Abstract:
The parameters in the cointegration vector and the loading parameters are not the only interesting parameters in a vector cointegration model. With a reformulation of the model the intercept parameters can be decomposed into growth parameters and cointegration mean parameters. These parameters have economic interpretations and are therefore also important. We show how these parameters can be estimated and restricted. The latter can be achieved by using a linear switching algorithm. Consumption and money demand applications illustrate the method.

Keywords: Johansen procedure, cointegrated VAR, growth rates, cointegration means, linear switching algorithm, consumption, money demand, savings ratio.

JEL classification: C32, C51, C52, E21, E41.

Acknowledgement: Thanks to Gunnar Bårdsen, Søren Johansen, Bjørn Naug, Terje Skjerpen and Anders Rygh Swensen for valuable comments on various versions of the paper. Earlier versions of the paper have been presented at conferences in Norges Bank, Oslo (March 21-22, 2001) and at the European University Institute, Florence (June 6-9, 2001).

Address: Håvard Hungnes, Statistics Norway, Research Department, P. O. Box 8131 Dep., N-0033 Oslo, Norway. E-mail: havard.hungnes@ssb.no
Discussion Papers comprise research papers intended for international journals or books. As a pre-print a Discussion Paper can be longer and more elaborate than a standard journal article by including intermediate calculation and background material etc.

Abstracts with downloadable PDF files of Discussion Papers are available on the Internet: http://www.ssb.no

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
N-2225 Kongsvinger

Telephone: +47 62 88 55 00
Telefax: +47 62 88 55 95
E-mail: Salg-abonnement@ssb.no
1 Introduction

Cointegrated vector autoregressive (VAR) models are a powerful tool in analysing time series. Granger’s representation theorem (see Engle and Granger, 1987) shows that cointegrated time series can be represented in an equilibrium correction vector autoregressive model. Furthermore, Johansen (1988) shows that canonical correlation technique combined with reduced rank regression technique can be used to estimate such models. These techniques are implemented in standard time series packages such as PcGive (see Doornik and Hendry, 2001) and Cats in Rats (see Hansen and Juselius, 1994). A lot of work has been accomplished in estimating long-run cointegration relationships in economics. The cointegration vectors can be tested against economic theory.

However, other parameters in a cointegrated VAR model also have economic interpretations. By rewriting the equilibrium correction form of the VAR model (VEqCM), we can identify the underlying growth of the variables as well as the long-run means of the cointegration relationships.

Within the VEqCM the intercepts can either be restricted to lie in the cointegration space, or not. If the intercepts are not restricted to lie in the cointegration space (‘unrestricted’), they allow the system to have both growth and cointegration means. If, however, the intercepts are restricted, there is no growth in the system, (see Johansen and Juselius, 1990).

The growth rates tell us how much to expect (unconditionally) the variables in the system to grow from one period to the next. If the system is used for forecasting, the vector of growth rates will be one of the most important ones in providing good forecasts. In fact, as the forecast horizon approaches infinity, the forecasts will rely on this vector only, see Clements and Hendry (1999, p. 49-51).

There are also variables we do not believe will grow over time. If the interest rate or the inflation rate is assumed to be $I(1)$, we may not want to allow them to grow. Especially not if we want to use the system for forecasting. However, restricting the intercepts to lie in the cointegration space may be too restrictive, as the system may include variables we believe do grow over time. We then want to restrict some of the variables to have no growth and let other variables in the system grow. We develop an estimation procedure in which we allow restrictions in the system on some or all of the growth rates.

The cointegration means may also have economic interpretations. In a system with (the logs of) consumption and income, the intercept in the cointegration vector can be interpreted as the equilibrium savings ratio (see example 1 below) if the income elasticity is unity. A system with nominal interest rate and inflation (both assumed to be $I(1)$), where the cointegration mean can be interpreted as the equilibrium real interest rate, is another example.

Sometimes we may want to restrict the cointegration mean. Assume we are testing the law of one price, and are analysing a system with an unrestricted intercept to
allow the prices to grow over time.\footnote{The law of one price states that one product shall have the same price in two different regions. Let $P$ be the price of the product in one of the regions and $P^*$ the price in the other region. Furthermore, let $E$ be the exchange rate (if the two regions lie in two different countries). Then the law of one price states that $P = A \cdot E \cdot P^*$, where $A$ is a constant capturing differences in the price level due to transportation costs etc. The strict version of the law of one price states that $A = 1$, i.e. there are no differences in the prices of the product in the two regions.} We may find that $p - e - p^*$ is the cointegration relationship, (where $p$ and $p^*$ are the domestic and foreign price respectively, and $e$ the exchange rate, all variables measured in logs,) and want to test the strict version of the law of one price. This implies testing if the cointegration mean is equal to zero. To achieve this, we have to decompose the intercepts in the system in growth rates and a cointegration mean, and test if the mean is equal to zero. This can also be achieved by the estimation procedure presented here.

The paper is organized as follows: In section 2 we show how the growth rates and cointegration means can be estimated, illustrating this with an example for the private savings ratio in Norway. In section 3 a linear switching algorithm is presented, together with an illustration on money demand in Denmark. The switching algorithm we derive here is an extension of the linear switching algorithm in Boswijk (1995).

\section{Growth rates and cointegration means}

In this section we look at some properties of the cointegrated VAR model. In particular, we focus on how the growth rates and cointegration means can be estimated. The method is illustrated by an analysis of income and consumption.

