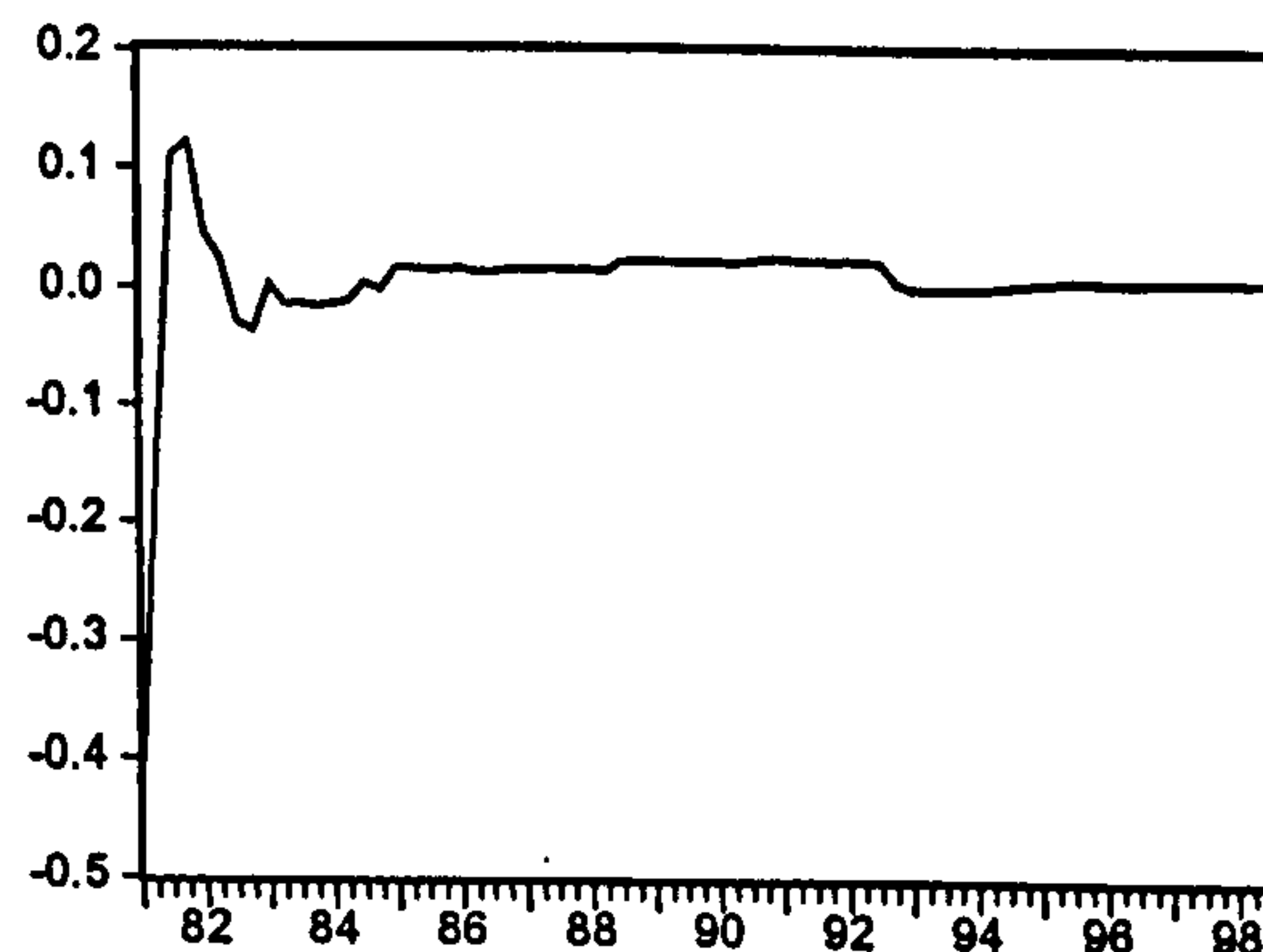
— UK-Canada



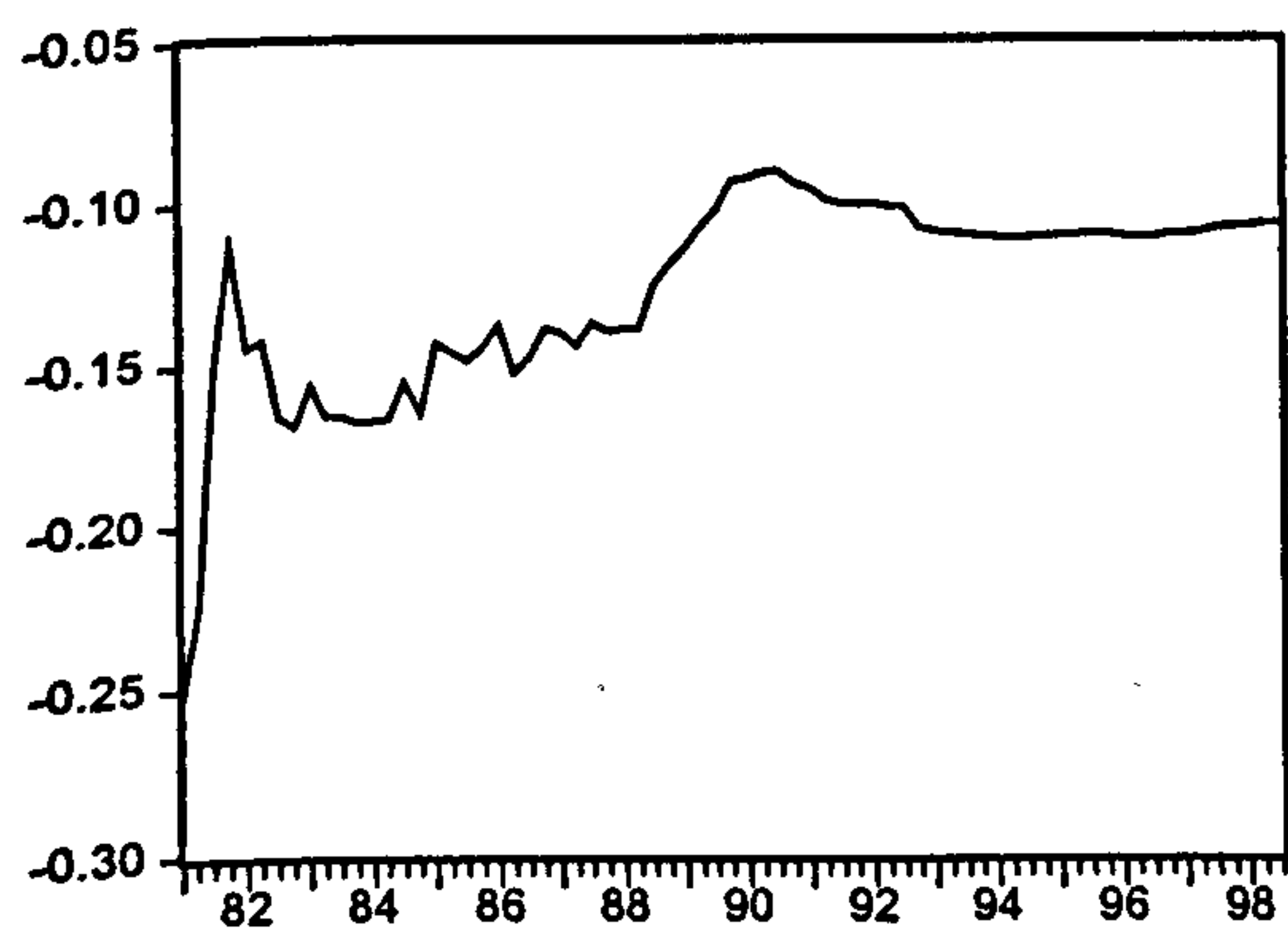
— UK-France



