

TESTING FOR CHANGES IN THE LONG-RUN CAUSAL STRUCTURE OF COINTEGRATED VECTOR AUTOREGRESSIONS

Marco R. Barassi
Centre for Monetary and Financial Economics
South Bank University, London

Guglielmo Maria Caporale
Centre for Monetary and Financial Economics
South Bank University, London

Stephen G. Hall
Management School
Imperial College

March 2001

Abstract

In this paper we examine, using Monte Carlo methods, several alternative approaches to detecting changes in the long-run causal structure of cointegrated VARs. We adopt recursive and rolling 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. Rolling estimation is found to perform better in detecting instability in the parameters and their true value after the breakpoint. However, it only detects a region where the break is likely to occur. To overcome the drawbacks of these techniques, we propose an OLS-based sequential-dummy test. We derive its critical values for different sample sizes and discuss its statistical properties. We show that it has a non-standard distribution and reasonable power even in small samples. It is also a useful graphical device, as it shows a kink at the breakpoint. However, it is not able to deal with the case of multiple breakpoints. Here a possible solution is represented by the use of the Kalman filter. We analyse its performance on the same DGPs. The results confirm its effectiveness as a graphical tool, but they also indicate that it exhibits relatively wide margins of error, suggesting that its efficiency is sometimes limited.

Keywords: Long-Run Causality, Structural Change, Monte Carlo Methods

JEL classification: C12, C15, C32

Financial support from ESCR grant no. R00222955, Interest Rate Linkages in a Changing World, is gratefully acknowledged. We are also grateful to Fabio Spagnolo, Nicola Spagnolo and Andrea Cipollini for useful comments.

Corresponding author: *Professor Guglielmo Maria Caporale, South Bank University, Business School, 103 Borough Road, London SE1 0AA, UK. Tel: +44 (0)20 7815 7012. Fax: +44 (0)20 7815 8226. E-mail: g.m.caporale@sbu.ac.uk*

1 Introduction

The underlying motivation for this paper 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 paper 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 \mathbf{a} matrix. It is structural change in this part of the system which we wish to focus on in this paper.

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. 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 extend the CUSUM test to the case of non-stationary regressors considering 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

$$1.1 \quad \Delta X_t = \Gamma_1 \Delta X_{t-1} + \dots + \Gamma_{k-1} \Delta X_{t-k+1} + \mathbf{ab}' X_{t-k} + \mathbf{e}_t$$

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 $\mathbf{e}_t \sim \text{iidN}(0, \Omega)$, \mathbf{b} is the cointegrating matrix and \mathbf{a} is the matrix of adjustment coefficients that determines the long-run causal structure, and $\mathbf{ab}' = \mathbf{P}$ has reduced rank $r < p$. It is convenient to rewrite the model as

$$1.2 \quad Z_{0t} = \Theta Z_{1t} + \mathbf{ab}' Z_{2t} + \mathbf{e}_t,$$

where $\mathbf{Q} = (\mathbf{G}_1, \dots, \mathbf{G}_{k-1})$, $Z_{0t} = \mathbf{D}X_t$, $Z_{1t} = (\Delta X_{t-1}', \dots, \Delta X_{t-k+1}')$, $Z_{2t} = (X_{t-k})$. 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 \mathbf{a} and \mathbf{b} (suitably normalised) as well as the rank of the matrix $\mathbf{P} = \mathbf{ab}'$.

Quintos (1995) uses the trace statistic to assess the constancy of the cointegrating rank against the hypothesis of a changing rank of $\Pi = \mathbf{ab}'$. 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.¹

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

$$1.3 \quad Z_{0t} = \Theta Z_{1t} + \mathbf{a}(t)\mathbf{b}(t)' Z_{2t} + \mathbf{e}_t,$$

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

- A) $\mathbf{a}(t)\mathbf{b}(t)' = (\mathbf{a}_1\mathbf{b}_1 I_{1t} + \mathbf{a}_2\mathbf{b}_2 I_{2t} + \dots + \mathbf{a}_k\mathbf{b}_k I_{kt})$
- B) $\mathbf{a}(t)\mathbf{b}(t)' = (\mathbf{a}_1\mathbf{b}_1 I_{1t} + \mathbf{a}_2\mathbf{b}_2 I_{2t} + \dots + \mathbf{a}_k\mathbf{b}_k I_{kt})$
- C) $\mathbf{a}(t)\mathbf{b}(t)' = (\mathbf{a}_1\mathbf{b}_1 I_{1t} + \mathbf{a}_2\mathbf{b}_2 I_{2t} + \dots + \mathbf{a}_k\mathbf{b}_k I_{kt})$

¹ 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.

