Discussion Papers
Department of Economics
University of Copenhagen

No. 21-07

Adjustment coefficients and exact rational expectations in cointegrated vector autoregressive models

by

Søren Johansen and Anders Rygh Swensen

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark
Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 1601-2461 (E)

Manuscript submitted to *The Econometrics Journal*, pp. 1–21.

Adjustment coefficients and exact rational expectations in cointegrated vector autoregressive models

Søren Johansen[†] and Anders RyghN Swensen[‡]

[†] *Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26 DK-1353 Copenhagen.*
E-mail: `soren.johansen@econ.ku.dk`
[‡] *Department of Mathematics, University of Oslo, P.O. Box 1053 NO-0316 Blindern Oslo, Norway.*
E-mail: `swensen@math.uio.no`

Summary
In cointegrated vector autoregressive models exact linear rational expectation relations can imply restrictions on the adjustment parameters. We show how such restrictions can be tested, in particular when the restrictions imply weak exogeneity of some variables.

Keywords: *Abstract,Exact rational expectations; Cointegrated VAR model; Reduced rank regression; Adjustment coefficients*

JEL Classification: *C32*

1. INTRODUCTION

In cointegrated vector autoregressive models CVAR, or reduced rank VAR models, the parameters of the cointegration vectors and the adjustment parameters have an important role. The adjustment coefficients describe the speed of convergence toward the equilibrium defined by the cointegration vectors, and therefore are essential in specifying the error correcting properties of these models.

The adjustment coefficients are also crucial for investigating whether some variables can be classified as exogenous and therefore the possibility of a model reduction. Exogeneity can be defined in several ways, but the classification weak, strong and super exogenous introduced by Engle, Hendry and Richard (1983) has been influential. Strong exogeneity is important for prediction and forecasting and super exogeneity can be given a causal interpretation postulating invariance with respect to certain policy interventions. Both strong and super exogeneity presuppose weak exogeneity which means that the joint distribution factorizes in a particular way. The joint distribution can always be written as a product of a conditional and a marginal distribution. In the presence of weak exogeneity the parameters of the joint distribution can in addition be divided in two parts, one describing the marginal distribution and another related to the marginal distribution, without any constraints between the two parts. The parameters of the conditional and marginal parts are therefore freely varying.

In the CVAR model weak exogeneity is equivalent to one or several rows in the matrix of adjustment coefficients, α, being zero. Test for this hypothesis has been developed. If the adjustment parameters corresponding to particular variables vanish, the error correcting effect of the cointegration vector will not be present for these particular variables.

Expectations are fundamental in economics. For interpretation a crucial question is how they are defined and implemented. There have been several proposals for how this can be done. In the approach introduced by Muth (1961), which is usually called rational expectations, expectations are essentially the same as predictions of the relevant economic theory. In the setup of this paper this is summarized in the CVAR and the predictions are the conditional expectations given the available observations at the time.

Present value models are a place where expectations occur. In an example from Campbell and Shiller (1987) the value of stocks at the beginning of period t, Y_t is expressed as a discounted sum of expected present and future dividends, $Y_t = \delta \sum_{j=0}^{\infty} \delta^j E_t[y_{t+j}]$, where y_t is the dividend from the period $t-1$. Present value models can often be expressed as a linear combination of observed values and conditional expectations of others. In the example an implication is the relation $S_t = [\delta/(1-\delta)]E_t[\Delta Y_{t+1}]$ where the spread S_t equals $Y_t - [\delta/(1-\delta)]y_t$.

As we can see the variables are not treated symmetrically. In Hansen and Sargent (1982) it is stressed that in models involving declining expected future forcing variables these are typically described as being beyond the economic agents control. An investigation whether they can be considered as exogenous in some sense is therefore natural and weak exogeneity is the reasonable place to start.

In this paper we will consider how one can test simultaneously for linear rational expectations relations of the type described and weak exogenity, or more generally that the error correction parameters satisfy the same restrictions, i.e. $\alpha = A\psi$ where A is a known matrix and ψ is unknown. Also the case where some of the error correction parameters are known in addition to the restrictions from the exact rational expectation hypothesis, will be considered. In both cases we discuss a convenient parameterization by means of variation of variation free parameters. In the first case this leads to a general procedure for maximum likelihood estimation and testing. The second case is more complicated and here we suggest a switching procedure for a particular case.

The paper is a follow up of three related papers dealing with various aspects of the theme: Johansen and Swensen (1999) where the CVAR contained an unrestricted constant, Johansen and Swensen (2004) where the constant and trend could be restricted and in particular Johansen and Swensen (2008) where simultaneous tests for restrictions on the parameters in the cointegration vectors and rational expectations were considered. We point out that the relations involving the conditional expectations we consider are exact as defined by Hansen and Sargent (1981) and (1991), i.e. do not contain additional stochastic terms. To see the problems occurring for non-exact specifications one can consult Boug *et al.* (2010) or Swensen (2014).

The organization of the paper is as follows. In Section 2 the rational expectation models are explained. In Section 3 we consider the case where the columns of the matrix of adjustment coefficients belong to a subspace. In Section 4 the case where parts of the adjustment parameters are known is treated and Section 5 contains an application.

2. THE RESTRICTIONS IMPLIED BY EXACT RATIONAL EXPECTATIONS

This section defines the cointegrated vector autoregressive model as the statistical model which is assumed to generate the data and formulates the parameter restrictions implied by the exact rational expectation hypothesis.

2.1. The cointegrated vector autoregressive model

Let the p-dimensional vectors of observations be generated according to the vector autoregressive model

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k} \Gamma_i \Delta X_{t-i} + \mu + \varepsilon_t,\ t = 1, \ldots, T \tag{2.1}$$

where $X_{-k+1}, \ldots, X_0$ are fixed and $\varepsilon_1, \ldots, \varepsilon_T$ are independent, identically distributed Gaussian vectors, with mean zero and covariance matrix Ω. We assume that $\{X_t\}_{t=1,2,\ldots}$ is $I(1)$ and that $\Pi = \alpha\beta'$ where the $p \times r$ matrices α and β have full column rank r. This implies that X_t is non-stationary, ΔX_t is stationary, and that $\beta' X_t$ is stationary. It is the stationary relations between non-stationary processes and the interpretation as long-run relations, that has created the interest in this type of model in economics. Also note that the columns of α and $\Pi = \alpha\beta'$ span the same space. As in Johansen (1996) we define the models.

$\mathcal{H}(r)$: The model is defined by equation (2.1), where α and β are $p \times r$ matrices and otherwise no further restrictions on the parameters. The number of identified parameters in the matrix $\alpha\beta'$ is $\#(\alpha\beta') = pr + r(p - r)$.

In the following we also assume that α is restricted either by homogeneity restrictions of the form $\alpha = A\psi$, $(sp(\alpha) \subset sp(A))$ or that some alpha vectors are known, $\alpha = (a, a_\perp \psi)$ $(sp(a) \subset sp(\alpha))$. This defines two sub-models of $\mathcal{H}(r)$.

$\mathcal{H}_1(r)$: The model is defined by equation (2.1) and the restriction $\alpha = A\psi$, where A is a known $p \times s$ matrix of rank s, and ψ is an $s \times r$ matrix of parameters, $r \leq s \leq p$. In this case the number of parameters is $\#(\alpha\beta') = sr + r(p - r)$.

$\mathcal{H}_2(r)$: The model is defined by equation (2.1) and the restriction $\alpha = (a, a_\perp \phi)$ where a is a known $p \times m$ matrix of rank $m > 0$, and ϕ is a $(p - m) \times (r - m)$ matrix of parameters, $m \leq r \leq p$ such that $\alpha\beta' = a\beta_1' + a_\perp \phi \beta_2'$. In this model $\#(\alpha\beta') = mp + (r - m)(2p - r)$.

2.2. Estimation of the cointegrated vector autoregressive models, $\mathcal{H}(r)$, $\mathcal{H}_1(r)$, and $\mathcal{H}_2(r)$

It is well known, see Johansen (1996), that the Gaussian maximum likelihood estimator of β is calculated by reduced rank regression, as in Anderson (1951), of ΔX_t on X_{t-1} corrected for the stationary regressors

$$\Delta X_{t-1}, \ldots, \Delta X_{t-k+1}.$$

Once β is determined, the other parameters are estimated by regression.

Model $\mathcal{H}_1(r)$ is estimated using the likelihood function obtained by defining two regression equations and using a conditionality argument. In the following we will use the usual notation that if a is an $n \times m$, $0 < m < n$ matrix of full rank, then $\bar{a} = a(a'a)^{-1}$ and satisfies $a'\bar{a} = I_m$, and $a_\perp$ is an $n \times (n - m)$ matrix such that $a_\perp' a = 0$ and the $n \times n$ matrix $(a, a_\perp)$ is nonsingular, finally $I_n = \bar{a}a' + \bar{a}_\perp a_\perp'$.

The model $\mathcal{H}_2(r)$ is estimated by first finding the parameter β_1 from the conditional equation for $\bar{a}'\Delta X_t$ given $\bar{a}'_\perp\Delta X_t$ and the past and then finding the parameters ϕ and β_2 by reduced rank regression in the marginal model for $\bar{a}'_\perp\Delta X_t$.

2.3. The model for exact rational expectations and some examples

The model formulates a set of restrictions on the conditional expectation of X_{t+1} given the information $\mathcal{O}_t$ in the variables up to time t, which we write in the form

$\mathcal{RE}$: The model based exact rational expectations formulates relations for conditional expectations

$$E[c'\Delta X_{t+1}|\mathcal{O}_t] = \tau d' X_t + \sum_{i=1}^{\ell} \tau_i d_i' \Delta X_{t+1-i} + d_\mu. \tag{2.2}$$

Here $E_t = E[\cdot|\mathcal{O}_t]$ denotes the conditional expectation in the probabilistic sense of model (2.1), given the variables $X_1, \ldots, X_t$. The matrices c of dimensions $p \times q$, d of dimensions $p \times n$ and d_i of dimensions $p \times n_i$, $i = 1, \ldots, \ell$ are known full rank matrices and $\tau(q \times n)$, $\tau_i(q \times n_i), i = 1, \ldots, \ell$ are parameters. The elements of the $q \times 1$ vector d_μ are either known or parameters. We assume that $n \leq q$ and $\ell \leq k$.