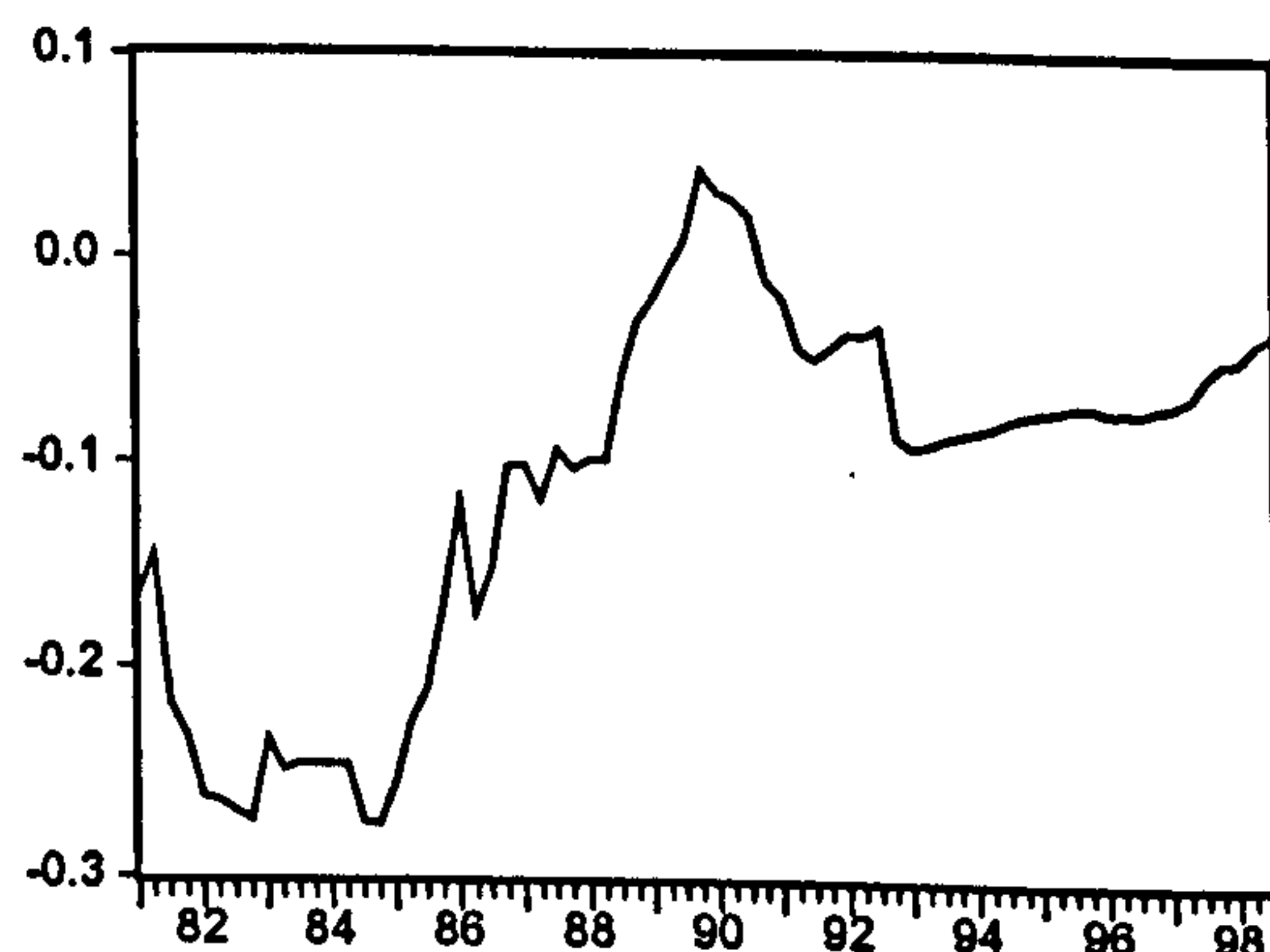
— UK-Germany



— UK-Italy

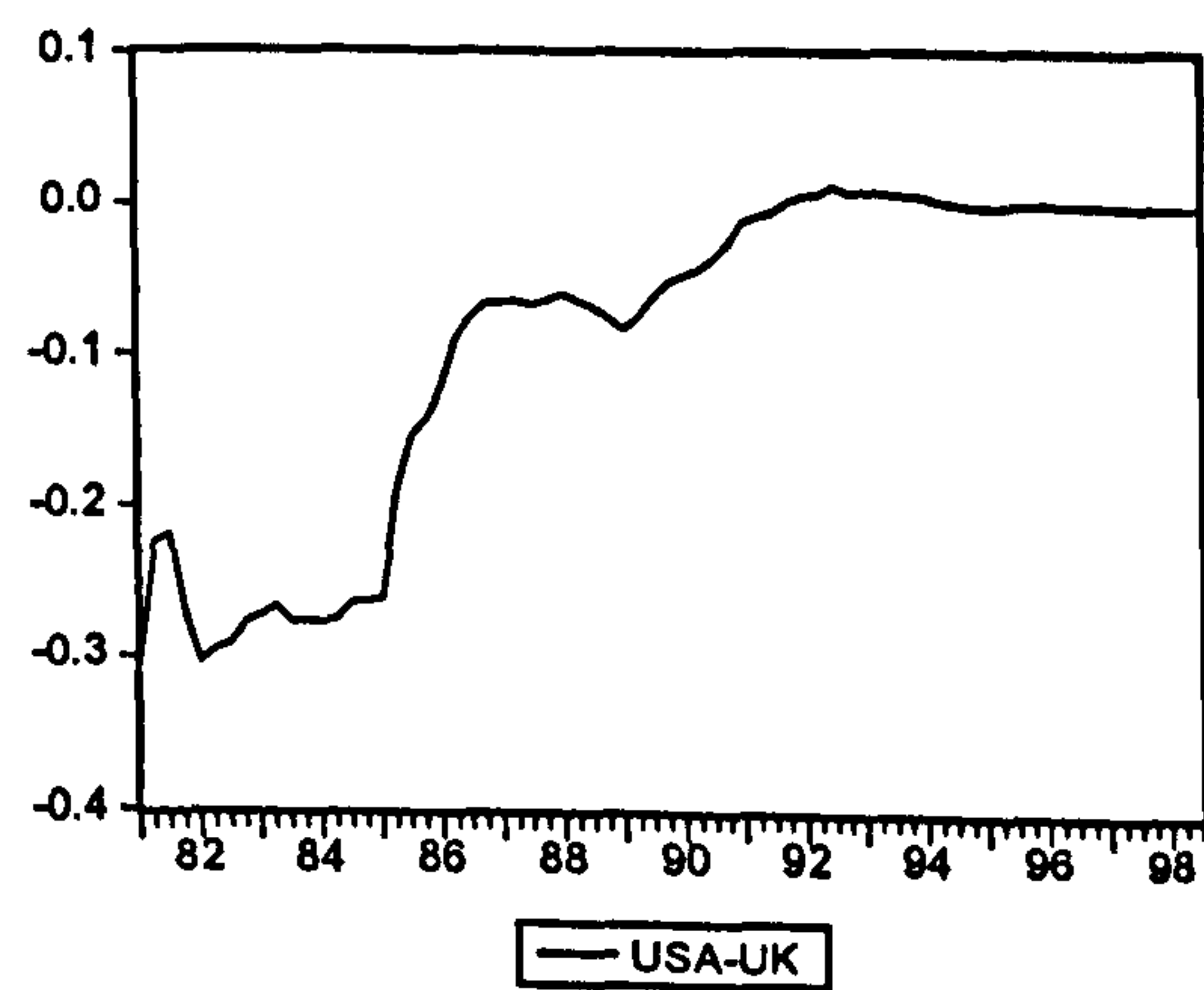