\subsection{Granger’s representation theorem}

In (1) $X_t$ is an $n$-dimensional vector of non-stationary $I(1)$ variables, $\delta$ is a vector of intercepts, $\alpha$ and $\beta$ are matrices of dimension $n \times r$ (where $r$ is the number of cointegration vectors) and $\beta'X_t$ is $I(0)$. Furthermore, $\Gamma_i$ is a $n \times n$ matrix of coefficients and $\Delta$ is the difference operator. $CS_t$ is a vector of centred seasonal dummies. The residual $\varepsilon$ is assumed to be white noise Gaussian ($\varepsilon_t \sim N(0, \Omega)$).

\begin{equation}
\Delta X_t = \delta + \alpha \beta'X_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + sCS_t + \varepsilon_t, \quad t = 1, 2, ..., T
\end{equation}

The system grows at the unconditional rates $E[\Delta X] = \gamma$ with long run (cointegration) means $E[\beta'X] = \mu$ apart from terms involving seasonal dummies. Then we can rewrite the relationship as

\begin{equation}
\Delta X_t - \gamma = \alpha (\beta'X_{t-1} - \mu) + \sum_{i=1}^{p-1} \Gamma_i (\Delta X_{t-i} - \gamma) + sCS_t + \varepsilon_t.
\end{equation}
Notice that $\beta' \gamma = 0$. To prove this, we premultiply the left hand side of (2) with $\beta'$.

$$\beta' (\Delta X_t - \gamma) = (\beta' X_t - \beta' X_{t-1}) - \beta' \gamma$$

We see that since $E[\Delta X] = \gamma$ the unconditional expectation of the left hand side of (3) equals zero. And since $E[\beta' X] = \mu$ the unconditional expectation of the right hand side of (3) will equal zero only if $\beta' \gamma = 0$.

We define a matrix $\beta_\perp$ with columns orthogonal to the columns in $\beta$. The matrix $\beta_\perp$ is a $n \times (n - r)$ matrix of full column rank, such that $\beta \beta_\perp = 0$. If we can write $\beta = (b_1, b_2)'$, where $b_1$ is $r \times r$ of full rank, then we can take $\beta_\perp = (-b_2 b_1^{-1}, I_{n-r})'$ (see Johansen (1995, p. 48) or Johansen and Swensen (1999, p. 80)).

Another way to write the relationship between the cointegration vectors and the growth rates are $\gamma = \beta_\perp \psi$, where $\psi$ is a vector with $n - r$ elements. To see this replace $\gamma$ in (3) with $H \psi$, where $H$ is an $n \times (n - r)$ matrix. Then $\beta' H \psi = 0$. However, $H \psi = \gamma$ and is generally not equal to zero. The only way $\beta' H \psi = 0$ is therefore when $\beta' H = 0$, and this holds when $H = \beta_\perp$.

For given values of $\alpha$ and $\beta$ we can now estimate $\gamma$ and $\mu$. By comparing equation (1) and (2) we see that

$$\delta = \Gamma \gamma - \alpha \mu,$$

where $\Gamma = (I - \sum_{i=1}^{p-1} \Gamma_i)$. After some rearrangement we find

$$\gamma = C \delta,$$

and

$$\mu = \alpha (\Gamma \gamma - \delta),$$

where $C = \beta_\perp (\alpha_\perp' \Gamma \beta_\perp)^{-1} \alpha_\perp'$ and $\alpha = (\alpha' \alpha)^{-1}$. These properties are known from Granger’s representation theorem, see Engle and Granger (1987) and Johansen (1991).

### 2.2 Example 1: Norwegian consumption

We use quarterly data for households’ total income and total consumption, measured on the natural logarithmic scale. The estimation period is 1991Q2-2000Q4. From figure 1 we see that both income ($y$) and consumption ($c$) fluctuate over the quarters. Centred seasonal dummies are therefore included in the empirical analysis.

---

2This expression will only be valid if $\text{Rank}(b_1) = r$. However, since $\text{Rank}(\beta) = r$, it is always possible to order the variables in the system such that $\text{Rank}(b_1) = r$.

3To find (5), premultiply (4) by $\alpha_\perp'$; $\alpha_\perp' \delta = (\alpha_\perp' \Gamma \beta_\perp \psi)$. The $n - r$ quadratic matrix $(\alpha_\perp' \Gamma \beta_\perp \psi)$ must have full rank, or some of the variables in the system are $I(2)$. Therefore, $\psi = (\alpha_\perp' \Gamma \beta_\perp)^{-1} \alpha_\perp' \delta$, or $\gamma = \beta_\perp \psi = \beta_\perp (\alpha_\perp' \Gamma \beta_\perp)^{-1} \alpha_\perp' \delta$. To find (6), rearrange (4) to $\alpha \mu = \Gamma \beta \gamma - \delta$ and premultiply with $\alpha'$.  

4The results are obtained by combining PcFiml 9.2 (see Doornik and Hendry, 1997) and Ox 2.1 (see Doornik, 1996).
In the upper part logs of real income and real consumption are plotted. The lower graph plots quarterly growth rates in real income and consumption. Data for 1999 and 2000 are preliminary. Source: Statistics Norway, National Accounts.

In the analysis we use 2 lags ($p = 2$). The cointegration rank tests yield the test statistics reported in table 1. Both the $\lambda$–max and the trace statistic support one cointegration vector at a 1 per cent level. The Reimers (1992) adjusted versions of the tests also indicate one cointegration vector.

Estimating the system under the restriction of one cointegration vector, and utilizing the results from Section 2.1 to calculate the growth rates and cointegration mean yields

\[
\begin{pmatrix}
\Delta y_t - 0.0078 \\
\Delta c_t - 0.0076
\end{pmatrix}
= 
\begin{pmatrix}
-0.464 \\
0.601
\end{pmatrix}
\begin{pmatrix}
y_{t-1} \cdot 1.023 \cdot c_{t-1} + 0.206 \\
0.166 \\
0.192
\end{pmatrix}
+ \Gamma_1 \begin{pmatrix}
\Delta y_{t-1} - 0.0078 \\
\Delta c_{t-1} - 0.0076
\end{pmatrix}
+ \delta C S_t + \varepsilon_t.
\]

(7)

From (7) we see that both income and consumption grow at a rate of about 0.8
Table 1: Consumption: Cointegration rank

<table>
<thead>
<tr>
<th>$H_0 : rank = r$</th>
<th>$\lambda$</th>
<th>$\lambda$ - max</th>
<th>95%</th>
<th>trace</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 0$</td>
<td>0.4776</td>
<td>25.32**</td>
<td>14.1</td>
<td>25.88**</td>
<td>15.4</td>
</tr>
<tr>
<td>$r \leq 1$</td>
<td>0.0142</td>
<td>0.56</td>
<td>3.8</td>
<td>0.56</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Two asterisks denotes significance at the 1 per cent level. The reported critical values are taken from Osterwald-Lenum (1992).