We give next two examples of a rational expectations model and discuss the relation to models $\mathcal{H}_1(r)$ and $\mathcal{H}_2(r)$.

Example 2.1. The variables real consumption, CP_t, real labour income, YL_t, and real capital income YK_t, are fundamental in models for aggregate consumption, both for those in the Keynesian tradition and for versions building on a permanent income hypothesis. Campbell (1987) studied a permanent income hypothesis for consumption of the form:

$$CP_t = \gamma(YK_t + \frac{r}{1+r}\sum_{i=0}^{\infty}(\frac{r}{1+r})^i E_t[YL_{t+i}])$$

where r is the expected real interest rate and $\gamma \leq 1$ is a proportionality factor. Current consumption is therefore a fraction of present and estimated future labour income and present capital income.
Savings is defined as $S_t = (YL_t + YK_t) - CP_t/\gamma$. Campbell showed that

$$S_t - \Delta YL_t - (1+r)S_{t-1} = -re_t \tag{2.3}$$

where

$$e_t = \frac{1}{1+r}\sum_{i=0}^{\infty}(\frac{1}{1+r})^i(E_t[YL_{t+i}] - E_{t-1}[YL_{t+i}]).$$

Then e_t is a martingale difference, such that $E_t[e_{t+1}] = 0$. Using iterated expectations, (2.3) therefore implies

$$E_t[S_{t+1} - \Delta YL_{t+1}] - (1+r)S_t = 0. \tag{2.4}$$

Expressed by the variables $X_t = (CP_t, YL_t, YK_t)'$, (2.4) can be written

$$E_t[\Delta YK_{t+1} - \Delta CP_{t+1}/\gamma] = -\frac{r}{\gamma}CP_t + r(YL_t + YK_t).$$

When the proportionality factor γ is known, this has the form (2.2) with $c = (-1/\gamma, 0, 1)'$, $d = (-1/\gamma, 1, 1)'$ and $\tau = r$ a parameter to be estimated.
An alternative to the permanent income hypothesis is that consumption is determined by current income as suggested by Keynes. This can be modeled using a VAR model for $(CP_t, Y_t, W_t)'$ of type $\mathcal{H}_1(r)$ with $\alpha = (1, 0, 0)'$ if the reduced rank is 1.

EXAMPLE 2.2. In Boug *et al.* (2017) the following model for inflation dynamics was studied

$$\Delta p_t = \gamma_f E_t[\Delta p_{t+1}] - \lambda(p_t - \delta_1 ulc_t - \delta_2 uic_t) + \gamma_b \Delta p_{t-1} + \delta_0$$

where ulc_t, uic_t denote unit labour cost and unit import cost, and p_t is the consumer price index, all in logarithms. Define $X_t = (ulc_t, uic_t, p_t)'$, $c = (0, 0, 1)'$ and $d = (-\delta_1, -\delta_2, 1)'$. Dividing by γ_f one gets

$$E_t[\Delta p_{t+1}] = (\lambda/\gamma_f)(p_t - \delta_1 ulc_t - \delta_2 uic_t) + (1/\gamma_f)\Delta p_t - (\gamma_b/\gamma_f)\Delta p_{t-1} - \delta_0/\gamma_f$$

which can be expressed as

$$c' E_t[\Delta X_{t+1}] = \tau d' X_t + \tau_1 d_1' \Delta X_t + \tau_2 d_2' \Delta X_{t-1} + \mu. \tag{2.5}$$

This is of the form (2.2) with $\ell = 2$ and $d_1 = d_2 = e_3$.

2.4. Combining the exact rational expectations and the vector autoregressive models

In this section we combine the exact rational expectations and the vector autoregressive models, $\mathcal{H}_1(r)$ and $\mathcal{H}_2(r)$ and express the exact rational expectations model (2.2) as restrictions on the coefficients of the statistical model (2.1). Taking the conditional expectation of $c'\Delta X_{t+1}$ given $X_1, \ldots, X_t$, we get by using (2.1),

$$c' E_t[\Delta X_{t+1}] = c'\alpha\beta' X_t + \sum_{i=1}^{k} c'\Gamma_i \Delta X_{t+1-i} + c'\mu.$$

Equating this expression to (2.2) implies that the following conditions must be satisfied

$$c'\alpha\beta' = \tau d',\ c'\Gamma_i = \tau_i d_i',\ i = 1, \ldots, \ell,\ c'\Gamma_i = 0,\ i = \ell + 1, \ldots, k, c'\mu = d_\mu.$$

This can be summarized as:

PROPOSITION 2.1. *The exact rational expectations restrictions (2.2) imply that the matrix $\alpha\beta'$ is restricted as*

$$c'\alpha\beta' = \tau d', \tag{2.6}$$

and the remaining parameters are restricted as

$$c'\Gamma_i = \tau_i d_i',\ i = 1, \ldots, \ell, c'\Gamma_i = 0,\ i = \ell + 1, \ldots, k,\ c'\mu = d_\mu. \tag{2.7}$$

Note that (2.6) implies that $\tau d'\beta_\perp = 0$, so that when $n \leq q$ we find $\bar{\tau}'\tau d'\beta_\perp = d'\beta_\perp = 0$, and hence $sp(d) \subset sp(\beta)$ and therefore $n \leq r$.

Consider the following two specifications of restriction (2.6) on α

$$c' A\psi\beta' = \tau d' \tag{2.8}$$

and

$$c'(a, a_\perp \phi)\beta' = \tau d'. \tag{2.9}$$

We define two sub-models of $\mathcal{H}_1(r)$ and $\mathcal{H}_2(r)$ respectively which satisfy the restrictions in $\mathcal{RE}$.

$\mathcal{H}_1^\dagger(r)$: The model is a submodel of $\mathcal{H}_1(r)$ which satisfies the restrictions (2.8) and (2.7),
$\mathcal{H}_2^\dagger(r)$: The model is a submodel of $\mathcal{H}_2(r)$ which satisfies the restrictions (2.9) and (2.7).

When estimating models $\mathcal{H}_1^\dagger(r)$ and $\mathcal{H}_2^\dagger(r)$ it is convenient to use a parametrization of freely varying parameters. Such a parametrization is given for the first model in Section 3 together with an analysis of the estimation problem. In Section 4 we discuss the model $\mathcal{H}_2^\dagger(r)$

3. THE SAME RESTRICTIONS ON ALL α, $\mathcal{H}_1^\dagger(R)$

We first give a representation in terms of freely varying parameters of the matrix $\alpha\beta'$, when restricted by $\alpha = A\psi$ and $c'\alpha\beta' = \tau d'$, see (2.8). In this case the relation between $sp(c)$ and $sp(A)$ must be taken into account.

3.1. A reparameterization of $\mathcal{H}_1^\dagger(r)$

Following Johansen and Swensen (2008) it is convenient to define the $s \times o$ matrix u and $(p-q) \times o$ matrix v such that $A'\bar{c}_\perp = uv'$ where o is equal to the rank of $A'\bar{c}_\perp$. The space $\mathbb{R}^p$ has the orthogonal decomposition $(c, \bar{c}_\perp \bar{v}, c_\perp v_\perp)$. Also

$$A = c\bar{c}'A + c_\perp \bar{c}_\perp' A = c\bar{c}'A + c_\perp v u'$$

so $sp(c, A) = sp(c, c_\perp v)$, and $c_\perp v_\perp$ spans the orthogonal complement of $sp(c, A)$. In particular it follows that if $c \in sp(A)$, $sp(A) = sp(c, A) = sp(c, c_\perp v)$ such that

$$c_\perp v \in sp(A) \text{ and } o = s - q \tag{3.1}$$

since A and c have rank s and q respectively.

Proposition 3.1. *Consider the matrix $\Pi = \alpha\beta'$ of rank r of the model defined in equation (1). Let c and d be known matrices of full rank and dimensions $p \times q$ and $p \times n$ respectively where $n \leq q$. Let A be a known $p \times s$ matrix such that $rank(A'\bar{c}_\perp) = o \leq s$, and $A'\bar{c}_\perp = uv'$ for matrix u of rank o and dimension $s \times o$ and matrix v of rank o and of dimension $(p-q) \times o$.*

Consider two sets of restrictions on the parameters of the matrix Π.

The first is formulated as

$$\alpha = A\psi, c'\alpha\beta' = \tau d' \text{ and } rank(\alpha\beta') = r, \tag{3.2}$$

see (2.8).

The second set of restrictions is formulated as

$$\alpha\beta' = \bar{c}\tau d' + c_\perp v\theta d' + c_\perp v\kappa\zeta' d_\perp', \tag{3.3}$$

where it is assumed that $rank(A'\bar{c}_\perp) = o \geq r - n$ and that there exist matrices κ and ζ of full rank and dimensions $o \times (r-n)$ and $(p-n) \times (r-n)$respectively, and where θ is $o \times n$.

It holds that (3.2) implies (3.3), and if further $c \in sp(A)$ then (3.3) implies (3.2).

We summarize the dimensions of the matrices introduced in the table

Table 1. Summary of the coefficient and parameter matrices and their dimensions, as used in the formulation of the model and its reparametrization in Proposition 3.1

	Model			Reparametrization			
Coefficients	$A_{p\times s}$	$c_{p\times q}$	$d_{p\times n}$	$u_{s\times 0}$	$v_{(p-q)\times o}$		
Parameters	$\psi_{s\times r}$	$\beta_{p\times r}$	$\tau_{q\times n}$	$\kappa_{o\times(r-n)}$	$\theta_{o\times n}$	$\zeta_{(p-n)\times(r-n)}$	$\nu_{s\times(p-n)}$

Remark 3.1. Note that the $p \times n$ matrix d contains the known cointegration relations, so $r-n$ is the number of freely varying cointegration relations. Similarly s is the dimension of the range of α, so $s - r$ is a measure of the indeterminacy of the adjustment vectors α.