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 paper 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 rely 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 paper 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.

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:

$$2.1 \quad A^0 \Delta X_t = \Gamma \Delta X_{t-1} + \mathbf{ab}' X_{t-1} + e_t$$

that can be rewritten as

$$2.2 \quad \Delta X_t = (A^0)^{-1} \Gamma \Delta X_{t-1} + (A^0)^{-1} \mathbf{ab}' X_{t-1} + (A^0)^{-1} e_t$$

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

$$2.3 \quad A^0 = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.3 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 0.25 \\ 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ -2 \end{bmatrix};$$

and allow the matrix A^0 to switch to

$$2.4 \quad A^1 = \begin{bmatrix} 1 & 0 \\ 0.5 & 1 \end{bmatrix}$$

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

$$2.5 \quad \Pi = \begin{bmatrix} -0.25 & 0.50 \\ 0 & 0 \end{bmatrix}, \quad \Pi^* = \begin{bmatrix} -0.25 & 0.50 \\ 0.125 & -0.25 \end{bmatrix}.$$

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 $T= 50, 100, 200, \text{ and } 400$. 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 A sequential test for structural change

Consider the following cointegrating VAR in error correction form,

$$\begin{aligned}
 3.1 \quad \Delta Z_{1t} &= \Gamma_{11} \Delta Z_{t-1} + \dots + \Gamma_{1k-1} \Delta Z_{t-k+1} + \mathbf{a}_1(\mathbf{b}' Z_{t-k}) + e_{1t} \\
 \Delta Z_{2t} &= \Gamma_{21} \Delta Z_{t-1} + \dots + \Gamma_{2k-1} \Delta Z_{t-k+1} + \mathbf{a}_2(\mathbf{b}' Z_{t-k}) + e_{2t}
 \end{aligned}$$

where $\mathbf{b}' Z_{t-k}$ 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,

$$3.2 \quad \mathbf{a} = \begin{bmatrix} \mathbf{a}_1 \\ 0 \end{bmatrix}, \quad \mathbf{a}^\circ = \begin{bmatrix} \mathbf{a}_1^\circ \\ \mathbf{a}_2^\circ \end{bmatrix}, \quad \alpha_2^\circ \neq 0.$$

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 sub-samples 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,

$$\begin{aligned}
 3.3 \quad \Delta Z_{1t} &= \Gamma_{11}\Delta Z_{t-1} + \dots + \Gamma_{1k-1}\Delta Z_{t-k+1} + \mathbf{a}_1(\mathbf{b}'Z_{t-k}) + \mathbf{a}_1^\circ D_i(\mathbf{b}'Z_{t-k}) + e_{1t} \\
 \Delta Z_{2t} &= \Gamma_{21}\Delta Z_{t-1} + \dots + \Gamma_{2k-1}\Delta Z_{t-k+1} + \mathbf{a}_2(\mathbf{b}'Z_{t-k}) + \mathbf{a}_2^\circ D_i(\mathbf{b}'Z_{t-k}) + e_{2t}
 \end{aligned}$$

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

$$3.4 \quad \tau = t - \max = \operatorname{argmax}\{t_{\alpha_i} = 0(n)\}, n = 2, T - 2.$$

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.

4 Kalman filtering

The Kalman (1960, 1963) filter technique is adopted to estimate linear models with time-varying coefficients². 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:

$$4.1 \quad y_t = x_t' \beta + \varepsilon_t \quad \varepsilon_t \sim N(0, H_t)$$

and the transition equation be:

$$4.2 \quad \beta_t = T\beta_{t-1} + \eta_t \quad \eta_t \sim N(0, Q_t)$$

with the initial conditions given by:

$$4.3 \quad \beta_0 \sim N(\beta_0, \sigma^2 P_0)$$

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

² For a more detailed account of Kalman filtering, see Cuthbertson *et al* (1992).