Table 2: Consumption: Likelihood ratio test of reductions

<table>
<thead>
<tr>
<th>Equation</th>
<th>log $L$</th>
<th>p-value</th>
<th>[d.f.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>339.09</td>
<td>0.36</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>338.66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

per cent each quarter. This corresponds to an annual growth rate of about 3.1 per cent.\(^5\) However, consumption and income do not grow at exactly the same rate, since the cointegration vector $\beta' \neq (1,-1)$.

In (7), the cointegration mean has no economic interpretation alone. However, we can interpret $0.023 \cdot c_{t-1} - 0.206$ as the long-run savings ratio. We see that the equilibrium savings ratio increases with the consumption level. On the other hand, the coefficient for consumption in the cointegration vector differs from unity with less than one standard deviation. We therefore restrict the cointegration vector to be $\beta' = (1,-1)$.

$$\begin{pmatrix}
\Delta y_t - 0.0077 \\
\Delta c_t - 0.0077
\end{pmatrix}
\begin{pmatrix}
-0.462 \\
0.563
\end{pmatrix}
\begin{pmatrix}
y_t - c_t - 0.0628 \\
y_t - c_t - 0.0628
\end{pmatrix}
\begin{pmatrix}
\Delta y_{t-i} - 0.0077 \\
\Delta c_{t-i} - 0.0077
\end{pmatrix}
+ \hat{\Gamma}_1
+ \tilde{\epsilon}_t,$$

(8)

When restricting the savings ratio to be stationary the growth rates become 0.008, which corresponds to an annual growth rate of 3.1 per cent.\(^5\) The estimate of the equilibrium savings ratio is 6.3 per cent.\(^6\)

The savings ratio is plotted in figure 2. The average savings ratio in the estimation period is 6.1 per cent, which is a bit smaller than the savings ratio in (8). The reason is that the latter is adjusted for system dynamics. If we had used (8) to forecast, the savings ratio would converge to 6.3 per cent (apart from seasonality) as the forecasting horizon approach infinity.

An underlying growth rate of 3.1 per cent is high. Our data sample covers a period with a boom in the economy. In 1991 (the beginning of the sample) Norway

\(^5\)The annual growth rate $g$ is calculated by $g = \exp(4 \cdot \gamma) - 1$.

\(^6\)More precisely, the savings ratio is defined as $s_t = (Y_t - C_t)/Y_t$. By using this definition $s = \exp(\mu) - 1$, which in the last estimation yields a saving ratio of 6.5 per cent.
was in a recession with a high unemployment rate. In 2000 Norway might have been on the top of a business cycle. The estimated trend is therefore probably cutting through the true underlying trend from below. Forecasts of consumption and income based on this model would probably be too high. If we had estimated $\alpha$ and $\beta$ only, we would not discovered this important source to forecasting failure.

Table 2 reports the log likelihood value of the different models together with likelihood ratio tests which are $\chi^2$-distributed asymptotically. Let $H_0$ be the null hypothesis and $H_A$ the alternative, where the latter is the model where only the cointegration rank is restricted (here equation (7)). The likelihood ratio value can be calculated by

$$2 \left[ \log L (H_A) - \log L (H_0) \right] \overset{\alpha}{\sim} \chi^2 (d.f),$$

where $d.f.$ is the degrees of freedom. From the table we see that the hypothesis of a stationary savings ratio is not rejected.

Figure 2: Savings Ratio, both seasonally adjusted and unadjusted. Seasonal adjustment is based on a static regression of the savings ratio on an intercept and centred seasonal dummies in the estimation period.
3 Testing restrictions on $\gamma$ and $\mu$

When no restrictions are imposed on the growth rates $\gamma$ and the cointegration means $\mu$, the vector of intercepts $\delta$ in (1) have $n$ variation free elements, and these could be estimated with ordinary least square. However, when we restrict $\gamma$ and/or on $\mu$, another estimation procedure must be used. Since we are imposing restrictions on the growth parameters and cointegration means as a part of the cointegration analysis, our estimation procedure must also allow for restrictions on $\alpha$ and $\beta$. We extend the linear switching algorithm in Boswijk (1995) to also involve restrictions on the growth rates. We define $\beta^* = (\beta', -\mu)'$ and $X_t^* = (X_t', 1)'$, so restrictions on the cointegration means can be imposed on $\beta^*$.

3.1 The maximum likelihood problem with restrictions

The algorithm in Boswijk (1995) allows for linear restrictions on $\alpha$ and $\beta$. The restrictions on the cointegration vectors can be written as $R_{\beta} \text{vec} \beta^* = c_{\beta}$ or

$$\text{vec} \beta^* = H_{\beta} \phi + h_{\beta},$$

(9)

where $H_{\beta} = (R_{\beta})_{\perp}$ and $h_{\beta} = R_{\beta} c_{\beta}$. Since we are stacking the cointegration vectors into one vector, we can allow for restrictions between the cointegration vectors as well as within them.

Similarly, restrictions on the loading parameters can be written $R_{\alpha} \text{vec} \alpha' = 0$ or

$$\text{vec} \alpha' = H_{\alpha} \phi,$$

(10)

where $H_{\alpha} = (R_{\alpha})_{\perp}$. Here the intercepts are excluded, since we normally only test exclusion restrictions on $\alpha$.\footnote{It is straightforward to include intercepts in (10).}

The restrictions on $\gamma$ are a bit more complex, since - in addition to the restrictions we want to place on $\gamma$ - the cointegration vector also imposes restrictions on $\gamma$. The restrictions we want to impose on $\gamma$ can be written as $R_{\gamma} \gamma = c_{\gamma}$, whereas the restrictions imposed by the cointegration vectors can be expressed as $\beta \gamma = 0$. In a compact notation, these restrictions involve $(\beta, R_{\gamma})' \gamma = (0', c_{\gamma})'$, which equivalently can be written as

$$\gamma = H_{\gamma} \psi + h_{\gamma},$$

(11)

where $H_{\gamma} = (\beta, R_{\gamma})_{\perp}$ and $h_{\gamma} = R_{\gamma} c_{\gamma}$.