Remark 3.2. From Proposition 3.1 one gets conditions where a reparameterization using τ, κ, θ and ζ is possible instead of $\alpha = A\psi$ and β restricted as described in equation (3.2). To be more specific: in equation (3.3) $\alpha\beta'$ is expressed by the parameters τ, θ and $\kappa\zeta'$. Conversely, to express τ, θ and $\kappa\zeta'$ by $\alpha = A\psi$ and β, note that from equation (3.4) below

$$\begin{pmatrix} c' \\ \bar{c}'_\perp \end{pmatrix} A\psi\beta'(\bar{d}, \bar{d}_\perp) = \begin{pmatrix} \tau & 0 \\ v\theta & v\kappa\zeta' \end{pmatrix}.$$

Hence $\tau = c'A\psi\beta'\bar{d}$ and $\theta = (v'v)^{-1}v'\bar{c}'_\perp A\psi\beta'\bar{d} = u'\psi\beta'\bar{d}$. Also $v\kappa\zeta' = \bar{c}'_\perp A\psi\beta'\bar{d}_\perp = vu'\psi\beta'\bar{d}_\perp$ such that $\kappa\zeta' = u'\psi\beta'\bar{d}_\perp$ after multiplication with $\bar{v}'$.

Remark 3.3. The number of parameters in the matrix Π restricted as in Model $H_1^\dagger(r)$: $\Pi = A\psi\beta'$ and by (2.8): $c'\Pi = c'A\psi\beta' = \tau d'$, can be found from the representation (3.3) in Proposition 2. The number is given by

$$\#\tau_{q\times n} + \#\theta_{o\times n} + \#\kappa_{o\times(r-n)}\zeta'_{(r-n)\times(p-n)} = qn + on + (r-n)(p-r+o),$$

which for $c \in sp(A)$, where $o = s - q$, reduces to $sr + (r-n)(p-r-q)$.

Proof of proposition 3.1. *Proof that (3.2) implies (3.3).* Pre-multiplying $\alpha\beta' = A\psi\beta'$ by $(c, \bar{c}_\perp)'$ and post-multiplying by $(\bar{d}, \bar{d}_\perp)$ we find using $c'\alpha\beta' = \tau d'$ that

$$\begin{pmatrix} c' \\ \bar{c}'_\perp \end{pmatrix} \alpha\beta'(\bar{d}, \bar{d}_\perp) = \begin{pmatrix} \tau d'\bar{d} & \tau d'\bar{d}_\perp \\ \bar{c}'_\perp\alpha\beta'\bar{d} & \bar{c}'_\perp\alpha\beta'\bar{d}_\perp \end{pmatrix} = \begin{pmatrix} \tau & 0 \\ \bar{c}'_\perp\alpha\beta'\bar{d} & \bar{c}'_\perp\alpha\beta'\bar{d}_\perp \end{pmatrix}.$$

Next use $\bar{c}'_\perp A = vu'$ to simplify the entries $\bar{c}'_\perp\alpha\beta'\bar{d}$ and $\bar{c}'_\perp\alpha\beta'\bar{d}_\perp$. First $\bar{c}'_\perp\alpha\beta'\bar{d} = \bar{c}'_\perp A\psi\beta'\bar{d} = vu'\psi\beta'\bar{d} = v\theta$ for $\theta = u'\psi\beta'\bar{d}$. Furthermore, let $\nu = \psi\beta'\bar{d}_\perp$ such that

$$\bar{c}'_\perp\alpha\beta'\bar{d}_\perp = \bar{c}'_\perp A\psi\beta'\bar{d}_\perp = vu'\nu.$$

Then

$$\begin{pmatrix} c' \\ \bar{c}'_\perp \end{pmatrix} A\psi\beta'(\bar{d}, \bar{d}_\perp) = \begin{pmatrix} \tau & 0 \\ v\theta & vu'\nu \end{pmatrix}. \tag{3.4}$$

Since the matrix $\alpha\beta'$ has rank r and τ has rank n, the matrix $vu'\nu$ must have rank $r - n \geq 0$. Also $r - n = rank(vu'\nu) \leq rank(vu') = rank(\bar{c}'_\perp A) = o$.

We apply Sylvester's inequality, see Horn and Johnson (2013) p.13, to the $(p-q)\times o$ matrix v and the $o\times(p-n)$ matrix $u'\nu$ and find

$$rank(v)+rank(u'\nu)-o\le rank(vu'\nu)\le min(rank(v),rank(u'\nu)).$$

Because $rank(v)=o$, this shows that $rank(u'\nu)\le rank(vu'\nu)$. The reverse inequality is obvious so we find $rank(u'\nu)=rank(vu'\nu)=r-n$.

Then, since $o\ge r-n$ there exists matrices κ and ζ of full rank and dimensions $o\times(r-n)$ and $(p-n)\times(r-n)$ respectively such that that $u'\nu=\kappa\zeta'$ and

$$A\psi\beta'=(\bar{c},c_\perp)\begin{pmatrix}\tau & 0\\ v\theta & v\kappa\zeta'\end{pmatrix}\begin{pmatrix}d'\\ d'_\perp\end{pmatrix}=\bar{c}\tau d'+c_\perp v\theta d'+c_\perp v\kappa\zeta' d'_\perp.$$

Proof that (3.3) implies (3.2). Sylvester's inequality applied to the $(p-q)\times o$ matrix v and the $o\times(r-n)$ matrix κ gives

$$rank(v)+rank(\kappa)-o\le rank(v\kappa)\le min(rank(v),rank(\kappa)).$$

or

$$o+(r-n)-o\le rank(v\kappa)\le min(o,r-n)=r-n.$$

such that the equality holds and $rank(v\kappa)=r-n$. The last inequality follows from the assumption that $r-n\le o$.

The relation

$$\begin{pmatrix}c'\\ \bar{c}'_\perp\end{pmatrix}\alpha\beta'(\bar{d},\bar{d}_\perp)=\begin{pmatrix}c'\\ \bar{c}'_\perp\end{pmatrix}(\bar{c}\tau d'+c_\perp v\theta d'+c_\perp v\kappa\zeta' d'_\perp)(\bar{d},\bar{d}_\perp)=\begin{pmatrix}\tau & 0\\ v\theta & v\kappa\zeta'\end{pmatrix}$$

shows that $\alpha\beta'$ has rank r because τ has rank n and $v\kappa\zeta'$ has rank $r-n$. This follows from an application of Sylvester's inequality since $rank(v\kappa)=r-n$ and

$$rank(v\kappa)+rank(\zeta')-(r-n)=(r-n)+(r-n)-(r-n)\le rank(v\kappa\zeta')\le(r-n).$$

Premultiplying the expression in (3.3) by c' and $A'_\perp$, implies that

$$\begin{aligned}c'\alpha\beta'&=c'\bar{c}\tau d'=\tau d',\\ A'_\perp\alpha\beta'&=A'_\perp(\bar{c}\tau d'+c_\perp v\theta d'+c_\perp v\kappa\zeta' d'_\perp).\end{aligned}$$

But the assumption $c\in sp(A)$ implies $A'_\perp\bar{c}\tau d'=0$. The assumption also implies $c_\perp v\in sp(A)$, see (3.1), such that $A'_\perp c_\perp v=0$ and $A'_\perp\alpha\beta'=0$. Therefore the space spanned by the columns of $\alpha\beta'$ is contained in $sp(A)$, i.e. $\alpha=A\psi$. ■

3.2. *Estimating model $\mathcal{H}_1^\dagger(r)$*

Notice that when $c\in sp(A_\perp)$ there is no model satisfying the constraints (3.2) except for the uninteresting case $d=0$. We therefore only consider the case where c has a component in $sp(A)$. If

$$\varepsilon_t(\psi,\beta,\Gamma_1,\dots,\Gamma_k,\mu)=\Delta X_t-A\psi\beta' X_{t-1}-\sum_{i=1}^k\Gamma_i\Delta X_{t-i}-\mu,\ t=1,\dots,T$$

the maximum likelihood estimates are found by maximizing

$$-\frac{T}{2}\log(\Omega)-\frac{1}{2}\sum_{t=1}^T\varepsilon_t(\psi,\beta,\Gamma_1,\dots,\Gamma_k,\mu)'\Omega^{-1}\varepsilon_t(\psi,\beta,\Gamma_1,\dots,\Gamma_k,\mu)$$

with respect to $\psi, \beta, \Gamma_1, \ldots, \Gamma_k, \mu, \Omega, \tau, \tau_1, \ldots, \tau_\ell$ under the constraints (2.7) and (2.8), in particular $c'A\psi\beta' = \tau d'$. A consequence of the following lemma is that it is not necessary to take into account the part of c that is not in $sp(A)$ because the condition $\tilde{\tilde{c}}'A\psi\beta' = 0$ is always satisfied. Here $\tilde{\tilde{c}}'A$ is the projection of c on $sp(A_\perp)$. Thus it suffices to find the maximum likelihood estimates for $c \in sp(A)$.

Lemma 3.1. *Let $\tilde{c} = \bar{A}A'c$ and $\tilde{\tilde{c}} = \bar{A}_\perp A'_\perp$. Then*

$$c'A\psi\beta' = \tau d' \tag{3.5}$$

is equivalent to

$$\tilde{c}'A\psi\beta' = \tau d' \text{ and } \tilde{\tilde{c}}'A\psi\beta' = 0. \tag{3.6}$$

Proof of Lemma 3.1. *Proof that (3.5) implies (3.6).* Assume (3.5). Then $\tilde{c}'A\psi\beta' = c'\bar{A}A'A\psi\beta' = c'A\psi\beta' = \tau d'$ and $\tilde{\tilde{c}}'A\psi\beta' = c'\bar{A}_\perp A'_\perp A\psi\beta' = 0$
Proof that (3.6) implies (3.5). Adding the terms, $\tau d' = (\tilde{c}' + \tilde{\tilde{c}}')A\psi\beta' = \tilde{c}'A\psi\beta' = c'A\bar{A}'A\psi\beta' = c'A\psi\beta'$. ■

Remark 3.4. If $X_t = (X'_{1t}, X'_{2t})'$ and $A = (I, 0)'$, X_{2t} is weakly exogenous for the parameters (β, ψ), Johansen (1996). Then $c \in sp(A)$ means that the elements in the lower rows of c are equal to 0. Generally, let $c = (c'_1, c'_2)'$ where c_1 and c_2 are $s \times q$ and $(p-s) \times q$ matrices respectively. Then $c'A = (c'_1, c'_2)A = c'_1$ such that the restriction $c'A\psi\beta' = \tau d'$ does not involve c_2, which represents the part of c not in $sp(A)$. Thus, in economic modelling, only c_1, which determines the non-null rows in $\Pi = A\psi\beta'$, and the restrictions $c'\Gamma_i = \tau_i d'_i,\ i = 1, \ldots, \ell,\ c'\Gamma_i = 0,\ i = \ell+1, \ldots, k$. must be specified.