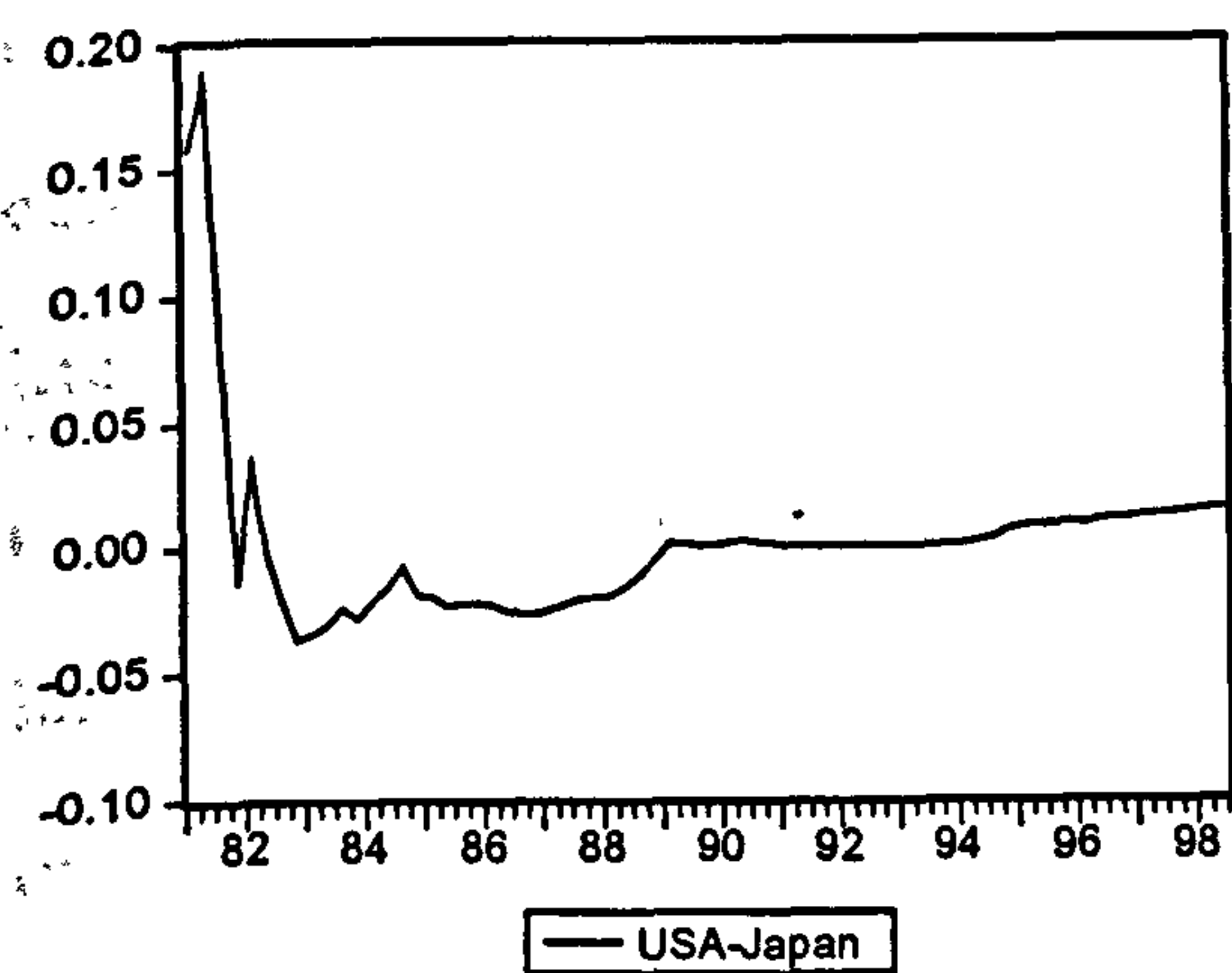
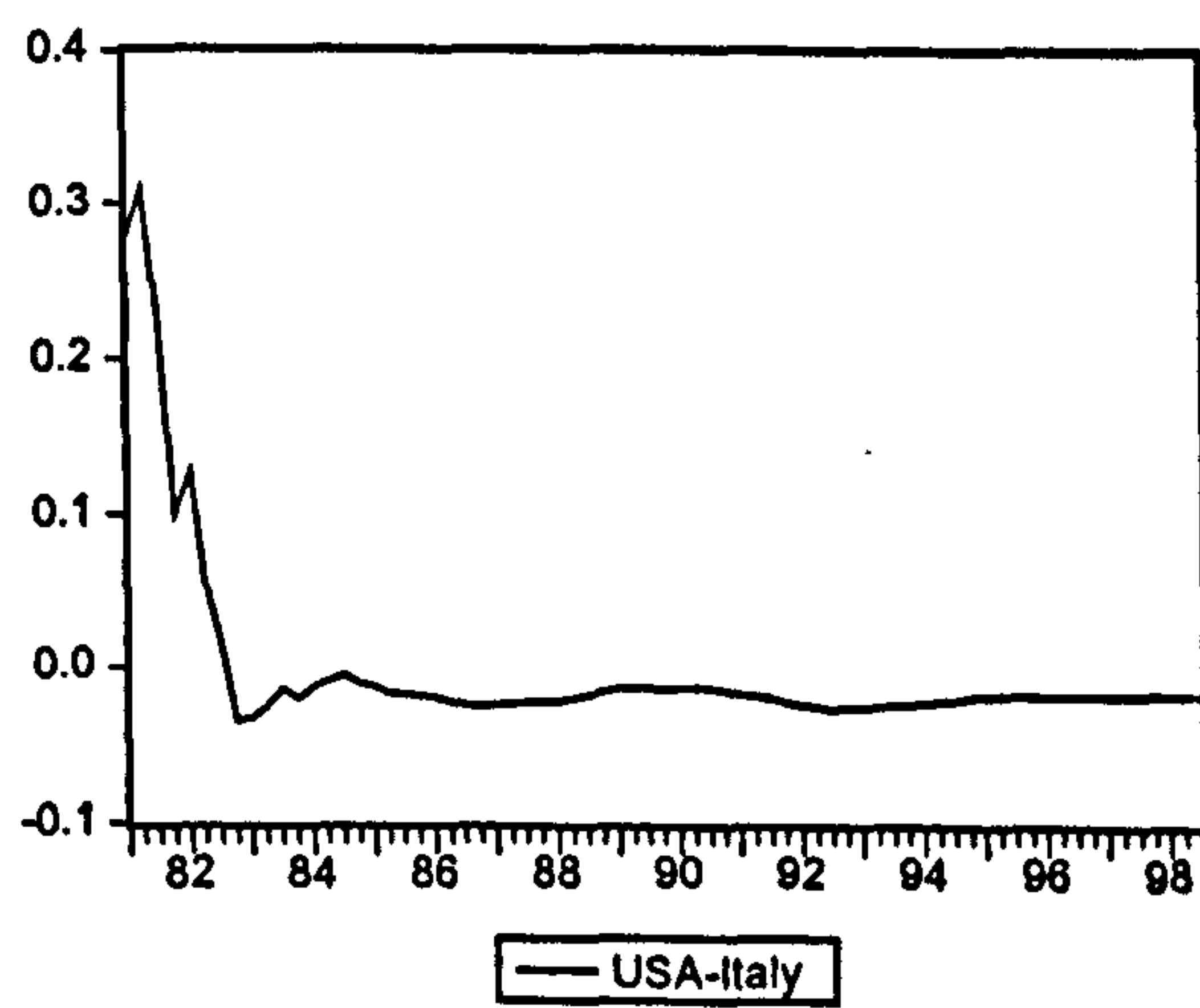
