Quantitative and Qualitative Analysis in Social Sciences
Volume 3, Issue 3, 2009, 43-77
ISSN: 1752-8925

Cointegrated Vector Autoregressive Models with Adjusted Short-Run Dynamics

Takamitsu Kurita[a]
Fukuoka University

Bent Nielsen[b]
University of Oxford

Abstract

A family of cointegrated vector autoregressive models with adjusted short-run dynamics is introduced. These models can describe evolving short-run dynamics in a more flexible way than standard vector autoregressions, and yet likelihood analysis is based on reduced rank regression using conventional asymptotic tables. The family of dynamics-adjusted vector autoregressions consists of three models: a model subject to short-run parameter changes, a model with partial short-run dynamics and a model with short-run explanatory variables. An empirical illustration using US gasoline prices is presented, together with some simulation experiments.

JEL Classifications: C31, C51, C52.

Keywords: Cointegration, adjusted short-run dynamics, parameter change, likelihood ratio test.

Acknowledgements: The numerical results were generated using *GiveWin / PcGive* (Doornik and Hendry, 2001), *PcGets* (Hendry and Krolzig, 2001) and *Ox* (Doornik, 2001). David F. Hendry and Katarina Juselius are gratefully acknowledged for providing us with the data. The second author received financial support from ESRC grant RES-000-27-0179.

[a] Faculty of Economics, Fukuoka University, Bunkei Center Building, 8-19-1 Nanakuma, Johnanku, Fukuoka, 814-0180, Japan; tel.: +81 (0) 92 871 6631, email: tkurita@fukuoka-u.ac.jp

[b] Department of Economics, University of Oxford, Manor Road, Oxford, OX1 3UQ, UK; tel.: +44 (0) 1865 278630, email: bent.nielsen@nuffield.ox.ac.uk

1 Introduction

A family of cointegrated vector autoregressive models with adjusted short-run dynamics is introduced. These models can describe evolving short-run dynamics in a more flexible way than standard vector autoregressions, and yet likelihood analysis is based on reduced rank regression as in Johansen (1988, 1996) using the same asymptotic tables. The family of dynamics-adjusted vector autoregressions consists of three models: *a model subject to short-run parameter changes*, *a model with partial short-run dynamics* and *a model with short-run explanatory variables*.

The new family of models are designed for situations where residuals of conventional vector autoregressions show signs of autocorrelation or autoregressive conditional heteroscedasticity (ARCH). If such model mis-specification is associated with known structural-break dates or certain types of stationary regressors, the proposed methods can work without altering the conventional asymptotic arguments. Compared with the use of modestly mis-specified vector autoregressions in empirical work, the proposed models are advantageous in terms of the flexibility in describing short-run dynamics. The results of cointegration rank tests are broadly the same, since usual cointegration rank tests are somewhat robust to model mis-specification. The new family of models can be combined further with cointegrated vector autoregressions allowing for deterministic shifts, developed by Johansen, Mosconi and Nielsen (2000).

To describe this new class of models, consider a vector autoregression of order k and dimension p for a time series, $X_{-k+1}, ..., X_0, X_1, ..., X_T$, given by the equation

$$\triangle X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \triangle X_{t-i} + \varepsilon_t, \quad \text{for } t = 1, ..., T, \tag{1}$$

where the innovations ε_t are independent $\mathsf{N}(0, \Omega)$-distributed. Cointegration arises when Π has reduced rank r and can be written as $\Pi = \alpha\beta'$ for some $(p \times r)$ matrices α and β. Following Johansen (1992a) and Juselius (2006, Section 4.2), the autoregressive model can be reparameterised as

$$\triangle^2 X_t = \Pi X_{t-1} - \Gamma \triangle X_{t-1} + \sum_{i=1}^{k-2} \Psi_i \triangle^2 X_{t-i} + \varepsilon_t, \tag{2}$$

where $\Gamma = I - \sum_{i=2}^{k-1} \Gamma_i$ and $\Psi_i = -\sum_{j=i+1}^{k-1} \Gamma_j$. Note that the models (1) and (2) are equivalent. There is a one-one mapping between the parameters of the two models, and the likelihood functions are equivalent since the innovation ε_t is not affected by the reparameterisation.

According to the Granger-Johansen representation theorem (see Johansen, 1996, Theorem 4.2), the cointegrated relation $\beta' X_{t-1}$ determines the *long-run dynamics* of the model. In this paper we refer to α and Γ as the parameters for *medium-run dynamics* as they describe how the process adjusts to changes in $\beta' X_{t-1}$ and $\triangle X_{t-1}$, respectively. We define $\Psi_1, ...\Psi_{k-2}$ as the parameters for the *short-run dynamics* of the model in that they are irrelevant to the evolution of the common stochastic trends. The short-run dynamics often correspond to acceleration rates of variables in a dynamic economic system. Some acceleration rates, such as inflation acceleration, can play an important role in an econometric model as determinants of agents' behaviour (see Galí and Gertler, 1999, for instance). Thus, equation (2) shows how acceleration rates react to long-run and medium-run dynamics, suggesting that the interpretation of $\triangle^2 X_t$ as mere surprise is only reasonable when Π and Γ turn out to be zero.

As shown below, changes in the medium and long term parameters, α, Γ and β, are reflected in the impact parameter of the common stochastic trends, thereby affecting the asymptotic distributions of cointegration rank tests. Also, if the innovation variance Ω is non-constant, the likelihood cannot be maximized using reduced rank regression. In other words, if the medium and long term parameters or the innovation variance are time-varying, cointegration analysis can neither be based on the reduced rank regression nor the conventional limiting distributions of rank tests, but is somewhat more involved, see Andrade, Bruneau, and Gregoir (2005), Cavaliere and Taylor (2006), and Hansen (2000, 2003).

A feature of changes or modifications of the short-run parameters, $\Psi_1, ...\Psi_{k-2}$, is that they do not have any effect on the impact parameter for the stochastic trends. As a consequence, the standard procedures for determining cointegration rank and hypothesis testing can be used. The class of dynamics-adjusted models therefore provides a more flexible framework for modelling the dynamics of the data than the usual cointegration models, while allowing the use of the standard cointegration tables and interpretations.

This paper is organised as follows. Section 2 presents a family of cointegrated vector autoregressive models with adjusted short-run dynamics, and Section 3 considers the statistical properties of the models centering on the Granger-Johansen representations and reduced rank regression. Section 4 provides the asymptotic analysis of a cointegration rank test in the family of the dynamics-adjusted vector autoregressive models, and it is shown that the asymptotic results are identical to those in Johansen (1996, ch. 10, 11). Section 5 gives several model extensions and a survey of the related literature. Section 6 presents an empirical illustration, while Section 7 conducts simulation experiments for one of the dynamics-adjusted vector autoregressive model. An overall summary and conclusion are provided in Section 8.

This paper uses the following notational conventions. For a certain matrix a with full column rank, $\overline{a} = a(a'a)^{-1}$ and so $a'\overline{a} = I$. An orthogonal complement $a_\perp$ is defined such that $a'_\perp a = 0$ with the matrix $(a, a_\perp)$ being of full rank. The symbols $\overset{w}{\to}$ and $\overset{p}{\to}$ are used to signify weak convergence and convergence in probability, respectively.

2 Models with Adjusted Short-Run Dynamics

A general expression of dynamics-adjusted vector autoregressive (DAVAR) models is presented first, before turning to a class of three sub-models.

Let us consider a class of vector autoregressive models with adjusted short-run dynamics based on equation (2):

$$\triangle^2 X_t = (\Pi, \Pi_l) \begin{pmatrix} X_{t-1} \\ t \end{pmatrix} - \Gamma \triangle X_{t-1} + \mu + \Phi V_t + \varepsilon_t, \tag{3}$$

where $V_t \in \mathbf{R}^q$ is a set of short-run dynamics, which will differ according to the model specifications considered below. Deterministic terms, t and μ, are also included in equation (3). The linear trend t is restricted in a possible cointegration space so as to avoid the generation of a quadratic trend (see Johansen, 1996, Ch.5). The innovations $\varepsilon_1, ..., \varepsilon_T$ have independent and identical normal $N(0, \Omega)$ distributions, and the starting values $X_{-k+1}, ..., X_0$ are conditioned upon. The parameters Π, Γ, $\Omega \in \mathbf{R}^{p\times p}$, Π_l, $\mu \in \mathbf{R}^p$ and $\Phi \in \mathbf{R}^{p\times q}$ vary freely and Ω is positive definite. The hypothesis of reduced cointegration rank is given by

$$\mathsf{H(r)}: \quad rank\,(\Pi, \Pi_l) \leq r \text{ or } (\Pi, \Pi_l) = \alpha(\beta', \gamma'),$$

where $\alpha, \beta \in \mathbf{R}^{p\times r}$ and $\gamma' \in \mathbf{R}^r$. For future reference we define $X^*_{t-1} = \left(X'_{t-1}, t\right)'$ and $\beta^* = (\beta', \gamma')'$. Specifying the general short-run dynamics ΦV_t in equation (3) generates three dynamics-adjusted models, which are presented in the following sub-sections.

2.1 Model Subject to Short-Run Parameter Changes

The first sub-model is a *cointegrated vector autoregressive model subject to short-run parameter changes*, where the idea is to allow different short-term dynamics in two or more separate regimes. Without loss of generality, the number of regimes is chosen to be two throughout the paper. The lengths of the first and second sub-samples are T_1 and T_2, respectively. The total sample is therefore given by $T = T_1 + T_2$, and we define $T_0 = 0$. Thus, the series are given by $X_1, ..., X_{T_1}$ and $X_{T_1+1}, ..., X_T$, and the parameters of the lagged second-order differences $\triangle^2 X_{t-i}$ are considered to be different between the two periods. Therefore, the adjusted short-run dynamics term ΦV_t is defined as

$$\Phi V_t = \sum_{i=1}^{k-2} \Psi_i^{(1)} \left\{ \triangle^2 X_{t-i} 1_{(0<t\leq T_1)} \right\} + \sum_{i=1}^{k-2} \Psi_i^{(2)} \left\{ \triangle^2 X_{t-i} 1_{(T_1<t\leq T)} \right\}, \tag{4}$$

where $\Psi_i^{(1)}$, $\Psi_i^{(2)} \in \mathbf{R}^{p\times p}$, and $1_{(\cdot)}$ is an indicator function. Note that the long and medium term parameters α, β^*, Γ, μ, Ω are common for the two periods, while the parameters for the short-term dynamics change from $\Psi_1^{(1)}, ..., \Psi_{k-2}^{(1)}$ to $\Psi_1^{(2)}, ..., \Psi_{k-2}^{(2)}$. A review of other cointegration models with parameter shifts is given in Section 5. This includes models allowing shifts in deterministic terms, along the lines of Johansen, Mosconi and Nielsen (2000). An empirical illustration using US gasoline price data is presented in Section 6.