Before we present the log likelihood function, we must define some variables. We first define $Z_t = \text{vec}(\Delta X_t, \Delta X_{t-1}, ..., \Delta X_{t-p+1})$, $\Phi = (I_n, -\Gamma_1, -\Gamma_2, ..., -\Gamma_{p-1})$ and
where $e_{p \times 1} = (1, 1, \ldots, 1)'$. This log likelihood function becomes

$$
\log L(\alpha, \beta^*, \gamma, \Gamma_1, \Gamma_2, \ldots, \Gamma_{p-1}, s, \Omega) = -\frac{T}{2} n \log (2\pi) - \frac{T}{2} \log |\Omega|
$$

$$
-\frac{1}{2} \sum_{t=1}^{T} \left[ (\Phi(Z_t - E\gamma) - \alpha \beta^* X_{t-1}^* - sCS_t)' \Omega^{-1} \right. 
$$

$$
\left. \times (\Phi(Z_t - E\gamma) - \alpha \beta^* X_{t-1}^* - sCS_t) \right].
$$

The maximisation problem is to maximize (12) under the restrictions (9) - (11). In the next subsection we derive a switching algorithm to deal with this maximisation problem.

### 3.2 The linear switching algorithm

It turns out to be convenient also to use a log likelihood function where we condition on the growth rates. For a given set of growth rates satisfying (11) we can define $Z_{0t}^*(\psi) = \Delta X_{t-1} - \gamma$, $Z_{1t}^* = X_{t-1}^*$, $Z_{2t}^*(\psi) = \text{vec}(\Delta X_{t-1} - \gamma, \Delta X_{t-2} - \gamma, \ldots, \Delta X_{t-(p-1)} - \gamma, CS_t)$ and $\Theta = (\Gamma_1, \Gamma_2, \ldots, \Gamma_{p-1}, s)$. The log likelihood function conditioned on the growth rates is

$$
\log L(\alpha, \beta^*, \Gamma_1, \Gamma_2, \ldots, \Gamma_{p-1}, s, \Omega; \gamma)
$$

$$
= -\frac{T}{2} n \log (2\pi) - \frac{T}{2} \log |\Omega|
$$

$$
-\frac{1}{2} \sum_{t=1}^{T} \left[ (Z_{0t}^* - \alpha \beta^* Z_{1t}^* - \Theta Z_{2t}^*)' \Omega^{-1} (Z_{0t}^* - \alpha \beta^* Z_{1t}^* - \Theta Z_{2t}^*) \right].
$$

Furthermore, we define

$$
M_{ij}^*(\psi) = T^{-1} \sum_{t=1}^{T} Z_{it}^* Z_{jt}^*, \quad i, j = 0, 1, 2, \quad (14)
$$

and

$$
S_{ij}^*(\psi) = T^{-1} \sum_{t=1}^{T} R_{it}^* R_{jt}^*, \quad i, j = 0, 1, \quad (15)
$$

where $R_{0t}^*(\psi)$ and $R_{1t}^*(\psi)$ are the residuals we obtain by regressing $Z_{0t}^*(\psi)$ and $Z_{1t}^*$ on $Z_{2t}^*(\psi)$ respectively. Finally, $\otimes$ is the Kronecker product.

**Theorem 1 (The conditional maximum likelihood estimators)** The conditional maximum likelihood estimators for $\Theta$, $\phi$, $\varphi$, $\Omega$ and $\psi$ in (12) under the restrictions (9) - (11) are given by
\[ \hat{\Theta}(\psi, \phi, \varphi) = M_{02}^* (M_{22}^*)^{-1} - \alpha \beta s'M_{12}^* (M_{22}^*)^{-1}, \]  
\[ \hat{\phi}(\psi, \varphi, \Omega) = \left[ H'_\beta \left( \alpha'\Omega^{-1} \alpha \otimes S_{11}^* \right) H_\beta \right]^{-1} \times H'_\beta \left[ \left( \alpha'\Omega^{-1} \alpha \otimes I_{n+1} \right) vec S_{10}^* - \left( \alpha'\Omega^{-1} \alpha \otimes S_{11}^* \right) \right], \]  
\[ \hat{\varphi}(\psi, \phi, \Omega) = \left[ H'_\alpha \left( \Omega^{-1} \otimes \beta s'S_{11}^* \right) H_\alpha \right]^{-1} \left[ H'_\alpha \left( \Omega^{-1} \otimes \beta s' \right) vec S_{10}^* \right], \]  
\[ \hat{\Omega}(\psi, \phi, \varphi) = S_{00}^* - \alpha \beta s'S_{10}^* - S_{01}^* \beta s' + \alpha \beta s'S_{11}^* \beta s', \]  
\[ \hat{\psi}(\phi, \varphi, \Theta, \Omega) = \left[ H'_\psi E' \Phi \Omega^{-1} \Phi EH_\gamma \right]^{-1} \times \left[ H'_\psi E' \Phi \Omega^{-1} \left( \Phi \bar{Z} - \alpha \beta s' \bar{X} - sCS - \Phi Eh_\gamma \right) \right], \]  
where \( \bar{Z} = T^{-1} \sum_{t=1}^T Z_t, \bar{X} = T^{-1} \sum X_{t-1}^* \) and \( CS = T^{-1} \sum CS_t \).

See the appendix for the proof.

The term \( sCS \) in (20) equals to zero if we have the same number of observations for each season in the calendar year. If, however, we have an estimation period with more observations from some seasons than from others, this term will generally not equal zero. In the example below the estimation period is 1974Q3-1987Q3, which means that we have one more observation from the third quarter than the others.

Note that \( \beta s' \bar{X} = \beta \bar{X} - \mu \) in (20). The first part is the average cointegration mean in the estimation period, and \( \mu \) is the system cointegration mean. These will not generally be equal, as we saw in the example above.