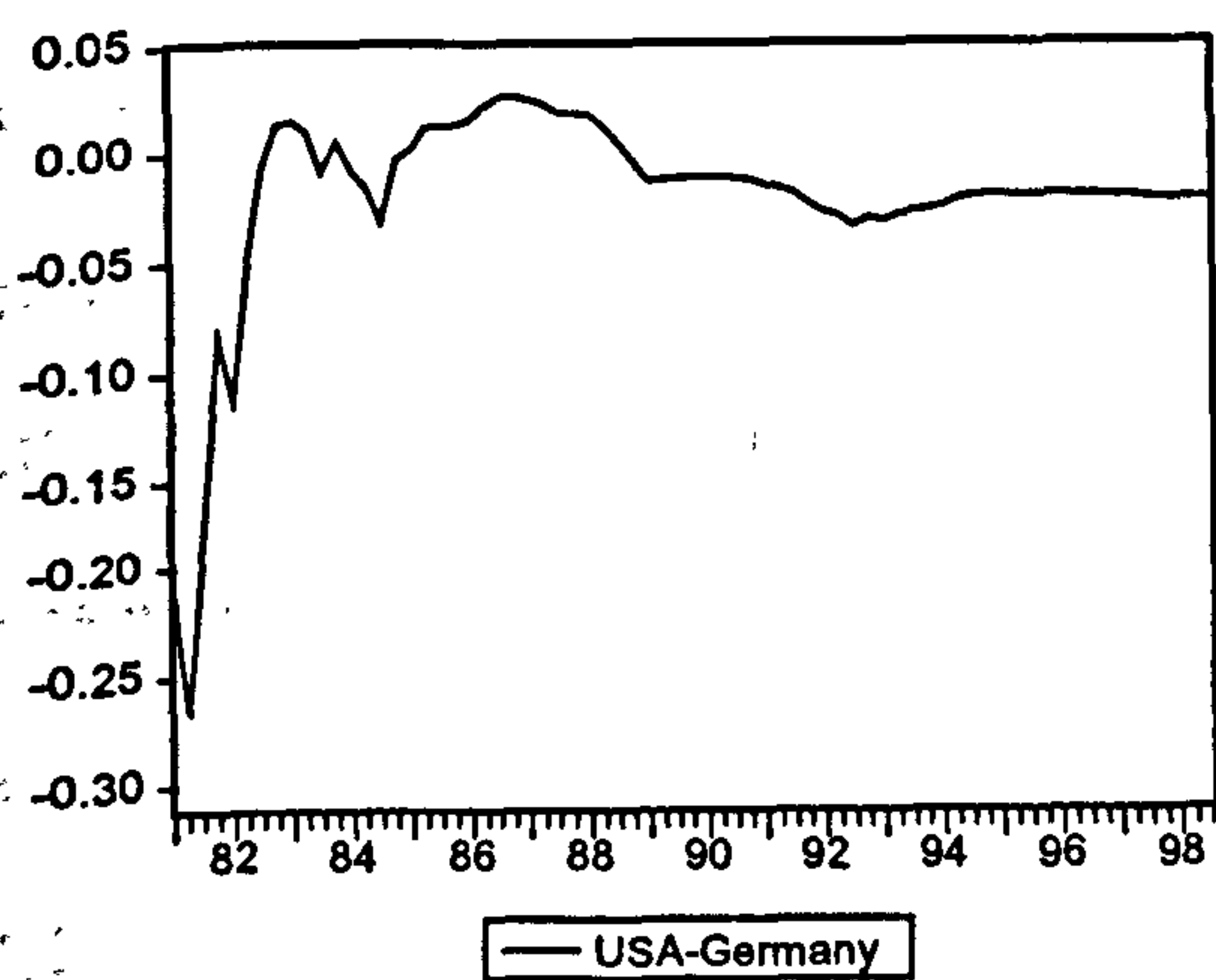
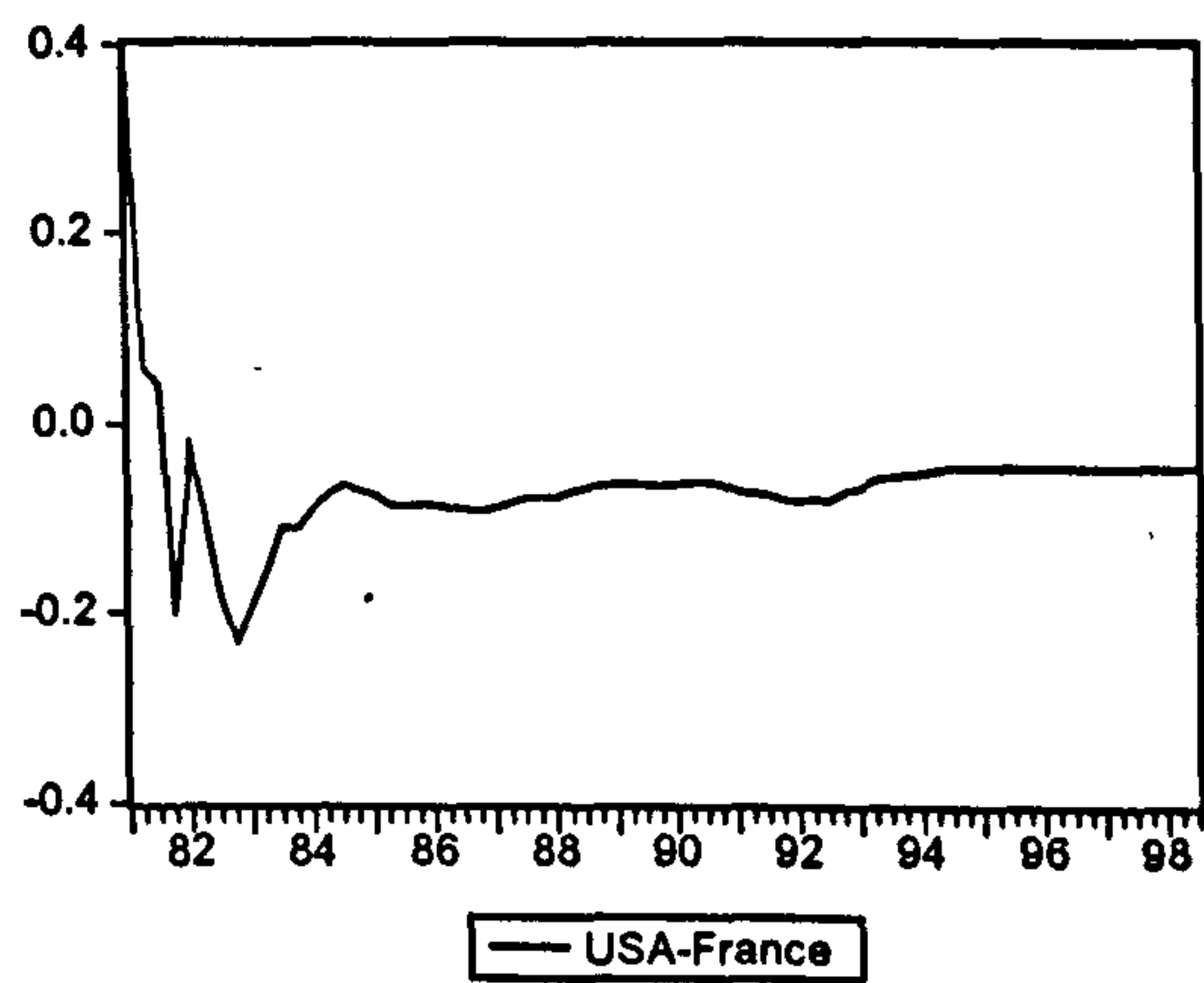
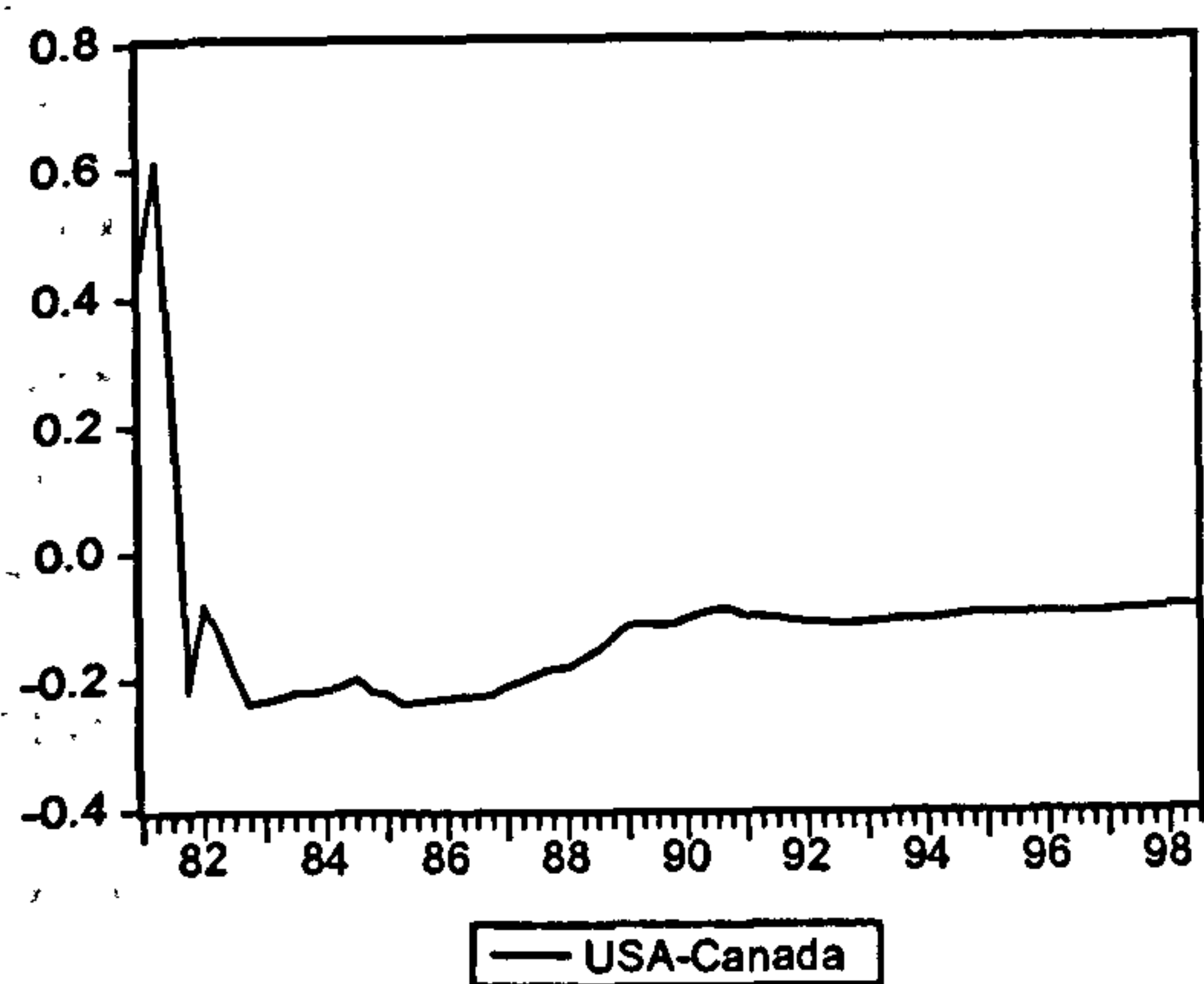