2.2 Model with Partially Reduced Short-Run Dynamics

The second sub-model is a *cointegrated vector autoregressive model with partially reduced short-run dynamics*, where the idea is to allow different lag lengths for the components of the second-order differenced terms. Decompose $X_t = \left(X'_{1,t}, X'_{2,t}\right)'$ and define ΦV_t as

$$\Phi V_t = \sum_{i=1}^{l-2} \Psi_i \triangle^2 X_{t-i} + \sum_{s=l+m-1}^{k-2} \Psi_s^{(block)} \triangle^2 X_{t-s}, \quad \text{for } 0 \leq m, \tag{5}$$

where $\Psi_s^{(block)}$ represents a matrix with zero elements in the block columns corresponding to $X_{2,t}$, that is,

$$\Psi_s^{(block)} = \begin{pmatrix} \Psi_{s,1} & 0 \\ \Psi_{s,2} & 0 \end{pmatrix}. \tag{6}$$

Note that $\triangle^2 X_{t-l-m+1}$ does not have to be consecutive to $\triangle^2 X_{t-l+2}$. Also note that (6) can be defined such that zero elements in the block columns correspond to $X_{1,t}$ rather than $X_{2,t}$. The adjusted short-run dynamics term (5) allows us to analyse models with complicated lag structure; for instance, a model for time series subject to multiplicative seasonal effects.

2.3 Model Augmented with Short-Run Explanatory Variables

The final specification is a *partially cointegrated vector autoregressive model augmented with short-run explanatory variables*, where ΦV_t is specified as

$$\Phi V_t = \sum_{i=1}^{k-2} \Psi_i \bigtriangleup^2 X_{t-i} + \sum_{s=0}^{n-2} \Lambda_s \bigtriangleup^2 Z_{t-s}, \tag{7}$$

where Z_{t-s} is a set of v-dimensional explanatory variables with a parameter $\Lambda_s \in \mathbf{R}^{p\times v}$. The process Z_t is assumed to be $I(1)$ or of smaller order and also independent of ε_t, so that Z_t is strongly exogenous. Formal conditions that Z_t needs to satisfy are given in the next sub-section. This model is mentioned by Rahbek and Mosconi (1999), although they focus on the case where $\bigtriangleup Z_t$, in our notation, is $I(1)$. Empirical examples of Z_t are oil prices used in Rahbek and Mosconi (1999).

Note that the model with equation (4) is a *parameter-changing* DAVAR model, while the models specified by equations (5) and (7) are *parameter-constant* DAVAR models.

3 Statistical Analysis of Cointegration

This section presents a statistical analysis of the three DAVAR models specified by equations (4), (5) and (7). The Granger-Johansen representation of each sub-model is investigated, followed by a discussion of the maximum likelihood analysis.

3.1 Granger-Johansen Representations

The Granger-Johansen representations of the three DAVAR models are closely related to those of Johansen (1996, Theorem 4.2) and Johansen, Mosconi and Nielsen (2000). A set of conditions required for the representation varies slightly according to the model specification, because the models have different characteristic polynomials. For the general form (3), it is assumed that the following conditions hold.

Assumption 1

1. *The matrices α and β have full column rank r.*
2. *The matrix $\alpha'_\perp \Gamma \beta_\perp$ has full rank $p-r$.*

The first condition implies that there are at least $p-r$ common stochastic trends and cointegration arises when $r \geq 1$. The second condition prevents the process from being $I(2)$ or of higher order. Note that these two conditions only involve medium and long term parameters. A combination of these conditions ensures that the number of common stochastic trends is exactly $p-r$. A further condition is needed to ensure all the characteristic roots, but the the $p-r$ unit roots are stationary. That condition will depend on the choice of V_t and will be discussed below.

Under Assumption 1 and a suitable assumption on the characteristic roots, the Granger-Johansen representation will have the appearance

$$X_t = C \sum_{s=1}^{t} \varepsilon_s + y_t + \tau_c + \tau_l t + A_t, \tag{8}$$

where $C = \beta_{\perp}(\alpha'_{\perp}\Gamma\beta_{\perp})^{-1}\alpha'_{\perp}$ is the impact parameter of the common stochastic trends, and the process y_t has zero mean and satisfies a Law of Large Numbers. The parameters τ_c and τ_l satisfy

$$\beta'\tau_c = \overline{\alpha}'\left(\Gamma C - I\right)\mu + \overline{\alpha}'\left(\Gamma C\Gamma - \Gamma\right)\overline{\beta}\gamma' - \gamma', \tag{9}$$

$$\tau_l = C\mu + \left(C\Gamma - I\right)\overline{\beta}\gamma', \tag{10}$$

and A_t depends on the initial values such that $\beta' A_t = 0$. The expression for $\beta'_{\perp}\tau_c$ is more lengthy (see Hansen, 2005). In particular, $\beta' X_t + \gamma' t$ and $\triangle X_t$ can also be given stationary initial distributions. In the representation (8), y_t and A_t vary according to the specifications (4), (5) and (7) due to the fact that the characteristic polynomials are different between the models. Therefore each of the sub-models is considered in turn.

3.1.1 Representation for the Model Subject to Short-Run Parameter Changes

Consider the vector autoregression with shifts in short-run parameters as given by equation (4). In addition to Assumption 1, an assumption on the characteristic polynomial is needed.

Assumption 2 *The characteristic polynomial for the model (3), in which the short-run dynamics ΦV_t satisfies equation (4), is given by*

$$A^{(j)}(z) = (1-z)^2 I_p - \alpha\beta' z + \Gamma(1-z)z - \sum_{i=1}^{k-2}\Psi_i^{(j)} z^i (1-z)^2,$$

for sub-samples $j = 1, 2$. *For each sub-sample, the characteristic roots solving the equation* $\det\{A^{(j)}(z)\} = 0$ *satisfy* $|z| > 1$ *or* $z = 1$.

Assumption 2 ensures that the process is neither explosive nor seasonally cointegrated. It follows immediately from Johansen (1996, Theorem 4.2) that, for each sub-sample $X_{T_{j-1}+1}, ..., X_{T_j}$, the initial values $X_{T_{j-1}-k+1}, ..., X_{T_{j-1}}$ can be given a distribution so that the representation (8) holds where $y_t^{(j)}$ are stationary processes described as infinite series of past innovations. This result is refined in the following theorem, which gives a solution with two additional features as in Nielsen (2001, Lemma A1). First, it holds throughout the full sample as a function of the two sets of initial values, thereby generalising the work of Hansen (2005) to a situation with two sub-samples. Secondly, the processes $y_t^{(j)}$ are expressed in terms of the observed regressors in the reduced rank regression equation (3). This property is convenient in that it helps interpreting estimators and test statistics, and in that it also holds without the stationarity in Assumption 2. The proof of the next theorem is based on Nielsen (2001) and is given in the Appendix.

Theorem 1 *Consider the model (3) satisfying equation (4). Suppose that Assumption 1 holds. Then the model equation has the solution (8) with* y_t *and* A_t *given by*

$$y_t = \sum_{j=1}^{2} y_t^{(j)} 1_{(T_{j-1}<t\leq T_j)} \quad \text{and} \quad A_t = A^{(1)} + A^{(2)} 1_{(t>T_1)},$$

where, for $j = 1, 2$,

$$\begin{aligned} y_t^{(j)} &= (I - C\Gamma)\overline{\beta}\beta' X_t + C(\Gamma - I) \triangle X_t + \sum_{i=1}^{k-2} C\Psi_i^{(j)} \triangle X_{t-i}, \\ A^{(1)} &= C\left\{\Gamma X_0 - (\Gamma - I) \triangle X_0 - \sum_{i=1}^{k-2} \Psi_i^{(1)} \triangle X_{-i}\right\}, \\ A^{(2)} &= C\sum_{i=1}^{k-2}\left\{\Psi_i^{(1)} - \Psi_i^{(2)}\right\} \triangle X_{T_1-i}. \end{aligned}$$

If in addition Assumption 2 holds, the initial values $X_{T_{j-1}-k+1}, ..., X_{T_{j-1}}$ *for each sub-sample can then be given a distribution so that the process* $y_{T_{j-1}+1}^{(j)}, ..., y_{T_j}^{(j)}$ *is stationary.*

Note that, in the representation (8), the stationary part $y_t^{(j)}$ is affected by the change in the short-term parameters $\Psi_i^{(j)}$, whereas both the impact matrix C and the linear trend parameter τ_l remain unchanged throughout the whole period. This isolation of the common trends from the parameter change plays a crucial role in the asymptotic analysis developed in Section 4 below.

3.1.2 Representation for the Case of Partially Reduced Short-Run Dynamics

Consider now the model with partially reduced short-run dynamics. As for the model with short-run parameter shifts, an assumption on the characteristic polynomial is required.

Assumption 3 *The characteristic polynomial for model (3) satisfying equation (5) is*

$$A^*(z) = (1-z)^2 I_p - \alpha\beta' z + \Gamma(1-z)z - \sum_{i=1}^{l-2} \Psi_i z^i (1-z)^2 - \sum_{s=l+m-1}^{k-2} \Psi_s^{(block)} z^s (1-z)^2 .$$

The characteristic roots solving the equation $\det\{A^*(z)\} = 0$ *satisfy* $|z| > 1$ *or* $z = 1$.

We can now turn to the representation theorem. Just as for the model with short-run parameter changes, this shows that the specification of the short-run dynamics has no bearing on the common trends.

Theorem 2 *Consider model (3) satisfying equation (5). Suppose that Assumption 1 holds. Then the model equation has the solution (8) with* y_t *and* A_t *given by*

$$y_t = (I - C\Gamma)\overline{\beta}\beta' X_t + C(\Gamma - I) \triangle X_t + \sum_{i=1}^{l-2} C\Psi_i \triangle X_{t-i} + \sum_{s=l+m-1}^{k-2} C\Psi_s^{(block)} \triangle X_{t-s}, \tag{11}$$

$$A_t = A = C\left\{\Gamma X_0 - (\Gamma - I)\triangle X_0 - \sum_{i=1}^{l-2} \Psi_i \triangle X_{-i} - \sum_{s=l+m-1}^{k-2} \Psi_s^{(block)} \triangle X_{-s}\right\}. \tag{12}$$

If in addition Assumption 3 holds, then the initial values $X_{-k+1}, ..., X_0$ *can be given a distribution so that the process* $y_1, ..., y_T$ *is stationary.*

Proof. See the Appendix. ■

3.1.3 Representation for the Model with Short-Run Explanatory Variables

Consider finally the model with short-run explanatory variables. Once again an assumption on the characteristic polynomial is required for the representation theorem.

Assumption 4 *The characteristic polynomial for the model (3) satisfying equation (7) is*

$$A^{**}(z) = (1-z)^2 I_p - \alpha\beta' z + \Gamma(1-z)z - \sum_{i=1}^{k-2} \Psi_i z^i (1-z)^2.$$

The characteristic roots solving the equation $\det\{A^{**}(z)\} = 0$ *satisfy* $|z| > 1$ *or* $z = 1$.

Since a set of additional regressors Z_t is now incorporated in the model, an additional condition on Z_t needs to be introduced. Following Rahbek and Mosconi (1999), the regressor Z_t is assumed to be a strongly exogenous linear process (see Phillips and Solo, 1992, for details of linear processes). This is presented as follows.

Assumption 5 *The first-order difference of the process* Z_t *satisfies*

$$\triangle Z_t = B(L)\eta_t, \quad \text{for } t = 1, ..., T,$$

where $\eta_1, \cdots, \eta_T$ *are independently* $N(0, \Sigma)$*-distributed conditional on the initial value* Z_0*, and the process* Z_t *is independent of* ε_t*. The polynomial* $B(z) = \sum_{i=0}^{\infty} B_i z^i$ *is convergent for* $|z| < 1 + \delta$ *for some* $\delta > 0$*.*

Note that $\det\{B(1)\} \neq 0$ is not assumed so Z_t can be either $I(1)$ or of smaller order. In Assumption 5 the normal distribution is adopted for the sake of simplicity, and the asymptotic results given below hold for a broader class of stationary processes. See Hansen (1992a) for details of such processes. These settings lead to the representation theorem, which once again shows that the short-term dynamics have no impact on the common trends.