We now suggest the following estimation procedure:

The maximum likelihood estimators of \( \psi, \phi, \varphi, \Theta, \) and \( \Omega \) may be obtained by the following iterative procedure, starting from a set of initial values \( \{ \psi_0, \phi_0, \varphi_0, \Theta_0, \Omega_0 \} \):

- **I** \( \hat{\psi}_j = \psi \left( \hat{\phi}_{j-1}, \hat{\varphi}_{j-1}, \hat{\Theta}_{j-1}, \hat{\Omega}_{j-1} \right) \)
- **II** \( \hat{\phi}_j = \phi \left( \hat{\psi}_j, \hat{\varphi}_{j-1}, \hat{\Omega}_{j-1} \right) \)
- **III** \( \hat{\varphi}_j = \varphi \left( \hat{\psi}_j, \hat{\phi}_j, \hat{\Omega}_{j-1} \right) \)
- **IV** \( \hat{\Theta}_j = \Theta \left( \hat{\psi}_j, \hat{\varphi}_j, \hat{\varphi}_j \right) \)
- **V** \( \hat{\Omega}_j = \Omega \left( \hat{\psi}_j, \hat{\phi}_j, \hat{\varphi}_j \right) \)

The iterative procedure needs a set of starting values. In fact, it only needs starting values for the free growth rates parameters (\( \psi \)), the cointegration vectors (\( \phi \)) and the loading parameters (\( \varphi \)) since starting values for the other parameters (\( \Theta \) and \( \Omega \)) can be calculated by (16) and (19).

It may be tempting to use the relations in theorem 1 with unrestricted parameters to compute starting values for \( \psi, \phi \) and \( \varphi \) too. However, this is not a good idea when there are more than one cointegration vector. The unrestricted estimator of
\( \beta^* \) is only unique up to a rotation which spans the same space. When restrictions are imposed on \( \beta \) these restrictions may lead to a rotation of this space. To take account of this, we use the method described in Doornik (1995).

Let
\[
\text{vec} \hat{\beta}^* = H_{\beta} \hat{\phi} + h_{\beta} = H_{\beta} \left( H_{\beta} \left( \beta_{\text{unr}}^* - h_{\beta} \right) \right) + h_{\beta},
\]
where \( H_{\beta} = H_{\beta} (H_{\beta}^T H_{\beta})^{-1} \) and the subscript \( \text{unr} \) indicates the parameters are revealed by the unrestricted cointegrated VAR model. Define \([ \cdot ]\) as dropping those rows which have no restrictions in them; if this yields less than \( r \) rows, then add rows back in, so that the \([ \cdot ]\) matrix is \( q \times r \), with \( q \geq r \). Then the least square estimator
\[
\hat{A} = \left( \left[ \beta_{\text{unr}}^* \right]' \left[ \beta_{\text{unr}}^* \right] \right)^{-1} \left( \left[ \beta_{\text{unr}}^* \right]' \left[ \beta_{\text{unr}}^* \right] \right)
\]
is used to derive
\[
\hat{\alpha}_{-1} = \alpha_{\text{unr}} \cdot \hat{A}^{-1}.
\]

Now the loading matrix \( \alpha \) is consistent with the restricted \( \beta \), and we can use the relations in theorem 1 to calculate starting values for \( \psi, \phi \) and \( \varphi \):
\[
\hat{\phi}_0 = \phi \left( \hat{\gamma}_{\text{unr}}, \hat{\alpha}_{-1}, \hat{\Omega}_{\text{unr}} \right),
\]
\[
\hat{\varphi}_0 = \varphi \left( \hat{\gamma}_{\text{unr}}, \hat{\phi}_0, \hat{\Omega}_{\text{unr}} \right),
\]
\[
\hat{\psi}_0 = \psi \left( \hat{\phi}_0, \hat{\varphi}_0, \hat{\Theta}_{\text{unr}}, \hat{\Omega}_{\text{unr}} \right).
\]

As discussed in Johansen (1991), the distribution of \( \hat{\beta} \) is mixed normal, i.e. the variance matrix is stochastic. The discussion there also indicates that inference on \( \beta \) may be done as if \( \alpha \) were known, and vice versa. Following this result, we compute the ‘variance’ of \( \hat{\beta} \) as
\[
V(\text{vec} \hat{\beta}^*) = \frac{T}{T - k} \left( H_{\beta} \left[ TH_{\beta} \left( \alpha' \Omega_{\text{unr}}^{-1} \alpha \otimes S_{11} \right) H_{\beta} \right]^{-1} H_{\beta}' \right),
\]
where the term inside the square brackets is (the negative of) the double derivative of (13) with respect to \( \phi \). The scale factor \( T/(T - k) \) (where \( k \) is the integer part of the ratio between the freely estimated parameters in the system and the number of the dependent variables in the system) is used to control for degrees of freedom, see Doornik (1995).

Similarly, the ‘variance’ of \( \hat{\alpha} \) is
\[
V(\text{vec} \hat{\alpha}') = \frac{T}{T - k} \left( H_{\alpha} \left[ TH_{\alpha}' \left( \Omega^{-1} \otimes \beta' S_{11} \beta \right) H_{\alpha} \right]^{-1} H_{\alpha}' \right).
\]

In (22) we use
\[
S_{11} = T^{-1} \sum_{t=1}^{T} R_{tt} R_{tt}',
\]
Table 3: Money demand: Cointegration rank

<table>
<thead>
<tr>
<th>$H_0: \text{rank} = r$</th>
<th>$\lambda$</th>
<th>$\lambda - \text{max}$</th>
<th>95%</th>
<th>95%†</th>
<th>trace</th>
<th>95%</th>
<th>95%†</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 0$</td>
<td>0.4169</td>
<td>28.59*</td>
<td>27.1</td>
<td>27.1</td>
<td>45.67</td>
<td>47.2</td>
<td>48.3</td>
</tr>
<tr>
<td>$r \leq 1$</td>
<td>0.1776</td>
<td>10.36</td>
<td>21.0</td>
<td>21.1</td>
<td>17.07</td>
<td>29.7</td>
<td>31.5</td>
</tr>
<tr>
<td>$r \leq 2$</td>
<td>0.1125</td>
<td>6.33</td>
<td>14.1</td>
<td>14.9</td>
<td>6.71</td>
<td>15.4</td>
<td>18.0</td>
</tr>
<tr>
<td>$r \leq 3$</td>
<td>0.0072</td>
<td>0.38</td>
<td>3.8</td>
<td>8.2</td>
<td>0.38</td>
<td>3.8</td>
<td>8.2</td>
</tr>
</tbody>
</table>