— UK-Japan



— UK-USA

Figure 4.16. Time path of adjustment coefficients in error correction equations for USA



Chapter 5

Conclusions

In this thesis we have investigated some issues related to the analysis of cointegrated time series. Specifically, we studied the identification of the structural cointegrating relations, and the identification of the long-run causal structure that links the variables with particular reference to the case in which this structure breaks down. We have reviewed the relevant literature in chapter 1. In chapter 2 we have introduced the extended Davidson Methodology, and we have used it to examine the structural linkages that exist between the G7 short- and long-term interest rates. Application of this method has shown the importance of testing for irreducibility as a diagnostic. The criterion of minimum variance has allowed us to distinguish between structural and solved IC vectors without any prior theoretical assumptions. Furthermore, we have performed Likelihood Ratio tests on the loading factors to all IC relations in order to gather information on the causal structure that links the rates.

The results obtained in chapter 2 can be summarised as follows. As far as the short rates are concerned, first, we found the rank of the system to be 6 when the EDM approach is taken, compared to 4, obtained using the Johansen (1988) procedure. Second, we found six irreducible structural relations, of which the two most significant ones involve the US and Canada, and Italy and France. Third, exogeneity tests seem to indicate a US world-wide leadership and reject the hypothesis of a German leadership in Europe, therefore confirming the findings of Caporale and Williams (1998b) and of other authors (see, e.g., Katsimbris and Miller, 1993). In brief, US and Canada appear to constitute the fundamental block, UK rates respond more to non-European rates than to other European countries, Italy is following France, and France and Germany respond to world rates rather than to each other. Last, it is

important to mention the role of Japan that acts as the link between European rates and the US one.

Going to the results concerning the long rates, we have obtained a rank of four for the system of long-term rates compared to a rank of two as indicated by the Johansen (1988) rank test on the whole group of series. More crucially we have been able to rule out the Italian and German rates as they appear to be non-cointegrated with the other series.

We have been able to isolate four irreducible structural relations that link the US and Canada, and the latter with the UK, France and Japan respectively. It appears that the driving force is Canada, which is linked to the US, UK and France in three out of the four fundamental relations, and which seems to be a reference point for the US that behave in an accommodating fashion. Notice that long-term rates in Italy and Germany, being non-cointegrated, seem to be determined by factors that are specific to these two countries. To investigate this matter further, we have analysed their expected inflation rates that are likely to affect the respective risk premia. The results obtained provide evidence that non-convergence of Italian and German rates is justifiable on the basis of non-convergence of their risk premia, and therefore interest rates in these two countries are independent from every other country's rates.

In chapter 3, we have examined, using Monte Carlo methods, several alternative approaches to detecting structural changes in the causality structure of cointegrated VARs. First, we have adopted recursive techniques as misspecification tests for the detection of instabilities in the adjustment coefficients and of the breakpoints. The results obtained show that recursive eigenvalues are fail to highlight a break in the causal structure of the models, whilst the recursive coefficients only show non-constancy of the parameters of interest.

We have then adopted a rolling FIML-based technique for the estimation of the model. The main result is that this technique shows more convincingly the instability of the parameters of

interest and the true value of after the breakpoint. However, rather than a breakpoint it only detect a region where the break is likely to occur.

To overcome the drawbacks of these techniques, we proposed an OLS-based sequential test. For this test we derived the critical values for different sample sizes and studied its statistical properties. We found that the test has a non-standard distribution and has reasonable power at the 90% level even for smaller samples (e.g. 80, 100). Also, it is relatively efficient graphically in that it shows a kink at the breakpoint. We also found that, at that point, the numerical value of the coefficient of the sequential dummy corresponds to the true value of the parameter after the break has occurred. The possible occurrence of multiple breaks invited us to consider the use of the Kalman Filter as a testing procedure. We tested the performance of a time-varying version of the Kalman filter on the same DGPs, showing the graphic usefulness of such a technique as far as the breakpoint and the magnitude of the break are concerned. Regardless of the location of the breakpoint, this well established technique proves in fact its usefulness in both detecting the breakpoints and providing the true value of the parameters after the break has occurred.

In chapter 4 we applied the proposed sequential technique as well as the Kalman Filter to the causal linkages of the G7 short-term rates with interesting results. As a general observation we can say that the two procedures perform well and do not provide conflicting results. However, some of the results are expectedly more clearly interpretable by examining the Kalman Filter estimates. To outline our findings- first we did not find any cases of reversal of causality, second we did not find evidence of substantial changes in the feedback that the US, Canadian and German rates receive. Clearly, the US and Canada have consistently constituted the fundamental block of the G-7 group, with the US economy being a constant point of reference. Third, the absence of breaks in the adjustment coefficients in the equations for the German rate therefore confirming the weak leadership of Germany within the

Eurozone. Moreover, we found some evidence of breaks in the equations for the French and Italian rates as it seems that they start following German rates more closely after the collapse of the ERM. Another result suggests that Japan link its rate to the US and Canadian ones soon after the FED was targeting the monetary base. Lastly, the sequential test evidence points to a break in the causal linkages between the UK and the German rates after the third-fourth quarter of 1992 implying the disengagement of UK policies from developments in the EU. This progressive disengagement of UK from the policies pursued in the EU is more evident from the Kalman filter estimates. In the following period, UK rates seem to be linked much more to world rates, as shown by the higher speed of adjustment parameter after 1992 in the equations linking UK and world rates.

Appendix A

```
/* Gauss Routine for Kalman Filter estimation of a system with Break at 1/2  
SAMPLE SIZE 100 */
```

```
cls;  
library coint, pgraph, cml;  
cmlset;  
graphset;
```

```
ofile = "c:\\marco\\kal100.out" ;  
output file = ^ofile reset;
```

```
output on ;  
screen on ;
```

```
an = 104;  
init = 50;  
n = init+an;  
nrep = 1000;
```

```
@parameters values@
```

```
lambda= eye(2);
```

```
@ loading weights matrix @
```

```
let alpha1[2,1]=-0.25, 0;
```

```
@ cointegrating coefficients @
```

```
let beta1[1,2]= 1, -2;
```

```
@ coefficients of dynamics @
```

```
let g[2,2]= 0.3, 0, 0, 0.3;
```

```
@ structural VECM Matrix @
```

```
let aa[2,2]= 1, 0.5, 0, 1;
```

```
aaa = inv (aa);
```

```
let aa1[2,2]= 1, 0, 0.5, 1;
```

```
aaa1 = inv(aa1);
```

```
@ seed for generating error term @
```

```
seede = 3592152 ;
```

```

aa2 = zeros(nrep,an-3);
aa2p = zeros(nrep,an-3);
aa2s = zeros(nrep,an-3);
aa1 = zeros(nrep,an-3);
aa1p = zeros(nrep,an-3);
aa1s = zeros(nrep,an-3);

```

@ start loop for number of replications @

```

jj=1; do while jj<=nrep;

```

```

x = zeros(2,n);

```

@ generate a multimormal random error @

```

fn
multnorm(n,sigma) = rndns(n,rows(sigma),seede)*chol(sigma);
e = multnorm(n,lambda);
u = e';
u1 = u[:,1]; u2 = u[:,2];

```

@ generate a cointegrated VAR with break in alphas @

```

t=3; do while t<=init+an/2;

```

```

x[:,t]=x[:,t-1]+
((aaa*g)*(x[:,t-1]-x[:,t-2]))+aaa*(alpha1*beta1)*x[:,t-1]+(aaa*u[:,t]);

```

```

t=t+1; endo;

```

```

t=(init+an/2)+1; do while t<=n;

```

```

x[:,t]=x[:,t-1]+
((aaa1*g)*(x[:,t-1]-x[:,t-2]))+aaa1*(alpha1*beta1)*x[:,t-1]+(aaa1*u[:,t]);

```

```

t=t+1;
endo;

```

@ dataset @

```

p=x';
pp = trimr(p,0,1);
z = pp[init+1:n-1,:];
dz= diff(z,1);
dz_1= lagn(dz,1);