Theorem 3 *Consider the model (3) satisfying equation (7). Suppose that Assumption 1 holds. Then the model equation has the solution (8) with* y_t *and* A_t *given by*

$$y_t = (I - C\Gamma)\overline{\beta}\beta' X_t + C(\Gamma - I) \triangle X_t + \sum_{i=1}^{k-2} C\Psi_i \triangle X_{t-i} + \sum_{s=1}^{n-2} C\Lambda_s \triangle Z_{t-s}, \quad (13)$$

$$A_t = A = C\left[\Gamma X_0 - (\Gamma - I) \triangle X_0 - \sum_{i=1}^{k-2} \Psi_i \triangle X_{-i} - \sum_{s=0}^{n-2} \Lambda_s \triangle Z_{-s}\right]. \quad (14)$$

If in addition Assumptions 4 and 5 hold, then the initial values $X_{-k+1}, ..., X_0$ *can be given a distribution so that the process* $y_1, ..., y_T$ *is stationary.*

Proof. See the Appendix. ■

3.2 Reduced Rank Regression

The maximum likelihood analysis of cointegration rank in model (1) is based on the reduced rank regression of ΔX_t and X^*_{t-1} corrected for ΔX_{t-1} and all the other regressors. The correction for ΔX_{t-1} implies that this is equivalent to analysing model (2) by the reduced rank regression of $\Delta^2 X_t$ and X^*_{t-1} corrected for all the other regressors. This equivalence also enables us to apply the standard reduced rank regression to the general form of DAVAR models given by equation (3).

Following Johansen (1996, ch. 6) it is convenient to start by introducing the notation:

$$Z_{0t} = \Delta^2 X_t, \quad Z_{1t} = X^*_{t-1} = \begin{pmatrix} X_{t-1} \\ t \end{pmatrix}, \quad Z_{2t} = \begin{pmatrix} \Delta X_{t-1} \\ 1 \end{pmatrix}, \quad Z_{3t} = V_t.$$

The likelihood function is then maximized by regressing Z_{0t} on Z_{2t} and Z_{3t} to obtain residuals R_{0t}, and also regressing Z_{1t} on Z_{2t} and Z_{3t} to get residuals R_{1t}. The sample product moment matrix for the residuals is defined as

$$\begin{pmatrix} S_{00} & S_{01} \\ S_{10} & S_{11} \end{pmatrix} = \frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} R_{0t} \\ R_{1t} \end{pmatrix} \begin{pmatrix} R_{0t} \\ R_{1t} \end{pmatrix}'. \tag{15}$$

The log-likelihood ratio for the hypothesis of at most r cointegrating relations, $H(r)$, against $H(p)$ is given by

$$LR\left\{H(r)|H(p)\right\} = -T\sum_{i=r+1}^{p} \log(1-\widehat{\lambda}_i), \tag{16}$$

where $1 \geq \widehat{\lambda}_1 \geq \cdots \geq \widehat{\lambda}_p \geq 0$ are solutions to the following generalised eigenvalue problem:

$$\det\left(\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}\right) = 0. \tag{17}$$

The problem (17) can be solved in a numerically stable way using singular value decompositions (see Doornik and O'Brien, 2002). The asymptotic analysis of the rank test statistic and the maximum likelihood estimators is discussed in the next section.

4 Asymptotic Analysis of Cointegration

It is first demonstrated for each of the three DAVAR models that the asymptotic distributions of the rank test statistics correspond to that of the conventional cointegration rank test. Subsequently it is argued that the asymptotic properties of estimators are also the same as those in the standard cointegrated vector autoregressive model.

Consider the general form (3) with Assumption 1 satisfied. It can be shown that, as $T \to \infty$, the $\log LR$ test statistic (16) has an asymptotic distribution given by

$$LR\left\{H(r)|H(p)\right\} \xrightarrow{w} tr\left\{\int_0^1 dB\left(u\right)F'\left(\int_0^1 FF'du\right)^{-1}\int_0^1 FdB\left(u\right)'\right\}, \qquad (18)$$

where $B\left(u\right)$ is a $(p-r)$ dimensional standard Brownian motion and F is a $(p-r+1)$ dimensional process consisting of

$$F = \begin{cases} B(u) - \int_0^1 B(u)du, \\ u - \frac{1}{2}, \end{cases}$$

for $u \in [0,1]$. The limiting distribution (18) is the same as that in Johansen (1996, ch. 6). This result holds irrespective of the three model specifications, (4), (5) and (7). In other words, adjusted short-run dynamics have no impact on the asymptotic distribution of the cointegration rank test. Therefore, the tables in Johansen (1996, ch.15), also see Doornik (1998), can be utilised. As a consequence, one can also make conventional χ^2-based asymptotic inferences for tests on the adjustment and cointegration spaces, as in Johansen (1996, ch.13).

All the proofs in this section follow Johansen (1996, ch. 10, 11), although some modifications are required for the model with equation (4) in order to take account of changes in the short-run parameters.

4.1 Rank Test in the Parameter-Changing DAVAR Model

As described above, the asymptotic distribution of the rank test statistic (16) for the parameter-changing DAVAR model given by equation (4) is given by equation (18). This result, together with the required conditions, are presented as the following theorem:

Theorem 4 *Consider the model (3) satisfying equation (4). Suppose (i) $T_j/T = a^{(j)}$ is fixed while $T \to \infty$ and (ii) Assumptions 1 and 2 are satisfied. Then, as $T \to \infty$, the asymptotic distribution of the rank test statistic (16) is given by equation (18).*

Proof. Follow the proof of Theorem 11.1 in Johansen (1996) using Lemmas 5 and 6 given below instead of Lemmas 10.1 and 10.3 in Johansen (1996). ■

The asymptotic properties of the product moment matrices, S_{00}, $S_{01}\beta^*$, $\beta^{*\prime}S_{11}\beta^*$, need to be investigated in order to adapt Lemmas 10.1 and 10.3 of Johansen (1996) to the present model. As a prerequisite, define the variance-covariance matrix of the stationary processes for each sub-sample

$$\widehat{Var}\left(\begin{array}{c|c} \triangle^2 X_t & \\ \beta^{*\prime}X^*_{t-1} & \triangle^2 X_{t-1}, \ldots, \triangle^2 X_{t-k+2} \\ \triangle X_{t-1} & \end{array}\right) \xrightarrow{p} \begin{pmatrix} \Sigma^{(j)}_{00.3} & \Sigma^{(j)}_{0\beta.3} & \Sigma^{(j)}_{02.3} \\ \Sigma^{(j)}_{\beta 0.3} & \Sigma^{(j)}_{\beta\beta.3} & \Sigma^{(j)}_{\beta 2.3} \\ \Sigma^{(j)}_{20.3} & \Sigma^{(j)}_{2\beta.3} & \Sigma^{(j)}_{22.3} \end{pmatrix}.$$

Furthermore, define

$$\Sigma_{lm.3} = \sum_{j=1}^{2} a^{(j)}\Sigma^{(j)}_{lm.3}, \quad \text{for } l, m = 0, 2, \beta, \tag{19}$$

$$\Sigma_{lm} = \Sigma_{lm.3} - \Sigma_{l2.3}\Sigma^{-1}_{22.3}\Sigma_{2m.3}, \quad \text{for } l, m = 0, \beta. \tag{20}$$

The asymptotic properties of the sample product moment matrices can then be stated.

Lemma 5 *Suppose the assumptions in Theorem 4 are satisfied. Then,*

$$\begin{pmatrix} S_{00} & S_{01}\beta^* \\ \beta^{*\prime}S_{10} & \beta^{*\prime}S_{11}\beta^* \end{pmatrix} \xrightarrow{p} \begin{pmatrix} \Sigma_{00} & \Sigma_{0\beta} \\ \Sigma_{\beta 0} & \Sigma_{\beta\beta} \end{pmatrix}, \tag{21}$$

where

$$\Sigma_{00} = \alpha\Sigma_{\beta 0} + \Omega, \quad \Sigma_{0\beta} = \alpha\Sigma_{\beta\beta}, \quad \Sigma_{00} = \alpha\Sigma_{\beta\beta}\alpha' + \Omega. \tag{22}$$

Proof. See the Appendix. ■

The derived results (21) and (22) match those in Johansen (1996, Lemmas 10.1 and 10.3), although the required proofs are more involved. As shown in the Appendix, this is because the sample product moments in each sub-sample converge to their population values, thereby their linear combinations using $a^{(j)}$ can also be defined accordingly. As shown above, Lemma 5 is required in the proof of Theorem 4.

Next, we investigate the asymptotic properties of non-stationary components with a view to adjusting Lemma 10.3 of Johansen (1996) to the present model. Let

$$B_T' X_{t-1}^* = \begin{pmatrix} \alpha_\perp' \Gamma & -\alpha_\perp' \Gamma \tau_l \\ 0 & T^{-1/2} \end{pmatrix} \begin{pmatrix} X_{t-1} \\ t \end{pmatrix}$$
$$= \begin{bmatrix} \alpha_\perp' \Gamma \left\{ C \sum_{s=1}^{t-1} \varepsilon_s + Y_{t-1}^{(j)} + \tau_c - \tau_l + A^{(1)} + A^{(2)} 1_{(t-1>T_1)} \right\} \\ T^{-1/2} t \end{bmatrix}, \qquad (23)$$

which isolates the deterministic trend in the final row of the vector. Then, we prove the following lemma:

Lemma 6 *Suppose the assumptions in Theorem 4 are satisfied. Then, as* $T \to \infty$*,*

$$T^{-1/2} \sum_{s=1}^{int(Tu)} \varepsilon_s \xrightarrow{w} W(u) \quad and \quad T^{-1/2} B_T' X_{int(Tu)}^* \xrightarrow{w} \begin{pmatrix} \alpha_\perp' W(u) \\ u \end{pmatrix},$$

where $W(u)$ *is a Brownian motion in* $p - r$ *dimensions with variance matrix* Ω*. The asymptotic distributions of the non-stationary product moments are*

$$B_T' S_{11} \beta^* \in O_p(1), \quad T^{-1} B_T' S_{11} B_T \xrightarrow{w} \int_0^1 G G' du, \quad B_T'(S_{10} - S_{11} \beta^* \alpha') \xrightarrow{w} \int_0^1 G (dW(u))',$$

where

$$G = \begin{bmatrix} \alpha_\perp' \left\{ W(u) - \int_0^1 W(u) du \right\} \\ u - \frac{1}{2} \end{bmatrix}.$$

Proof. The stationary processes $Y_{t-1}^{(j)}$ and the terms τ_c, τ_l, $A^{(1)}$, $A^{(2)} 1_{(t-1>T_1)}$ appearing in equation (23) are all of order $O_p(1)$ uniformly in t. Therefore,

$$B_T' X_{t-1}^* = \begin{pmatrix} \alpha_\perp' \Gamma C \sum_{s=1}^{t-1} \varepsilon_s + O_p(1) \\ T^{-1/2} t \end{pmatrix}. \qquad (24)$$

The desired results follow as in the proofs for Lemmas 10.2 and 10.3 in Johansen (1996) since the random walk term is of order $T^{1/2}$ and thus dominates other terms of order $O_p(1)$. ■

The limiting distributions in Lemma 6 are identical to those in Johansen (1996, Lemma 10.3). This is because all the stationary processes are irrelevant asymptotically and the parameter change has no impact on C as shown in Theorem 1.