One asterisk denotes significance at the 5 per cent level. The rows labelled ‘95%’ contain the standard critical values, and in the rows labelled ‘95%†’ the critical values for the case where the true model has no deterministic trends are reported. The critical values are taken from Osterwald-Lenum (1992).

where $R_{11}$ are the residuals we obtain by regressing $Z_{1t} = \{X_{t} - 1, \Delta X_{t-1}, \Delta X_{t-2}, \ldots, \Delta X_{t-p+1}, CS_t\}'$. The reason for using $\beta' S_{11} \beta$ in (22) instead of $\beta' S_{11}^* \beta^*$ (which we would obtain if we used the double derivatives of (13) with respect to $\psi$) is to take account to the covariance between the intercepts in the cointegration relations and the other coefficients outside the cointegration vectors. The expression in (22) is used to compute the ‘variance’ of $\hat{\alpha}$ in the standard literature, see e.g. Johansen (1991).

Since the estimator of $\gamma$ is on expectation orthogonal to the other coefficients (see appendix), we compute this ‘variance’ as

$$V(\gamma) = \frac{T}{T - k} \left( H_\gamma \left[ TH_\gamma' \left( E' \Phi^{-1} \Phi E \right) H_\gamma \right]^{-1} H_\gamma' \right).$$

### 3.3 Example 2: Danish money demand

To illustrate the estimation method we use data for money demand in Denmark. This is the data used by Johansen and Juselius (1990) to illustrate how one can restrict the intercepts to lie in the cointegration space. Restricting the intercepts to lie in the cointegration space implies restricting the variables in the system not to grow over time. This might be realistic for the bond rate ($i^b$) and the deposit rate ($i^d$), but not for (the logs of) real money ($m_2$) and real income ($y$). The data are plotted in figure 3.

Centred seasonal dummies are included in the empirical analysis. We use the same estimation period as Johansen and Juselius (1990): 1974Q3-1987Q3. In the VEqCM 2 lags are included. In contrast to Johansen and Juselius (1990) we include the intercepts unrestricted.

Table 3 includes two columns of critical values for each of the two tests. The first row (labelled 95%) contains the standard critical values in a system with the intercepts unrestricted. In the second row (labelled 95%†) the critical values for the case where the true model has no deterministic trends are reported, see Osterwald-Lenum (1992). The latter set of critical values is reported since we cannot reject the hypothesis that there is no growth in the system.
The rank test indicates that there is one or zero cointegration vectors in the data. The $\lambda$-max test supports one cointegration vector at a five per cent significance level and the trace test supports one cointegration vector at a 10 per cent level (independent of which of the two tables of critical values we use).\footnote{In Johansen and Juselius (1990) the $\lambda$-max test is significant at five per cent. However, their trace test is not significant even at the 10 per cent level (though very close to be so).} We continue the analysis by assuming that there is one cointegration vector among the variables. Estimating the system with one cointegration vector yields the following equilibrium relation:

$$m_2 = 1.04y - 5.22i^b + 4.23i^d - 6.02$$

where $m_2$ is real money, $y$ is real income, $i^b$ is the bond rate, and $i^d$ is the deposit rate.

We follow Johansen and Juselius (1990) by restricting the income elasticity to equal unity and the money demand is homogenous of degree zero in the two interest rates,
i.e.

\[
\beta^* = H_\beta \phi + h_\beta = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.
\]

This yields

\[
\begin{pmatrix} \Delta m2 - 0.0081 \\ \Delta y - 0.0038 \\ \Delta i^b + 0.0012 \\ \Delta i^d + 0.0005 \end{pmatrix}_t = \begin{pmatrix} -0.166 \\ 0.101 \\ 0.016 \\ 0.032 \end{pmatrix} \begin{pmatrix} m2 - y + 5.91 (i^b - i^d) - 6.19 \end{pmatrix}_{t-1} + \hat{\Gamma}_1 \begin{pmatrix} \Delta m2 - 0.0081 \\ \Delta y - 0.0038 \\ \Delta i^b + 0.0012 \\ \Delta i^d + 0.0005 \end{pmatrix}_{t-1} + \hat{CS}_t + \hat{\varepsilon}_t.
\]

The results indicate a positive growth in money and income. In annual terms these growth rates are 3.3 and 1.5 per cent respectively. The results also indicate a negative growth in the interest rates; a 0.5 percentage points annual decrease in the bond rate and a 0.2 percentage points decrease annually in the deposit rate. However, most of the growth rates parameters are insignificant (measured with the t-value).

We now impose the restriction that there is no underlying growth in the two interest rates. These restrictions imply

\[
R'_\gamma = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
\]

In addition we have the restriction \( \beta' \gamma = 0 \). With the restriction we have imposed on the cointegration vector, this restriction involves \( (1, -1, b_1, -b_1) \gamma = 0 \). Therefore, the total set of restrictions on \( \gamma \) can be written as

\[
\begin{pmatrix} \beta' \\ R'_\gamma \end{pmatrix} \gamma = \begin{pmatrix} 1 & -1 & b_1 & -b_1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.
\]
Table 4: Money demand: Likelihood ratio test of reductions

<table>
<thead>
<tr>
<th>Equation</th>
<th>log L</th>
<th>− log</th>
<th>p-value</th>
<th>d.f.</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>970.92</td>
<td>36.6386</td>
<td>0.64</td>
<td>[2]</td>
</tr>
<tr>
<td>25</td>
<td>970.47</td>
<td>36.6214</td>
<td>0.80</td>
<td>[4]</td>
</tr>
<tr>
<td>26</td>
<td>970.08</td>
<td>36.6070</td>
<td>0.32</td>
<td>[6]</td>
</tr>
<tr>
<td>27</td>
<td>967.42</td>
<td>36.5065</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The restrictions can also be expressed as

$$\gamma = H_\gamma \psi = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \psi,$$

where $\psi$ is a scalar.\(^9\)