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:

$$4.4 \quad \hat{\beta}_{t/t-1} = T\hat{\beta}_{t-1}$$

and the covariance matrix is defined as:

$$4.5 \quad P_{t/t-1} = TP_{t-1}T' + Q_t$$

Finally, the updating formulae are given by:

$$4.6 \quad \hat{\beta}_t = \hat{\beta}_{t/t-1} + P_{t/t-1}x(y_t - x'\hat{\beta}_{t/t-1})/(x'P_{t/t-1}x + H_t)$$

and

$$4.7 \quad P_t = P_{t/t-1} - P_{t/t-1}x'xP_{t/t-1}' / (x'P_{t/t-1}x + H_t)$$

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:

$$4.8 \quad \Delta Z_t = \Gamma_1 \Delta Z_{t-1} + \dots + \Gamma_{k-1} \Delta Z_{t-k+1} + \mathbf{a} \mathbf{b}' Z_{t-k} + e_t,$$

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

$$4.9 \quad \mathbf{a}_t = \mathbf{a}_{t-1} + \mathbf{u}_t, \quad \mathbf{u}_t \text{ is } N(0, \mathbf{S}^2 I)$$

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.

5 Results

We start with the FIML estimation using the whole sample. The first result concerns the estimates of the cointegrating coefficients denoted as \mathbf{b}_1 and \mathbf{b}_2 respectively in table 1. These coefficients (standardised with respect to \mathbf{b}_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 \mathbf{a}_2 are instead biased being a weighted average of the two values that \mathbf{a}_2 takes before and after the break. These are never too close to 0.125, that is the value that \mathbf{a}_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 breaks³ 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

³By 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.

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 on \mathbf{a}_2 would be easily rejected⁴. 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.

Table 1. Full Sample FIML Estimates (1000 replications)						
	DGP1		DGP2		DGP3	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
sample size 50						
a1	-0.278	0.1919	-0.2663	0.056	-0.294	0.2117
a2	0.0982	0.1105	0.0633	0.0623	0.0191	0.117
b1	1	#	1	#	1	#
b2	-1.9674	0.3113	-2.0089	0.1417	-1.9573	0.3775
sample size 100						
a1	-0.2558	0.0375	-0.2542	0.0381	-0.2555	0.038
a2	0.0894	0.043	0.0539	0.0411	0.0259	0.0376
b1	1	#	1	#	1	#
b2	-2.0096	0.0913	-2.0069	0.073	-2.0035	0.0627
sample size 200						
a1	-0.2508	0.0253	-0.2495	0.0252	-0.2508	0.0258
a2	0.0852	1.0298	0.0523	0.0289	0.0248	0.0266
b1	1	#	1	#	1	#
b2	-2.0016	0.0437	-2.0036	0.0369	-2.0025	0.0319
sample size 400						
a1	-0.2495	0.0182	-0.2482	0.0183	-0.2492	0.0184
a2	0.0845	0.0209	0.0515	0.0204	0.0241	0.0184
b1	1	#	1	#	1	#
b2	-2.0016	0.0247	-2.0011	0.198	-2.0006	0.017

With a sample size of 200 we would not be able to reject the hypothesis of a zero coefficient on \mathbf{a}_2 at the 95% level, while the same hypothesis would be rejected for a sample of 400, where we would find a coefficient on \mathbf{a}_2 that is significantly different

⁴ Notice that the rejection of the hypothesis of a zero in favour of a non-zero coefficient would not account

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 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 chart 1 and 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 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

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

estimation and the inspection of the t-values for the \mathbf{a}_2 coefficient provide us with results that are different depending on the sample size and the breakpoint. They indicate (see table 2) that when \mathbf{a}_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 \mathbf{a}_2 is only significant at a 90% level in DGP1. For the samples of size 100, we see that it 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 \mathbf{a}_2 that takes values significantly different from zero after 100 points have been added from the recursive procedure. Notice that the \mathbf{a}_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.

Table 2 t-values of recursive estimates of \mathbf{a}_2							
DGP1	N/4	(3/8)N	N/2	(5/8)N	(3/4)N	(7/8)N	N
<i>ss-50</i>	#	0.849822	1.108077	1.302761	1.493226	1.632482	1.730184*
<i>ss-100</i>	0.266668	0.82887	1.203619	1.548965	1.826124*	2.08068**	2.34645**
<i>ss-200</i>	0.193132	0.957559	1.522429	2.010072**	2.435999**	2.822665**	3.17873**
<i>ss-400</i>	0.177762	1.253367	2.083647**	2.785881**	3.395936**	3.95377**	4.46386**
DGP2	N/4	(3/8)N	N/2	(5/8)N	(3/4)N	(7/8)N	N
<i>ss-50</i>	#	0.382371	0.45922	0.687334	0.90886	1.072572	0.247567
<i>ss-100</i>	0.266668	0.236806	0.195866	0.615602	0.937328	1.225415	1.519585
<i>ss-200</i>	0.193132	0.155975	0.150348	0.749731	1.232245	1.667304*	2.059217**
<i>ss-400</i>	0.177762	0.152599	0.12635	0.967444	1.667098*	2.288966**	2.853594**
DGP3	N/4	(3/8)N	N/2	(5/8)N	(3/4)N	(7/8)N	N
<i>ss-50</i>	#	0.382371	0.332199	0.282985	0.377341	0.553404	0.739095
<i>ss-100</i>	0.266668	0.236806	0.195866	0.156649	0.128573	0.496831	0.811484
<i>ss-200</i>	0.193132	0.155975	0.150348	0.137181	0.106946	0.628267	1.062327
<i>ss-400</i>	0.177762	0.152599	0.12635	0.111431	0.097988	0.812283	1.428967
* denotes a coefficient \mathbf{a}_2 significantly different from zero at 90% level							
** denotes a coefficient \mathbf{a}_2 significantly different from zero at 95% level							
<i>N = sample size (ss)</i>							
<i>The bold indicates the breakpoints in the DGPs</i>							