4.2 Rank Test in the Parameter-Constant DAVAR Models

We now turn to the parameter-constant models specified by equations (5) and (7). Once again the asymptotic distribution of the rank test statistic (16) is given by equation (18). The proofs given in Johansen (1996) are now more directly adaptable, since no parameter shifts are involved.

Theorem 7 *Consider the two parameter-constant DAVAR models with Assumption 1 satisfied. Suppose that either of the following two cases holds:*
1. V_t *is specified by equation (5) with Assumption 3 fulfilled,*
2. V_t *is specified by equation (7) with Assumptions 4 and 5 fulfilled.*
Then, as $T \to \infty$*, the asymptotic distribution of the rank test statistic (16) is given by equation (18).*

Proof. Since there is no parameter shift in either model, the following variance-covariance matrix of the stationary processes can be defined:

$$\widehat{Var}\left(\begin{array}{c} \triangle^2 X_t \\ \beta^{*\prime} X^*_{t-1} \end{array} \middle| \triangle X_{t-1}, V_t \right) \xrightarrow{p} \left(\begin{array}{cc} \Sigma_{00} & \Sigma_{0\beta} \\ \Sigma_{\beta 0} & \Sigma_{\beta\beta} \end{array} \right),$$

to which Lemma 10.1 in Johansen (1996) is applicable. As in the proof of Lemma 6, the stationary processes Y_t and the terms τ_c and A_t in the representation (8) are all of order $O_p(1)$ uniformly in t, so equation (24) holds. Lemmas 10.2 and 10.3 in Johansen (1996) also hold, since the random walk term is of order $T^{1/2}$ and thus dominates other terms of order $O_p(1)$. Finally, the proof of Theorem 11.1 in Johansen (1996) can be used so as to reach the asymptotic distribution given by equation (18). ■

4.3 Properties of the Estimators

Finally, let us consider the asymptotic properties of the estimators, $\widehat{\alpha}$ and $\widehat{\beta}$, for the adjustment vectors and for the cointegrating relations. In the standard cointegrated vector autoregression, the results in Lemmas 10.1 and 10.3 in Johansen (1996) provide the basis for the limiting properties of the estimators. Based on these two lemmas, Johansen (1996, ch. 13) shows that the estimators $\widehat{\alpha}$ and $\widehat{\beta}$ have asymptotic normal and mixed normal distributions, respectively.

For the parameter changing DAVAR model given by equation (4), Lemmas 10.1 and 10.3 in Johansen (1996) were replaced by Lemmas 5 and 6 in Section 4.1. These lemmas have the same appearance. It can then be shown that Lemmas 13.1 and 13.2 in

Johansen (1996) hold based on these new lemmas instead of Lemmas 10.1 and 10.3 in Johansen (1996). Thus $\widehat{\alpha}$ and $\widehat{\beta}$ have limiting normal and mixed normal distributions respectively, so conventional χ^2-based asymptotic inferences for $\widehat{\alpha}$ and $\widehat{\beta}$ apply.

For the parameter constant DAVAR models given by equation (5) or (7), Lemmas 10.1 and 10.3 in Johansen (1996) apply as argued in the proof of Theorem 7. This again leads to conventional χ^2-based asymptotic inferences for $\widehat{\alpha}$ and $\widehat{\beta}$.

5 Model Extensions and Related Issues

This section considers several extensions of the vector autoregressive model with short-run parameter changes, and also presents a survey of the literature related to parameter shifts in cointegrated processes.

Johansen, Mosconi and Nielsen (2000) investigate shifts in deterministic terms (a linear trend and constant) in cointegrated vector autoregressive models. Thus, the combination of the analysis in this paper and Johansen, Mosconi and Nielsen (2000) enables us to extend the model with equation (4) to

$$\triangle^2 X_t = \alpha \begin{pmatrix} \beta \\ \gamma^{(j)} \end{pmatrix}' \begin{pmatrix} X_{t-1} \\ t \end{pmatrix} - \Gamma \triangle X_{t-1} + \mu^{(j)} + \sum_{i=1}^{k-2} \Psi_i^{(j)} \triangle^2 X_{t-i} + \varepsilon_t, \qquad (25)$$

where structural breaks occur in the deterministic trend and constant as well as in the short-term dynamics. The same type of asymptotic results as developed above also apply to a model with a restricted constant:

$$\triangle^2 X_t = \alpha \begin{pmatrix} \beta \\ \mu^{(j)} \end{pmatrix}' \begin{pmatrix} X_{t-1} \\ 1 \end{pmatrix} - \Gamma \triangle X_{t-1} + \sum_{i=1}^{k-2} \Psi_i^{(j)} \triangle^2 X_{t-i} + \varepsilon_t, \qquad (26)$$

in which broken constant levels are allowed in addition to the changes in the short-run parameters. Economies which have experienced relatively large regime changes could be described by these models. An empirical illustration using model (26) is presented in Section 6. We should note that parameter break points in the deterministic terms and short-run parameters do not need to coincide, as shown in the empirical illustration.

In order to perform rank tests in equations (25) and (26) one needs to use the response surfaces in Johansen, Mosconi and Nielsen (2000), which are conducted for the cases of deterministic shifts embodied in $\gamma^{(j)}$ and $\mu^{(j)}$. Using the same argument as above, it can be argued that the changing short-term dynamics do not affect the asymptotic properties of the rank test statistics, so we are also able to use estimated response surfaces given in Johansen, Mosconi and Nielsen (2000) for the analysis of the extended models (25) and (26). In practice, additional indicator variables need to be incorporated in the models for the likelihood function to be conditional on the initial values of each sub-sample; see Johansen, Mosconi and Nielsen (2000) and the empirical illustration given below for details.

The cointegrated vector autoregressive model subject to short-run parameter shifts belongs to a class of structural changes in cointegrated models, which have been extensively discussed in the literature. Hansen (1992b), Quintos and Phillips (1993), and Campos, Ericsson and Hendry (1996) would be the earliest papers, although not necessarily within a vector autoregressive framework. Parameter changes in cointegrated vector autoregressions are studied in several papers such as Seo (1998), Hansen and Johansen (1999), and Hansen (2000, 2003). The latter paper, in particular, considered a number of possible patterns of parameter changes using generalised reduced rank regression. However, the existing literature is limited in terms of testing cointegration rank in the presence of parameter changes, *i.e.* the number of cointegrating vectors is often assumed to be given. Exceptions are Inoue (1999), Johansen, Mosconi and Nielsen (2000) and Andrade, Bruneau and Gregoir (2005). The first two papers address rank tests with breaks in deterministic terms such as a linear trend, and the paper by Johansen, Mosconi and Nielsen (2000), as reviewed above, provides a general framework for cointegration analysis in such cases.

While this paper considers cointegration analysis in the presence of short-run parameter changes, extensions could allow changes in the medium and long term parameters. Changes in α and β are investigated in Andrade, Bruneau and Gregoir (2005). These types of changes, however, affect the limiting distributions of the common stochastic trends through the corresponding changes in the C matrix. Therefore, the asymptotic arguments in such cases can be much more involved than those of changes in Ψ_i. Andrade, Bruneau and Gregoir (2005) address these issues using principal components analysis, but as a consequence the conventional reduced rank procedure is no longer applicable. The other extension would be to allow the covariance matrix, Ω, to be shifted. However, changes in Ω lead to multiple reduced rank conditions with the consequence that the ordinary procedure can no longer be used. See Hansen (2000, 2003), Cavaliere and Taylor (2006), for the cases where Ω is time-varying. From a viewpoint of operational applied work, there would be room for further research on cointegration rank tests in the presence of changes in other parameters than the short-run parameters.

It is worth noting that there are practical situations where the locations of parameter shifts are unknown to researchers. The issue of unknown change points in cointegrated systems is discussed by Seo (1998), Hansen and Johansen (1999), Hansen (2000), and Andrade, Bruneau and Gregoir (2005). The identification of change points with unknown cointegration rank would be one of the most important issues to be further addressed. Identifying parameter break points is also an issue for the DAVAR model subject to short-run parameter shifts (4). This issue is discussed using a simulation study in Section 7.

6 Empirical Illustration

This section provides an empirical illustration of the dynamics-adjusted model with short-run parameter changes as defined by equations (3) and (4). The data set is composed of two weekly gasoline prices observed at different locations in the United States over the period 1987.24 to 1998.29. The number of observations is therefore 578. These data were previously analysed by Hendry and Juselius (2001). While their paper is mainly methodologically oriented, their discussion (p. 79) indicates that the data have been used in a study of non-competitive behaviour. This is presumably the reason why no more details are available about the origin of the data. Hunter and Burke (2007) have recently followed up on the point that cointegration analysis can potentially reveal non-competitive behaviour.

Three models are presented in this section. The first model is a second order vector autoregression. This model does not quite capture the temporal dependence in the data. The second and third models seek to describe the temporal dependence more accurately. The second model employs a fifth order vector autoregression with structural shifts in the intercept as in equation (26). The second model is similar to that used by Hendry and Juselius (2001). This model addresses the autocorrelation issue found in the first model, but leaves some autoregressive conditional heteroscedasticity (ARCH) unmodelled. The ARCH effects are addressed in the third model, which is a VAR model adjusted with short-run parameter changes. The latter two models give similar results for the cointegration analysis, but they give a different way of addressing the short-run temporal dependence, with the third model quite possibly giving a more accurate data-description. This will matter in situations where the short-run dependence is of interest in addition to the long-run analysis.

6.1 A Standard Vector Autoregression

The logs of the prices ($p_{a,t}$ and $p_{b,t}$) are displayed in Figure 1(a)-(b), and residuals (u_a and u_b) from a simple unrestricted second order vector autoregression model with a constant are presented with $\pm 2 \times$ standard error bands in Figure 1(c)-(d). Upsurges in both of the series, together with corresponding outliers in the residuals, are evident at the 31st week in 1990 corresponding to the start of the Gulf war.

The application of residual autocorrelation tests for the pre-Gulf war period (1987.26 - 1990.30) and for the post-war period (1990.33 - 1998.29), as reported in Table 1, indicate unmodelled and changing temporal dependence. Section 7 reports a simulation analysis of the properties of such residual autocorrelation tests for the sub-samples.

Figure 1

US Gasoline Prices and Residuals from the
Constant Parameter second order Vector Autoregression

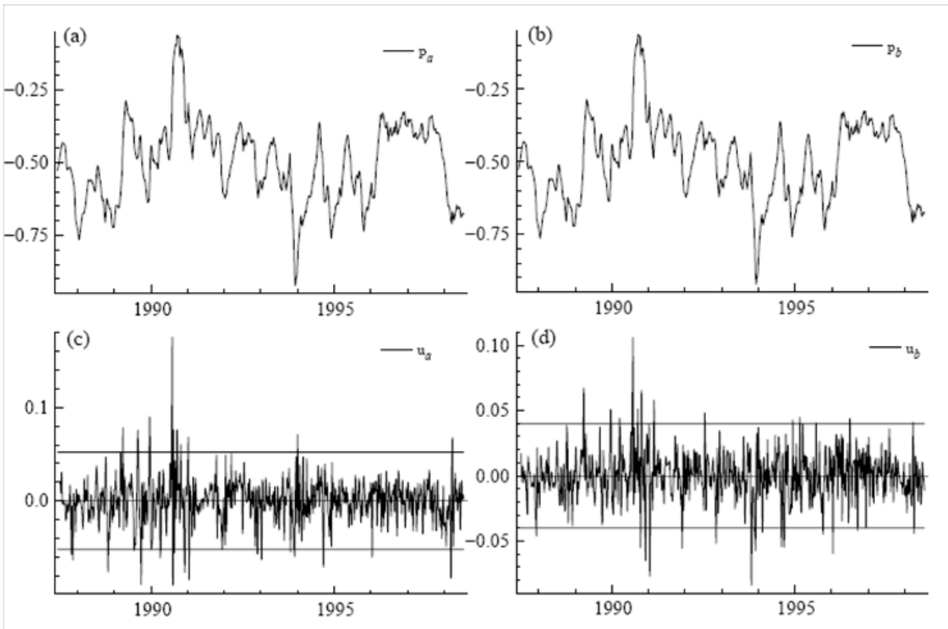


Table 1: Tests for residual autocorrelation in two sub-samples.