We see that the restrictions imposed on the growth rates imply that real money and real income grow at the same rate, i.e. $\gamma_1 = \gamma_2$. Imposing the restrictions on the growth rates we get the following results:

$$m2 - y + 5.89 \left( i^b - i^d \right) - 6.21 \left( i^b - i^d \right)_{t-1} + \hat{\Gamma}_1 + \hat{\varepsilon}_t$$

The common estimated growth rate for money and income corresponds to an annual growth rate of 1.6 per cent. From the estimated model we see that the two

\(^9\)In our example $H_\gamma$ is independent of $b$, which means we do not have to update $H_\gamma$ for each iteration. Generally, however, $H_\gamma$ will change when the unrestricted parameters in $\beta$ changes, and $H_\gamma$ must therefore be updated for each iteration.
Table 5: Cointegration coefficient estimates for different restrictions on $\alpha$ and $\gamma$

<table>
<thead>
<tr>
<th>$\gamma'$</th>
<th>$\alpha' = (\ast, \ast, \ast, \ast)$</th>
<th>$\alpha' = (\ast, \ast, 0, 0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\ast, \ast, \ast, \ast)$</td>
<td>$b_1 = 5.907, b_2 = -6.193 \ (0.531, 0.037)$</td>
<td>$b_1 = 5.808, b_2 = -6.189 \ (0.560, 0.038)$</td>
</tr>
<tr>
<td>$(\ast, \ast, 0, 0)$</td>
<td>$b_1 = 5.889, b_2 = -6.209 \ (0.523, 0.037)$</td>
<td>$b_1 = 5.805, b_2 = -6.204 \ (0.559, 0.039)$</td>
</tr>
<tr>
<td>$(0, 0, 0, 0)$</td>
<td>$b_1 = 5.884, b_2 = -6.214 \ (0.523, 0.038)$</td>
<td>$b_1 = 5.811, b_2 = -6.207 \ (0.560, 0.040)$</td>
</tr>
</tbody>
</table>

Asterisk denote that the parameter is unrestricted.

Interest rates may be weakly exogenous. Imposing weak exogeneity yields equation (27).

$$
\begin{pmatrix}
\Delta m_2 - 0.0047 \\ 
\Delta y - 0.0047 \\ 
\Delta i^b \\ 
\Delta i^d
\end{pmatrix}_t =
\begin{pmatrix}
-0.126 \\ 
0.146 \\ 
0 \\ 
0
\end{pmatrix}
\begin{pmatrix}
m_2 - y + 5.80 (i^b - i^d) - 6.20 \\
(0.56) \\ 
(0.04)
\end{pmatrix}_t - 1 + \Gamma_1
\begin{pmatrix}
\Delta m_2 - 0.0047 \\ 
\Delta y - 0.0047 \\ 
\Delta i^b \\ 
\Delta i^d
\end{pmatrix}_t - 1 + \hat{s}CS_t + \hat{e}_t
$$

(27)

Table 4 shows that none of the restrictions imposed are rejected. (The unrestricted system (24) is always the alternative hypothesis.) From equation (27) we see that the growth rates for money and income is probably not significant (t-value of 1.3). Imposing the restriction that there is no growth in the system (the restrictions imposed by Johansen and Juselius, 1990), we get a log likelihood value of 966.56 and a corresponding p-value of 0.27 (with 7 degrees of freedom). We can therefore not reject that all the growth rates equals zero.

In table 5 we see how the estimates of the parameters in the restricted cointegrated vector $\beta' = (1, -1, b_1, -b_1, b_2)$ change with different restrictions on the loading parameters ($\alpha$) and growth rates ($\gamma$). From the table we see that the restrictions on the loading parameters change the estimates of the cointegration vector (and particularly $b_1$) more than restrictions on the growth rates do. There can be two reasons for this result. First, the restrictions on the loading vector are more binding, as can be seen from the relatively large drop in the log likelihood value as the restrictions of weak exogeneity are imposed. More binding restrictions will normally change the other
parameters more. Second, the multiplicative relationship between $\alpha$ and $\beta$ may lead to that restrictions on $\alpha$ will be more important than restrictions on $\gamma$ with respect to the cointegration vector.

4 Conclusions and suggestions for further work

Sometimes it is relevant to estimate and restrict growth rates and cointegration means in VAR models. These parameters may have economic interpretations, and in particular restrictions on the growth rates are interesting to test. We show that this can be achieved by using an iterative procedure.

When restricting growth rates and cointegration means, the degrees of freedom increases. If these restrictions are valid, the estimates of the other parameters in the system will be more precise. On the other hand, the parameters in the cointegration vectors are superconsistent, and the gain may not be large. In the Danish data we see that the estimates hardly changes by including these restrictions. However, this may be so as these restrictions are barely binding. More research will be needed in order to learn how important restrictions on growth rates and cointegration means are for the estimates of the cointegration vectors.

The method presented here can easily be extended to models including more deterministic variables. A deterministic trend is often included in the cointegration vectors. When a trend is included, the growth rates are no longer orthogonal to the cointegration vectors. However, if $\rho$ is the vector of trend coefficients in the cointegration vectors, $\beta'\gamma = \rho$ (or $\gamma = \beta\rho$) will capture the restrictions between the coefficients.

Sometimes we also want to include step dummies in the system. If included, these will pick up changes in the growth rates as well as changes in the cointegration means. However, we may also want to know how the step dummy influences the system: Does the step dummy change the cointegration means, or the growth rates only? Do all growth rates change when the step dummy is included? Such questions can be answered by applying the same method to the step dummy as for the intercepts.

As an example take testing of purchasing power parity: In many countries the inflation rate was higher in the 1980s than in the 1990s. If we test for purchasing power parity, we may include a shift dummy to take account of the shift in the growth rate for prices. However, the shift dummy may pick up shift in the real exchange rate as well as shift in the growth rates. Utilizing the estimation procedure presented here we can test whether the shift in the real exchange rate is significant or not.
5 References


6 Appendix

Proof of Theorem 1. To prove the theorem we use $trAB = trBA = (vecA')'vecB$ and $vec(AXB) = (B' \otimes A)vecX = ((A \otimes B')vecX)'$, where $tr$ is the trace operator.

Deriving equation (16) is straightforward, see e.g. Johansen (1995, p. 90).