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 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 and 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).

Table 3							
t-values of rolling estimates of α_2							
Size of rolling window	N/4	(3/8)N	N/2	(5/8)N	(3/4)N	(7/8)N	N
30	0.134419	0.421034	0.827222	1.20794	1.529823	1.499496	1.505702
40	0.358436	0.582036	0.911967	1.292864	1.64591*	1.848319*	1.862825*
50	0.500187	0.748444	0.946235	1.265746	1.510874	1.850584*	2.130953**
N = 100							
Breakpoint is at N/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							

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, 4 and 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 4.

Table 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

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 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 allows us to detect the break quite accurately as far the timing and the magnitude of the break are concerned. The 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, while the Kalman filter may be graphically useful to detect the presence of breaks and their magnitude, it is not always very efficient. We can see this from the reported graphs (graphs 7 and 8) that refer to coefficients estimated for samples of size 100 generated according to DGP2. Whilst the break points and their magnitude are graphically observable, there are relatively wide margins of error, as shown by the confidence interval bands. These highlight the fact that the coefficients estimated with such a technique might often be not significantly different

from zero after the break has occurred, suggesting that the efficiency of the Kalman filter is relatively limited. This result applies to all the DGPs and sample sizes.

6 Conclusions

In this paper 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. Nevertheless, even if

well established, this technique shows relatively wide margins of error, suggesting that its efficiency is sometimes limited.

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
- 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
- Cuthbertson, K., Hall, S.G., and M.P. Taylor (1992), *Applied Econometric Techniques*, Harvester Wheatsheaf.
- Gregory, A. W., Nason, J. M. and D. Watt (1996) "Testing for structural breaks in cointegrated relationships", *Journal of Econometrics*, 71, 324-341
- Gregory, A. and B. E. Hansen, (1996) "Residual –based tests for cointegration in models with regime shifts", *Journal of Econometrics*, 70, 99-126
- 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
- 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.
- 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.
- 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

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

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

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

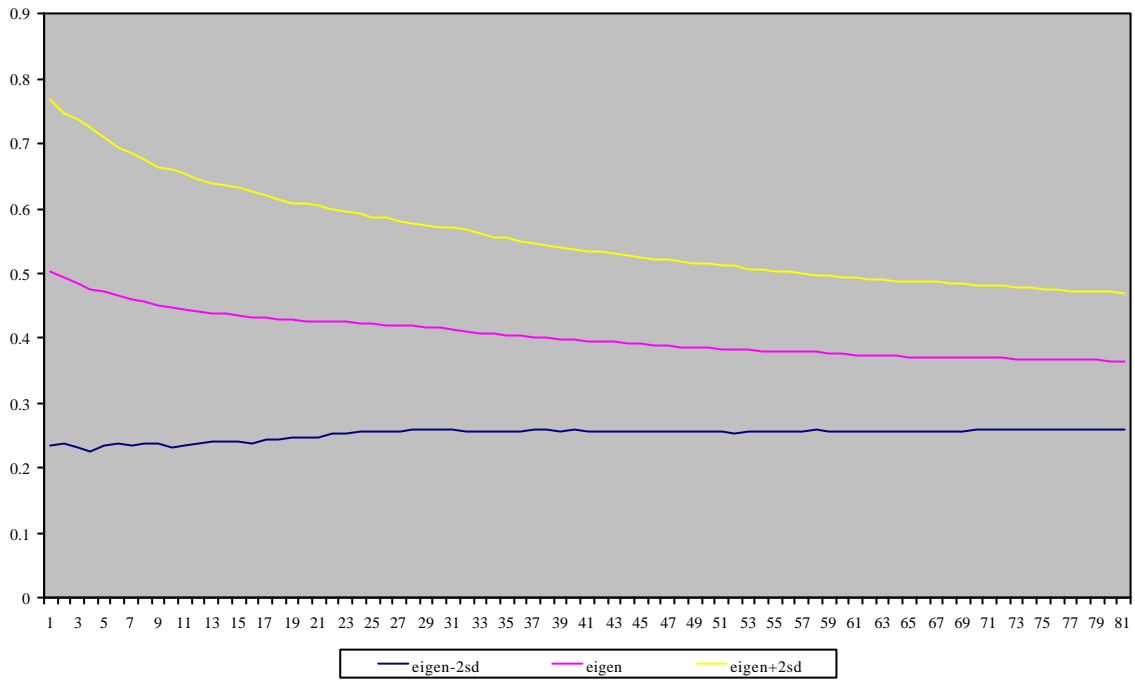
Quintos, C. E. and P. C. B. Phillips (1993) "Parameter constancy in cointegrating regressions", *Empirical Economics*, 18, 675-706

