Identifying the Deterministic Components in Cointegrated VAR Models using GRaM for Ox Professional - User Manual and Documentation

Håvard Hungnes[*]

April 29, 2005

Abstract

GRaM (Growth Rates and cointegration Means) is a program for estimating long-run properties in cointegrated VAR models. As other program packages for cointegration analysis, the cointegrating vectors and their corresponding adjustment parameters can be estimated. In addition, the program decomposes the parameters for the deterministic terms into interpretable counterparts such as growth rate parameters and cointegration mean level parameters. These parameters express long-run properties of the model. For example, the growth rate parameters tell us how much to expect (unconditionally) the variables in the system to grow from one period to the next, and will therefore be especially important to identify if the system is to be used for forecasting.

GRaM also allows for different types of hypotheses testing. Both linear tests on the cointegrating space and on the 'growth rates' are allowed for, i.e. we can test for both the long-run relationship between the time series and the long-run growth in the different time series.

This document describes the program. A tutorial shows how to use the program.

[*]E-mail: havard.hungnes@ssb.no Homepage: http://folk.ssb.no/hhu

1 Introduction

1.1 System requirements

Requirements:
- Windows
- Ox Professional 3.30
- Simulated Annealing (Ox package)

1.2 Installation

You need Ox Professional, (c) J. A. Doornik, version 3.3. (or newer) in order to run the program. (Please note that you cannot use the console version of Ox.) If you do not have Ox Professional you can buy if from the following address:

- http://www.timberlake.co.uk/software/Oxmetrics/Ox/ox.html

If you have purchased Ox Professional 3.0, 3.1 or 3.2, and have a valid licence code, then you can download the upgrade to version 3.3 from this address:

- http://www.doornik.com/download_oxpro.html

It is important that you upgrade Ox Professional to the version 3.3 (or newer). This is because the program applies a function (MaxSQP) that is not implemented in earlier versions of Ox.

You must also run Windows to run GRaM for Ox. (I apologize for this...) The reason is that the program use the OxPack interface in Ox Professional, and this runs (unfortunately) only under Windows.[1]

Finally, you need to include the MaxSA package in the directory /Ox/Packages/. (The MaxSA package provides an implementation of Simulated Annealing, see Goffe

---

[1]In the next version of Ox, Ox Professional 4.x, OxPack will probably run under any platform (and not just Windows). GRaM will therefore probably also work on all platforms.

et al. (1994) for a documentation of the algorithm.) You can download this package from Charles Bos' home page:

- http://www.tinbergen.nl/˜cbos/software/maxsa.html

The next step is to download GRaM. You can download it from the following page: [Not yet available for downloading!]

- http://folk.ssb.no/hhu

Unzip the files and save them in the directory /Ox/Packages/. To run the program start OxPack, which is a part of Ox Professional. In OxPack, choose 'Add/Remove Package..' under the 'Package' menu. Choose 'Browse' to find gram.oxo. (If you followed the instruction above, you will find it in the directory /Ox/Packages/.) Finally, press the button 'Add'. Now you can choose 'GRaM' in the 'package' menu. And that is it!

1.3   User license and citation

GRaM (Growth Rates and cointegtartion Means) 0.99 is copyright (c) Håvard Hungnes 2003-2004. It is free to use under the following conditions:

1. Its use in published research is acknowledged. Please cite this document in your list of references.

2. (Disclaimer) The code is distributed "as is", with no warranties as to fitness for any purpose. Use it at your own risk.

Ox Professional 3.3 (or newer) is required to use the package, see 'Installation'. Ox should be cited whenever it is used. For example you could say in the text: "the results are generated using Ox version 3.30 (see Doornik, 2001)" and then give the references: "Doornik (2001), Object-Oriented Matrix Programming Using Ox, 4th ed. London: Timberlake Consultants Press and Oxford: www.nuff.ox.ac.uk/Users/Doornik."

2 Theoretical introduction

2.1 Introduction

In analysing a dynamic econometric model we are often interested in identifying and testing its long run properties. The cointegrating vectors are examples of long run relationships between different variables. However, also the underlying growth rates (i.e. steady state growth rates) can be identified in cointegrated vector autoregressive (VAR) models.

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 very important in providing good forecasts. In fact, as the forecasting horizon approaches infinity, the forecast will rely on this vector only.

GRaM decomposes all deterministic terms in a cointegrated VAR model into interpretable counterparts. The corresponding coefficients describe the long run (steady state) growth rates for the variables, and possibly shifts in level and growth rates (the latter depending on the type of deterministic variables that are included in the system). Combined with the coefficients for the cointegrating vectors, they also describe level and trends (and possibly shifts in these) in the cointegrating vectors.

Examples of applications:

- Hungnes (2002); cointegrated VAR model where some variables were allowed to grow and some were not.

- Hungnes (2005); coitnegrated VAR with structural breaks.

Another example of application for the procedure is to test for "zero-mean" convergence. "Zero-mean" convergence implies that the difference between two variables (say; output per capita in two countries) is stationary with zero mean. To test if the mean is zero, the deterministic terms must be decomposed.

Throughout the paper we define the orthogonal complement of the full column rank matrix A as $A_\perp$ such that $A'_\perp A = 0$ and $(A, A_\perp)$ has full rank. (The orthogonal complement of a nonsingular matrix is 0, and the orthogonal complement of a zero matrix is an identity matrix of a suitable dimension.)Furthermore, for a matrix A with dimension $n \times m$ ($m \leq n$) with full column rank, we define $\overline{A} = A (A'A)^{-1}$.

2.2 Model formulation and Granger's representation theorem

Let Y_t be an n-dimensional vector of variables that are integrated of order one at most. α and β are matrices of dimension $n \times r$ (where r is the number of cointegrating vectors) and $\beta'Y_t$ comprises r cointegrating $I(0)$ relations. Furthermore, Γ_i ($i = 1, 2, ..., p - 1$) are $n \times n$ matrices of coefficients, where p is the number of lags. Δ is the difference operator. D_t^* is a vector of deterministic variables. The errors ε_t are assumed to be Gaussian white noise ($\varepsilon_t \sim N(0, \Omega)$).

$$\Delta Y_t = \alpha \left(\beta' Y_{t-1} \right) + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \delta D_t^* + \varepsilon_t, \quad t = 1, 2, ..., T. \tag{1}$$

It is common to distinguish between deterministic variables that are restricted to lie in the cointegration space and those which are not. Let $\delta D_t^* = \delta_0 D_{0,t}^* + \delta_1 D_{1,t}^*$, where $D_{0,t}^*$ includes the deterministic variables restricted to lie in the cointegrating space (i.e such that $\delta_0 = \alpha \overline{\alpha}' \delta_0$ or equivalently $\alpha'_\perp \delta_0 = 0$). Disregarding different types of dummies (such as impulse dummies, shift dummies and seasonal dummies), the most common two specifications for these deterministic variables are $\left(D_{0,t}^*, D_{1,t}^* \right) = (1, \varnothing)$ (i.e. restricted constant, excluding a linear drift in Y_t, labelled H_c) and $\left(D_{0,t}^*, D_{1,t}^* \right) = (t, 1)$ (i.e. restricted linear trend, excluding a quadratic trend in Y_t, labelled H_l). If there are trends in the data H_l is recommended, and in systems without trends H_c is recommended.

Let us assume that the process in (1) is generated by hypothesis H_l. The system grows at the unconditional rate $E_t [\Delta Y_t] = \gamma$ with long run (cointegration) mean levels $E_t [\beta' (Y_t - \gamma)] = \mu$. We can rewrite the system as

$$\Delta Y_t - \gamma = \alpha \left(\beta' Y_{t-1} - \mu - \rho (t - 1) \right) + \sum_{i=1}^{p-1} \Gamma_i \left(\Delta Y_{t-i} - \gamma \right) + \varepsilon_t, \tag{2}$$

where

$$\rho \equiv \beta' \gamma \tag{3}$$

is the vector of the trends in the cointegrating vectors. For the system to be stable, the

following restriction must hold:

Condition 2.1 *Assume that $n - r$ of the roots of the characteristic polynomial*

$$A(z) = (1 - z) I_n - \alpha \beta' z - \sum_{i=1}^{p-1} \Gamma_i (1 - z) z^i$$

are equal to 1 and the remaining roots are outside the complex unit circle.

In GRaM we formulate the cointegrated system slightly differently than in (1). We do this in order to make the deterministic components easier to interpret, such as in (2).

Let D_t be a vector of q deterministic variables, such as trend and seasonally dummies. The system can then be written as

$$\Delta Y_t - \gamma \Delta D_t = \alpha \left(\beta' (Y_{t-1} - \gamma D_{t-1}) - \mu \right) + \sum_{i=1}^{p-1} \Gamma_i [\Delta Y_{t-i} - \gamma \Delta D_{t-i}] + \varepsilon_t, \qquad (4)$$

where γ is now an $n \times q$ matrix.

If $D_t = t$, the system in (4) is equal to the system in (2). This is the case with linear trend in the variables, i.e. H_l.

In the case where there are no trends in the variables, D_t vanishes from (4), and the system can be written as

$$\Delta Y_t = \alpha \left(\beta' Y_{t-1} - \mu \right) + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \varepsilon_t. \qquad (5)$$

In both these cases there is a one-to-one correspondence between the system written in the normal way, as in (1), and in the alternative way, as in (4) or (5). If the system is estimated as in (1) with $\left(D_{0,t}^*, D_{1,t}^* \right) = (t, 1)$, we can always identify the coefficients in (4) by applying Theorem 2.1.

Also when seasonal dummies are included, (1) and (4) are two alternative ways of writing the same system. Generally, however, there is no such one-to-one relationship between the formulations in (1) and (4).