	$p_{a,t}$	$p_{b,t}$
Pre-war	$F_{ar}(1,155) = 6.31[0.01]$	$F_{ar}(1,155) = 10.16\ [0.00]$
Post-war	$F_{ar}(1,407) = 0.01\ [0.93]$	$F_{ar}(1,407) = 0.003\ [0.96]$

Notes: $F_{ar}(k,\cdot)$ is a test for kth order serial correlation reported as an F statistic (see Godfrey, 1978; Nielsen, 2007).

6.2 A Model with Intercept Break

The temporal dependence found above is not adequately described by a standard vector autoregression. Thus, Hendry and Juselius (2001) adjusted the vector autoregression with a step dummy variable $1_{(t\geq 1990.31)}$, as well as the impulse dummies

$$1_{(t=89.13)},\ 1_{(t=89.39)},\ 1_{(t=89.51)},\ 1_{(t=90.31)},\ 1_{(t=90.49)},\ 1_{(t=91.03)},\ 1_{(t=93.43)},$$

and the blip dummy

$$\Delta 1_{(t=98.11)} = 1_{(t=98.11)} - 1_{(t=98.12)}.$$

Following the suggestion of Johansen, Mosconi and Nielsen (2000), the sample is broken up in two, $t < 1990.31$ and $t \geq 1990.31$. For each of the two sub-samples the likelihood is conditioned on the first five observations, which is equivalent to introducing a further set of five dummies:

$$1_{(t=90.31)},\ 1_{(t=90.32)},\ \cdots,1_{(t=90.35)},$$

of which the first is already included above so it is redundant. Both the constant and the step dummy are to be restricted to the cointegrating space as described in equation (26).

The inclusion of these 12 dummies appears to lead to a well-specified model. The starting model is a 5-lag model so the sample period runs from 1987.29 to 1998.29. F-type tests for lag-length reduction support a 2-lag model, as reported in Table 2. A simulation study reported in Section 7 will shed more light on these lag reduction tests in the context of the third model presented below.

Table 2: Lag-length reduction tests for vector autoregression with broken intercept.

4 vs. 5:	F = 2.03 [0.09]	3 vs. 4:	F = 1.50 [0.19]
2 vs. 3:	F = 2.14 [0.07]	2 vs. 4:	F = 1.83 [0.07]

A 2-lag model has been selected, so that the likelihood is conditioned on the first two observations of each sub-sample. Mis-specification analysis is then conducted as reported in Table 3. Given the large number of observations and the multiple testing, it would seem reasonable to adopt a significance level of 1% or even 0.5%. The issue of temporal dependence discussed in Section 6.1 is not evident in the autocorrelation tests, but still seems to appear in the ARCH tests. The normality tests also seem somewhat unsatisfactory, with some excess kurtosis present. This level of mis-specification is tolerable in the context of a cointegration analysis according to the studies of Rahbek, Hansen and Dennis (2002) for ARCH deviations and Cheung and Lai (1993) and Gonzalo (1994) for non-normality. A cointegration analysis can therefore be carried out, with critical values derived from the response surfaces in Johansen, Mosconi and Nielsen (2000). The results are largely similar to those for the dynamics-adjusted model presented in Section 6.3, and therefore omitted.

Table 3: Mis-specification tests for vector autoregression with intercept break.

	$F_{ar}(1,557)$	$F_{ar}(4,554)$	$F_{arch}(1,556)$	$F_{arch}(4,550)$	$\chi^2_{nd}(2)$	$Sk.$	$Ku.$
$p_{a,t}$	2.09[0.15]	2.57[0.03]	9.72[0.002]	3.48[0.008]	19.92[0.00]	-0.14	4.43
$p_{b,t}$	0.21[0.65]	2.26[0.06]	6.75[0.010]	2.47[0.043]	18.41[0.00]	-0.12	3.94
	$F_{ar}(4,1110)$	$F_{ar}(16,1098)$	$\chi^2_{nd}(4)$				
System	0.64[0.63]	1.26[0.22]	56.03[0.00]				

Notes: $F_{arch}(k,\cdot)$ is a test for kth order ARCH (see Engle, 1982). $\chi^2_{nd}(\cdot)$ is a test for residual normality (see Doornik and Hansen, 1994[2008]), which is presented together with residual skewness (*Sk.*) and kurtosis (*Ku.*).

6.3 A Model with Changing Short-Run Dynamics

The model with an intercept break did not capture the short-run dynamics adequately as seen by the ARCH tests in Table 3. As an alternative, a new model with short-run parameter changes defined by equations (3) and (4) is therefore applied. This DAVAR model also allows for a break in intercept, so it is an empirical representation based on equation (26).

Given the discussion above, it is natural to allow a break in the short-term dynamics at the start of the war at 1990.31. In addition, Hendry and Juselius (2001) suggested a shift in the residual variance in late 1992, which we have dated to 1992.31. On the basis of the above reasoning, a 5-lag model of the following form is adopted:

$$\begin{aligned}\triangle^2 X_t = \left(\Pi, \Pi_c^{(1)}, \Pi_c^{(2)}\right) \begin{pmatrix} X_{t-1} \\ 1_{(t<1990.31)} \\ 1_{(t\geq 1990.36)} \end{pmatrix} - \Gamma \triangle X_{t-1} + \sum_{i=1}^{3} \Psi_i^{(1)} 1_{(t<1990.31)} \triangle^2 X_{t-i} \\ + \sum_{i=1}^{3} \Psi_i^{(2)} 1_{(1990.31\leq t<1992.31)} \triangle^2 X_{t-i} + \sum_{i=1}^{3} \Psi_i^{(3)} 1_{(t\geq 1992.31)} \triangle^2 X_{t-i} + \Lambda D_t + \varepsilon_t,\end{aligned} \tag{27}$$

where D_t represents the set of dummy variables in Section 6.2.

A lag-reduction analysis was carried out as reported in Table 4. The reduction to

4 lags appears least problematic, whereas the reductions to 3 or 2 are more marginal when using the 1% level. A reduction from a 4 lag model with changing short-term dynamics to a model with constant short-term dynamics in Section 6.2 is most clearly rejected, which provides evidence against constant short-term dynamics. We settle for a 4-lag model with changing short-term dynamics. Thus the likelihood is conditioned on the first four observations of each sub-sample, and the dummy $1_{(t \geq 1990.36)}$ allowing a break in the intercept is replaced by $1_{(t \geq 1990.35)}$. Mis-specification analysis is reported in Table 5, in which the ARCH effect is now eliminated, although the non-normality remains.

Table 4: Lag-length reduction tests for DAVAR with changing short-term dynamics.

4 vs. 5:	F = 1.65 [0.07]	3 vs. 4:	F = 1.79 [0.04]
2 vs. 3:	F = 1.83 [0.04]	4* vs. 4:	F = 2.20 [0.004]

Notes: The model indicated by 4* is a model without changing short-term dynamics as discussed in Section 6.2.

Table 5: Mis-specification tests for DAVAR(4) with 1 structural break in intercept and 2 in short-term dynamics.

	$F_{ar}(1,543)$	$F_{ar}(4,540)$	$F_{arch}(1,542)$	$F_{arch}(4,536)$	$\chi^2_{nd}(2)$	$Sk.$	$Ku.$
$p_{a,t}$	0.09[0.76]	1.90[0.11]	0.85[0.34]	0.57[0.68]	20.31[0.00]	-0.23	4.46
$p_{b,t}$	1.59[0.21]	1.28[0.28]	5.01[0.03]	2.96[0.02]	22.40[0.00]	-0.13	4.02
	$F_{ar}(4,1082)$	$F_{ar}(16,1070)$	$\chi^2_{nd}(4)$				
System	1.25[0.29]	1.44[0.11]	58.33[0.00]				

The results of a cointegration analysis are reported in Table 6. The upper panel gives rank tests, with critical values derived from the response surfaces in Johansen, Mosconi and Nielsen (2000). It is clearly rejected that the rank could be zero, whereas the hypothesis that the rank is at most one is not rejected at the 1% level. That decision is marginal, although consistent with an interpretation that the price level moves in a non-stationary way but coordinated across different locations. The lower panel of Table 6 gives an analysis of the cointegrating vector β and the adjustment vector α. A three dimensional hypothesis is considered, consisting of price homogeneity, weak exogeneity of $p_{b,t}$ (see Engle, Hendry and Richard, 1983; Johansen, 1992b) and a zero intercept for the period prior to the war. The first two restrictions were also made by Hendry and Juselius (2001), whereas the intercept restriction perhaps has a weaker substantive interpretation. The hypothesis cannot be rejected. The weak exogeneity of $p_{b,t}$ indicates a form of price leadership of $p_{b,t}$ over $p_{a,t}$, as the former drives the latter in the long run. Thus the gasoline price data are judged to be a simple example of non-competitive behaviour; see Hunter and Burke (2007) for details.

Table 6: Cointegration analysis.

$LR\,(0|2) = 71.61\ [0.00] \quad (c_{99}^{RS}(0) = 31.18)$
$LR\,(1|2) = 15.06\ [0.02] \quad (c_{99}^{RS}(1) = 16.53)$

	$p_{a,t}$	$p_{b,t}$	$\mu^{(1)}$	$\mu^{(2)}$		$p_{a,t}$	$p_{b,t}$	$\chi^2_{(df)}$
β	1	−1	0	−0.016 [0.007]	α	−0.12 [0.02]	0	$0.38[0.94]_{(df=3)}$

Notes: Upper panel reports rank determination. Lower panel reports first α, β with restrictions for homogeneity and weak exogeneity, and then for zero level in the first period.