Equations (17) and (18) (see Boswijk, 1995 Theorem 2): The derivatives of (13) with respect to $\phi$ and $\varphi$ (under restrictions (10) and (9)) are

\[
\frac{\partial \log L}{\partial \phi} = TH'_{\beta'}vec \left[ S_{10}^* \Omega^{-1} \alpha - S_{11}^* \beta^* \alpha' \Omega^{-1} \alpha \right],
\]

\[
\frac{\partial \log L}{\partial \varphi} = TH'_{\alpha'}vec \left[ \beta'' S_{10}^* \Omega^{-1} - \beta'' S_{11}^* \beta^* \alpha' \Omega^{-1} \right].
\]

Setting (28) equal zero, substituting (9) in (28) and solving for $\phi$ yields (17). Similarly, setting (29) equal to zero and using (29) and (10) leads to (18).

Equation (19): Solving (13) with respect to $\Omega$ and using $\partial \log |\Omega| / \partial \Omega = \Omega^{-1}$ together with (14) and (15) leads to (19).

Equation (20): The derivative of (12) with respect to $\psi$ (under the restriction (11)) is

\[
\frac{\partial \log L}{\partial \psi} = TH'_{\gamma'}vec \left[ (Z' \Phi' - X' \beta^* \alpha' - CS s') - \gamma' E' \Phi' \right] \Omega^{-1} \Phi E
\]

Setting (30) equal to zero and using (11) lead to (20). ■

Proof of the claim that the estimator for $\gamma$ is expectationally orthogonal to the estimators for $\beta$ and $\Theta$. To prove this claim we must prove that the derivatives of (30) with respect to $\varphi$ and $vec \Theta$ has expectation zero. The derivative of (30) with respect to $\varphi$ is

\[
\frac{\partial \log L}{\partial (vec \Theta')' \partial \psi} = -TH'_{\gamma'} \left( E' \Phi' \Omega^{-1} \otimes X' \beta^* \right) H_{\alpha}.
\]

Since $X' \beta^* = (\beta^* X')' = (\beta' X - \mu)'$ has unconditional expectation equal to zero, (31) will have unconditional expectation equal to zero as well.

Since $\Phi = (I_n, -\Theta)$ proving $E \left[ \partial^2 \log L / \partial (vec \Theta')' \partial \psi \right] = 0$ is equivalent to proving $E \left[ \partial^2 \log L / \partial (vec \Theta')' \partial \psi \right] = 0$. From (30) we have

\[
\frac{\partial \log L}{\partial \psi} = TH'_{\gamma'}vec \left\{ \left[ (Z' - \gamma' E') \Phi' - X' \beta^* \alpha' - CS s' \right] \Omega^{-1} \Phi E \right\}.
\]

Note that the expression inside the square brackets is zero. Therefore we get an expression equal to zero when we take the derivative with respect to the last $\Phi$. Note also that the element within the normal parentheses has expectation zero, so the expectation of the derivative with respect to the first $\Phi$ is also equal to zero. ■
Recent publications in the series Discussion Papers


221 J.K. Dagsvik (1998): Choice among Lotteries when Preferences are Stochastic


233 M. Søberg (1998): Uncertainty and International Negotiations on Tradable Quota Treaties

234 J.K. Dagsvik and L. Brubakk: Price Indexes for Elementary Aggregates Derived from Behavioral Assumptions


238 J.K. Dagsvik, A.S. Flaatten and H. Brunborg: A Behavioral Two-Sex Model

239 K.A. Brekke, R.B. Howarth and K. Nyborg (1998): Are there Social Limits to Growth?


244 J. Sexton and A. R. Swensen (1999): ECM-algorithms that converge at the rate of EM

245 E. Berg, S. Kverndokk and K.E. Rosendahl (1999): Optimal Oil Exploration under Climate Treaties


247 R. Johansen and J.K. Dagsvik (1999): The Dynamics of a Behavioral Two-Sex Demographic Model

248 M. Søberg (1999): Asymmetric information and international tradable quota treaties. An experimental evaluation


252 R. Aaberge (1999): Sampling Errors and Cross-Country Comparisons of Income Inequality


254 A. Langårgen and R. Aaberge: A Structural Approach for Measuring Fiscal Disparities


258 L. Lindholt (1999): Beyond Kyoto: CO2 permit prices and the markets for fossil fuels


265 Y. Li (2000): Modeling the Choice of Working when the Set of Job Opportunities is Latent
E. Holmøy and T. Hægeland (2000): Aggregate Productivity and Heterogeneous Firms


R. Bjørnstad (2000): The Effect of Skill Mismatch on Wages in a small open Economy with Centralized Wage Setting: The Norwegian Case

R. Aaberge (2000): Ranking Intersecting Lorenz Curves


A. Bruvoll and H. Medin (2000): Factoring the environmental Kuznets curve. Evidence from Norway


K. Nyborg and M. Rege (2000): The Evolution of Considerate Smoking Behavior

M. Sober (2000): Imperfect competition, sequential auctions, and emissions trading: An experimental evaluation


L.-C. Zhang (2000): On dispersion preserving estimation of the mean of a binary variable from small areas


A. Langørgen (2000): Revealed Standards for Distributing Public Home-Care on Clients


A. Raknerud and R. Golombek: Exit Dynamics with Rational Expectations


J.T. Lind (2001): Tout est au mieux dans ce meilleur des ménages possibles. The Pangloss critique of equivalence scales


T. Hegeland (2001): Experience and Schooling: Substitutes or Complements

T. Hegeland (2001): Changing Returns to Education Across Cohorts. Selection, School System or Skills Obsolescence?

R. Bjørnstad (2001): Learned Helplessness, Discouraged Workers, and Multiple Unemployment Equilibria in a Search Model

K. G. Salvanes and S. E. Firre (2001): Job Creation, Heterogeneous Workers and Technical Change: Matched Worker/Plant Data Evidence from Norway


B. Bye and T. Åvåland (2001): The welfare effects of housing taxation in a distorted economy: A general equilibrium analysis


T. Kornstad (2001): Are Predicted Lifetime Consumption Profiles Robust with respect to Model Specifications?