An alternative way to write the system, is to write the system where the deterministic components are removed. Let Y^d be defined as Y where the deterministic components are removed, i.e.

$$Y_t^d = Y_t - \gamma D_t$$

with D_t as the vector of deterministic variables and γ as the corresponding matrix of

coefficients. The system can therefore alternatively be written as

$$\Delta Y_t^d = \alpha \left(\beta' Y_{t-1}^d - \mu \right) + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i}^d + \varepsilon_t. \tag{6}$$

Now we have the following theorem:

Theorem 2.1 (Granger's representation theorem with deterministic variables) *Under Condition 2.1, Y_t in (4) has the representation*

$$Y_t = C \sum_{i=1}^{t} \varepsilon_i + \iota + \gamma D_t + B_t, \tag{7}$$

where $C = \beta_\perp \left(\alpha'_\perp \Gamma \beta_\perp \right)^{-1} \alpha'_\perp$ with $\Gamma = I_n - \sum_{i=1}^{p-1} \Gamma_i$. The process B_t is stationary with zero expectation. The level coefficients ι depends on initial values in such a way that

$$\mu = \beta' \iota. \tag{8}$$

Proof. By using $Y_t^d = Y_t - \gamma D_t$ (i.e. the system in (6)), the proof follows from the proof of Theorem 4.2. in Johansen (1995). ∎

The formulation of the system as in (2) involves that the representation of the process in (7) is valid in the whole data sample and independent of the type of deterministic variables included in D_t.

3 GRaM tutorial

3.1 Introduction

To start GRaM you need to start OxPack and choose Gram under 'Packages'. (If you can not find Gram under this menu you have not installed GRaM properly. See Section 1.2 on how to install GRaM.)

In order to follow this tutorial you need to load the data file luetkepohl.in7. (You can find this file /Ox/Packages/Gram/Data/.) Lütkepohl and Wolters (1998) and Saikkonen and Lütkepohl (2000) use a data set covering the unification to estimate a model for money demand in Germany.[2]

First we look at the data. Choose 'Graphics...' under the 'Tools' menu in GiveWin. (Alternatively you could click on the graphics button or press Alt+G.) Highlight the following data series: m, y, r and Dp. Then press '$<<$' and choose 'Apply separate actual values plots' in the top left part of the window. Now you should get a graph which looks something like Figure 1 on the following page.

The four variables are; (log of) real money M3 (m), (log of) real GNP (y), an opportunity cost of money (r) and inflation (Dp). The opportunity cost of money is defined as the difference between long term interest rate and the own rate on M3.

In the data series for money and income there is a significant shift in the level from 1990q3. This corresponds to the monetary (re-) unification of Germany, which took place July 1, 1990.[3] For the opportunity cost of money and the inflation rate there is no obvious shift in level or trend.

To take account for the shift in level for money and income, a shift dummy is constructed. This shift dummy, D1990q3, is zero until 1990q2 and unity thereafter. A

---

[2]The data are available at: ftp://141.20.100.2/pub/econometrics/germanm3.zip. Source: Lütkepohl and Wolters (1998).

[3]The official date of unification is October 3, 1990. However, the monetary unification took place about three months earlier.

Figure 1: German money demand data
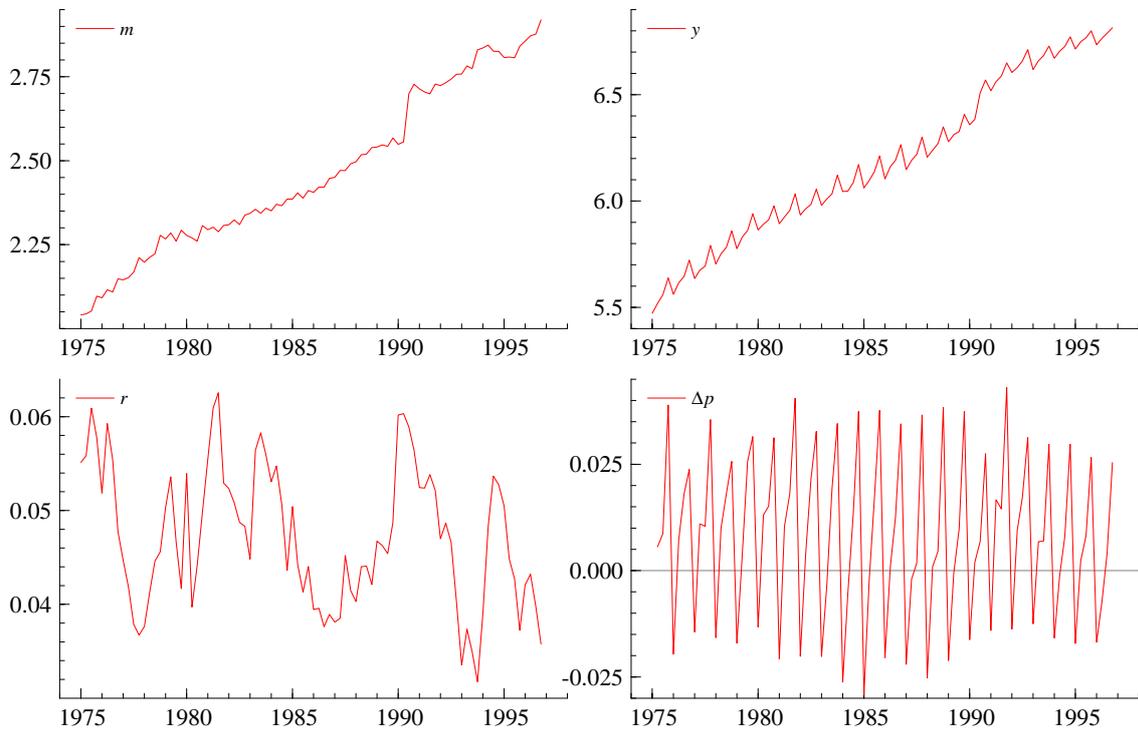
The money stock (m), (real) income/gdp (y), the opportunity cost of money (r) and inflation quarter/quarter (Dp) in Germany (West Germany until 1990q2 and unified Germany thereafter).

corresponding broken trend, defined as the cumulate of the shift dummy, is also included to allow for testing shifts in the growth rates, see Hungnes (2005).

There is (possibly) a seasonal pattern in the income and inflation rate series. Therefore, some seasonal dummies are constructed; CS1, CS2 and CS3.

3.2 Formulation

Now, let us turn to the estimation of this data set.

If you have started GRaM, OxPack will look something like Figure 2 on the next page. Choose 'Formulate...' under 'Model' (or click on the highlighted 'Formulate' button, or press CTL+Y). You will then get to the 'Data selection' menu.

In the 'Data selection' menu you choose the data series to be included in the model. First choose the number of lags; here 2 lags. Then choose the variables to be included (both endogenous and deterministic variables); here all variables in the 'Database'. Finally, choose which variables to be endogenous and which to be deterministic. GRaM

Figure 2: GRaM in OxPack


Figure 3: Data selection
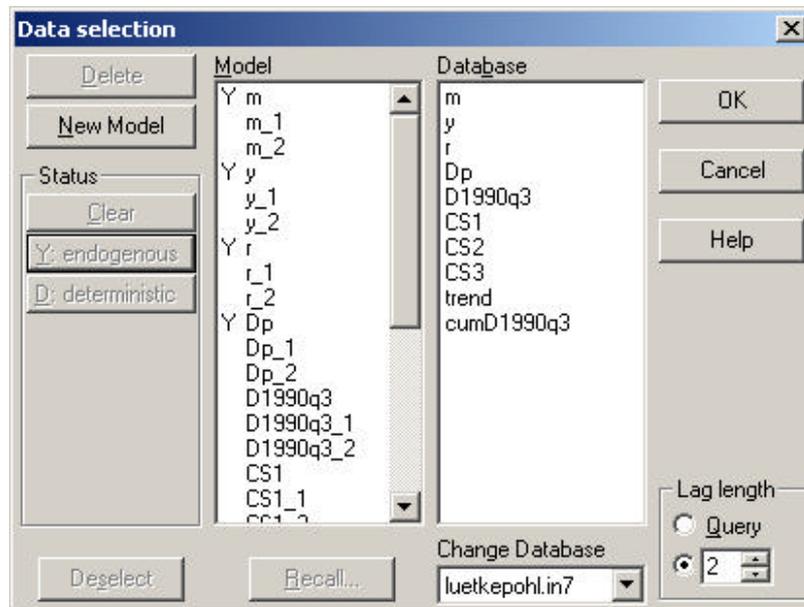
automatically suggests the first variable only to be endogenous, and the remaining being deterministic. Her we want the first four variables (money, income, opportunely cost of money, and the inflation rate) to be endogenous. Therefore, highlight the the variables form 'y' to 'Dp' in the 'Model', and click the button 'Endogenous'.[4] The 'Data selection' menu should now look something like Figure 3.

The next menu is the 'Estimate model' menu, see Figure 4 on the next page. Here

---

[4]You do not have to highlight the deterministic variables and choose 'Deterministic'; if a variable is not chosen to be 'Endogenous', GRaM will assume it is 'Deterministic'.

Figure 4: Estimate Model
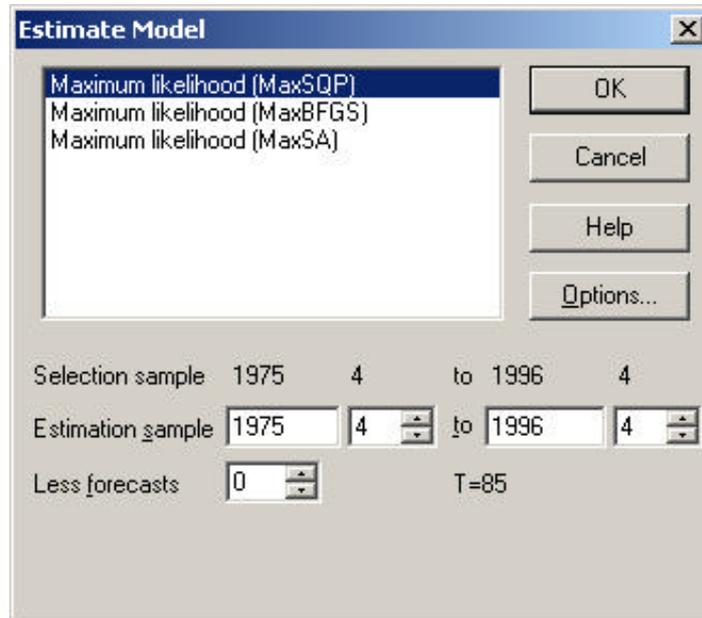
you can choose different types of simulation algorithms, and you can adjust the estimation period. The different types of simulation algorithms will be explained in more details later. Now, use 'Maximum likelihood (MaxSQP)'.