A parsimonious representation can now be pursued with the aim of reducing the rather large number of parameters arising from the 4 lags and the changing short-term dynamics. Although the finding of $p_{b,t}$ being weakly exogenous allows us to model a single equation for $p_{a,t}$ given $p_{b,t}$, we estimate a joint bivariate system with a view to revealing time-varying structure of both variables. The series are mapped to $I(0)$ space by differencing and also using the restricted cointegrated relation, which leads to an equilibrium correction term in the $I(0)$ representation. *PcGets*, see Hendry and Krolzig (2001), helps us to reduce the general $I(0)$ model, which contains a number of short-run dynamic terms. We select the following reduced equilibrium correction model:

$$\begin{aligned}
\triangle^2 \widehat{p}_{a,t} &= \underset{(0.16)}{0.68}\,\triangle^2 p_{b,t} - \underset{(0.03)}{0.53}\,\triangle p_{a,t-1} + \underset{(0.08)}{0.49}\,\triangle p_{b,t-1} - \underset{(0.02)}{0.12} ecm_{t-1} \\
&\quad + \underset{(\mathbf{0.05})}{\mathbf{0.16}}\,\triangle^2 p_{a,t-1}\mathbf{1}_{(t<1990.31)} - \underset{(\mathbf{0.04})}{\mathbf{0.10}}\,\triangle^2 p_{a,t-2}\mathbf{1}_{(t\geq 1992.31)} + \widehat{\Lambda}_a D_t \qquad (28) \\
\triangle^2 \widehat{p}_{b,t} &= -\underset{(0.03)}{0.46}\,\triangle p_{b,t-1} + \underset{(\mathbf{0.08})}{\mathbf{0.28}}\,\triangle^2 p_{b,t-1}\mathbf{1}_{(t<1990.31)} + \widehat{\Lambda}_b D_t,
\end{aligned}$$

where

$$ecm_t = p_{a,t} - p_{b,t} - \mathbf{0.016} \times \mathbf{1}_{(t\geq 1990.35)}.$$

The first equation represents a model for $p_{a,t}$ conditional on $p_{b,t}$ and the second corresponds to a marginal model for $p_{b,t}$. No lagged values of $p_{a,t}$ are significant in the marginal model, so $p_{b,t}$ is considered to be not only weakly exogenous but also strongly exogenous (see Engle, Hendry and Richard, 1983, for strong exogeneity). Three short-run dynamics are judged to be significant, two of them are in the first sub-period (the pre-Gulf war period) and one of them is in the third period. Note that these are highlighted in bold. Other short-term dynamic terms are insignificant and therefore excluded from the model. It is also checked that the medium-run parameter Γ in the preferred model is constant throughout the whole sample.

The preferred model (28) shows that the short-run dynamic terms in the pre-war period, $\triangle^2 p_{a,t-1} 1_{(t<1990.31)}$ and $\triangle^2 p_{b,t-1} 1_{(t<1990.31)}$, are highly significant. These terms indicate a possible impact of the Gulf war on the underlying data generating mechanism, together with the level shift observed in the constant term. The other short-run dynamics term in the first equation, $\triangle^2 p_{a,t-2} 1_{(t\geq 1992.31)}$, is a lagged value at $t-2$ and has a negative coefficient, in contrast to the term $\triangle^2 p_{a,t-1} 1_{(t<1990.31)}$ with a positive coefficient. The changes in sign and lag also suggest that a fairly large shift seems to have taken place in the short-run dynamics around the middle of the sample period.

7 Simulation Experiments

The analysis of the model with changing short-run dynamics in Section 6.3 suggested that a reduction to a 4-lag model with constant short-run dynamics as used in Section 6.2 is inappropriate. Yet, in the analysis of the model with constant short-run analysis it was found that a 4-lag model could be reduced further to a 2-lag model. However, the analysis of Section 6.1 tentatively suggested that residual autocorrelation could be found in sub-samples. These findings are supported here by a small-scale simulation experiment of recursive tests for residual autocorrelation.

The design for an artificial data generation process is based on the empirical analysis of the US gasoline price data in Section 6. The data generating process for $X_t = (X_{1,t}, X_{2,t})'$ is given by the following set of equations:

$$
\begin{aligned}
\triangle^2 X_{1,t} &= 0.7 \triangle^2 X_{2,t} - 0.5 \triangle X_{1,t-1} + 0.5 \triangle X_{2,t-1} - 0.1(X_{1,t-1} - X_{2,t-1} - 0.02) \\
&\quad + 0.2 \triangle^2 X_{1,t-1} 1_{(t<1990.31)} - 0.1 \triangle^2 X_{1,t-2} 1_{(t\geq 1992.31)} + \varepsilon_{1,t}, \\
\triangle^2 X_{2,t} &= -0.5 \triangle X_{2,t-1} + 0.3 \triangle^2 X_{2,t-1} 1_{(t<1990.31)} + \varepsilon_{2,t}, \\
(\varepsilon_{1,t}, \varepsilon_{2,t})' &\sim N(0, 0.02^2 \times I), \qquad t = 6, \cdots, T,
\end{aligned}
$$

where the innovations $(\varepsilon_{1,t}, \varepsilon_{2,t})'$ are replaced by pseudo-random independent drawings. The parameters for the artificial processes for $X_{1,t}$ and $X_{2,t}$ are round-off equivalents to the coefficients of the equations for $p_{a,t}$ and $p_{b,t}$ in model (28). A change in a constant term is omitted from the data generation process, as the main interest lies in shifts in the short-term dynamics. The following five initial values for both $X_{1,t}$ and $X_{2,t}$ are taken from the data set of the gasoline prices:

$$\begin{pmatrix} X_{1,1}, \cdots, X_{1,5} \\ X_{2,1}, \cdots, X_{2,5} \end{pmatrix} = \begin{pmatrix} -0.549 & -0.530 & -0.522 & -0.508 & -0.475 \\ -0.527 & -0.515 & -0.505 & -0.491 & -0.466 \end{pmatrix}.$$

The number of observations and the locations of the parameter shifts are the same as those in the empirical study. That is, the effective simulation sample could be treated as if it spanned from 1987.29 to 1998.29 and parameter changes took place in 1990.31 and 1992.31. This gives breaks at time 164 and 267 in the full sample length 578.

Figure 2

Recursive Rejection Frequencies of Residual Autocorrelation Tests

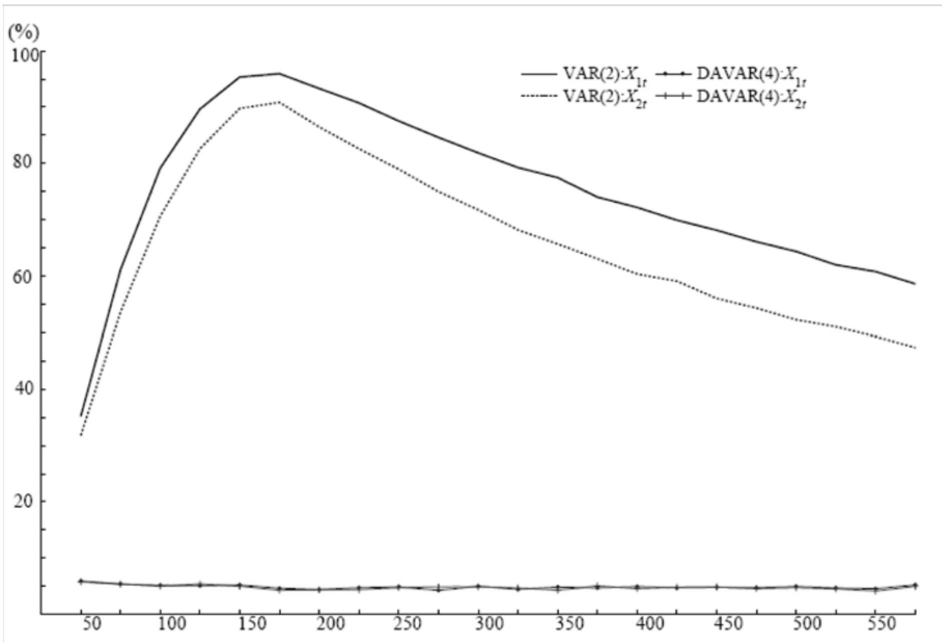


Figure 2 shows recursive graphics of the rejection frequencies of 1st-order serial autocorrelation tests at the 5% level for two models: a well-specified 4-lag DAVAR with changing short-term dynamics, denoted DAVAR(4), and a mis-specified 2-lag constant parameter vector autoregression, denoted VAR(2). The sample size for the recursive calculation starts at $T = 50$, and increases by 25 observations ($T = 75, 100, 125, \cdots$) until it reaches the final point $T = 578$. The number of replications at each sample size is 10,000. In Figure 2 the rejection frequencies for both $X_{1,T}$ and $X_{2,T}$ in DAVAR(4) uniformly correspond to the chosen 5% level. However, the rejection rates with respect to VAR(2) are much larger than 5%. The rates for both $X_{1,T}$ and $X_{2,T}$ in VAR(2) reach their peaks (95% and 90% respectively) at the location which is the closest to the first parameter break point ($T = 164$), but then reduces towards the nominal level. Therefore, in the case of misleading model reduction being performed, testing residual autocorrelation in several sub-samples could be useful in identifying potential shifts in the short-run parameters, as used tentatively in Section 6.1.

Clements and Hendry (1999), and Hendry (2000) demonstrate that parameter shifts are difficult to detect by the break-point Chow test for parameter constancy, except for changes in equilibrium-mean. Thus, it is not always straightforward to identify where the short-run parameters change. However, if misleading model reduction is performed, the effect of the omitted parameter shifts is then reflected in the residuals of the reduced model, inducing temporal correlation in the residuals. Thus recursive residual autocorrelation tests seem to be informative in identifying the change points.

Finally, we investigate size properties of the rank test statistics estimated from the VAR(2) and DAVAR(4) models. As indicated in the empirical illustration, the rank test seems to be fairly robust to the omission of the short-run parameter changes. However, it is expected from the existing simulation studies such as Gonzalo (1994) that the mis-specified VAR(2) model shows larger size distortions than the well-specified DAVAR(4) model. Using the same setting as the simulation experiment given above, we recursively calculate the rank test statistics for the null of a single cointegrating relation and their rejection frequencies at the 5% level. Figure 3 presents the rejection frequencies with the corresponding confidence intervals. The figure also displays the confidence interval for the 5% level. As expected, size distortions are rather obvious in VAR(2) as compared with DAVAR(4). Furthermore, the estimated quantiles of the rank test statistics using the full sample ($T = 578$) are plotted against the corresponding simulated asymptotic quantiles in Figure 4. Using the methodology in Johansen (1996, ch. 15), we tabulate the asymptotic quantiles using 1,000 observations with 10,000 replications. The horizontal axis corresponds to the asymptotic quantiles, while the vertical axis to the estimated quantiles. The baseline asymptotic quantile-quantile plots are represented by the straight line of 45 degree, while the quantile-quantile plots for VAR(2) and DAVAR(4) are given by the dotted thick and thin lines, respectively. Size distortions are again conspicuous in VAR(2), even though we use a fairly large number of observations. Size control is usually treated as a fundamental requirement in conventional statistical inferences. The simulation study thus indicates the potential importance of correctly specifying the underlying dynamic structure of the data in conducting inferences for the cointegrating rank.