```

```

y1 = z[:,1];
y2 = z[:,2];
dy1 = dz[:,1];
dy2 = dz[:,2];
dy1_1 = dz_1[:,1];
dy2_1 = dz_1[:,2];

{ev,vec,lr1,lr2,s00,s0k,sk0,skk} = sj2(z,-1,1);

proc (8) = sj2(x,p,k);
    local dx, z, q, r0t, rkt, sig,a,c,b,d,e,lr1,lr2, zz ;
    local s00, s0k, sk0, skk, tmp, crit,i,lx, s, alpha ;

    dx  = diff(x,1) ;
    q   = 2 ;
    z   = lagn(dx,1) ;
    do until (q >= k);
    z = z~lagn(dx,q);
        q = q + 1 ;
    endo ;

    if (_NoDet NE 1) ;
        z   = detrend(trimr(z,k,0),p);
        dx  = detrend(trimr(dx,k,0),p);
        r0t = dx - z*(dx/z);
        lx  = detrend(trimr(lagn(x,k),k+1,0),p);
        rkt = lx - z*(lx/z);
    else ;
        z   = trimr(z,k,0);
        dx  = trimr(dx,k,0);
        r0t = dx - z*(dx/z);
        lx  = trimr(lagn(x,k),k+1,0);
        lx  = lx~ptrend(p,rows(lx)) ;
        rkt = lx - z*(lx/z);
    endif ;

    skk  = rkt'rkt/rows(rkt) ;
    sk0  = rkt'r0t/rows(rkt) ;
    s00  = r0t'r0t/rows(r0t) ;
    s0k  = sk0' ;
    sig  = sk0*inv(s00)*(sk0');

    {a,b,d,c} = eigrg2(inv(skk)*sig);

    d = d*inv(chol(d'skk*d)) ;

    if (_eigerr NE 0) ;
        "WARNING: _eigerr NOT zero! Eigerr = " _eigerr ;
    endif ;

```

```
/* Order the eigen values and eigen vectors */
```

```
zz = (-a)^(d');
zz = sortc(zz,1);
a = -zz[:,1];
d = zz[:,2:cols(zz)]'; /* Done */
```

```
/* Compute the trace statistics */
```

```
lr1 = zeros(rows(a),1);
i = 1;
do while (i <= rows(a));
    lr1[i,1] = -rows(rkt)*sumc(trimr(ln(1-a),i-1,0));
    i = i + 1;
endo;
```

```
/* Max eigen value statistics */
```

```
lr2 = -rows(rkt)*ln(1-a);
```

```
retp(-sortc(-a,1),d,lr1,lr2, s00, s0k,sk0,skk);
endp;
```

```
beta = evec[:,1]/(evec[1,1]);
alpha = s0k*beta*inv(beta'*skk*beta);
lrpi = alpha*beta';
om = s00-alpha*beta'*sk0;
cv = beta'*x;
cr = cv';
dx1 = dy1[2:an-2,];
dx2 = dy2[2:an-2,];
dx1_1 = dy1[1:an-3,];
dx2_1 = dy2[1:an-3,];
ect_1 = (cr[init+2:n-2,])';
zz = dx1_1~dx2_1~ect_1';
data = zz;
```

```
/******
```

Define the global constants

n = number of measured elements (rows of y)

k = number of states

s = number of shocks affecting the state

obs = number of observations (including predicted values)

```
*****/
```

```
_n = 2;
```

```
_k = 2;
```

```

_s = 2;
_obs = cols(ect_1);

@ Include the code to do the Kalman Filtering @

#include filter.src;

@ Specify the initial values of the hyper parameters @

parms =
1.15|
1.15|
0.03|    @ sigmaq1 @
0.03;    @ sigmaq2 @

@Specify the data@

_y = (dx2~dx1)'; /*dep var dy1|dy2*/

_Z[1,.] = ect_1;
_Z[4,.] = ect_1;
_R[1,.] = ones(1,_obs);
_R[4,.] = ones(1,_obs);
_T[1,.] = ones(1,_obs);
_T[4,.] = ones(1,_obs);
_P[1,1] = _infinity;
_P[4,1] = _infinity;
_H[1,.] = (parms[1]^2)*ones(1,_obs);
_H[4,.] = (parms[2]^2)*ones(1,_obs);

proc(0) = SysMat(parms);

    @ Define z @

_Z = _Z;

@ Define d @

_d[1,.] = (0.3*dx2_1') @+(-0.15*dx1_1')@;
_d[2,.] = (0.3*dx1_1') @+(-0.15*dx2_1')@;

@ Define H @

@ Define T @

@ Define c @

```

```
_c[1,] = zeros(1,_obs);
_c[2,] = zeros(1,_obs);
```

```
@ Define R @
```

```
_R[1,] = ones(1,_obs);
_R[4,] = ones(1,_obs);
```

```
@ Define Q @
```

```
_Q[1,] = parms[3]^2* ones(1,_obs);
_Q[4,] = parms[4]^2* ones(1,_obs);
```

```
@ Define the prior state vector a0 if it varies with parameters @
_a0 = zeros(2,1);
```

```
@ Define the prior precision matrix P0 using the @
```

```
_P0 = eye(2) .* 1e5;
```

```
endp;
```

```
@ Do the filtering by calling the sysmat and filter procs. @
```

```
call SysMat(parms);
```

```
call Filter;
```

```
call Smoother;
```

```
call cmlprt(cml(data',0,&LL,parms));
```

```
@ Show the filtering output graphically @
```

```
aaaaa2 = _a[1, .];
aaaaa2p = _ap[1, .];
aaaaa2s = _as[1, .];
```

```
aa2[jj,] = aaaaa2;
aa2p[jj,] = aaaaa2p;
aa2s[jj,] = aaaaa2s;
```

```
aaaaa1 = _a[2, .];
aaaaa1p = _ap[2, .];
aaaaa1s = _as[2, .];
```

```
aa1[jj,] = aaaaa1;
```

```

aalp[jj,] = aaaaalp;
aals[jj,] = aaaaals;
jj= jj+1; endo;

ava2= trimr(meanc(aa2),2,0);
ava2p= trimr(meanc(aa2p),2,0);
ava2s= trimr(meanc(aa2s),2,0);
sdava2= trimr(stdc(aa2),2,0);
sdava2p= trimr(stdc(aa2p),2,0);
sdava2s= trimr(stdc(aa2s),2,0);

ssdava2 = sdava2*2;
ssdava2p = sdava2p*2;
ssdava2s = sdava2s*2;

a2cihi = (ava2 + ssdava2);
a2pcihi = (ava2p + ssdava2p);
a2scihi = (ava2s + ssdava2s);
a2cilo = (ava2 - ssdava2);
a2pcilo = (ava2p - ssdava2p);
a2scilo = (ava2s - ssdava2s);

ava1= trimr(meanc(aa1),2,0);
ava1p= trimr(meanc(aa1p),2,0);
ava1s= trimr(meanc(aa1s),2,0);
sdava1= trimr(stdc(aa1),2,0);
sdava1p= trimr(stdc(aa1p),2,0);
sdava1s= trimr(stdc(aa1s),2,0);

ssdava1 = sdava1*2;
ssdava1p = sdava1p*2;
ssdava1s = sdava1s*2;

a1cihi = (ava1 + ssdava1);
a1pcihi = (ava1p + ssdava1p);
a1scihi = (ava1s + ssdava1s);
a1cilo = (ava1 - ssdava1);
a1pcilo = (ava1p - ssdava1p);
a1scilo = (ava1s - ssdava1s);

begwind;
window(2,2,0);
xlabel("time");
ylabel("coefficient value");
figure=seqa(1,1,rows(ava2));

setwind(1);

title("Alpha2 one step ahead forecast");
xy(figure,(a2pcilo~ava2p~a2pcihi));

```

```

nextwind;
  title("Alpha2 smoothed");
  xy(ffigure,(a2scilo~ava2s~a2scihi));
nextwind;
  title("Alpha1 one step ahead forecast");
  xy(ffigure,(a1pcilo~ava1p~a1pcihi));

nextwind;
  title("Alpha1 smoothed");
  xy(ffigure,(a1scilo~ava1s~a1scihi));
endwind;

```

```

print "      a2cilo      ava2      a2cihi" a2cilo~ava2~a2cihi;
print;
print "      a2pcilo      ava2p      a2pcihi" a2pcilo~ava2p~a2pcihi;
print;
print "      a2scilo      ava2s      a2scihi" a2scilo~ava2s~a2scihi;
print;

```

@ Initialise the global system matrices and filtration
output matrices @

```

_y = zeros(_n,_obs);
_Z = zeros(_k*_n,_obs);
_d = zeros(_n,_obs);
_H = zeros(_n*_n,_obs);

_T = zeros(_k*_k,_obs);
_c = zeros(_k,_obs);
_R = zeros(_k*_s,_obs);
_Q = zeros(_s*_s,_obs);

_a = zeros(_k,_obs);
_a0 = zeros(_k,1);
_ap = _a;
_as = _a;

_P = zeros(_k*_k,_obs);
_P0 = zeros(_k*_k,1);
_Pp = _P;
_Ps = _P;

_LL = zeros(1,_obs);