We will now show how estimation of $\mathcal{H}_1^\dagger(r)$ defined in Subsection 2.4 can be performed by reduced rank regression and regression under the assumption that $c \in sp(A)$. Then $o = s - q$, see (3.1). Two cases need to be distinguished.

First consider the case where $o < r - n$. The restricted model can then be estimated by premultiplying model (2.1) with $(c, \bar{c}_\perp)'$ and incorporating the restrictions. Then one can reparameterize by conditioning $\bar{c}'_\perp \Delta X_t$ on $c'\Delta X_t$ and the past. The conditional equation can be estimated by a combination of reduced rank and ordinary least squares (OLS) regressions, and the parameters in marginal equation for $c'\Delta X_t$ can be estimated by OLS regressions. Details can be found in Appendix A.

Next consider the case where $o \geq r - n$. Then a more elaborate argument is needed. The matrix $\alpha\beta'$ can according to Proposition 2 be reparameterized as

$$\alpha\beta' = \bar{c}\tau d' + c_\perp v\theta d' + c_\perp v\kappa\zeta' d'_\perp, \tag{3.7}$$

and $c'\Gamma_i,\ i = 1 \ldots, k$ are restricted as

$$c'\Gamma_i = \tau_i d'_i,\ i = 1, \ldots, \ell,\ c'\Gamma_i = 0,\ i = \ell+1, \ldots, k. \tag{3.8}$$

Remark that $rank(\tau d') = n$ and $rank(v(\theta d' + \kappa\zeta' d'_\perp)) = o$. Hence, since the matrix (3.7) has rank r, $o + n \geq r$, i.e. $rank(A'\bar{c}_\perp) \geq r - n$.

We first define the three processes X^*_{1t}, X^*_{2t}, X^*_{3t}, by

$$X^*_t = \begin{pmatrix} X^*_{1t} \\ X^*_{2t} \\ X^*_{3t} \end{pmatrix} = \begin{pmatrix} \bar{v}'\bar{c}'_\perp X_t \\ v'_\perp \bar{c}'_\perp X_t \\ c'X_t \end{pmatrix} = GX_t, \tag{3.9}$$

of dimensions $(s-q, p-s, q)$ respectively. Define $\alpha^*, \Gamma_i^*,\ i=1,\ldots,k, \mu^*$ and ε_t^* similarly, i.e. $\alpha^* = (\alpha_1^{*\prime}, \alpha_2^{*\prime}, \alpha_3^{*\prime})' = G\alpha, \Gamma_i^* = (\Gamma_{1i}^{*\prime}, \Gamma_{2i}^{*\prime}, \Gamma_{3i}^{*\prime})' = G\Gamma_i, \mu^* = G\mu$ and $\varepsilon_t^* = G\varepsilon_t$. Let $\Sigma = G\Omega G'$. The equations then become

$$\Delta X_t^* = \alpha^* \beta' X_{t-1} + \sum_{i=1}^{k} \Gamma_i^* \Delta X_{t-i}^* + \mu^* + \varepsilon_t^*. \tag{3.10}$$

We next want the conditional equations and define $\varepsilon_t^{**} = K\varepsilon_t^*$ where

$$K = \begin{pmatrix} I_{s-q} & -\omega_{12.3} & -\omega_{13.2} \\ 0 & I_{p-s} & -\omega_{2.3} \\ 0 & 0 & I_q \end{pmatrix} \tag{3.11}$$

with $\omega_{12.3} = \Sigma_{12.3}\Sigma_{22.3}^{-1}$, $\omega_{13.2} = \Sigma_{13.2}\Sigma_{33.2}^{-1}$ and $\omega_{2.3} = \Sigma_{23}\Sigma_{33}^{-1}$. Then $\varepsilon_{1t}^{**}, \varepsilon_{2t}^{**}, \varepsilon_{3t}^{*}$ are independent, and we find the equation for ΔX_{1t}^* given $(\Delta X_{2t}^*, \Delta X_{3t}^*)$, the equation for ΔX_{2t}^* given ΔX_{3t}^* and the equation for ΔX_{3t}^* by pre-multiplying the right hand side of equation (3.10) by the matrix K. Thus

$$\Delta X_t^* = \begin{pmatrix} \omega_{12.3} & \omega_{13.2} & \alpha_1^{**}\beta' & \Gamma_{11}^{**} & \ldots & \Gamma_{1k}^{**} & \mu_1^{**} \\ 0 & \omega_{2.3} & \alpha_2^{**}\beta' & \Gamma_{21}^{**} & \ldots & \Gamma_{2k}^{**} & \mu_2^{**} \\ 0 & 0 & \alpha_3^{*}\beta' & \Gamma_{31}^{*} & \ldots & \Gamma_{3k}^{*} & \mu_3^{*} \end{pmatrix} Z_t^* + \begin{pmatrix} \varepsilon_{1t}^{**} \\ \varepsilon_{2t}^{**} \\ \varepsilon_{3t}^{*} \end{pmatrix},$$

$$Z_t^{*\prime} = \left(\Delta X_{2t}^{*\prime}, \Delta X_{3t}^{*\prime}, X_{t-1}', \Delta X_{t-1}^{*\prime}, \ldots, \Delta X_{t-k}^{*\prime}, 1\right),$$

where $\alpha_1^{**} = \alpha_1^* - \omega_{12.3}\alpha_2^* - \omega_{13.2}\alpha_3^*$, $\alpha_2^{**} = \alpha_2^* - \omega_{2.3}\alpha_3^*$ with similar expressions for $\Gamma_{mi}^{**}, \mu_m^{**}, \varepsilon_m^{**}, m = 1, 2$, $E(\epsilon_{1t}^{**}\epsilon_{1t}^{**\prime}) = \Sigma_{11.23}$, $E(\epsilon_{2t}^{**}\epsilon_{2t}^{**\prime}) = \Sigma_{22.3}$ and $E(\epsilon_{3t}^{**}\epsilon_{3t}^{**\prime}) = \Sigma_{33}$.

We now introduce the restrictions. Under (3.7), $\alpha\beta' = \bar{c}\tau d' + c_\perp v(\theta d' + \kappa\zeta' d_\perp')$, we find

$$\begin{pmatrix} \bar{v}'\bar{c}_\perp' \\ v_\perp'\bar{c}_\perp' \\ c' \end{pmatrix} \alpha\beta' = \begin{pmatrix} \bar{v}'\bar{c}_\perp' \\ v_\perp'\bar{c}_\perp' \\ c' \end{pmatrix} [(\bar{c}\tau + c_\perp v\theta)d' + c_\perp v\kappa\zeta' d_\perp']$$

$$= \begin{pmatrix} \theta \\ 0 \\ \tau \end{pmatrix} d' + \begin{pmatrix} \kappa\zeta' \\ 0 \\ 0 \end{pmatrix} d_\perp'.$$

After pre-multiplication by the matrix K defined in (3.11) we have

$$\begin{pmatrix} \alpha_1^{**} \\ \alpha_2^{**} \\ \alpha_3^{*} \end{pmatrix} \beta' = \begin{pmatrix} \theta - \omega_{13.2}\tau \\ -\omega_{2.3}\tau \\ \tau \end{pmatrix} d' + \begin{pmatrix} \kappa\zeta' \\ 0 \\ 0 \end{pmatrix} d_\perp'.$$

With the definitions $\Gamma_{j1}^{***} = (\Gamma_{j1}^{**} \ldots \Gamma_{j,\ell}^{**}), j = 1, 2$, $\Gamma_{31}^{***} = (\Gamma_{31}^{*} \ldots \Gamma_{3,\ell}^{*}) = (\tau_1 d_1', \ldots, \tau_\ell d_\ell')$

and $\Gamma_{j2}^{***} = (\Gamma_{j,\ell+1}^{**} \ldots \Gamma_{j,k}^{**}), j = 1, 2$ the conditional equations may be written

$$\Delta X_t^* = \begin{pmatrix} \omega_{12.3} & \omega_{13.2} & \theta - \omega_{13.2}\tau & \kappa\zeta' & \Gamma_{11}^{***} & \Gamma_{12}^{***} & \mu_1^{**} \\ 0 & \omega_{2.3} & -\omega_{2.3}\tau & 0 & \Gamma_{21}^{***} & \Gamma_{22}^{***} & \mu_2^{**} \\ 0 & 0 & \tau & 0 & \Gamma_{31}^{***} & 0 & \mu_3^{*} \end{pmatrix} Z_t^{**} + \begin{pmatrix} \varepsilon_{1t}^{**} \\ \varepsilon_{2t}^{**} \\ \varepsilon_{3t}^{*} \end{pmatrix},$$

where

$$Z_t^{**\prime} = \left(\Delta X_{2t}^{*\prime}, \Delta X_{3t}^{*\prime}, X_{t-1}' d, X_{t-1}' d_\perp, \Delta X_{t-1}^{*\prime}, \ldots, \Delta X_{t-k}^{*\prime}, 1\right).$$

The parameters in the equation for ΔX_{1t}^* using $\theta^* = \theta - \omega_{13.2}\tau$, are variation independent of the parameters in the equations for ΔX_{2t}^* and ΔX_{3t}^*. The coefficient $\omega_{2.3}\tau$, however, represents a cross equation restriction as a product of $\omega_{2.3}$ from the equation for ΔX_{2t}^*, and τ from the equation for ΔX_{3t}^*. In the analysis below we therefore first assume τ is known and introduce the variable $\Delta X_{3t}^* - \tau d' X_{t-1}$, then estimate by reduced rank regression and finally optimize over τ.