In the 'General restrictions' menu (Figure 5 on the following page) you choose the number cointegrating vectors, and the number of restrictions on the different coefficient matrixes. Choose '2' for 'Cointegrating rank', and zero for all the restrictions. (In the next subsections you will learn about imposing restrictions.)

After you click on 'OK' on the 'General restrictions' menu, the program starts to simulate in order to try to find the maximum likelihood value and the corresponding estimates. The program will (probably) write some intermedin results. After a few seconds (depending on how fast your computer is), you will get the final result, see Figure 6 on page 13.

The first few lines reports intermediate results from the simulation. We will ignore this part of the output here.

The next lines reports the estimated coefficient matrixes, with the corresponding standard deviations. First, γ' (gamma') is reported (with standard deviations), followed by α (alpha), β (beta) and ρ' (rho). (The standard deviation for β can not be computed when no restrictions are imposed.)

11

Figure 5: General restrictions
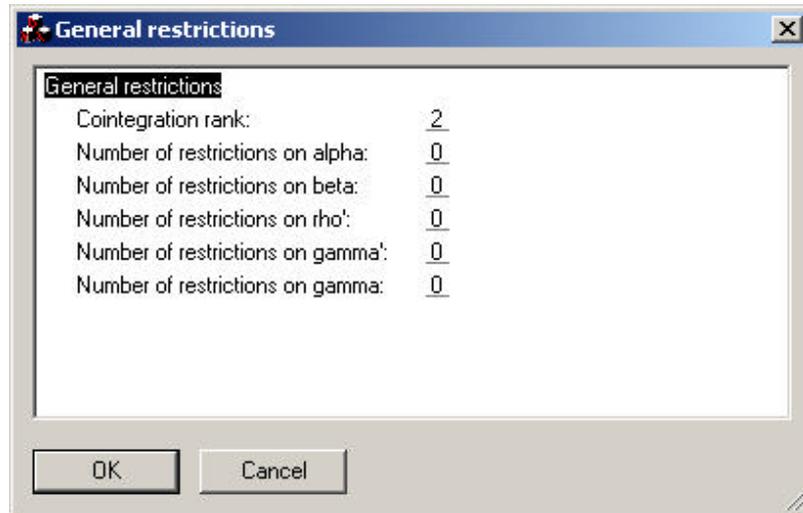
The eigenvalues of the companion matrix are also reported. These should be either on or inside the unit circle, i.e. the modulus should be unity or less. (The number of unit roots should equal the number of endogenous variable minus the cointegrating rank, see Condition 2.1 on page 6.) These eigenvalues are reported to evaluate the reliability of the results; if one or more of the eigenvalues are outside the unit circle results are not reliable (and the corresponding estimated system is explosive).

Finally, the likelihood value and some other information about the system are reported. Also the results form the simulation algorithm is reported. The output here is similarly to the output in 'Multiple-Equation Dynamic Modelling...' package in Pc-Give, see Doornik and Hendry (2001).

Now, consider the estimated γ' in Figure 6 on the next page. The second last row of the estimated γ', labelled 'trend', reports the trend in the different variables in the system. Since both money and income are measured in logs, the corresponding coefficients can be interpreted as growth rates. The growth rate for money is therefore 0.53 per cent each quarter. This corresponds to an annual growth rate of 2.13.[5] The estimated quarterly growth rate for income is 1.37 per cent, which corresponds to an annual growth rate of 5.61 per cent.

The coefficients for the trend in the variables for opportunity cost and inflation

---

[5]The annual growth rate can be calculated as $\gamma_a = \left(1 + \gamma_q\right)^4 - 1$, where γ_a is the annual growth rate and γ_q is the quarterly growth rate. However, when the growth rate is close to zero, the simplification $\gamma_a \approx 4 \cdot \gamma_q$ works well.

Figure 6: GiveWin output

```
Ox version 3.30 (Windows) (C) J.A. Doornik, 1994-2003
GRaM package version 0.99, (c) H. Hungnes, 2004, object created on 14-10-2004

GRaM( 1) Estimating luetkepohl.in7 using M_MAXSQP

STEP 0 : Starting values

STEP 1 : Simulation ('C.I. space restrictions')
it0    f=      39.82048 e2=6.209e+004 step=1
it100  f=      40.14324 e2= 0.0001371 step=1
it106  f=      40.14324 e2=2.054e-007 step=1
Strong convergence

STEP 3 : Switching
Sw0:  f = 40.14324      ->  1223.6487411012531
Sw1:  f = 40.14324      ->  1223.6487411012531

gamma'
                        m             y             r            Dp
D1990q3           0.14638       0.10704  6.5850e-005     0.0031323
CS1             -0.012079     -0.091688    0.0011336     -0.052181
CS2            -0.0085290     -0.069543    0.0024576     -0.029688
CS3             -0.021426     -0.055743    0.0020736     -0.020349
trend           0.0052940      0.013730  8.6271e-005   3.6537e-005
cumD1990q3      0.0059167    -0.0030122   -0.0011189   -0.00050455


st.d.
D1990q3         0.0084791      0.012040    0.0040269     0.0017932
CS1             0.0019897      0.0030763   0.00088277    0.0016195
CS2             0.0025453      0.0032858    0.0010042    0.0019274
CS3             0.0019685      0.0030466   0.00087032    0.0016218
trend           0.0013092      0.0012990   0.00037854   0.00013691
cumD1990q3      0.0026274      0.0024518   0.00075454   0.00026426

alpha
          m             y             r            Dp
    0.035864      0.018976     0.0042251     -0.042662
   -0.062164     0.00098746    0.0068308    -0.0050408

st.d.
    0.0090993     0.015770     0.0051025     0.0055971
    0.011059      0.019167     0.0062016     0.0068027

beta
m                 1.0000        2.2378
y                -2.9040        1.0000
r                 1.1931        8.2963
Dp                33.856        5.5940
Constant          13.689       -10.861

rho'
D1990q3          -0.058340       0.45268
CS1              -1.5111        -0.40122
CS2              -0.80875       -0.23431
CS3              -0.54601       -0.20032
trend            -0.033237       0.026497
cumD1990q3       -0.0037529     -0.0018769

Eigenvalues of companion matrix
        real          imag        modulus
      0.83777       0.00000       0.83777
     -0.038338      0.52297       0.52437
     -0.038338     -0.52297       0.52437
     -0.23264       0.16179       0.28337
     -0.23264      -0.16179       0.28337
      0.059922      0.00000       0.059922
       1.0000       0.00000        1.0000
       1.0000       0.00000        1.0000


log-likelihood    1223.64874  -T/2log|Omega|    1706.08784
no. of observations        85  no. of parameters       54
rank of long-run matrix    2  no. long-run restr.       0
cointegration space is not identified

SQP using analytical derivatives (eps1=1e-005; eps2=0.0005):
Strong convergence
```

13

rate are approximately zero, implying no growth in these variables. (Based on the t-values, i.e. the ratio between the estimated coefficient and its corresponding standard deviation, the null hypothesis of each of these two growth rates to be zero would not be rejected.)

In the last line of the estimated γ' in Figure 6 on the preceding page the corresponding coefficients for $cumD1990q3$ are reported. Since $cumD1990q3$ is defined as the cumulated value of the step dummy $D1990q3$, this variable is zero until 1990q2, and increases by one unit each quarter thereafter. The corresponding coefficients therefore picks up the shift in the trend in the different variables. Therefore, the quarterly growth rate for money is 1.12 per cent after 1990q2 (i.e. 0.53 + 0.59). The reported coefficient for the trend is therefore the trend (or growth rate) in the variable before 1990q3.

In the first line of the estimated γ' in Figure 6 on the page before reports the level shift in the different variables. As expected, the shift is large for both money and income, and the shift is much less for the opportunity cost and the inflation rate.

From the reported γ we can also see how much seasonality there is in the different variables. The reported coefficients indicates that there is most seasonality in income and inflation, which we also guessed based on the graphs of the series.

3.3 Imposing restrictions on γ'

3.3.1 Restrictions on all variables

In the previous subsection γ was estimated without restrictions. But we may also want to impose restrictions on γ (or some other matrixes of coefficients). One interesting hypothesis is to test if there is a significant shift in the trend (or the growth rate) in the variables from 1990q3. This implies testing if the coefficients in the last row of the estimated γ' in Figure 6 on the preceding page are zero.

In GRaM restrictions on γ' are imposed by applying restriction matrixes. The restrictions we want to impose on γ' is written as

$$R'_{\gamma'}\gamma' = 0. \tag{9}$$

The hypothesis that the coefficients for the trend shift are zero can be tested with

Figure 7: General restrictions (2)
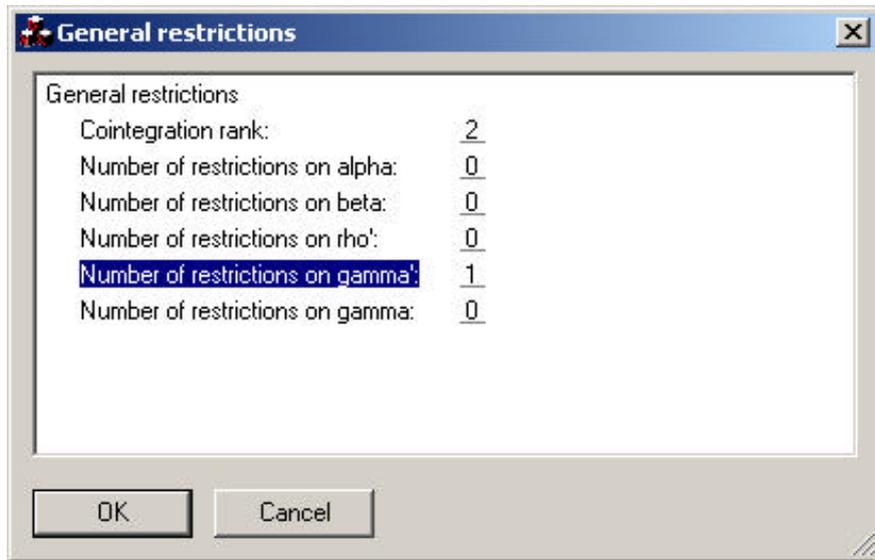
the restriction matrix $R' = (0,0,0,0,0,1)$, since

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma_{m,D1990q3} & \gamma_{y,D1990q3} & \gamma_{r,D1990q3} & \gamma_{Dp,D1990q3} \\ \gamma_{m,CS0} & \gamma_{y,CS0} & \gamma_{r,CS0} & \gamma_{Dp,CS0} \\ \gamma_{m,CS1} & \gamma_{y,CS1} & \gamma_{r,CS1} & \gamma_{Dp,CS1} \\ \gamma_{m,CS0} & \gamma_{y,CS2} & \gamma_{r,CS2} & \gamma_{Dp,CS2} \\ \gamma_{m,trend} & \gamma_{y,trend} & \gamma_{r,trend} & \gamma_{Dp,trend} \\ \gamma_{m,cumD90q3} & \gamma_{y,cumD90q3} & \gamma_{r,cumD90q3} & \gamma_{Dp,cumD90q3} \end{pmatrix}$$

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

implies $\gamma_{i,cumD1990q3} = 0$, $i = \{m,y,r,Dp\}$.