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

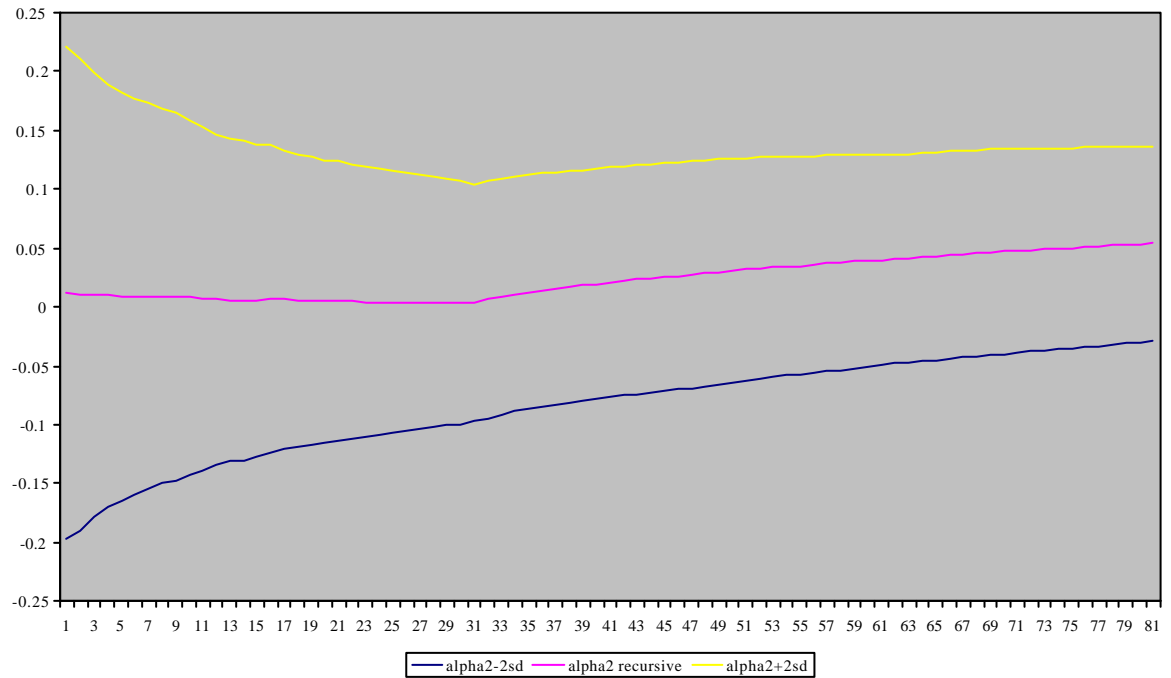
Stock, J. and M. Watson (1993) " A simple estimator of cointegrating vectors in higher order integrated systems", *Econometrica*. 61, 783-820