We summarize the estimation procedure when $rank(A'\bar{c}_\perp) \geq r - n$ as

PROPOSITION 3.2. *Estimation of the model $\mathcal{H}_1^\dagger(r)$ when $rank(A'\bar{c}_\perp) \geq r - n$ can be conducted in three steps when τ is known:*

1. *By reduced rank regression of ΔX_{1t}^* on $d'_\perp X_{t-1}$ corrected for the remaining regressors in Z_t^{**} by OLS, find estimates $\hat{\omega}_{12.3}$, $\hat{\omega}_{13.2}$, $\hat{\theta}^*$, $\hat{\kappa}$, $\hat{\zeta}$, $\hat{\Gamma}_{11}^{**}, \dots, \hat{\Gamma}_{1k}^{**}$, $\hat{\mu}_1^{**}$, and $\hat{\Sigma}_{11.23}$.*
2. *For fixed value of τ impose the restrictions $c'\Gamma_i = \Gamma_{3i}^* = \tau_i d'_i$, $i = 1, \dots, \ell$, $c'\Gamma_i = \Gamma_{3i}^* = 0$, $i = \ell + 1, \dots, k$, see (3.8), introduce the variable $\Delta X_{3t}^* - \tau d' X_{t-1}$ and find the system*
$$\begin{pmatrix} \Delta X_{2t}^* \\ \Delta X_{3t}^* - \tau d' X_{t-1} \end{pmatrix} = \begin{pmatrix} \omega_{2.3} & & \Gamma_{21}^{***} & & \Gamma_{22}^{***} & \mu_2^{**} \\ 0 & \tau_1 d'_1 & \dots & \tau_\ell d'_\ell & 0 & \mu_3^* \end{pmatrix} Z_t^{***} + \begin{pmatrix} \varepsilon_{2t}^{**} \\ \varepsilon_{3t}^* \end{pmatrix},$$
$$Z_t^{***\prime} = \left((\Delta X_{3t}^* - \tau d' X_{t-1})', \Delta X_{t-1}^{*\prime}, \dots, \Delta X_{t-k}^{*\prime}, 1 \right).$$
This system can be estimated by OLS regression. From the equation for ΔX_{2t}^* a regression gives estimates $\hat{\omega}_{2.3}$, $\hat{\Gamma}_{2i}^{**}$, $i = 1, \dots, k$, $\hat{\mu}_2^*$, $\hat{\Sigma}_{22.3}$. From the equation for $\Delta X_{3t}^* - \tau d' X_{t-1}$, an regression gives the estimates $\hat{\tau}_i$, $i = 1, \dots, \ell$, $\hat{\mu}_3$, $\hat{\Sigma}_{33}$ all depending on τ.
3. *Finally, the maximal value of the likelihood equals, apart from constants,*
$$L_{\max}^{-2/T}(\tau) = |\hat{\Sigma}_{11.23}||\hat{\Sigma}_{22.3}||\hat{\Sigma}_{33}|/|c'c||\bar{v}'\bar{c}'_\perp \bar{c}_\perp \bar{v}||v'_\perp \bar{c}'_\perp \bar{c}_\perp v_\perp|.$$

We can therefore calculate the maximized likelihood for the three equations for a given τ, and maximize it respect to τ, by a general maximization algorithm if τ is not known.

REMARK 3.5. The case $r - n = 0$ needs a special comment. In this case the last term in equation (3.7) disappears which has the implication that the reduced rank regression of ΔX_{1t}^* on $d_\perp X_{t-1}$ is not necessary in step 1 in Proposition 3.2, and OLS regression is sufficient. For the case where $A = I$ and $n = q$ see Johansen and Swensen (1999) where the procedure is written out in detail.

REMARK 3.6. Another possibility for estimating the equations for $\Delta X_{2t}^*, \Delta X_{3t}^*$ when τ is unknown is to notice that the model

$$\Delta X_{2t}^* = \sum_{i=1}^{k} \Gamma_{2i}^* \Delta X_{t-i}^* + \mu_2^* + \varepsilon_{2t}^*, \tag{3.12}$$
$$\Delta X_{3t}^* = \tau d' X_{t-1} + \sum_{i=1}^{\ell} \Gamma_{3i}^* \Delta X_{t-i}^* + \mu_3^* + \varepsilon_{3t}^*.$$

with the restriction (3.8) is linear in the conditional mean and hence can be estimated

by generalized least squares for fixed variance matrix. For fixed linear parameters the variance can be estimated from the residuals, such that an iteration procedure can be defined. This is an example of a coordinate search method, see e.g Nocedal and Wright (2006). The model is a special case of seemingly unrelated regressions, SUR. In Oberhofer and Kmenta (1974) it is shown that for such models the sequence has a limit point which is a solution of the likelihood equations. In general, for iterative methods for maximization there is no guarantee that they will converge toward the global maximum if there are several local maxima. Drton and Richardson (2004) contains a discussion of multi modularity of the likelihood in bivariate SUR models.
Let the residual from this fit be $R_{23,t}$ and let $S_{23} = \frac{1}{T}\sum_{t=1}^{T} R_{23,t}R'_{23,t}$. Then the maximal value of the likelihood, apart from constants, can be expressed as

$$L_{\max}^{-2/T} = |\hat{\Sigma}_{11.23}||S_{23}|/|c'c||\bar{v}'\bar{c}'_{\perp}\bar{c}_{\perp}\bar{v}||v'_{\perp}\bar{c}'_{\perp}\bar{c}_{\perp}v_{\perp}|.$$

REMARK 3.7. For the case where all elements in the matrix d_μ are known, a small modification is necessary in the equation for the variable $\Delta X^*_{3,t}$. Instead of regressing $\Delta X^*_{3,t} - \tau d'X_{t-1}$ on $d'_1\Delta X^*_{t-1}, \ldots, d'_\ell \Delta X^*_{t-\ell}$ and 1, regress $\Delta X^*_{3,t} - \tau d'X_{t-1} - d_\mu$ on $d'_1\Delta X^*_{t-1}, \ldots, d'_\ell \Delta X^*_{t-\ell}$ only. In particular, if $d_\mu = 0$ the response is $\Delta X^*_{3,t} - \tau d'X_{t-1}$ and the regressor 1 is dropped.

REMARK 3.8. There is an interesting modification of the estimation procedure described above. If the coefficient $-\omega_{2.3}\tau$ of $d'X_{t-1}$ in the equation for ΔX_{2t} in (B.1) is replaced by a freely varying parameter the new system will contain $(p-s)q$ extra parameters. It will, however, have a structure so that the expanded parameter set can be estimated by ordinary least squares.

4. SOME α ASSUMED KNOWN, $\mathcal{H}_2^{\dagger}(R)$

We consider the situation where the freely varying parameters of the matrix $\alpha\beta'$ are restricted by $\alpha = (a, a_\perp\phi)$ such that $c'\alpha\beta' = c'(a\beta'_1 + a_\perp\phi\beta'_2) = \tau d'$ where a is a known $p \times m$ matrix of full rank, see (2.9). We start with two special cases, **A**: $c'a = 0$, $0 < m \leq r - n$ and β_1 known and **B**: $0 < m = r$. Thereafter we consider how the more general model where also β_1 is unknown can be estimated.

4.1. Estimating some special cases of the model $\mathcal{H}_2^{\dagger}(r)$

A: when $c'a = 0$, $0 < m \leq r - n$ and β_1 known it is possible to apply Proposition 3.1. The constraints on the matrix Π are now

$$c'\Pi = c'\alpha\beta' = c'(a\beta'_1 + a_\perp\phi\beta'_2) = \tau d' \text{ or } c'\alpha\beta' = c'a_\perp\phi\beta'_2 = \tau d'$$

such that the model can be written

$$\Delta X_t - a\beta'_1 X_{t-1} = a_\perp\phi\beta'_2 X_{t-1} + \sum_{i=1}^{k}\Gamma_i\Delta X_{t-i} + \mu + \varepsilon_t \tag{4.1}$$

with constraints $c'a_\perp\phi\beta'_2 = \tau d'$ and (2.7).

These are analogous to the restrictions (2.8) and (2.7) with $rank(a_\perp\phi\beta'_2) = r_1$ equal to $r - m$ and the known $p \times (p-m)$ matrix $a_\perp$ corresponding to A. In addition $c'a = 0$ is equivalent to $c \in sp(a_\perp)$ which implies that the number of columns of $a_\perp$ minus the number of constraints equals the rank of $a'_\perp c$, i.e. $p - m - q = rank(a'_\perp\bar{c}_\perp)$, see (3.1). Also, we deal only with the situations where $0 < m \leq r - n$. The case where $r - m = n$, i.e. $m = r - n$ is special as noted in Remark 3.5. In particular since $d \subset sp(\beta_2)$, as explained after Proposition 2.1, and rank $sp(\beta_2) = r - m = n = rank(d)$, $sp(d) = sp(\beta_2)$ so β_2 is known up to a normalization in this case.

But there are also two distinct cases, as in section 3, depending on the value of $rank(a'_{\perp}\bar{c}_{\perp})$. If $rank(a'_{\perp}\bar{c}_{\perp}) < r_1 - n = r - m - n$, the argument based upon the conditional equation of $\bar{c}'_{\perp}\Delta X_t$ given $c'\Delta X_t$ and the past must be applied. If $rank(a'_{\perp}\bar{c}_{\perp}) \geq r_1 - n$, the conditions of Proposition 3.1 are satisfied, and the parameters can be estimated as described in Proposition 3.2. This can be done either by first keeping τ fixed and then using a general optimizing algorithm to find the maximizing value of τ or using the SUR procedure described in Remark 3.6.

B: when $0 < m = r$, $\alpha = a$ all adjustment parameters are known. After premultiplying the model (2.1) with $(c, \bar{c}_{\perp})'$ and incorporating the restrictions the parameters of the coefficient of the level does not have a multiplicative structure. The reason is that the parameters in β are the only unknowns. A direct application of Proposition 3.2 is therefore not possible, and a small modification of the arguments used there is necessary. The details can be found in Appedix B.

4.2. Estimating the model $\mathcal{H}_2^{\dagger}(r)$ when $c'a = 0$ and $0 < m \leq r - n$

Now, consider the situation where $c'(a\beta_1' + a_{\perp}\phi\beta_2') = \tau d'$ and $c'a = 0$. Then the model with the restrictions imposed can be written

$$\bar{c}'_{\perp}\Delta X_t = \bar{c}'_{\perp}a\beta_1'X_{t-1} + \bar{c}'_{\perp}a_{\perp}\phi\beta_2'X_{t-1} + \sum_{i=1}^{k}\bar{c}'_{\perp}\Gamma_i\Delta X_{t-i} + \bar{c}'_{\perp}\mu + \bar{c}'_{\perp}\varepsilon_t,$$

$$c'\Delta X_t = \tau d'X_{t-1} + \sum_{i=1}^{\ell}\tau_i d_i'\Delta X_{t-i} + c'\mu + c'\varepsilon_t.$$

The model is unidentified if $rank(\bar{c}'_{\perp}a) < m$ and $rank(\bar{c}'_{\perp}a) > m$ is not possible since the matrix a has rank m. Hence $rank(\bar{c}'_{\perp}a) = m$ if we assume that the model is identified.