Since you have already formulated the model, you do not have to do that again. Therefore, choose 'Estimate...' under 'Model' (or click on the highlighted 'Estimate' button, or press CTL+L). You will then get to the 'Estimate Model' menu. Press 'OK'.

In the 'General restrictions' menu changes the number of restrictions on γ' in the second last line to 1, see Figure 7. (We choose 1 because we only impose one restriction on the γ' matrix.) Press 'OK'.

Now you come to the menu 'Restrictions on gamma''. For each restrictions, there

15

Figure 8: Restrictions on γ'
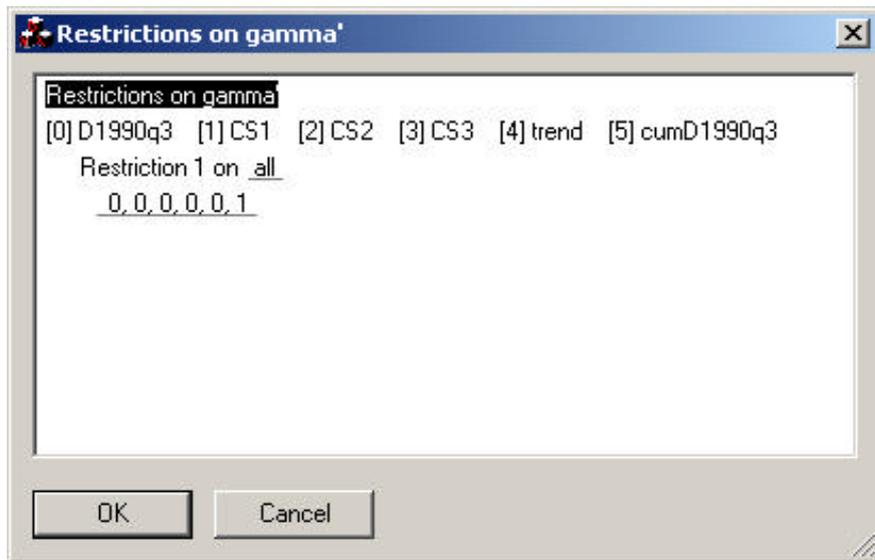
are two choices. The first choice is whether you want the restrictions to be restrictions on all the (endogenous) variables in your system, or on only one of them. If you press on the underlined 'all', a roll-down menu appears. Here you can chose between $'all'$, $'m'$, $'y'$, $'r'$, and $'Dp'$. We want to impose the restriction on all the variables in our system, and we therefore choose $'all'$ (which is the default).

The second choice is the restriction you want to impose. To make the specification of the restriction vector simpler, GRaM generates a row of equally many zeros as elements in the restriction vector. In addition, in the upper part of the menu, a row of the corresponding deterministic variables are reported. This should make it easier to to construct the restriction vectors.

We want to restrict the coefficients for $cumD1990q3$ to be zero for all the variables in the system. According to the line in the upper part of the menu, the variable $cumD1990q3$ corresponds to the last element in the restriction vector. Therefore, we changes the last element in the restriction vector to unity (or any non-zero number), see Figure 8. Press 'OK', and GRaM starts to simulate. After a few seconds, the results are reported to the screen. The last part of the print-out should look something like Figure 9 on the next page.

In the system we have estimated 4 independent coefficient restrictions were imposed, i.e. one restriction for each of the four endogenous variables. This is also reported in the output (no. long-run restr.: 4). Since this system is a restricted

16

Figure 9: GiveWin output (part of)

```
log-likelihood      1221.55518  -T/2log|Omega|      1703.99428
no. of observations          85  no. of parameters          54
rank of long-run matrix       2  no. long-run restr.          4
cointegration space is not identified

SQP using analytical derivatives (eps1=1e-005; eps2=0.0005):
Strong convergence
```

version of the previous, the log likelihood value is less (1221.56 vs. 1223.65). We can apply a likelihood ratio test to test the (null) hypothesis that there is no trend-break in the (endogenous) variables in the system. The likelihood ratio is twice the difference between the likelihood values, i.e. $2 \cdot (1223.65 - 1221.56) = 4.19$. To find the corresponding probability value, we choose 'Progress...' under 'Model'. In the 'Progress' menu, just choose 'OK'. The program reports information about the general and restricted system. It also reports an χ^2 test of the restriction. (We assume this ratio is χ^2 distributed, see Section 4.8 on page 38.) The reported probability value is 0.381, and with this probability value you would not reject the null hypothesis (that there is no shift in the trend - or 'growth rates' - in the variables).

Note that we would have got the same estimation results if we had deleted $cum1990q3$ as a deterministic variable in our system, and estimated that system without restrictions.

3.3.2 Restriction on one variable only

In the previous 'subsubsection' we imposed the restriction that there were no trend-breaks in any of the endogenous variables. Here we will show how to impose the additional restrictions that there is no trend in the opportunity cost or in the inflation rate. In the 'General restrictions' menu we therefore choose 3 restrictions on gamma' in the second last line.

In GRaM restrictions can also be imposed on one vector in γ'. If we want to impose

17

Figure 10: Restrictions on γ' (2)
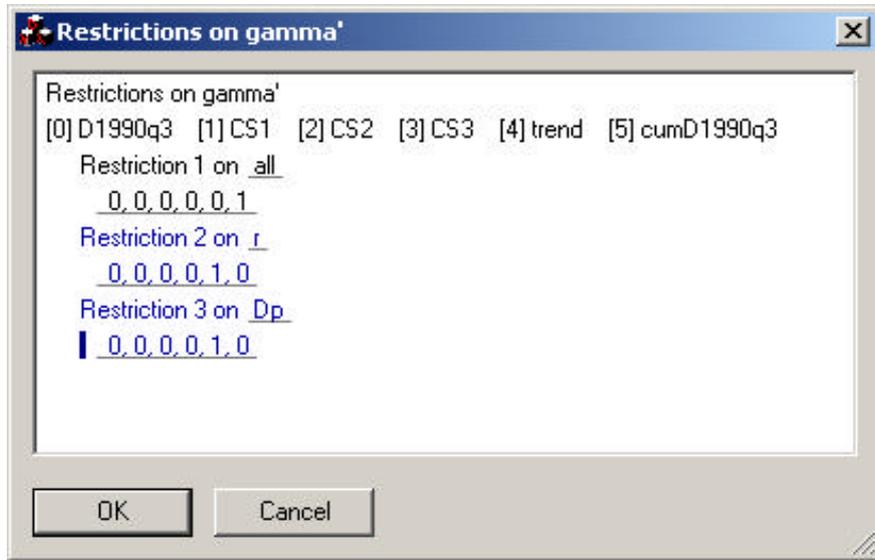
the restriction that there is no trend in the opportunity cost, this can be written as

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma_{r,D1990q3} \\ \gamma_{r,CS0} \\ \gamma_{r,CS1} \\ \gamma_{r,CS2} \\ \gamma_{r,trend} \\ \gamma_{r,cumD90q3} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}$$

In the 'Restrictions on gamma'' menu we shall now formulate 3 restrictions. The first restriction is the same as above. Since GRaM remembers the restrictions, we do not have to change the first restriction.

The second restriction is to restrict the trend for the opportunity cost to be zero. Change therefore 'all' to 'r' in the line 'Restriction 2 on ..'. The restriction vector should read $R' = (0,0,0,0,1,0)$. Similarly, for the third restriction we change 'all' to 'Dp' and replaces the second last zero with unity, see Figure 10.

In the output GRaM reports that there are now 6 independent coefficient restrictions, and the likelihood is now 1221.28. Comparing it to the unrestricted system yields $\chi^2 = 4.9[0.56]$, or comparing it to the system with restrictions only on the broken trend yields $\chi^2 = 0.7[0.70]$.

18

3.4   Imposing other (coefficient) restrictions

3.4.1   Restrictions on γ

Not all interesting restrictions on γ can be imposed by the formulation above. For example, in the system we look at, we might want to test if the 'growth rates' for money and income are equal. To allow for restrictions between the coefficients for the deterministic variables in γ, GRaM also allows for restrictions of the form[6]

$$R'_\gamma \gamma = 0, \qquad (10)$$

if the restriction is to hold for all deterministic variables, or

$$R'_{\gamma,i} \gamma_i = 0, \qquad (11)$$

if it refers to only one deterministic variable. (Here γ_i is the i'th (column) vector in γ.)

The restriction that the 'growth rates' for money and income are equal, can therefore be imposed by the formulation

$$\begin{pmatrix} 1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \gamma_{m,trend} \\ \gamma_{y,trend} \\ \gamma_{r,trend} \\ \gamma_{Dp,trend} \end{pmatrix} = \begin{pmatrix} 0 \end{pmatrix},$$

which implies $\gamma_{m,trend} = \gamma_{y,trend}$. This is an example of restrictions of the form in (11).