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

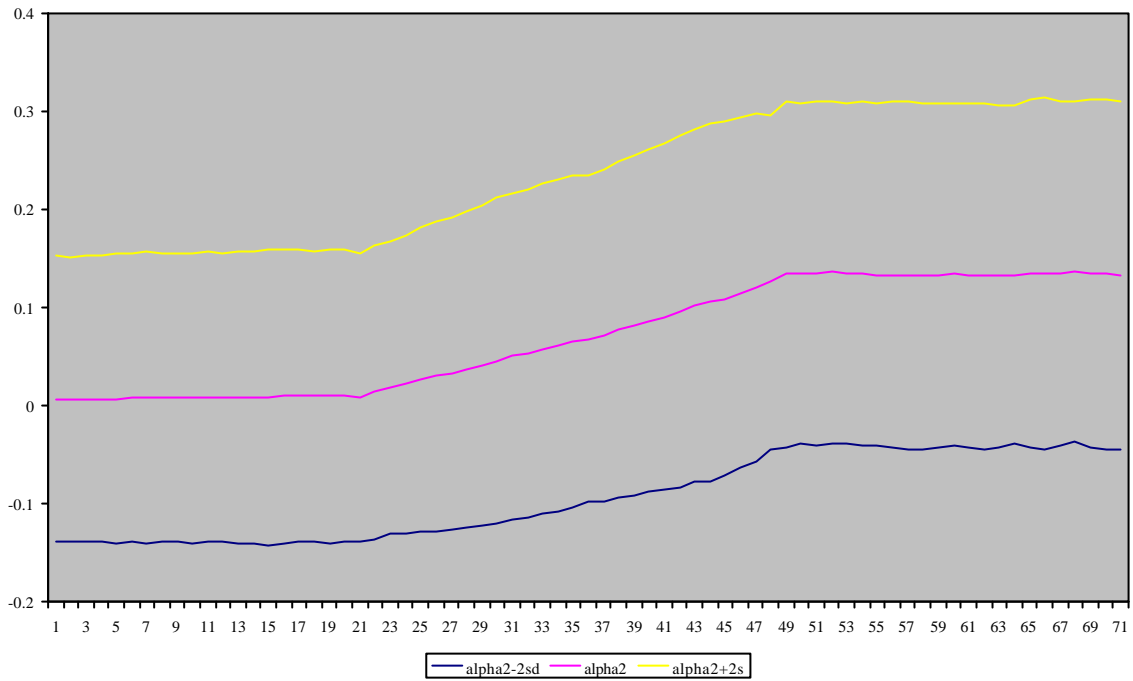
graph 1. Recursive eigenvalues for DGP2 of size 100 (average of 1000 replications)



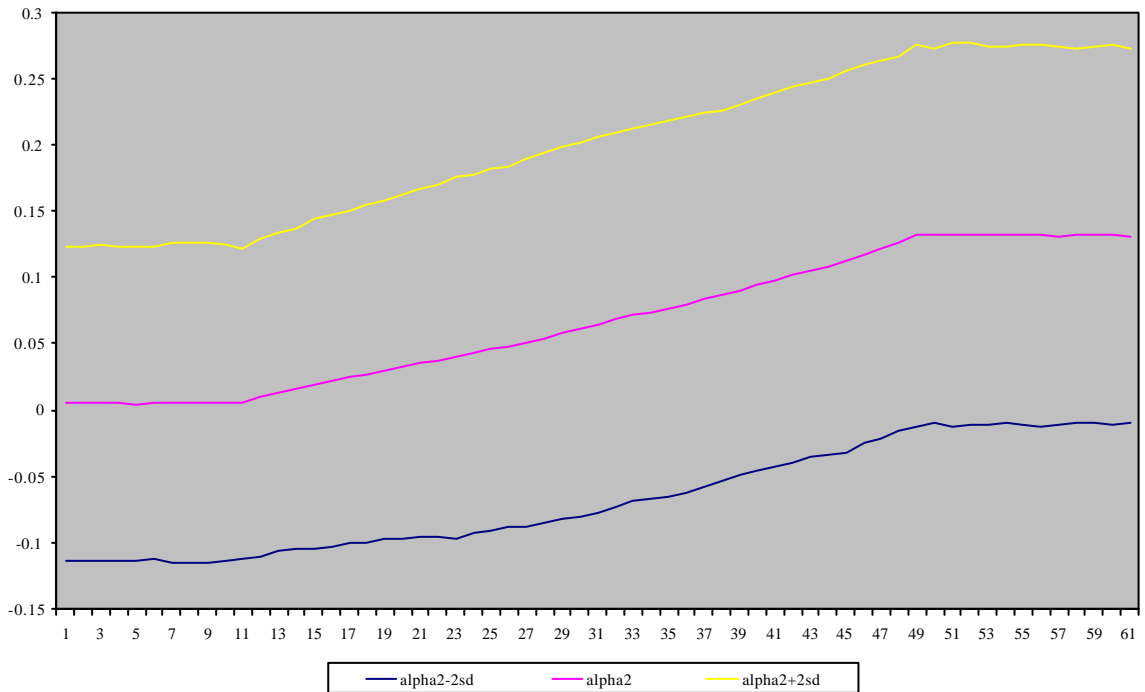
graph 2. Recursive estimates of alpha2 for DGP2 of size 100 (average of 1000 replications)