We propose the following iterative procedure for estimating the parameters. Assume first that τ is fixed and consider the situation where $rank(a'_{\perp}\bar{c}_{\perp}) \geq r_1 - n$.

Step 1. Keep ϕ and β_2 fixed. Writing the model

$$\bar{c}'_{\perp}(\Delta X_t - a_{\perp}\phi\beta_2'X_{t-1}) = \bar{c}'_{\perp}a\beta_1'X_{t-1} + \sum_{i=1}^{k}\bar{c}'_{\perp}\Gamma_i\Delta X_{t-i} + \bar{c}'_{\perp}\mu + \bar{c}'_{\perp}\varepsilon_t, \quad (4.2)$$

$$c'\Delta X_t - \tau d'X_{t-1} = \sum_{i=1}^{\ell}\tau_i d_i'\Delta X_{t-i} + c'\mu + c'\varepsilon_t$$

estimate the parameters β_1, $\bar{c}'_{\perp}\Gamma_i$, $i = 1, \ldots, k$, μ, $\tau_1, \ldots, \tau_\ell$ and Σ as described in **B** in the previous subsection.

Step 2. Keep β_1 fixed. Remembering that $rank(a_{\perp}\phi\beta_2') = r_1 = r - m$ and $c'a = 0$

write the model

$$\bar{c}_\perp'(\Delta X_t - a\beta_1' X_{t-1}) = \bar{c}_\perp' a_\perp \phi\beta_2' X_{t-1} + \sum_{i=1}^{k} \bar{c}_\perp' \Gamma_i \Delta X_{t-i} + \bar{c}_\perp'\mu + \bar{c}_\perp'\varepsilon_t, \quad (4.3)$$

$$c'\Delta X_t - \tau d' X_{t-1} = \sum_{i=1}^{\ell} \tau_i d_i' \Delta X_{t-i} + c'\mu + c'\varepsilon_t.$$

Estimate the parameters ϕ, β_2 (when they are unknown, i.e. $r-n>m$), $\bar{c}_\perp'\Gamma_i,\ i = 1,\dots,k,\ \mu,\ \tau_1,\dots,\tau_\ell$ and Σ as described in **A** of the previous subsection.

STEP 3. Repeat calculations N times.

The value of the likelihood increases for each iteration even though convergence toward a global maximum, or even convergence cannot be guaranteed. Also convergence, when it occurs, can be slow. But such coordinate search methods can still be useful, see Nocedal and Wright (2006), page 230.

If $\hat\beta_1, \hat\phi, \hat\beta_2, \widehat{\bar{c}_\perp'\Gamma_i},\ i = 1,\dots,k, \hat\mu, \hat\tau_1,\dots,\hat\tau_\ell$ and $\hat\Sigma$ are the maximum likelihood estimates when τ is fixed and

$$R_t = \begin{pmatrix} \bar{c}_\perp'\Delta X_t & - & \bar{c}_\perp' a\hat\beta_1 X_{t-1} - \bar{c}_\perp' a_\perp \hat\phi\hat\beta_2' X_{t-1} - \sum_{i=1}^k \widehat{\bar{c}_\perp'\Gamma_i}\Delta X_{t-i} - \bar{c}_\perp'\hat\mu \\ c'\Delta X_t & - & \tau d' X_{t-1} - \sum_{i=1}^{\ell} \hat\tau_i d_i' \Delta X_{t-i} - c'\hat\mu \end{pmatrix},\ t = 1\dots,T$$

are the residuals, the maximal value of the likelihood is

$$L_{\max}^{-2/T}(\tau) = |\sum_{t=1}^{T} R_t R_t'|/|c'c\ ||\bar{c}_\perp'\bar{c}_\perp|.$$

For unknown τ we can find the maximum likelihood estimators using a general numerical optimization procedure. Another possibility is to insert a SUR step to estimate an unknown τ in the conditional distributions described in A and B in the previous subsection.

The other situation, where $rank(a_\perp'\bar{c}_\perp) < r_1 - n$, is simpler. Then it is not needed to keep τ fixed in both steps. One can proceed as follows. In step 1 τ is kept fixed in addition to ϕ and β_2. The model (4.2) can be reparameterized by conditiong on the marginal equation. The other parameters can then be estimated by OLS regression. In step 2 the model can be reparameterized by conditiong on the marginal equation. The parameters, now including τ, can then be estimated by a reduced rank and OLS regression when $r-n-m>0$. When $r-n-m=0$ only OLS regression is needed.

REMARK 4.1. The number of parameters in the matrix Π satisfying $\mathcal{H}_2^\dagger(r): \Pi = a\beta_1' + a_\perp\phi\beta_2'$ and the restriction $c'(a\beta_1' + a_\perp\phi\beta_2') = \tau d'$ can be found counting the parameters estimated by the recursive procedure. Remember $rank(\bar{c}_\perp' a) = m$ and that $c'a = 0$ is assumed. We take τ as fixed for the moment so $\tau d'$ is a known $q \times p$ matrix. First, β_1 in step 1 contains $p \cdot rank(\bar{c}_\perp' a) = pm$ parameters. Secondly, in step 2 we have to consider the two cases.

If $rank(a_\perp'\bar{c}_\perp) \geq r_1 - n$, it follows from Remark 3.3, since $rank(\tau d' - c'a\beta_1') = q$ and $c'a = 0$, that the number of parameters in $a_\perp\phi\beta_2'$ satisfying $c'a_\perp\phi\beta_2' = \tau d'$ is $(p-m)(r-m) + (r-m)(p-r+m-q)$. We have used that s corresponds to $p-m$ and that the rank of $\phi\beta_2'$ is $r-m$. Thus the total number of parameters is $pm + (r-m)(2p-r-q)$.

If $o = rank(a_\perp'\bar{c}_\perp) < r_1 - n$ the matrix $\bar{c}_\perp' a_\perp\phi\beta_2'$ of dimension $(p-q)\times p$ of rank o contains $(p-q)o + o(p-o) = o(2p-q-o)$ parameters. In this case the total number of

parameters is therefore $pm + o(2p - q - o)$.
If also τ is unknown, the number of parameters is increased by nq.

5. AN APPLICATION

We consider the inflation model discussed in Example 2.2 In Boug *et al.* a data set covering the period 1982:1-2005:4 was analysed. A reduced rank vector autoregressive model with three lags, i.e. $k = 2$, an unrestricted constant, seasonal dummies and five impulse dummies was fitted to the time series. Rank equal to 1 was found to yield a satisfactory fit, the cointegration vector was estimated as

$$0.649ulc_t + 0.340uic_t = p_t,\ 2\log L_{\max}(\mathcal{H}(r)) = 2536.72$$

and the adjustment parameters as $\hat{\alpha} = (0.175, 0.143, -0.056)'$. Concerning the model $\mathcal{H}_1(1)$, Tables 3 amd 4 below show the result of testing whether ulc_t and uic_t are jointly weakly exogenous for α_3 and β and of testing whether uic_t is weakly exogenous for α_1, α_3 and β. The p-value of the former is 0.02.

Now we turn to testing the model $\mathcal{H}_1^{\dagger}(1)$ against the model $\mathcal{H}_1(1)$. First we estimate the model $\mathcal{H}_1^{\dagger}(1)$ under the assumption that $(ulc_t, uic_t)'$ is weakly exogenous for α_3 and β in addition to satisfying the exact rational expectations hypothesis defined in equation (2.5). Then $c = (0, 0, 1)'$ and $A = (0, 0, 1)'$ such that $c \in sp(A)$. Furthermore, $c_{\perp} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $A'c_{\perp} = (0, 0)$. Then $u = 0$ and $v = (0, 0)'$ such that $v_{\perp} = I_2$, $G = I_3$ and $X_t = X_t^*$, see (3.9).

The equations described in Section 3.2 become

$$v_{\perp}'\bar{c}_{\perp}'\Delta X_t = \Delta\begin{pmatrix} ulc_t \\ uic_t \end{pmatrix} = \sum_{i=1}^{2}\bar{c}_{\perp}'\Gamma_i\Delta X_{t-i} + \bar{c}_{\perp}'\Phi D_t + \bar{c}_{\perp}'\varepsilon_t,$$

$$c'\Delta X_t = \qquad \Delta p_t \qquad = \tau d'X_{t-1} + \sum_{i=1}^{2}\tau_i d_i'\Delta X_{t-i} + c'\Phi D_t + \varepsilon_{3t},$$

where d contains some fixed values of δ_1, δ_2 and D_t denotes the deterministic terms, i.e. the unrestricted constant, the seasonal dummies and the six impulse dummies. If τ is known the system can be estimated by first regressing Δulc_t and Δuic_t on Δulc_{t-i}, $i = 1, 2$, Δuic_{t-i}, $i = 1, 2$, Δp_{t-i}, $i = 1, 2$, and $\Delta p_t - \tau d'X_{t-1}$ and D_t, and then regressing $\Delta p_t - \tau d'X_{t-1}$ on $d_1'\Delta X_{t-1}$, $d_2'\Delta X_{t-2}$ and D_t.

When τ is an unknown parameter we can compute the profile likelihood. The maximum likelihood estimator is the value that maximizes this profile likelihood. Alternatively one may employ the method based on generalized least squares described in Section 3.2 to find the maximal value of the likelihood. Under the usual regularity condition the rational expectation hypothesis restrictions and the weak exogeneity can be tested by a likelihood ratio test. The number of parameters in the reduced rank VAR model is $3 + 2 + 18 = 23$ in addition to the coefficients of constants and dummies. The corresponding number after imposing the restrictions is $1 + 12 + 2 = 15$. The appropriate degrees of freedom is therefore 8 when τ must be estimated.