To impose this restriction, go to the 'General restrictions' menu. Now, change the number of restrictions on γ in the last line to 1 (and keep the 3 restrictions on γ'). After you click on 'OK', you get to the menu for 'Restrictions on gamma''. Do not changes these restrictions, but press 'OK'.

Now you get to the menu 'Restrictions on gamma'. The restriction we want to impose is on the 'trend' variable only. Therefore, change '*all*' to '*trend*'. The restriction vector we want to impose on the trend is $R' = (1, -1, 0, 0)$. Therefore, change the first element in the row of zeros to 1 and the second to -1. The restriction that the 'growth rates' for money and income are equal is now imposed, and you can let GRaM simulate by clicking 'OK'.

The resulting likelihood value is 1214.45, and the likelihood ratio of this additional restriction is 13.66 with an probability value of 0.00 (with one degree of freedom). Therefore, the hypothesis of equal 'growth rates' for money and income is rejected.

---

[6]In (10) the restrictions are imposed on γ, whereas in (9) the restrictions are imposed on the transformed of γ, i.e. on γ'.

3.4.2 Restrictions on ρ'

Restrictions on γ involves restrictions on the variables in the system. Restrictions on ρ involves restrictions on the cointegrating vectors in the system. We will now consider restrictions on the deterministic variables in the cointegrating vectors. The test we will impose, is if there is a significant shift in the trend (or the 'growth rates') in the two cointegrating vectors from 1990q3.

The restrictions are imposed similarly to the restrictions on γ, that is of the form

$$R'_\rho \rho' = 0 \tag{12}$$

if the restriction is to be imposed on 'all' cointegrating vectors.[7]

The hypothesis that the coefficients for the trend shift are zero in both cointegrating vectors can be tested with the restriction matrix $R'_\rho = (0,0,0,0,0,1)$, since

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \rho_{CI1,D1990q3} & \rho_{CI2,D1990q3} \\ \rho_{CI1,CS0} & \rho_{CI2,CS0} \\ \rho_{CI1,CS1} & \rho_{CI2,CS1} \\ \rho_{CI1,CS0} & \rho_{CI2,CS2} \\ \rho_{CI1,trend} & \rho_{CI2,trend} \\ \rho_{CI1,cumD90q3} & \rho_{CI2,cumD90q3} \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$$

implies $\gamma_{i,cumD1990q3} = 0$, $i = \{CI1, CI2\}$. (Here $CI1$ refers to the first cointegrating vector and $CI2$ refers to the second cointegrating vector.)

In the 'General restrictions' choose 1 in the row for 'Restrictions on rho'', and choose 0 for the restrictions on γ' and γ. (I.e. we are only considering restrictions on ρ.)

In the 'Restrictions on rho'' menu, keep 'all' after 'Restriction 1 on ..', since we will impose the restriction on all (or both, since there are only two cointegrating vectors) cointegrating vectors.[8] Next, change the last 0 to 1, so the restriction vector becomes $R'_\rho = (0,0,0,0,0,1)$. Press 'OK'.

The likelihood value with this restriction is 1223.03, which yields $\chi^2 = 1.23[0.54]$. The restriction is therefore not rejected. This we would expect, since the restriction of no trend break in any of the variables (which is a stronger restriction) was not rejected either.

We can test for co-breaking when we test restrictions on ρ. The concept of co-

---

[7]In the present version of GRaM it is not possible to impose restrictions on the individual cointegrating vectors.

[8]In the present version 'all' is the only possible choice, see footnote 7.

breaking was introduced by Hendry and Mizon (1998). If deterministic breaks in a system of equation can be removed by taking linear combination of the system variables, the variables are said to co-break. With GRaM we can test if the co-breaking vectors coincides with the cointegrating vectors. For example, the restriction of no level shift in the variables ($R'_{\gamma'} = (0, 0, 0, 0, 0, 1)$) are rejected in our system. However, if the restriction of no level shift in the two cointegrating vectors ($R'_{\rho} = (0, 0, 0, 0, 0, 1)$) is not rejected, the (space spanned by the) cointegarting vectors equals the (space spanned by the) co-breaking vectors. Then we can say that the cointegrating space also represents a co-breaking space.

3.4.3 Restrictions on β^*

As other programs, GRaM allows for restrictions among the endogenous variables in the cointegrating space, β. In addition, GRaM allows for restrictions on the intercept of the cointegrating vectors, i.e. μ in (2) on page 5. These coefficients (μ) may be interpreted as the cointegration mean levels, at least if $\rho = 0$. The coefficient restrictions on β and μ may be formalized as

$$R'_{\beta}\beta^* = c_{\beta}, \tag{13}$$

where $\beta^* = (\beta, -\mu')$. In (13) we have included a level variable, c_{β}, which makes it possible to normalise the cointegrating vectors.

3.5 Graphical output

With GRaM you can also get some graphical output of your estimation results. Estimate the system with cointregation rank 2 and without restrictions. Choose 'Graphical Analysis' under the 'Test' menu. You will then get a graphical output similarly to Figure 11 on the following page. In the first column the true time series are plotted together with one period predictions ('fitted'). In the second column the residuals are plotted. In the last column the 'de-trended' variables are plotted. The 'de-trended' variables are defined by

$$Y_t^{det} = Y_t - \gamma \cdot D_t, \tag{14}$$

i.e. where the deterministic components are removed from the series. In these 'de-trended' data series there should be no deterministics such as trend and seasonality. (If there is, this indicates that you have not included all necessary deterministic variables in the system.) From the graph we see that the trending is removed in the series for money and income. And in the graph for 'de-tended' inflation much of the variation is removed, which indicates that much of the variation in inflation were seasonal related.

Figure 11: Graphical output
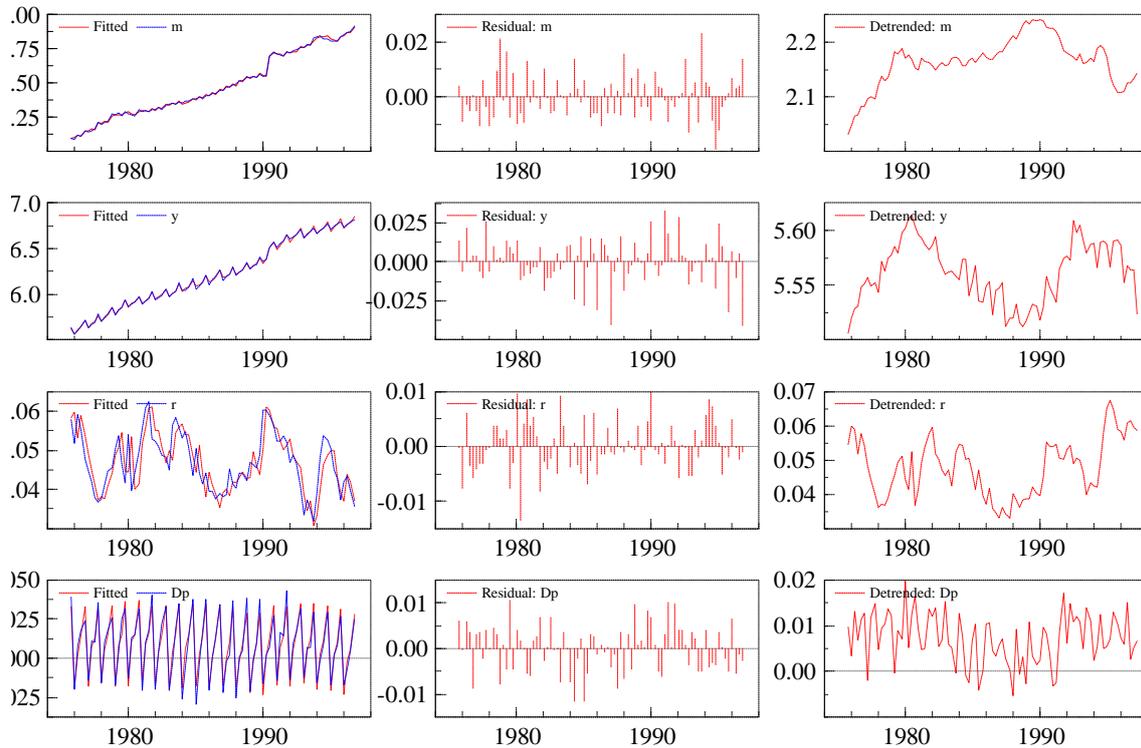
3.6 Using GRaM together with PcGive and PcGets

In GRaM you can estimate the long run properties of your system, and construct the corresponding 'de-trended' variables, see (14). These 'de-trended' variables can be saved in the database and then be used in other programs such as PcGive or PcGets. To save the 'de-trended' variables choose 'Store in database...' under the 'Test' menu, and cross off for 'De-trended variables'. Then choose names for your 'de-trended' variables. (GRaM suggests a prefix 'Det' before the original labels of your data series.)

3.7 The Batch editor

All estimation in GRaM can also be achieved through batch commands in GiveWin. Here only the new commands necessary to apply with GRaM is documented. For other batch commands, please look up in the GiveWin documentation.

`rank(`*`rank`*`)`: Sets the cointegration rank. Must appear before `restrictions` and `estimate`.

`restrictions(`*`label,...`*`)`: Impose restrictions. *label* must be one of the following; *beta* (if the restrictions shall be imposed on β), *rhoT* (if the restrictions shall be

imposed on ρ'), *gammaT* (if the restrictions shall be imposed on γ' or *gamma* (if the restrictions shall be imposed on γ). The next numbers represents the restriction (or restrictions). The first number after the label express if the restriction is to be imposed on the whole matrix of coefficients or on only one of the vectors. The number 0 represents a restriction on the whole matrix, otherwise the number corresponds to the vector the restriction is to be imposed on. The next numbers corresponds to the restriction vector as they are imposed in GRaM.