```

@ Make forward declarations of proceddures @
external proc SysMat;

```

_infinity = 1e5;

@ Define the filtration

proc(0) = Filter;

local t; @ loop counter @

@ local versions of the relevant system matrices @
local Ppt,Tt,Ht,Qt,Rt,Zt,Xt,Ft,Ft_inverse,vt;

Zt = reshape(_Z[:,1],_n,_k);
Ht = reshape(_H[:,1],_n,_n);
Tt = reshape(_T[:,1],_k,_k);
Rt = reshape(_R[:,1],_k,_s);
Qt = reshape(_Q[:,1],_s,_s);

_ap[:,1] = Tt * _a0 + _c[:,1];
Ppt = Tt * _P0 * Tt' + Rt * Qt * Rt';
_Pp[:,1] = vecr(Ppt);

local Wt,mask;
mask = _y[:,1] ./= miss(0,0);
Wt = selif(diagrv(zeros(_n,_n),mask),mask);

Ft = Wt * (Zt * Ppt * Zt' + Ht) * Wt';
Ft_inverse = invpd(Ft);
vt = Wt * (missrv(_y[:,1],0) - Zt * _ap[:,1] - _d[:,1]);

_a[:,1] = _ap[:,1] + Ppt * Zt' * Wt' * Ft_inverse * vt;
_P[:,1] = vecr(Ppt - Ppt' * Zt' * Wt' * Ft_inverse * Wt * Zt * Ppt);
_LL[:,1] = -0.5*(sumc(mask)*ln(2*pi)+ln(det(Ft))+vt'*Ft_inverse*vt);

for t (2,cols(_y),1);

Zt = reshape(_Z[:,t],_n,_k);
Ht = reshape(_H[:,t],_n,_n);
Tt = reshape(_T[:,t],_k,_k);
Rt = reshape(_R[:,t],_k,_s);
Qt = reshape(_Q[:,t],_s,_s);

_ap[:,t] = Tt * _a[:,t-1] + _c[:,t];
Ppt = Tt * reshape(_P[:,t-1],_k,_k) * Tt' + Rt * Qt * Rt';
_Pp[:,t] = vecr(Ppt);

mask = _y[:,t] ./= miss(0,0);
Wt = selif(diagrv(zeros(_n,_n),mask),mask);

if not scalmiss(Wt); @ we have some current data @

```

```

Ft = Wt * (Zt * Ppt * Zt' + Ht) * Wt';
Ft_inverse = invpd(Ft);
vt = Wt * ( missrv(_y[:,t],0) - Zt * _ap[:,t] - _d[:,t]);

_a[:,t] = _ap[:,t] + Ppt * Zt' * Wt' * Ft_inverse * vt;
_P[:,t] = vecr(Ppt - Ppt' * Zt' * Wt' * Ft_inverse * Wt * Zt * Ppt);
_LL[:,t] = -0.5 * (sumc(mask)*ln(2*pi) + ln(det(Ft)) + vt' * Ft_inverse * vt);

else; @ we have no data for the current period @

_a[:,t] = _ap[:,t];
_P[:,t] = _Pp[:,t];
_LL[:,t] = 0;

endif;

endfor;

endp;

@ Define the fixed interval smoother (Assumes that the filtering has been completed!) @

proc(0) = Smoother;

_as[:,cols(_y)] = _a[:,cols(_y)];
_Ps[:,cols(_y)] = _P[:,cols(_y)];

@ local versions of the relevant system matrices @
local Pstar,Pt,Ppt_lead;

local t; @ loop counter @

for t (cols(_y)-1,1,-1);

Pt = reshape(_P[:,t],_k,_k);
Ppt_lead = reshape(_Pp[:,t+1],_k,_k);
Pstar = Pt * reshape(_T[:,t+1],_k,_k)' * invpd(Ppt_lead);

_Ps[:,t] = vecr(Pt + Pstar * (reshape(_Ps[:,t+1],_k,_k) - Ppt_lead) * Pstar');

_as[:,t] = _a[:,t] + Pstar * (_as[:,t+1] - _ap[:,t+1]);

endfor;

endp;

@ Loglikelihood computation procedure @

proc(1) = LL(p,d);

```

```
call SysMat(p);  
call Filter;  
retp(_LL');
```

```
endp;
```

```
end;
```

References

- Andrews D.W.K. and W. Ploberger (1994) "Optimal tests when a nuisance parameter is present only under the alternative", *Econometrica*, 62, 1383-1414
- Artis, M.J. and W. Zhang (1998), "The linkage of interest rates within the EMS", *Weltwirtschaftliches Archiv*, 134, 1, 117-132.
- Banerjee, A., Lazarova, S. and G. Urga (1998), "Bootstrapping sequential tests for multiple structural breaks", D.P.no.17-98, Centre for Economic Forecasting, London
- Banerjee, A., Lumsdaine R. L. and J. H. Stock (1992) "Recursive and sequential tests of the unit-root and trend break hypotheses: theory and international evidence", *Journal of Business and Economic Statistics*, 10, 271-287
- Beeby, M., Funke, M. and S.G. Hall (1998), "The Relationship Between Money and Prices: An Econometric Appraisal Based on Cointegration and Causality", in Heji C., Hans S., Hanzon B. and Praagman K. (eds), *System Dynamics in Economic and Financial Models*, John Wiley Series in Financial Economics and Quantitative analysis, Chichester.
- Begg, D., Giavazzi, F., De Grauwe, P., Uhlig, H., and C. Wyplosz (1998), *The ECB: Safe at Any Speed?*, CEPR Report on Monitoring the European Central Bank.
- Boswijk, H.P.,(1992) "Cointegration, identification and exogeneity: Inference in structural error correction models", (Thesis Publishers Amsterdam)

Boswijk, H.P.,(1994) "Testing for an unstable root in conditional and structural error correction models", *Journal of Econometrics* 63, 37-60

Boswijk, H.P.,(1995) "Efficient inference on cointegration parameters in structural error correction models", *Journal of Econometrics* 69, 133-158

Bruneau, C. and E. Jondeau (1999), "Long-run causality, with an application to international links between long-term interest rates", *Oxford Bulletin of Economics and Statistics*, 61, 4, 545-568.

Caporale, G.M. and G. Williams (1998a), "Bond markets and macroeconomic performance", D.P. no.13-98, Centre for Economic Forecasting, London Business School.

Caporale, G.M. and G. Williams (1998b), "International linkages in short- and long-term interest rates", D.P. no.14-98, Centre for Economic Forecasting, London Business School.

Caporale, G.M. and G. Williams (1998c), "Can interest rate policy be independent?", *Economic Outlook*, 22, 4, 10-15.

Caporale, G.M. and N. Pittis (1999), "Efficient estimation of cointegrating vectors and testing for causality in vector autoregressions", *Journal of Economic Surveys*, 13, 1, 1-32.

Caporale, G.M., Hall, S.G., Urga G. and G. Williams (1997), "Aggregate money demand functions in five industrial countries: are they cointegrated? ", D.P. no.19-97, Centre for Economic Forecasting, *London Business School*.

Caporale, G.M., Kalyvitis, S. and N. Pittis (1996), "Interest rate convergence, capital controls, risk premia and foreign exchange rate market efficiency in the EMS", *Journal of Macroeconomics*, 18, 4, 693-714.

Cavaglia, S. (1992), "The persistence of real interest differentials - A Kalman filtering approach", *Journal of Monetary Economics*, 29, 429-443.

Cuthbertson, K., Hall, S.G., and M.P. Taylor (1992), *Applied Econometric Techniques, Harvester Wheatsheaf*.

Dahlquist, M. and S.F. Gray (2000), "Regime-switching and interest rates in the European Monetary System", *Journal of International Economics*, 50, 399-419.

Davidson, J. (1994), "Identifying cointegrating regressions by the rank condition", *Oxford Bulletin of Economics and Statistics*, 56, 1, 103-108.

Davidson, J. (1998a), "Structural relations, cointegration and identification: some simple results and their application", *Journal of Econometrics*, 87, 87-113.

Davidson, J. (1998b), "A Wald test of restrictions on the cointegrating space based on Johansen's estimator", *Economics Letters*, 59, 2, 183-187.

Davidson, J. and S.G. Hall (1991), "Cointegration in recursive systems", *Economic Journal* 101, 239-251

Dolado ,J.J., and Lutkepohl, H (1996) "Making Wald Tests Work for Cointegrated VAR Systems", *Econometrics Reviews*,15 , 369-386.