Figure 3

Recursive Rejection Frequencies of Cointegrating Rank Tests

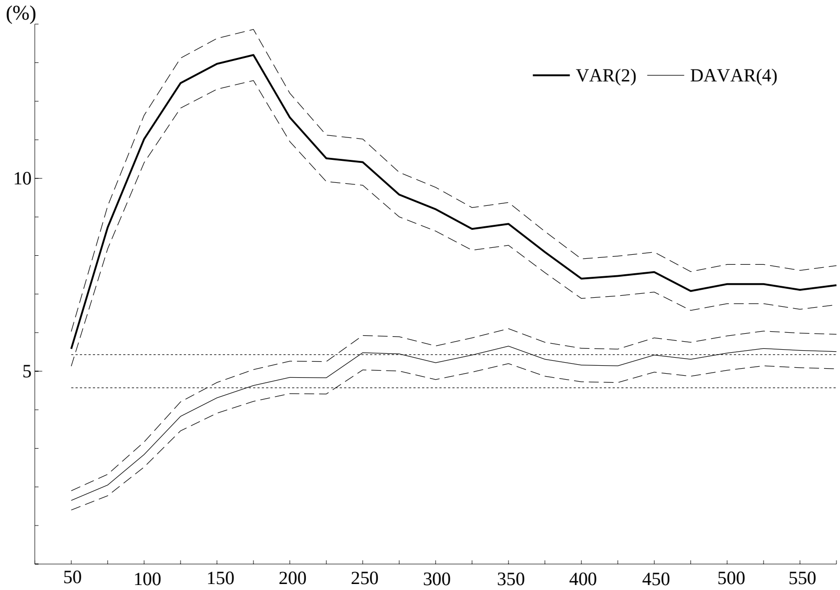
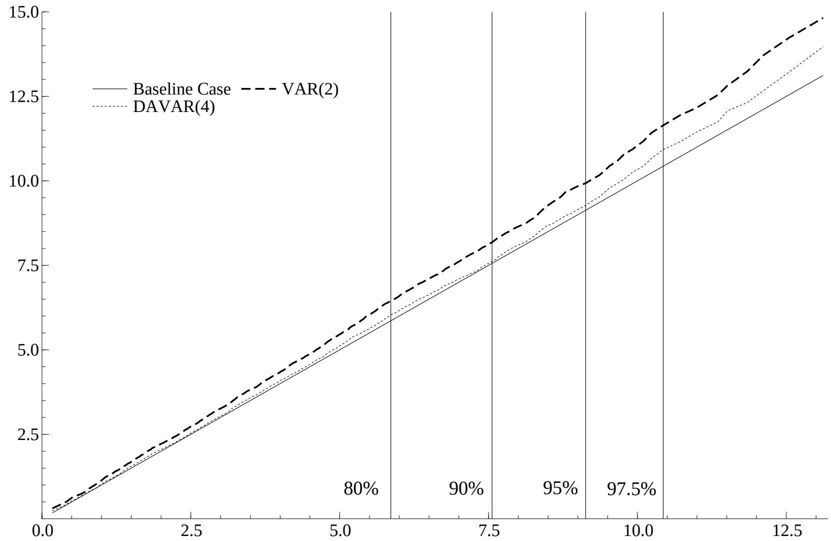


Figure 4

Quantile-Quantile Plots for Cointegrating Rank Tests



8 Summary and Conclusion

This paper introduced a family of cointegrated vector autoregressive models with adjusted short-run dynamics. The family comprises three models: a model subject to short-run parameter changes, a model with partial short-run dynamics and a model with short-run explanatory variables. The statistical analysis of these three models was presented, and it was demonstrated that the likelihood ratio test statistics for cointegration rank are based on reduced rank regression, as in the ordinary cointegrated vector autoregressive model. The asymptotic analysis of the three models was then presented, and it was proved that rank test statistics have the conventional limiting distribution. Thus, ordinary asymptotic quantiles for cointegration rank can be applied to all the three dynamics-adjusted models. The family of DAVAR models can give a congruent representation of some economic data, as shown in the empirical illustration using US gasoline prices. The simulation experiments shed some light on the issue of identifying a location where short-run parameters shift and on size distortions stemming from model mis-specifications. The issue of identifying parameter-changing points needs to be further addressed.

Appendix

Proof of Theorem 1

Let us consider first the homogenous case where μ and γ are both zero. For each period the usual Granger-Johansen representation theorem can be derived from Johansen (1996, Theorem 4.2). Thus, we define

$$Z_t = (X_t'\beta, \triangle X_t', ..., \triangle X_{t-k+1}')',$$

which implies that the processes $\left(Z_{T_{j-1}+1}, ..., Z_{T_{j-1}+T_j}\right)$ for $j = 1, 2$ can be given mean-zero stationary initial distributions under Assumptions 1 and 2.

It is therefore left to check that the individual representation for each period can be combined as stated. The proof of Nielsen (2001) is generalised. In the model (3) specified by equation (4) with the assumption of a homogenous equation, we replace $\triangle X_{t-1}$ by $\triangle X_t - \triangle^2 X_t$, rearrange it, multiply both sides by $\alpha_\perp'$ and then arrive at the following equation:

$$\alpha_{\perp}'\Gamma \triangle X_t = \alpha_{\perp}' \left\{ \varepsilon_t + (\Gamma - I) \triangle^2 X_t + \sum_{i=1}^{k-2} \Psi_i^{(j)} \triangle^2 X_{t-i} \right\}.$$

Summing up $\alpha_{\perp}'\Gamma \triangle X_s$ over $s = 1, ..., t$ yields

$$\alpha_{\perp}'\Gamma (X_t - X_0) = \alpha_{\perp}' \left\{ \sum_{s=1}^{t} \varepsilon_s + (\Gamma - I)(\triangle X_t - \triangle X_0) + \sum_{i=1}^{k-2} \Psi_i^{(1)} (\triangle X_{t-i} - \triangle X_{-i}) \right\},$$
$$\text{for } 0 < t \leq T_1,$$
$$\alpha_{\perp}'\Gamma (X_t - X_0) = \alpha_{\perp}' \left\{ \sum_{s=1}^{t} \varepsilon_s + (\Gamma - I)(\triangle X_t - \triangle X_0) + \sum_{i=1}^{k-2} \Psi_i^{(1)} (\triangle X_{T_1-i} - \triangle X_{-i}) \right.$$
$$\left. + \sum_{i=1}^{k-2} \Psi_i^{(2)} (\triangle X_{t-i} - \triangle X_{T_1-i}) \right\}, \quad \text{for } T_1 < t \leq T. \qquad (29)$$

The term $\alpha_{\perp}'\Gamma X_0$ is moved to the right-hand side. Regarding $\alpha_{\perp}'\Gamma X_t$, $\alpha_{\perp}'\Gamma$ is post-multiplied by the orthogonal projection identity $\beta_{\perp}\overline{\beta}_{\perp}' + \overline{\beta}\beta' = I$ such that we can obtain $\alpha_{\perp}'\Gamma \left(\beta_{\perp}\overline{\beta}_{\perp}' + \overline{\beta}\beta'\right) X_t$. Pre-multiplying equation (29) by $\beta_{\perp}(\alpha_{\perp}'\Gamma\beta_{\perp})^{-1}$ noting that $\alpha_{\perp}'\Gamma\beta_{\perp}$ is invertible by Assumption 1, and then re-arranging (29), we reach the following general formulation covering both of the above two cases:

$$\beta_{\perp}\overline{\beta}_{\perp}' X_t = C \left[\sum_{s=1}^{t} \varepsilon_s - \Gamma\overline{\beta}\beta' X_t + (\Gamma - I) \triangle X_t + \sum_{i=1}^{k-2} \Psi_i^{(j)} \triangle X_{t-i} \right.$$
$$\left. + \left\{ \Gamma X_0 - (\Gamma - I) \triangle X_0 - \sum_{i=1}^{k-2} \Psi_i^{(1)} \triangle X_{-i} + \sum_{i=1}^{k-2} \left(\Psi_i^{(1)} - \Psi_i^{(2)} \right) \triangle X_{T_1-i} 1_{(t>T_1)} \right\} \right].$$

Adding $\overline{\beta}\beta' X_t$ on both sides and noting the orthogonal projection identity, we arrive at the desired Granger-Johansen representation with $\tau_c = \tau_l = 0$ and

$$y_t^{(j)} = (I - C\Gamma)\overline{\beta}\beta' X_t + C(\Gamma - I) \triangle X_t + \sum_{i=1}^{k-2} C\Psi_i^{(j)} \triangle X_{t-i},$$
$$A^{(1)} = C \left\{ \Gamma X_0 - (\Gamma - I) \triangle X_0 - \sum_{i=1}^{k-2} \Psi_i^{(1)} \triangle X_{-i} \right\},$$
$$A^{(2)} = C \sum_{i=1}^{k-2} \left(\Psi_i^{(1)} - \Psi_i^{(2)} \right) \triangle X_{T_1-i}.$$

Note that $Y_t^{(j)}$ is a function of Z_t, showing that for each j the process $(y_{T_{j-1}+1}^{(j)}, ..., y_{T_{j-1}+T_j}^{(j)})$ can be given mean-zero, stationary initial distributions.

Consider next the non-homogenous case where μ and γ can be different from zero, and replace X_t by $\widetilde{X}_t + \tau_c + \tau_l t$ in the model (3) specified by equation (4). It is seen that if

$$\alpha\beta'(\tau_c - \tau_l) - \Gamma\tau_l + \mu = 0 \quad \text{and} \quad \beta'\tau_l + \gamma' = 0, \tag{30}$$

then a homogenous equation for $\widetilde{X}_t$ arises and the result derived above can be used for $\widetilde{X}_t$. The equations in (30) do not depend on the period j and therefore have the solutions (9) and (10) as found in Johansen, Mosconi and Nielsen (2000). ■

Proof of Theorems 2 and 3

The same proof as that of Theorem 1 given above can be applied to Theorems 2 and 3 without introducing the regime change index j, as the model (3) specified by either equation (5) or (7) is free from parameter shifts. ■

Proof of Lemma 5

To give a proof of Lemma 5, we need to use sub-sample reduced rank regression as a theoretical device. The regressor Z_{3t} introduced in Section 3.2 is replaced by

$$Z_{3t} = \left(Z_{3t}^{(1)\prime},\ Z_{3t}^{(2)\prime}\right)',$$

where

$$Z_{3t}^{(j)} = \left(\triangle^2 X'_{t-1}, \ldots, \triangle^2 X'_{t-k+2}\right)' 1_{(T_{j-1}<t\leq T_j)}, \quad \text{for} \quad j = 1, 2.$$

For each sub-sample period, Z_{0t}, Z_{1t} and Z_{2t} are regressed on $Z_{3t}^{(j)}$ to obtain sub-sample residuals $R_{0.3,t}^{(j)}$, $R_{1.3,t}^{(j)}$ and $R_{2.3,t}^{(j)}$. Secondly, for $h = 0, 1, 2$, concatenating the residuals $R_{h.3,t}^{(1)}$ and $R_{h.3,t}^{(2)}$ leads to a connected series $R_{h.3,t}$. Thirdly, $R_{0.3,t}$ and $R_{1.3,t}$ are regressed on $R_{2.3,t}$ to generate residuals R_{0t} and R_{1t}, which are used to calculate the sample product moment matrix defined by equation (15).