As an example, consider the batch code when estimating the German Money data set with no trend break for money and income and no trend for the opportunity cost of money and for the inflation rate, i.e. as the restrictions as in Figure 10 on page 18. Click the batch icon in GiveWin. Then the batch code in Figure 12 on the following page appears. We now consider the restrictions command in the batch code. The *label* is *gammaT* implies that the restrictions are imposed on γ'. The first number is 0, implying that the following restriction shall be imposed on all columns in γ'. The next numbers, 0, 0, 0, 0, 0, 1, corresponds to the restriction vector. In the next restriction, the number 3 tells us that the next restriction shall be imposed on the third column in γ', whereas the next numbers, 0, 0, 0, 0, 1, 0 is the restriction vector. Finally, the numbers in the final restriction implies that a similar restriction is also imposed in the forth column in γ'.

`startingvalues(iStart)`: Method for providing starting values, see Section 4.4 for a description of the different alternbatives. Use 0 for *Zeros (or almost zero)*; use 1 for *As for rank = 0*; use 2 for *OLS - Static regression (gamma restrictions imposed)*; use 3 for *OLS - Static regression (rho restrictions imposed)*; or 4 for *JMN*.

`switching(cSwitch, iSwitch, dSwitch)`: These commands are used to adjust settings in the switching algorithm. `cSwitch`; the maximum number of iteratins (default is 5000). `iSwitch`; print every `iSwitch`'th iteration, default is 100, 0 corresponds to not writing intermediate results. `dSwitch`; is the convergence criteria (default is 1.0e-15). The algorithm stops if the increase in likelihood from one iteration to the nest is less than this value.

`MaxControl(mxIter, iPrint, bCompact)` and `MaxControlEps(dEps1, dEps2)`: These commands can be used to adjust the settings of the iteration parameters for MaxSQP and MaxBFGS. `mxIter` is the maximum number of iteration; default is 1000, use -1 to leave the current value unchanged. `iPrint`; print every `iPrint`'th iteration, default is 100, 0 corresponds to not writing intermediate results, use -1 to leave unchanged. `bCompact`; if 1 uses compact form for iteration results. `dEps1` and `dEps2` changes the convergence criteria, see Doornik (2001).

`MaxSAControl(mxEval, iPrint)`, `MaxSAControlEps(dEps, iNEps)` and `MaxSACon-`

Figure 12: Batch code

```
module("OxPack");
package("Gram");
usedata("luetkepohl.in7");
system
{
    Y = m, y, r, Dp;
    D = m_1, m_2, y_1, y_2, r_1, r_2, Dp_1, Dp_2, D1990q3, D1990q3_1,
        D1990q3_2, CS1, CS1_1, CS1_2, CS2, CS2_1, CS2_2, CS3, CS3_1,
        CS3_2, trend, trend_1, trend_2, cumD1990q3, cumD1990q3_1,
        cumD1990q3_2;
}
rank(2);
restrictions( gammaT,0, 0, 0, 0, 0, 0, 1);
restrictions( gammaT,3, 0, 0, 0, 0, 1, 0);
restrictions( gammaT,4, 0, 0, 0, 0, 1, 0);
estimate("M_MAXSQP", 1975, 4, 1996, 4);
```

trolStep (iNS, iNT, dRT, vM, VC): These commands adjusts settings for MaxSA, see the corresponding documentation. Note that only one element (and not vectors) can be chosen for vtM and vC. Fianlly, MaxControlInitTemp(adT) is used to adjust the initial temperature. In GRaM the chosen initial temperature is much smaller than the default in the MaxSA package, and results in a warning when MaxSA is used. The choice of initial temperature in GRaM that small is done in order to prevent explosive solutions and to increase the speed until the maximising algorithm converges.

4 Technical documentation

4.1 The estimation problem without restrictions

In this section a conditional estimators for γ and β^* (respectively) are presented. Furthermore, the estimation problem is formulated.

First, consider the estimator for γ. By defining $Y_t^* = (Y_t', 1)'$ and $\beta^{*\prime} = (\beta', -\mu)$ the system can be written as

$$
\Delta Y_t - \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i}
$$

$$
= \alpha \beta^{*\prime} Y_{t-1}^* + \left(I_n, -\Gamma_1^* - \alpha \beta', -\Gamma_2^*, ..., -\Gamma_p^* \right) \left(I_{p+1} \otimes \gamma \right) \begin{pmatrix} D_t \\ D_{t-1} \\ \vdots \\ D_{t-p} \end{pmatrix} + \varepsilon_t \quad (15)
$$

where $\Gamma_1^* = \Gamma_1 + I_n$, $\Gamma_i^* = \Gamma_i - \Gamma_{i-1}$ $(i = 2, 3, ..., p - 1)$ and $\Gamma_p^* = -\Gamma_{p-1}$.

Define $Z_t = vec\left(\Delta Y_t, ..., \Delta Y_{t-p+1} \right)$, $D_t^v = vec\left(D_t, ..., D_{t-p} \right)$, $\Phi = \left(I_n, -\Gamma_1, ..., -\Gamma_{p-1} \right)$, and $\Phi^* = \left(I_n, -\Gamma_1^* - \alpha \beta', -\Gamma_2^*, ..., -\Gamma_p^* \right)$. The cointegrated VAR becomes

$$
\Phi Z_t = \alpha \beta^{*\prime} Y_{t-1}^* + \Phi^* \left(I_{p+1} \otimes \gamma \right) D_t^v + \varepsilon_t. \quad (16)
$$

Define

$$
\begin{aligned}
S_{DD} &= T^{-1} \sum D_t^v D_t^{v\prime}, \\
S_{ZD} &= T^{-1} \sum Z_t D_t^{v\prime}, \\
S_{YD} &= T^{-1} \sum Y_{t-1}^* D_t^{v\prime},
\end{aligned}
$$

and similarly for S_{ZZ} and $S_{YZ} = S_{ZY}'$. Furthermore, implicitly define the matrix M

with dimension $q(p+1)^2 n \times qn$ by[9]

$$vec\left(I_{p+1} \otimes \gamma\right) = Mvec\gamma'. \tag{18}$$

Theorem 4.1 *With given estimates of α, β^*, $\Gamma_1, ..., \Gamma_{p-1}$ and Ω, the conditional estimator for γ becomes*

$$
\begin{aligned}
\widehat{vec\gamma}'\left(\alpha, \beta^*, \Gamma_1, ..., \Gamma_{p-1}, \Omega\right) &= \left[M'\left(S_{DD} \otimes \Phi^{*\prime}\Omega^{-1}\Phi^*\right)M\right]^{-1} \\
&\times \left[M'vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right)\right],
\end{aligned}
$$

and corresponding variance matrix

$$V\left(\widehat{vec\gamma}'\right) = T^{-1}\left[M'\left(S_{DD} \otimes \Phi^{*\prime}\Omega^{-1}\Phi^*\right)M\right]^{-1}.$$

Proof. To prove the theorem we use $tr\left(AB\right) = tr\left(BA\right) = \left(vecA'\right)' vecB$ and $vec\left(AXB\right) = \left(B' \otimes A\right)vecX$ where tr is the trace operator. The log-likelihood function (when a constant is disregarded) can be written as

$$
\log L = -\frac{T}{2}\left[
\begin{array}{c}
tr\Omega^{-1}\Phi S_{ZZ}\Phi' + tr\Omega^{-1}\alpha\beta^{*\prime}S_{YY}\beta^*\alpha' \\
-tr\Omega^{-1}\Phi S_{ZY}\beta^*\alpha' - tr\Omega^{-1}\alpha\beta^{*\prime}S_{YZ}\Phi' \\
+N + tr\Omega^{-1}\Phi^*\left(I_{p+1} \otimes \gamma\right)S_{DD}\left(I_{p+1} \otimes \gamma'\right)\Phi^{*\prime} + N'
\end{array}
\right], \tag{19}
$$

where

$$
\begin{aligned}
N &= -tr\Omega^{-1}\Phi S_{ZD}\left(I_{p+1} \otimes \gamma'\right)\Phi^{*\prime} + tr\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\left(I_{p+1} \otimes \gamma'\right)\Phi^{*\prime} \\
&= -tr\left(I_{p+1} \otimes \gamma'\right)\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right) \\
&= -\left(vec\left(I_{p+1} \otimes \gamma\right)\right)' vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right) \\
&= -\left(vec\gamma'\right)' M'vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right).
\end{aligned}
$$

---

[9] The explicit form of this matrix is given as

$$M = \left[\left(\left(I_{p+1} \otimes K_{q,p+1}\right)\left(vecI_{p+1} \otimes I_q\right)\right) \otimes I_n\right]K_{q,n}, \tag{17}$$

where $K_{q,n}$ is the commutation matrix, see e.g. Magnus and Neudecker (1988, pp. 46-48). This is defined as $K_{q,n} = \sum_{i=1}^{q}\sum_{j=1}^{n}\left(J_{i,j} \otimes J'_{i,j}\right)$, where $J_{i,j}$ is an $q \times n$ matrix with all elements equal to zero except the $(i,j)'$th element, which is unity (see e.g. Lütkepohl, 1996, p. 116).