We next estimate the model $\mathcal{H}_1^{\dagger}(1)$ under the assumption that that uic_t is weakly exogenous for $(\alpha_1, \alpha_3)'$ and β in addition to satisfying the exact rational expectations hypothesis defined in equation (2.5). Then

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$$

and $c = (0, 0, 1)'$ as before such that also now $c \in sp(A)$. Also

$$A'c_\perp = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1, 0) = uv'.$$

Thus $\bar{v}'\bar{c}_\perp' = (1, 0, 0) = e_1'$ and $v_\perp'\bar{c}_\perp' = (0, 1, 0) = e_2'$. Also in this case therefore $G = I_3$ and the equations described in Section 3.2 now become

$$\begin{aligned} \bar{v}'\bar{c}_\perp'\Delta X_t = \; & \Delta ulc_t = \theta d'X_{t-1} + \sum_{i=1}^{2} e_1'\Gamma_i\Delta X_{t-i} + e_1'\Phi D_t + \varepsilon_{1t}, \\ v_\perp'\bar{c}_\perp'\Delta X_t = \; & \Delta uic_t = \sum_{i=1}^{2} e_2'\Gamma_i\Delta X_{t-i} + e_2\Phi D_t + \varepsilon_{2t}, \\ c'\Delta X_{3t} = \; & \Delta p_t = \tau d'X_{t-1} + \sum_{i=1}^{2} \tau_i d_i'\Delta X_{t-i} + e_1'\Phi D_t + \varepsilon_{3t}. \end{aligned}$$

The system can be estimated by first regressing Δulc_t on $d'X_{t-1}$, $\Delta ulc_{t-i},\ i = 1, 2$, $\Delta uic_{t-i},\ i = 0, 1, 2$, $\Delta p_{t-i},\ i = 1, 2$, $\Delta p_t - \tau d'X_{t-1}$ and D_t, then regressing Δuic_t on $\Delta ulc_{t-i},\ i = 1, 2$, $\Delta uic_{t-i},\ i = 1, 2$, $\Delta p_{t-i},\ i = 1, 2$, $\Delta p_t - \tau d'X_{t-1}$ and D_t and finally by regressing $\Delta p_t - \tau d'X_{t-1}$ on $d_1'\Delta X_{t-1}$, $d_2'\Delta X_{t-2}$ and D_t,

In this case there is an additional parameter under the hypothesis so the likelihood ratio test has 7 degrees of freedom. As before τ is treated as an unknown parameter and the maximum likelihood estimates can be obtained from the profile likelihood. Alternatively the approach based on generalized least squares can be used to find the maximal value of the likelihood.

The maximal values of 2 log likelihood from fitting the two models are displayed in Table 2 and Table 3 when $d = (-\delta_1, -\delta_2, 1)'$ is fixed as $d = (-2/3, -1/3, 1)'$. Choosing the known d as $d = (-2/3, -1/3, 1)$ is a sensible choice. As mentioned in Boug et al. (2017), Aukrust (1977) pointed out that for Norway the direct effect on consumer prices of a proportionate increase in import prices is around 0.33 percent. The p-value of the hypothesis that uic_t is weakly exogenous for α_1, α_3, and β is 0.10. The further hypothesis imposing in addition rational expectations, i.e. (2.5), is also not rejected.

The maximum likelihood estimate for τ is denoted by $\hat{\tau}$. Ignoring the cross equation restrictions, as pointed out in Remark 3.8, implies that τ is estimated by regression from the marginal equation for $\Delta X_{3,t}$. As one can see this estimate, denoted by $\tilde{\tau}$, is numerically quite similar to the result when the cross equation restrictions are taken into account. The estimated standard error is 0.012 such that an approximate 95% confidence interval is $(-0.075, -0.028)$.

Often the matrix d contains unknown parameters. Then an additional maximization is necessary for testing weak exogeneity and the rational expectations hypothesis simultaneously as proposed here. For the situation where $d = (-\delta, -(1-\delta), 1)'$ the profile likelihood, maximizing also over τ is shown in Figure 1 using the procedure `nlm` in the software package `R`, R (2020), for two situations. In one only unit import cost is considered as weakly exogenous in addition to satisfying the rational expectation hypothesis

Table 2. Summary of tests of ulc_t and uic_t jointly weakly exogenous and of the restriction (2.5) with $\alpha_1 = \alpha_2 = 0$. For $\mathcal{H}_1^\dagger(1)$ $d' = (-2/3, -1/3, 1)'$ is fixed. LR is the likelihood ratio.

Model	2 log $L_{\max}$(Model)	$-2\log LR$	df	p-value	$\hat{\tau}$	$\tilde{\tau}$
$\mathcal{H}(1)$	2536.72	-	-	-	-	-
$\mathcal{H}_1(1)$	2528.81	7.91	2	0.02	-	-
$\mathcal{H}_1^\dagger(1)$	2524.91	3.90	6	0.69	-0.052	-0.051

Table 3. Summary of tests of uic_t weakly exogenous and of the restriction (2.5) with $\alpha_2 = 0$. For $\mathcal{H}_1^\dagger(1)$ $d' = (-2/3, -1/3, 1)'$ is fixed. LR is the likelihood ratio.

Model	2 log $L_{\max}$(Model)	$-2\log LR$	df	p-value	$\hat{\tau}$	$\tilde{\tau}$
$\mathcal{H}(1)$	2536.72	-	-	-	-	-
$\mathcal{H}_1(1)$	2534.06	2.66	1	0.10	-	-
$\mathcal{H}_1^\dagger(1)$	2531.03	3.03	6	0.81	-0.048	-0.051

and in the other both unit labour cost and unit import cost are similarly treated. The maxima are 2531.06 and 2526.16 respectively corresponding to $\hat{\tau} = 0.65$ and $\hat{\tau} = 0.58$. The test statistics for testing weak exogeneity and the rational expectation hypothesis simultaneously agaist $\mathcal{H}(1)$ are therefore 5.66 and 10.56 with 6 and 7 degrees of freedom, which correspond to p-values 0.46 and 0.16 respectively. For the case where only uic_t is weakly exogenous in addition to satisfying the rational expectation hypothesis a 95% confidence interval for δ is (0.48,0.81).

Remark 5.1. The model fitted in this application is more general than model (2.1). The linear rational expectation relation (2.2) will for the more general model also imply restrictions on the coefficients of the deterministic variables. Regressing on these variables as we did means that the additional restriction are not taken into account. Hence, the test is for a more general model than the one described by the linear rational expectation relation. It only pertains to restrictions on the coefficients on the stochastic variables.

6. CONCLUSION

The theme of this paper has been analyzing cointegrated vector autoregressive models with restrictions on the error correction parameters and in addition restrictions from exact rational expectations imposed. We considered estimation and testing in such models where the error correction parameters satisfied the same restrictions, i.e. $\alpha = A\psi$, and also for some special cases the situation where some of the error correction parameters were known.

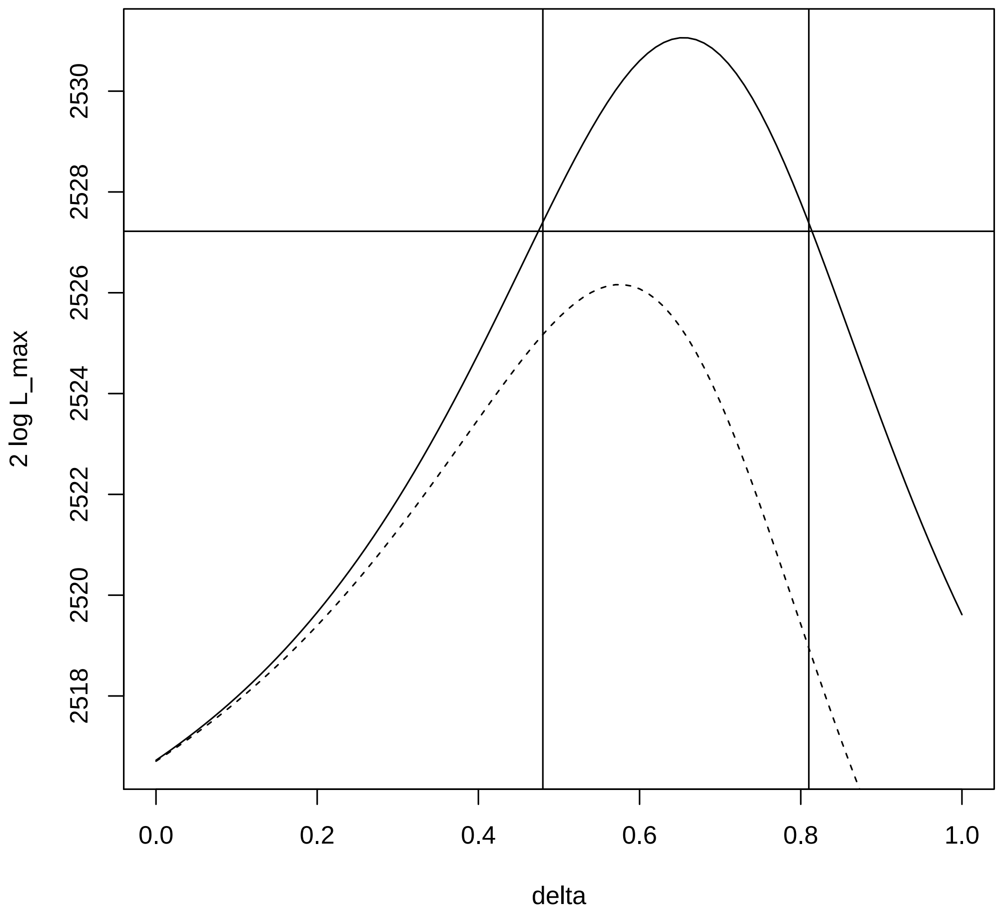

Figure 1. Profile likelihood for unit import cost weakly exogenous in addition to satisfying the rational expectation hypothesis (solid) and for unit labour cost and unit import cost weakly exogenous in addition to satisfying (2.5) (dashed). A 95% confidence interval is indicated in the former case.

ACKNOWLEDGMENTS

Søren Johansen acknowledges support from Center for Research in Econometric Analysis of Time Series, CREATES, funded by the Danish National Research Foundation.

REFERENCES

Anderson, T.W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Annals of Mathematical Statistics 22*, 327-351.