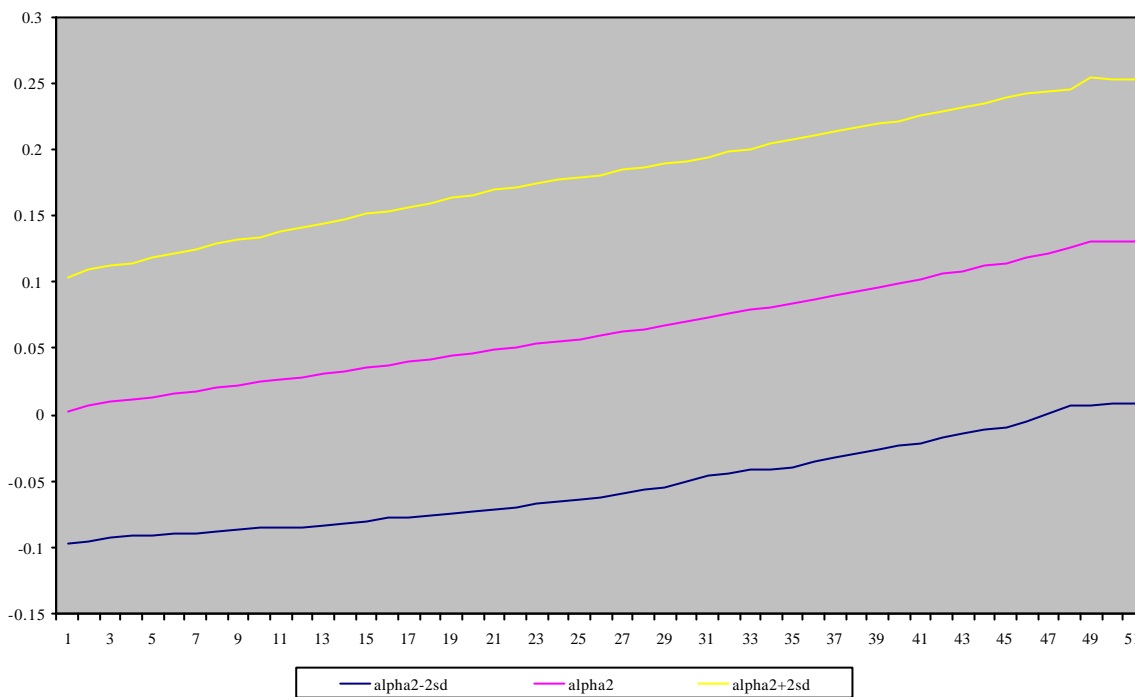
graph 3. Rolling estimates of alpha2 for DGP2 of size 100 with a rolling window of 30 (average of 1000 replications)



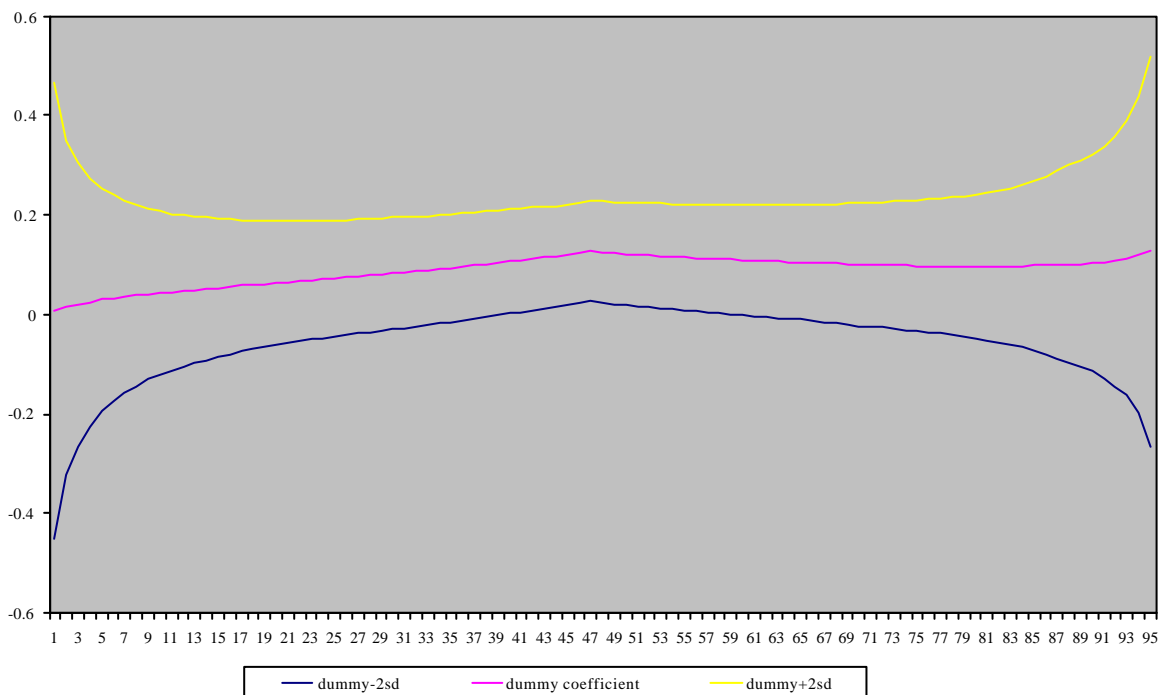
graph 4. Rolling estimates of alpha2 for DGP2 of size 100 with rolling window of 40 (average of 1000 replications)