The derivative of (19) with respect to $vec\gamma'$ is

$$\frac{\partial \log L}{\partial vec\gamma'} \tag{20}$$
$$= TM' \left[vec \left(\Phi^{*'}\Omega^{-1}\Phi S_{ZD} - \Phi^{*'}\Omega^{-1}\alpha\beta^{*'}S_{YD} - \Phi^{*'}\Omega^{-1}\Phi^* \left(I_{p+1} \otimes \gamma \right) S_{DD} \right) \right].$$

Disregarding the T, the last part of the right hand side of (20) can be written as

$$M'vec \left(\Phi^{*'}\Omega^{-1}\Phi^* \left(I_p \otimes \gamma \right) S_{DD} \right) = M' \left(S_{DD} \otimes \Phi^{*'}\Omega^{-1}\Phi^* \right) vec \left(I_p \otimes \gamma \right)$$
$$= \left[M' \left(S_{DD} \otimes \Phi^{*'}\Omega^{-1}\Phi^* \right) M \right] vec\gamma', \tag{21}$$

and combining (20) and (21) yields

$$\frac{\partial \log L}{\partial vec\gamma'} = T'M' \left[vec \left(\Phi^{*'}\Omega^{-1}\Phi S_{ZD} - \Phi^{*'}\Omega^{-1}\alpha\beta^{*'}S_{YD} \right) \right]$$
$$-T \left[M' \left(S_{DD} \otimes \Phi^{*'}\Omega^{-1}\Phi^* \right) M \right] vec\gamma'. \tag{22}$$

Setting this equal to zero and solving for $vec\gamma'$ yields the estimator, and taking the derivative of (22) with respect to $vec\gamma$ yields the inverted negative of the 'variance' matrix. ∎

Let $Z_{0,t}(\gamma) = \Delta Y_t - \gamma\Delta D_t$, $Z_{1,t}(\gamma) = \left[(Y_{t-1} - \gamma D_{t-1})', 1 \right]'$ and $Z_{2,t}(\gamma) = vec \left(\Delta Y_{t-1} - \gamma D_{t-1}, ..., \Delta Y_{t-p+1} - \gamma D_{t-p+1} \right)$. The system becomes

$$Z_{0,t} = \alpha\beta^{*'}Z_{1,t} + \Theta Z_{2,t} + \varepsilon, \tag{23}$$

where $\Theta = \left(\Gamma_1, ..., \Gamma_{p-1} \right)$.

An estimator for β in the equation above could be found by applying reduced rank regression, see e.g. Anderson (1951) or Johansen (1995). However, here only a conditional estimator for β is considered.

The equation above can be estimated by considering the concentrated log-likelihood function, which up to a constant term is given by

$$logL = -\frac{T}{2}tr\Omega^{-1} \left(S_{00} - S_{01}\beta^*\alpha' - \alpha\beta^{*'}S_{10} + \alpha\beta^{*'}S_{11}\beta^*\alpha' \right), \tag{24}$$

where

$$S_{ij} = T^{-1} \left(Z_i'Z_i - Z_i'Z_2 \left[Z_2'Z_2 \right]^{-1} Z_2'Z_j \right), i, j = 0, 1, \tag{25}$$

with Z_0, Z_1 and Z_2 the data matrices of $Z_{0,t}$, $Z_{1,t}$ and $Z_{2,t}$, respectively.

Theorem 4.2 *The conditional estimator for β is*

$$\widehat{\beta^*} = \left[\alpha' \Omega^{-1} \alpha \otimes S_{11} \right]^{-1} \left[\alpha' \Omega^{-1} \otimes I_{n+1} \right] vec S_{10}.$$

Proof. The derivative of (34) (with respect to $vec\beta^*$) is

$$\frac{\partial \log L}{\partial vec\beta^*} = Tvec \left[S_{10} \Omega^{-1} \alpha - S_{11} \beta^* \alpha' \Omega^{-1} \alpha \right] \tag{26}$$

Setting this equal to zero, and solving for β^* yields the (conditional) estimator in the theorem. ∎

Now, consider how to estimate the system. First, suppose β^* and γ were known. Then the remaining coefficients could be estimated by applying ordinary least square in (4). Let $l \left(\widehat{\alpha} \left(\beta^*, \gamma \right), \beta^*, \gamma', \widehat{\Gamma}_1 \left(\beta^*, \gamma \right), ..., \widehat{\Gamma}_{p-1} \left(\beta^*, \gamma \right), \widehat{\Omega} \left(\beta^*, \gamma \right) \right)$ be the corresponding likelihood value.

Problem 4.1 *The maximum likelihood estimates for β^* and γ are the coefficient matrices that solves*

$$\max_{\beta^*, \gamma} \left\{ l \left(\widehat{\alpha} \left(\beta^*, \gamma \right), \beta^*, \gamma', \widehat{\Gamma}_1 \left(\beta^*, \gamma \right), ..., \widehat{\Gamma}_{p-1} \left(\beta^*, \gamma \right), \widehat{\Omega} \left(\beta^*, \gamma \right) \right) \right\}.$$

The solution to this problem must imply that (the empirical counterparts to) (20) and (26) are equal to zero, i.e. that the first order conditions are satisfied.

Problem 4.1 describes how GRaM considers the estimation problem. The program tries to maximize the log-likelihood, and applies the two first order conditions (20) and (26) to increase the speed and probability to reach the maximum.

4.2 The estimation problem with restrictions

In this section the estimation problem with restrictions is described.

On the γ matrix all linear restrictions among the elements are allowed. Therefore, the restrictions on γ can be written as

$$R_\gamma^{v\prime} vec\gamma' = 0. \tag{27a}$$

The restrictions on γ can alternatively be written as

$$vec\gamma' = H_\gamma^v \cdot \phi_\gamma^v, \tag{27b}$$

where $H_\gamma^v = \left(R_\gamma^v \right)_\perp$.

In the present version of GRaM only restrictions on the space spanned by the coin-

tegration space are allowed. Define $\beta^* = (\beta', -\mu)'$ and $X_t^* = (X_t', 1)'$, so restrictions on the cointegration mean levels also can be imposed. The restrictions on β^* can be written as[10]

$$R_\beta^{m\prime} \beta^* = 0. \tag{28a}$$

Equivalently, we can write

$$\beta^* = H_\beta^m \cdot \phi_\beta^m, \tag{28b}$$

where $H_\beta^m = \left(R_\beta^m \right)_\perp$.

We may also want to impose restrictions on ρ. Let these restrictions be written as

$$R_\rho^{m\prime} \rho' = 0, \tag{29a}$$

which can be transformed into restrictions on β as

$$\left[R_\rho^{m\prime} \gamma' J \right] \beta^* = 0, \tag{29b}$$

where $J = (I_n, 0_{n\times 1})$, or into restrictions on γ as

$$\left[\beta' \otimes R_\beta^{m\prime} \right] vec\gamma' = 0. \tag{29c}$$

Now, consider how to estimate the system. First, suppose β^* and γ were known. Then the remaining coefficients could be estimated by applying ordinary least square on (4) on page 6. Let $l\left(\widehat{\alpha}\left(\phi_\beta^m, \phi_\gamma^v \right), H_\beta^m \cdot \phi_\beta^m, H_\gamma^v \cdot \phi_\gamma^v, \widehat{\Gamma}_1\left(\phi_\beta^m, \phi_\gamma^v \right), ..., \widehat{\Gamma}_{p-1}\left(\phi_\beta^m, \phi_\gamma^v \right), \right.$ $\left. \widehat{\Omega}\left(\phi_\beta^m, \phi_\gamma^v \right) \right)$ be the corresponding likelihood value.

Problem 4.2 *(General problem with matrix restrictions) The maximum likelihood coefficients for β, γ and ρ can be found by (27b), (28b) and $\rho = \beta'\gamma$ respectively, where we use the coefficient matrices that solves*

$$\max_{\phi_\beta^m, \phi_\gamma^v} \left\{ l\left(\widehat{\alpha}\left(\phi_\beta^m, \phi_\gamma^v \right), H_\beta^m \cdot \phi_\beta^m, H_\gamma^v \cdot \phi_\gamma^v, \widehat{\Gamma}_1\left(\phi_\beta^m, \phi_\gamma^v \right), ..., \widehat{\Gamma}_{p-1}\left(\phi_\beta^m, \phi_\gamma^v \right), \widehat{\Omega}\left(\phi_\beta^m, \phi_\gamma^v \right) \right) \right.$$
$$\left. subject\ to\ \left[\phi_\beta^{m\prime} \cdot H_\beta^{m\prime} \cdot J' \otimes R_\rho^{m\prime} \right] H_\gamma^v \phi_\gamma^v = 0 \right\}.$$

The solution must imply that (the empirical counterparts to) the following first order con-

---

ditions holds:

$$H_\gamma^{v\prime} M' \left[vec \left(\Phi^{*\prime} \Omega^{-1} \Phi S_{ZD} - \Phi^{*\prime} \Omega^{-1} \alpha \beta^{*\prime} S_{YD} - \Phi^{*\prime} \Omega^{-1} \Phi^* \left(I_{p+1} \otimes \gamma \right) S_{DD} \right) \right]$$
$$- \lambda \left[\left(\beta^{*\prime} J' \otimes R_\rho^{m\prime} \right) H_\gamma^{v\prime} \right] \tag{30a}$$
$$= 0$$

$$vec \left[H_\beta^{m\prime} S_{10} \Omega^{-1} \alpha + H_\beta^{m\prime} S_{11} \beta^* \alpha' \Omega^{-1} \alpha \right]$$
$$- \lambda' \left[I_r \otimes R_\rho^{v\prime} \gamma' J H_\beta^m \right] = 0 \tag{30b}$$

$$R_\rho^{m\prime} \left(\gamma' \beta \right) = 0 \tag{30c}$$

where λ is the vector of Lagrange multipliers.

To apply Solution 4.2 we need an algorithm that allows restrictions. Alternatively, we could transform the restrictions on ρ to restrictions on γ or β.

Remark 4.2.1 *Restrictions on ρ can be transformed to restrictions on γ by applying (29c). The joint set of restrictions imposed on γ is therefore*

$$\begin{pmatrix} R_\gamma^{v\prime} \\ \beta' \otimes R_\rho^{m\prime} \end{pmatrix} vec\gamma' = 0, \tag{31}$$

and we can apply Solution 4.2 (without the restriction) with the modification that H_γ^v is the orthogonal compliment to the transposed matrix in the parenthesis in (31).

Remark 4.2.2 *Restrictions on ρ can be transformed to restrictions on β by applying (29b). The joint set of restrictions imposed on β is therefore*

$$\begin{pmatrix} R_\beta^{m\prime} \\ R_\rho^{m\prime} \gamma' J \end{pmatrix} \beta^* = 0. \tag{32}$$

where $J = (I_n, 0_{n \times 1})$. The system can now be maximized by applying Solution 4.2 (without the restriction) where H_β^m is the orthogonal compliment to the transposed matrix in the parenthesis in (32).