For each sub-sample, the residuals satisfy the following equation:

$$R_{0.3,t}^{(j)} = \alpha\beta^{*\prime} R_{1.3,t}^{(j)} + \Gamma R_{2.3,t}^{(j)} + \widehat{\varepsilon}_t. \tag{31}$$

Under the satisfaction of Theorem 1, the initial values can be given stationary distributions. Thus, each term in this equation is a stationary and ergodic process, leading to the following result by the law of large numbers:

$$\frac{1}{T_j}\sum_{t=T_{j-1}+1}^{T_{j-1}+T_j}\begin{pmatrix} R_{0.3,t}^{(j)} \\ \beta^{*\prime}R_{1.3,t}^{(j)} \\ R_{2.3,t}^{(j)} \end{pmatrix}\begin{pmatrix} R_{0.3,t}^{(j)} \\ \beta^{*\prime}R_{1.3,t}^{(j)} \\ R_{2.3,t}^{(j)} \end{pmatrix}' =$$

$$\begin{pmatrix} S_{00.3}^{(j)} & S_{0\beta.3}^{(j)} & S_{02.3}^{(j)} \\ S_{\beta 0.3}^{(j)} & S_{\beta\beta.3}^{(j)} & S_{\beta 2.3}^{(j)} \\ S_{20.3}^{(j)} & S_{2\beta.3}^{(j)} & S_{22.3}^{(j)} \end{pmatrix} \xrightarrow{p} \begin{pmatrix} \Sigma_{00.3}^{(j)} & \Sigma_{0\beta.3}^{(j)} & \Sigma_{02.3}^{(j)} \\ \Sigma_{\beta 0.3}^{(j)} & \Sigma_{\beta\beta.3}^{(j)} & \Sigma_{\beta 2.3}^{(j)} \\ \Sigma_{20.3}^{(j)} & \Sigma_{2\beta.3}^{(j)} & \Sigma_{22.3}^{(j)} \end{pmatrix}.$$

Next, we consider the asymptotic properties of the product moment matrices over the whole sample. Define the moment matrix $S_{00.3}$ as

$$S_{00.3} = \frac{1}{T}\sum_{j=1}^{2}\sum_{t=T_{j-1}+1}^{T_{j-1}+T_j}\left(R_{0.3,t}^{(j)}\right)\left(R_{0.3,t}^{(j)}\right)'.$$

Slutsky's theorem yields

$$S_{00.3} = \sum_{j=1}^{2}\frac{T_j}{T}\frac{1}{T_j}\sum_{t=T_{j-1}+1}^{T_{j-1}+T_j}\left(R_{0.3,t}^{(j)}\right)\left(R_{0.3,t}^{(j)}\right)' \xrightarrow{p} \sum_{j=1}^{2} a^{(j)}\Sigma_{00.3}^{(j)} \stackrel{def}{=} \Sigma_{00.3},$$

where the final equality follows equation (19). The same argument is applied to the remaining moment matrices, deriving the following asymptotic result:

$$\begin{pmatrix} S_{00.3} & S_{01.3}\beta^* & S_{02.3} \\ \beta^{*\prime}S_{10.3} & \beta^{*\prime}S_{11.3}\beta & \beta^{*\prime}S_{12.3} \\ S_{20.3} & S_{21.3}\beta^* & S_{22.3} \end{pmatrix} \xrightarrow{p} \begin{pmatrix} \Sigma_{00.3} & \Sigma_{0\beta.3} & \Sigma_{02.3} \\ \Sigma_{\beta 0.3} & \Sigma_{\beta\beta.3} & \Sigma_{\beta 2.3} \\ \Sigma_{20.3} & \Sigma_{2\beta.3} & \Sigma_{22.3} \end{pmatrix}. \quad (32)$$

The final formulation of the sample product moment matrices is given by

$$\begin{pmatrix} S_{00} & S_{01}\beta^* \\ \beta^{*\prime}S_{10} & \beta^{*\prime}S_{11}\beta^* \end{pmatrix} = \begin{pmatrix} S_{00.3} & S_{01.3}\beta^* \\ \beta^{*\prime}S_{10.3} & \beta^{*\prime}S_{11.3}\beta^* \end{pmatrix} - \begin{pmatrix} S_{02.3} \\ \beta^{*\prime}S_{12.3} \end{pmatrix} S_{22.3}^{-1}\begin{pmatrix} S_{20.3} & S_{21.3}\beta^* \end{pmatrix}. \quad (33)$$

Applying equation (32) to (33) and using the definition given by equation (20) proves equation (21). In order to prove (22), we introduce the Yule-Walker equations for each period corresponding to the residual equation (31):

$$\Sigma_{00.3}^{(j)} = \alpha\Sigma_{\beta 0.3}^{(j)} + \Psi\Sigma_{20.3}^{(j)} + \Omega, \quad (34)$$

$$\Sigma_{0\beta.3}^{(j)} = \alpha\Sigma_{\beta\beta.3}^{(j)} + \Psi\Sigma_{2\beta.3}^{(j)}, \quad (35)$$

$$\Sigma_{02.3}^{(j)} = \alpha\Sigma_{\beta 2.3}^{(j)} + \Psi\Sigma_{22.3}^{(j)}. \quad (36)$$

Inserting equations (35) and (36) into (34) gives us

$$\Sigma_{00.3}^{(j)} = \alpha\Sigma_{\beta\beta.3}^{(j)}\alpha' + \alpha\Sigma_{\beta 2.3}^{(j)}\Psi' + \Psi\Sigma_{2\beta.3}^{(j)}\alpha' + \Psi\Sigma_{22.3}^{(j)}\Psi' + \Omega. \quad (37)$$

Then substituting equations (35)-(37) into equation (19) yields

$$\begin{aligned}\Sigma_{0\beta.3} &= \alpha\Sigma_{\beta\beta.3} + \Psi\Sigma_{2\beta.3},\\ \Sigma_{02.3} &= \alpha\Sigma_{\beta 2.3} + \Psi\Sigma_{22.3},\\ \Sigma_{00.3} &= \alpha\Sigma_{\beta\beta.3}\alpha' + \alpha\Sigma_{\beta 2.3}\Psi' + \Psi\Sigma_{2\beta.3}\alpha' + \Psi\Sigma_{22.3}\Psi' + \Omega.\end{aligned}$$

Combining these with equation (20) proves equation (22). ■

References

[1] Andrade, P., Bruneau C., and S. Gregoir (2005). Testing for the Cointegration Rank when some Cointegrating Directions are Changing. Journal of Econometrics, 124, 269-310.

[2] Campos, J., Ericsson N. R., and D. F. Hendry (1996). Cointegration Tests in the Presence of Structural Breaks. Journal of Econometrics, 70, 187-220.

[3] Cavaliere, G., and R. Taylor (2006). Testing the Null of Cointegration in the Presence of Variance Breaks. Journal of Time Series Analysis, 27, 613-636.

[4] Cheung, Y. W., and K. S. Lai (1993). Finite-sample Sizes of Johansen's Likelihood Ratio Tests for Cointegration. Oxford Bulletin of Economics and Statistics, 55, 313-328.

[5] Clements, M. P., and D. F. Hendry (1999). Forecasting Non-stationary Economic Time Series. MIT press, Cambridge, Massachusetts.

[6] Doornik, J. A. (1998). Approximations to the Asymptotic Distribution of Cointegration Tests. Journal of Economic Surveys, 12, 573-593.

[7] Doornik, J. A. (2001). Ox: An Object-Oriented Matrix Language. Timberlake Consultants Press, London.

[8] Doornik, J. A., and H. Hansen (1994)[2008]. An Omnibus Test for Univariate and Multivariate Normality. Economics Group WP W4&91, Nuffield College, University of Oxford. (Also published in Oxford Bulletin of Economics and Statistics, 70, 927-939.)

[9] Doornik, J. A., and D. F. Hendry (2001). Modelling Dynamic Systems Using PcGive, Volume 2. Timberlake Consultants Press, London.

[10] Doornik, J. A., and R. J. O'Brien (2002). Numerically Stable Cointegration Analysis. Computational Statistics and Data Analysis, 41, 185-193.

[11] Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50, 987-1007.

[12] Engle, R. F., Hendry D. F., and J. F. Richard (1983). Exogeneity. Econometrica, 51, 277-304.

[13] Galí, J., and M. Gertler (1999). Inflation Dynamics: A Structural Econometric Analysis. Journal of Monetary Economics, 44, 195-222.

[14] Godfrey, L. G. (1978). Testing for Higher Order Serial Correlation in Regression Equations when the Regressors Include Lagged Dependent Variables. Econometrica, 46, 1303-1313.

[15] Gonzalo, J. (1994). Five Alternative Methods of Estimating Long Run Equilibrium Relationships. Journal of Econometrics, 60, 1-31.

[16] Hansen, B. E. (1992a). Convergence to Stochastic Integrals for Dependent Heterogenous Processes. Econometric Theory, 8, 489-500.

[17] Hansen, B. E. (1992b). Tests for Parameter Instability in Regression with I(1) processes. Journal of Business and Economic Statistics, 10, 321-335.

[18] Hansen, H., and S. Johansen (1999). Some Tests for Parameter Constancy in Cointegrated VAR-models. Econometrics Journal, 2, 306-333.

[19] Hansen, P. R. (2000). Structural Changes in Cointegrated Processes. Ph.D. thesis, University of California at San Diego.

[20] Hansen, P. R. (2003). Structural Changes in the Cointegrated Vector Autoregressive Model. Journal of Econometrics, 114, 261-295.

[21] Hansen, P. R. (2005). Granger's Representation Theorem: A Closed-form Expression for the I(1) Process. Econometrics Journal, 8, 23-38.

[22] Hendry, D. F. (2000). On Detectable and Non-detectable Structural Change. Structural Change and Economic Dynamics, 11, 45-65.

[23] Hendry, D. F., and K. Juselius (2001). Explaining Cointegration Analysis: Part II. Energy Journal, 22, 75-120.

[24] Hendry, D. F., and H. M. Krolzig (2001). Automatic Econometric Model Selection Using PcGets. Timberlake Consultants Press, London.

[25] Hunter, J., and S. Burke (2007). Common Trends, Cointegration and Competitive Price Behaviour. Presented at the ESRC Econometrics Study Group Annual Conference 2007.

[26] Inoue, A. (1999). Tests of Cointegration Rank with a Trend Break. Journal of Econometrics, 90, 215-237.

[27] Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. Journal of Economic Dynamics and Control, 12, 231-254.

[28] Johansen, S. (1992a). A Representation of Vector Autoregressive Processes Integrated of Order 2. Econometric Theory, 8, 188-202.

[29] Johansen, S. (1992b). Cointegration in Partial Systems and the Efficiency of Single-equation Analysis. Journal of Econometrics, 52, 389-402.

[30] Johansen, S. (1996). Likelihood-Based Inference in a Vector Autoregressive Model. Oxford University Press, 2nd printing.

[31] Johansen, S., Mosconi R., and B. Nielsen (2000). Cointegration Analysis in the Presence of Structural Breaks in the Deterministic Trend. Econometrics Journal, 3, 216-249.

[32] Juselius, K. (2006). The Cointegrated VAR Model. Oxford University Press, Oxford.

[33] Nielsen, B. (2001). The Asymptotic Distribution of Unit Root Tests of Unstable Autoregressive Processes. Econometrica, 69, 211-219.

[34] Nielsen, B. (2007). Order Determination in General Vector Autoregressions, in Time Series and Related Topics: In Memory of Ching-Zong Wei; H.-C Ho, C.-K. Ing and T. L Lai (Eds). IMS Lecture Notes and Monograph Series, 52, 93-112.

[35] Phillips, P. C. B., and V. Solo (1992). Asymptotics for Linear Processes. Annals of Statistics, 20, 971-1001.

[36] Quintos, C. E, and P.C.B. Phillips (1993). Parameter Constancy in Cointegrating Regressions. Empirical Economics, 18, 675-706.

[37] Rahbek, A., and R. Mosconi (1999). The Role of Stationary Regressors in the Cointegration Test. Econometrics Journal, 2, 82-97.

[38] Rahbek, A., Hansen E., and J. G. Dennis (2002). ARCH Innovations and Their Impact on Cointegration Rank Testing. Centre for Analytical Finance WP 22, University of Copenhagen.

[39] Seo, B. (1998). Tests for Structural Change in Cointegrated Systems. Econometric Theory, 14, 222-259.