Doornik, J.A. and D.F. Hendry (1997), "Modelling dynamic systems using Pc-Fiml 9.0, *International Thomson Business Press*, London

Engle, R.F. and C.W.J. Granger (1987), "Cointegration and error correction: representation, estimation and testing", *Econometrica*, 55, 2, 251-276.

Engle, R.F., Hendry, D.F. and J-F. Richards, (1983) "Exogeneity", *Econometrica*, 51, 277-304

Fratianni, M. and J. von Hagen (1990), "German dominance in the EMS: evidence from interest rates", *Journal of International Money and Finance*, 9, 358-375.

Granger C.W.J., (1983) "Cointegrated variables and error correction models", working paper, UCSD

Granger, C.W.J., (1969) "Investigating causal relationships by econometric models and cross-spectral methods", *Econometrica*, 37, 424-438

Granger, C.W.J., (1988) "Some recent developments in a concept of causality." *Journal of Econometrics*, 39, 199-212.

Gregory, A. and B. E. Hansen, (1996) "Residual –based tests for cointegration in models with regime shifts", *Journal of Econometrics*, 70, 99-126

Gregory, A. W., Nason, J. M. and D. Watt (1996) "Testing for structural breaks in cointegrated relationships", *Journal of Econometrics*, 71, 324-341

Haldane, A.G. and S.G. Hall (1991), "Sterling's relationship with the dollar and the Deutschmark: 1976-1989", *Economic Journal*, 101, 436-443.

Hall, S.G. (1991), "The effect of varying length VAR models on the maximum likelihood estimates of cointegrating vectors", *Scottish Journal of Political Economy*, 38, 4, 317-323.

Hall, S.G. and M.R.Wickens (1993), "Causality in integrated systems", D.P. no. 27-93, Centre for Economic Forecasting, London Business School.

Hall, S.G., Robertson, D. and M.R. Wickens (1992), "Measuring convergence of the EC economies", *Manchester School*, 60, Supplement, 99-111.

Hamilton, J.D. (1994) "Time Series Analysis", *Princeton University Press*: Princeton, NJ

Hansen, B. E., (1992) "Tests for parameter instability in regressions with I(1) processes", *Journal of Business and Economic Statistics*, 10, 321-335.

Hansen, H. and S. Johansen (1999) "Some tests for parameter constancy in cointegrated VAR models", *Econometrics Journal*, 2, 306-333.

Hao, K. and B. Inder (1996) "Diagnostic test for structural change in cointegrated regression models", *Economics Letters*, 50, 179-187.

Harbo, I., Johansen, S., Nielsen, B.G. and Rahbek, A.C. (1998) "Asymptotic inference on cointegrating rank in partial systems", *Journal of Business and Economic Statistics*, Vol.16, No.4, 388-399,

Hendry, D.F. (1995), *Dynamic Econometrics*, Oxford University Press, Oxford.

Johansen, S. (1988), "Statistical analysis of cointegration vectors", *Journal of Economic Dynamics and Control*, 12, 231-254

Johansen, S. (1991), "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models", *Econometrica*, 59, 1551-1580.

Johansen, S. (1995a), "Identifying restrictions of linear equations with applications to simultaneous equations and cointegration", *Journal of Econometrics*, 69, 111-132.

Johansen, S. (1995b), "Likelihood based inference in cointegrated vector autoregressive models", Oxford University Press.

Johansen, S. and K. Juselius (1990), "Maximum likelihood estimation and inference on cointegration: with application to the demand of money", *Oxford Bulletin of Economics and Statistics*, 52, 169-210

Johansen, S. and K. Juselius (1994), "Identification of the long run and the short run structure: An application to the ISLM model", *Journal of Econometrics*, 63, 7-36.

Kalman, R.E. (1960), "A new approach to linear filtering and prediction problems", *Journal of Basic Engineering, Transactions of the ASME*, 82, 35-45.

Kalman, R.E. (1963), "New methods in Wiener filtering theory", in J.L. Bogdanoff and F. Kozin (eds), *Proceedings of the First Symposium of Engineering Applications of Random Function Theory and Probability*, 270-288, *John Wiley & Sons Inc.*, New York.

Karfakis, J.C. and D.M. Moschos (1990), "Interest rate linkages within the European Monetary System: a time series analysis", *Journal of Money, Credit and Banking*, 22, 3, 388-394.

Katsimbris, G.M. and S.M. Miller (1993), "Interest rate linkages within the European Monetary System: further evidence", *Journal of Money, Credit and Banking*, 25, 4, 771-779.

Kirchgassner, G. and J. Wolters (1993), "Does the DM dominate the Euro market? An empirical investigation", *Review of Economics and Statistics*, 75, 773-778.

Kremers, J.J.M., Ericsson, N.R. and J.J. Dolado (1992) "The power of cointegration test" *Oxford Bulletin of Economics and Statistics* 54, 325-348

Lucas, R. E. (1988) “ Money demand in the United States: a quantitative review”, *Carnegie–Rochester Conference Series on Public Policy*, 29, 137-168.

Lumsdaine R. L. and D. H. Papell (1997) “Multiple trend breaks and the unit root hypothesis”, *Review of Economics and Statistics*, 79, 212-218

Lutkepohl H., and H.E. Reimers, (1992), “Impulse response analysis of cointegrated systems”, *Journal of Economic Dynamics and Control*, 16, 53-78.

Martin, P. (1997), “The exchange rate policy of the Euro: A matter of size?”, CEPR D.P. no.1646.

Mosconi, R and C. Giannini (1992). “Non-causality in cointegrated systems: representation, estimation and testing”, *Oxford Bulletin of Economics and Statistics*, 54, 3, 399-417.

Nyblom J., (1989) “Testing for the constancy of parameters over time”, *Journal of the American Statistical Association*, 84, 223-230

Perron, P. (1989) “The great crash, the oil-price shock and the unit root hypothesis”, *Econometrica*, 57, 1361-1401

Pesaran, M.H. and Y. Shin (1994), “Long-run structural modelling”, *Working Paper*, Department of Applied Economics, University of Cambridge. Updated (1999)

Phillips, P. C. B. and B. E. Hansen (1990), "Statistical inference in instrumental variable regression with I(1) processes" *Review of Economic Studies*, 57, 99-125

Phillips, P. C. B., (1991), "Optimal inference in cointegrated systems", *Econometrica* 60, 119-143

Phillips, P. C. B., (1995), "Fully modified least squares and vector autoregression", *Econometrica* 63, 1023-1078

Ploberger, W., Kramer, W. and K. Kontrus (1989) "A new test for structural stability in the linear regression model", *Journal of Econometrics*, 40, 307-318

Portes, R. and H. Rey (1998), "The emergence of the Euro as an international currency", *Economic Policy*, 26, April 1998.

Quintos, C. E. and P. C. B. Phillips (1993) "Parameter constancy in cointegrating regressions", *Empirical Economics*, 18, 675-706.

Quintos, C. E., (1995) "Sustainability of the deficit process with structural shifts", *Journal of Business and Economic Statistics*, 13, 409-417

Reimers, H. E., (1992) " Comparison of tests for multivariate cointegration", *Statistical Papers*, 33, 335-359

Saikkonen, P., (1991) "Asymptotically efficient estimation of cointegrated regressions", *Econometric Theory*, 7, 1-21

Saikkonen, P., (1993) "Estimation of cointegration vectors with linear restrictions", *Econometric Theory*, 9, 19-35

Sargan, D., (1988), "Lectures on advanced econometric theory", *Blackwell Publishers*.

Seo, B. (1998) "Tests for structural change in cointegrated systems", *Econometric Theory*, 14, 222-259.

Sims, C.A. (1980) "Macroeconomics and reality", *Econometrica*, 48, 1-48

Stock, J. and M. Watson (1993) "A simple estimator of cointegrating vectors in higher order integrated systems", *Econometrica*. 61, 783-820

Toda, H.Y. and P.C.B. Phillips (1993), "Vector autoregression and causality" *Econometrica*, 61, 6, 1367-1393.

Toda, H.Y. and T. Yamamoto (1995), "Statistical inference in vector autoregressions with possibly integrated processes", *Journal of Econometrics*, 66, 225-250.

Urbain, J-P. (1992) "On weak exogeneity in error correction models", *Oxford Bulletin of Economics and Statistics* 54, 187-207

Zapata, H.O., and A.N. Rambaldi (1997), "Monte Carlo Evidence on Cointegration and Causation," *Oxford Bulletin of Economics and Statistics*, 59, 2, 285-298.

Zivot, E. and D.W.K. Andrews (1992) "Further evidence on the great crash, the oil-price shock and the unit root hypothesis", *Journal of Business and Economic Statistics*, 10, 251-270