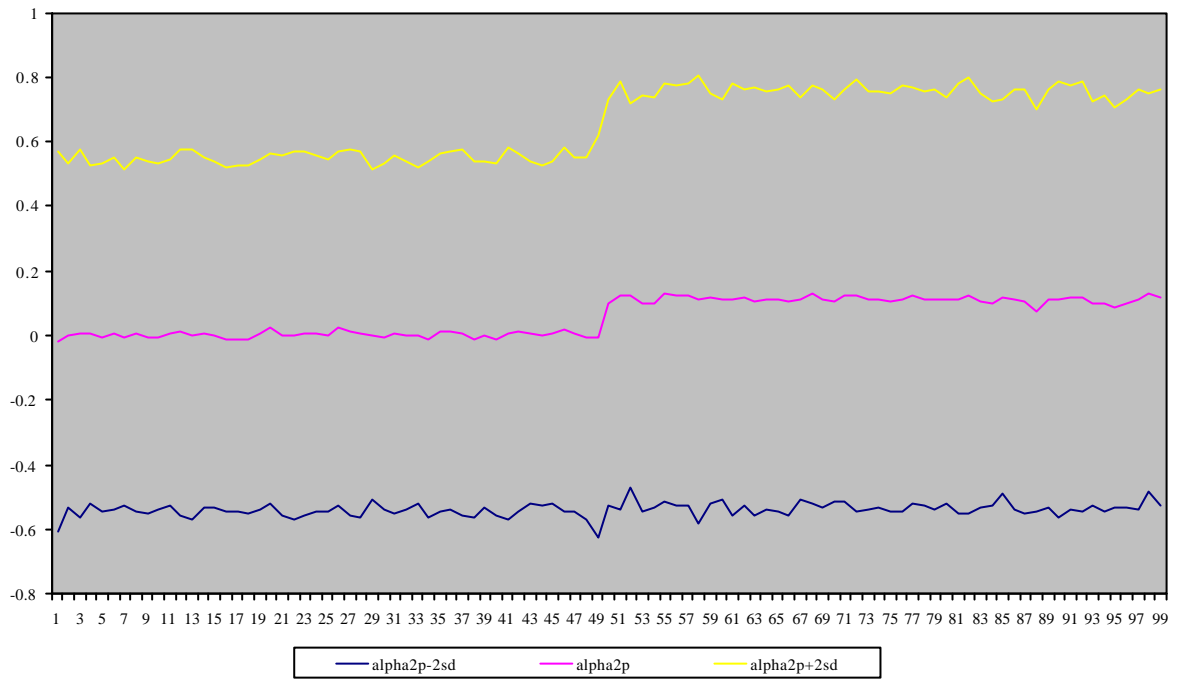
graph 5. Rolling estimates of alpha2 for DGP2 of size 100 with rolling window of 50 (average of 1000 replications)



graph 6. Coefficient of sequential dummy for DGP2 of size 100 (average of 2500 replications)



graph 7. One step ahead Kalman Filter predictions of alpha2 for DGP2 of size 100 (average of 1000 replications)



graph 8. Smoothed Kalman Filter estimates of alpha2 for DGP2 of size 1000 (average of 1000 replications)