Aukrust, O. (1977). Inflation in the Open Economy: A Norwegian Model. In L. B. Krause and W. S. Salânt (Eds.). *Worldwide Inflation*, Washington DC: Brookings.

Boug, P., Å. Cappelen and A.R. Swensen (2010). The new Keynesian Phillips curve revisited. *Journal of Economic Dynamics and Control 34,* 858-874.

Boug, P., Å. Cappelen and A.R. Swensen (2017). Inflation dynamics in a small open economy. *Scandinavian Journal of Economics 119,* 1010-1039.

Campbell, J.Y. (1987). Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis. *Econometrica 55*, 1249-1273.

Campbell, J.Y. and R.J. Shiller (1987). Cointegration tests and present value models. *Journal of Political Economy 95*, 1062-1088.

Drton, M. and T.S. Richardson (2004). Multimodality of the likelihood in the bivariate seemingjly unrelated regressions model. *Biometrika* **91**, 383-392.

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

Hansen, L.P. and T.J. Sargent (1981). Exact linear rational expectations models: Specification and estimation. Research Department Staff Report 71, Federal Reserve Bank of Minneapolis.

Hansen, L.P. and T.J. Sargent (1982). Instrumental variables procedures for estimating rational expectations models. *Journal of Monetary Econonomics 9*, 263-296.

Hansen, L.P. and T.J. Sargent (1991). Exact linear rational expectations models: Specification and estimation. In L.P. Hansen and T.J. Sargent (Eds.) *Rational Expectations Econometrics.* Boulder, CO: Westview Press,

Horn and Johnson (2013). *Matrix Analysis* (2nd. ed.). Cambridge: Cambridge University Press,

Johansen, S. (1996). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford: Oxford University Press,

Johansen, S. and A.R. Swensen (1999). Testing exact rational expectations in cointegrated vector autoregressive models. *Journal of Econometrics 93*, 73-91.

Johansen, S. and A.R. Swensen (2004). More on testing exact rational expectations in cointegrated vector autoregressive models: Restricted constant and linear term. *Econometrics Journal 7*, 389–397.

Johansen, S. and A.R. Swensen (2008). Exact rational expectations, cointegration, and reduced rank regression. *Journal of Statistical Planning and Inference 138*, 2738-2748.

Muth, J.F. (1961). Rational expectations and the theory of price movements. *Econometrica 29*, 315-335.

Nocedal, J. and and S. J. Wright (2006). *Numerical Optimization,* (2nd. ed.). New York, NY:Wiley.

Oberhofer, W. and J. Kmenta (1974). A general procedure for obtaining maximum likelihood estimates in generalized regression models. *Econometrica 42*, 579-590.

R Core Team (2020). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna: R Foundation for Statistical Computing. URL https://www.R-project.org/.

Swensen, A.R. (2014). Some exact and inexact linear rational expectations models in vector autoregressive models. *Economics Letters 123,* 216-219.

APPENDIX A: DETAILS OF $RANK(A'\bar{C}_\perp) < R - N$ OF SUBSECTION 3.2

The model, taking the restrictions into account, can be written

$$\bar{c}_\perp' \Delta X_t = vu'\psi\beta' X_{t-1} + \sum_{i=1}^{k} \bar{c}_\perp' \Gamma_i \Delta X_{t-i} + \bar{c}_\perp' \mu + \bar{c}_\perp' \varepsilon_t,$$

$$c' \Delta X_t = \tau d' X_{t-1} + \sum_{i=1}^{\ell} \tau_i d_i' \Delta X_{t-i} + c'\mu + c'\varepsilon_t.$$

The conditional equation, conditioning on $c'\Delta X_t$ and the past, of $\bar{c}_\perp' \Delta X_t$ and the marginal equation of $c'\Delta X_t$ are

$$\begin{aligned} \bar{c}_\perp' \Delta X_t = {} & vu'\psi\beta' X_{t-1} + \sum_{i=1}^{k} \bar{c}_\perp' \Gamma_i \Delta X_{t-i} + \bar{c}_\perp' \mu \\ & - \omega_{\bar{c}_\perp \cdot c}(c'\Delta X_t - \tau d' X_{t-1} - \sum_{i=1}^{\ell} \tau_i d_i' \Delta X_{t-i} - c'\mu) + \bar{c}_\perp' \varepsilon_t - \omega_{\bar{c}_\perp \cdot c} c'\varepsilon_t \end{aligned} \tag{A.1}$$

$$c' \Delta X_t = \tau d' X_{t-1} + \sum_{i=1}^{\ell} \tau_i d_i' \Delta X_{t-i} + c'\mu + c'\varepsilon_t. \tag{A.2}$$

where $\omega_{\bar{c}_\perp \cdot c} = E[\bar{c}_\perp' \varepsilon_t \varepsilon_t' c](E[c'\varepsilon_t \varepsilon_t' c])^{-1}$ and $\omega_{cc} = E[c'\varepsilon_t \varepsilon_t' c]$.

The coefficient matrix $vu'\psi\beta'$ in equation (A.1) has dimension $(p-q) \times p$ and rank o. Since $o = s - q \leq p - q$, the equation (A.1) can be estimated by a combination of reduced rank and ordinary OLS regressions. The parameters in equation (A.2) can be estimated by OLS regressions.

APPENDIX B: CASE B OF SUBSECTION 4.1

When $0 < m = r$, $\alpha = a$ all adjustment parameters are known. The transformed model with the restrictions incorporated is

$$\bar{c}_\perp' \Delta X_t = \bar{c}_\perp' a\beta' X_{t-1} + \sum_{i=1}^{k} \bar{c}_\perp' \Gamma_i \Delta X_{t-i} + \bar{c}_\perp' \mu + \bar{c}_\perp' \varepsilon_t,$$

$$c' \Delta X_t = \tau d' X_{t-1} + \sum_{i=1}^{\ell} \tau_i d_i' \Delta X_{t-i} + c'\mu + c'\varepsilon_t.$$

The only parameters of the matrix Π are the elements of β. When $r > rank(\bar{c}_\perp' a)$ the model is not identified since the elements of β cannot be distinguished. The possibility $r < rank(\bar{c}_\perp' a)$ is impossible since $rank(\bar{c}_\perp' a) \leq rank(a) = m = r$. Therefore, for identified models $v_2 = \bar{c}_\perp' a$ has dimensions $(p-q) \times r$ and full rank r. Note also that $r = m = rank(\bar{c}_\perp' a) \leq \min(rank(\bar{c}_\perp'), rank(a)) = \min(p-q, r) \leq p - q$.

The restricted model, by multiplying the equation for $\bar{c}_\perp' \Delta X_t$ with $(\bar{v}_2, v_{2\perp})'$, can be decomposed into three parts

$$\bar{v}_2'\bar{c}_\perp' \Delta X_t = \beta' X_{t-1} + \sum_{i=1}^{k} \bar{v}_2'\bar{c}_\perp' \Gamma_i \Delta X_{t-i} + \bar{v}_2'\bar{c}_\perp' \mu + \bar{v}_2'\bar{c}_\perp' \varepsilon_t,$$

$$v_{2\perp}'\bar{c}_\perp' \Delta X_t = \sum_{i=1}^{k} v_{2\perp}'\bar{c}_\perp' \Gamma_i \Delta X_{t-i} + v_{2\perp}'\bar{c}_\perp' \mu + v_{2\perp}'\bar{c}_\perp' \varepsilon_t,$$

$$c' \Delta X_t = \tau d' X_{t-1} + \sum_{i=1}^{\ell} \tau_i d_i' \Delta X_{t-i} + c'\mu + c'\varepsilon_t.$$

Define the parameters as in Section 3.2 and premultiply $(\beta', 0, d\tau')'$ with the matrix K, see (3.11), to get

$$K \begin{pmatrix} \beta \\ 0 \\ \tau d' \end{pmatrix} = \begin{pmatrix} \beta' - \omega_{13.2}\tau d' \\ -\omega_{2.3}\tau d' \\ \tau d' \end{pmatrix}.$$

The conditional equations of $(\Delta X_t'\bar{c}_\perp \bar{v}_2, \Delta X_t'\bar{c}_\perp v_{2\perp}, \Delta X_t' c)'$ are therefore given by

$$\begin{pmatrix} \bar{v}_2'\bar{c}_\perp' \Delta X_t \\ v_{2\perp}'\bar{c}_\perp' \Delta X_t \\ c'\Delta X_t \end{pmatrix} = \begin{pmatrix} \omega_{12.3} & \omega_{13.2} & \beta' - \omega_{13.2}\tau d' & \Gamma_{11}^{***} & \Gamma_{12}^{***} & \mu_1^{**} \\ 0 & \omega_{2.3} & -\omega_{2.3}\tau d' & \Gamma_{21}^{***} & \Gamma_{22}^{***} & \mu_2^{**} \\ 0 & 0 & \tau d' & \Gamma_{31}^{***} & 0 & \mu_3^{*} \end{pmatrix} Z_t^{**} + \begin{pmatrix} \varepsilon_{1t}^{**} \\ \varepsilon_{2t}^{**} \\ \varepsilon_{3t}^{*} \end{pmatrix},$$

where now

$$Z_t^{**\prime} = \left(\Delta X_{2t}^{*\prime}, \Delta X_{3t}^{*\prime}, X_{t-1}', \Delta X_{t-1}^{*\prime}, \ldots, \Delta X_{t-k}^{*\prime}, 1 \right).$$

The parameters in the equation for $\bar{v}_2'\bar{c}_\perp' \Delta X_t$ using $\beta' - \omega_{13.2}\tau d'$, are variation independent of the parameters in the equations for $v_{2\perp}'\bar{c}_\perp' \Delta X_t$ and $c'\Delta X_t$. The coefficient $-\omega_{2.3}\tau d'$, however, represents a cross equation restriction as a product of $\omega_{2.3}$ from the equation for $v_{2\perp}'\bar{c}_\perp' \Delta X_t$, and τ from the equation for $c'\Delta X_t$. Arguing as in Section 3.2 one can first assume that τ is known, introduce the variable $c'\Delta X_t - \tau d' X_{t-1}$, estimate by ordinary least squares and finally optimize over τ.

An alternative procedure is also in this case to estimate τ using the SUR procedure of Remark 3.6.