4.3 Conditional Reduced Rank Regression

Here it will be shown how the remaining coefficients may be estimated under the restrictions on β and ρ, given γ. When γ is known, possible restrictions on ρ must be

transformed into restrictions on β. Therefore, the total set of restrictions are given as in (32). The H matrix in (28b) is therefore defined as the orthogonal compliment of the transformed of the parentheses in (32).

Let $Z_{0,t}(\gamma) = \Delta Y_t - \gamma \Delta D_t$, $Z_{1,t}(\gamma) = \left[(Y_{t-1} - \gamma D_{t-1})', 1\right]'$ and $Z_{2,t}(\gamma) = vec\left(\Delta Y_{t-1} - \gamma D_{t-1}, ..., \Delta Y_{t-p+1} - \gamma D_{t-p+1}\right)$. The system becomes

$$Z_{0,t} = \alpha \beta^{*\prime} Z_{1,t} + \Theta Z_{2,t} + \varepsilon, \tag{33}$$

where $\Theta = \left(\Gamma_1, ..., \Gamma_{p-1}\right)$.

An estimator for β^* in the equation above could be found by applying reduced rank regression, see e.g. Anderson (1951) or Johansen (1995). However, here only a conditional estimator for β^* is considered.

The equation above can be estimated by considering the concentrated log-likelihood function, which up to a constant term is given by

$$logL = -\frac{T}{2} tr\Omega^{-1}\left(S_{00} - S_{01}\beta^*\alpha' - \alpha\beta^{*\prime}S_{10} + \alpha\beta^{*\prime}S_{11}\beta^*\alpha'\right), \tag{34}$$

where

$$S_{ij} = T^{-1}\left(Z_i'Z_i - Z_i'Z_2\left[Z_2'Z_2\right]^{-1}Z_2'Z_j\right), i,j = 0,1, \tag{35}$$

with Z_0, Z_1 and Z_2 the data matrices of $Z_{0,t}$, $Z_{1,t}$ and $Z_{2,t}$, respectively.

As shown in Johansen (1995, p. 91) the estimator for α and Ω (conditioned on β and the supposed known γ) are given by

$$\widehat{\alpha} = S_{01}\beta^*\left(\beta^{*\prime}S_{11}\beta^*\right)^{-1}, \tag{36}$$

$$\widehat{\Omega} = S_{00} - S_{01}\beta^*\left(\beta^{*\prime}S_{11}\beta^*\right)^{-1}\beta^{*\prime}S_{10}. \tag{37}$$

With the restrictions $\beta^* = H_\beta \cdot \phi_\beta$ imposed, minimizing the determinant of Ω corresponds to solving the following eigenvalue problem (see Johansen (1995, pp. 91-92 and Theorem 7.2 on p. 107))

$$\left|\lambda^* H_\beta' S_{11} H_\beta - H_\beta' S_{10} S_{00}^{-1} S_{01} H_\beta\right| = 0, \tag{38}$$

for the r highest eigenvalues $1 > \lambda_1^* > ... > \lambda_r^*$ with the corresponding eigenvectors $(v_1, ..., v_r)$. Then we use $\widehat{\beta^*} = H_\beta \cdot \widehat{\phi_\beta}$ with $\widehat{\phi_\beta} = (v_1, ..., v_r)$ as the estimator.

4.4 Starting values

In GRaM you can choose between different methods for computing starting values. The default method is descried in Section 4.4.3.

Some, but not all, coefficient restrictions are taken into account when the starting values are chosen. All restrictions imposed on γ are taken into account. This means that we are really considering starting values of ϕ_γ^v, and the corresponding starting values for γ is therefore given by (27b).

Also the restrictions on β as given in (28a) are taken into account.

Restrictions on ρ are not taken into account when starting values are computed if the chosen simulation algorithm is MaxSQP. However, if MaxBFGS or MaxSA are used as starting values, the restrictions on ρ are transformed into restrictions on the cointegration space (if possible). Therefore the joint set of restrictions that are taken into account when the starting values for β are calculated are given by (28b), where H_β^m is given by the orthogonal compliment to the transformed matrix in the parenthesis in (32).

4.4.1 Zeros (or almost zero)

Here, the starting values of the elements in the matrix γ are zero or close to zero.[11] (If restrictions are imposed on γ, the elements in ϕ_γ^v are set equal to (or close to) zero, and (27b) is used to calculate the corresponding starting values for γ.)

Based on these starting values for γ, GRaM apply (14) on page 21 to de-trend the data series. With these de-trended variables estimates for β are identified by normal reduced rank. Restrictions on β formed as in (28b) are imposed when estimating starting values for β. Depending on the simulation algorithm that is used, also restrictions via ρ may be imposed.

4.4.2 As for rank = 0

Here GRaM calculates starting values in two steps. In the first step the starting values to γ are set equal to (or close to, se above) zero (but no starting values for β are calculated. In the second step GRaM use the MaxBFGS algorithm to simulate for new starting values for γ in the system where cointegration rank equal to zero is imposed.

---

[11]Sometimes it can be problematic to use starting values of the elements in γ equal to zero. When there are restrictions on ρ, these restrictions are allays fulfilled when γ consists of zeros only, see (29a). Therefore, when γ equals zero restrictions on ρ will not involve any restrictions on β. However, when the simulation algorithm tries to change value of the elements in γ slightly, the restrictions on ρ will imply restrictions on β. The algorithm may therefore not move from the situation where all the elements in γ equals zero. This problem is solved by choosing starting values not equal (but close) to zero.

(Therefore are restrictions on β and ρ not taken into account. However, the restrictions on γ are imposed.)

Based on the starting values for γ from the first step, starting values for β are computed by reduced rank. Her the restrictions on the cointegration space (i.e. equation (28b)) or the joint set of restrictions on β and ρ given by Remark 4.2.2 are imposed, depending on which simulation algorithm is used. (The former is used when MaxSQP is chosen and the latter is used when MasBFGS or MaxSA are chosen.)

4.4.3 OLS - Static regression (gamma restrictions imposed)

Here the starting values of γ are found by estimating

$$Y_t = constant + \gamma \cdot D_t + u_t \tag{39}$$

when the restrictions given by (27a) are imposed on γ. Only a first round estimate is found, where the identity matrix is used as a proxy for the covariance matrix of the errors.

4.4.4 OLS - Static regression (rho restrictions imposed)

Also here starting values of γ are found by estimating (39), but here the restrictions on γ are combined with restrictions on the space of ρ, i.e.

$$\begin{pmatrix} R_\gamma^{v'} \\ I_n \otimes R_\rho^{m'} \end{pmatrix} vec\gamma' = 0. \tag{40}$$

Since we are imposing such strong restrictions on γ, we only need to impose the restrictions on the cointegration space (see equation (28b)) independent of which simulation algorithm we use. The restrictions implied by (29a) are already imposed on γ.

4.4.5 JMN/SL - Johansen, Mosconi, Nilsen and Saikkonen, Lütkepohl

The system in (4) on page 6 can be rewritten as

$$\Delta y_t = \alpha \left(\beta' Y_{t-1} - \mu - \rho D_{t-1} \right) + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \left[\gamma \Delta D_t - \sum_{i=1}^{p-1} \Gamma_i \gamma \Delta D_{t-i} \right] + \varepsilon_t \tag{41}$$

Use $\beta^* = H_\beta^m \cdot \phi_\beta^m$, $\gamma' = H_\gamma^m \cdot \phi_\gamma^m$, and $\rho' = H_\rho^m \cdot \phi_\rho^m$, where $H_\beta^m = \left(R_\beta^m \right)_\perp$, $H_\gamma^m =$

$\left(R_\gamma^m \right)_\perp$, and $H_\rho^m = \left(\left(R_\rho^{m\prime}, R_\gamma^{m\prime} \right)^\prime \right)_\perp$. Note that the special construction of H_ρ^m secures that all restrictions on γ also are imposed on ρ, i.e. $span\left(H_\rho^m \right) \subset span\left(H_\gamma^m \right)$.[12]

With these restrictions the system can be rewritten as

$$
\Delta Y_t = \alpha \left(\phi_\beta^{m\prime}, \phi_\rho^{m\prime} \right) \begin{pmatrix} H_\beta^{m\prime} Y_{t-1}^* \\ \left[H_\rho^{m\prime} D_{t-1} \right]_1 \end{pmatrix}
$$

$$
+ \left(\Gamma_1, ..., \Gamma_{p-1}, \Psi_0, \Psi_1, ..., \Psi_{p-1} \right) \begin{pmatrix} \Delta Y_{t-1} \\ \vdots \\ \Delta Y_{t-p+1} \\ \begin{bmatrix} H_\gamma^\prime \Delta D_t \\ H_\gamma^\prime \Delta D_{t-1} \\ \vdots \\ H_\gamma^\prime \Delta D_{t-p+1} \end{bmatrix}_0 \end{pmatrix} + \varepsilon_t,
$$

with some restrictions on the coefficient matrices Ψ_i (which are ignored when computing the starting values).

When the coefficient restrictions in the Ψ's are ignored, the system above may include linear dependent variables. Therefore, let $[...]_0$ only include linearly independent variables (i.e. exclude variables that are linearly dependent). Furthermore, in $[...]_1$, exclude all variables that are linearly dependent of other variables in $[...]_0$ or $[...]_1$.

Now α, β, Γ_i ($i = 1, ..., p - 1$) and Ω can be estimated by applying reduced rank technics, see e.g. Johansen (1995, pp. 89-93). Based on these estimates, γ can be estimated by[13]

$$
\widehat{vec\gamma}^\prime \left(\alpha, \beta^*, \Gamma_1, ..., \Gamma_{p-1}, \Omega \right) = H_\gamma^v \left[H_\gamma^{v\prime} M^\prime \left(S_{DD} \otimes \Phi^{*\prime} \Omega^{-1} \Phi^* \right) M H_\gamma^v \right]^{-1} \tag{42}
$$
$$
\times \left[H_\gamma^{v\prime} M^\prime vec \left(\Phi^{*\prime} \Omega^{-1} \Phi S_{ZD} - \Phi^{*\prime} \Omega^{-1} \alpha \beta^{*\prime} S_{YD} \right) \right],
$$

where the same notation as in Theorem 4.1 on page 26 is used. The proof is similar to the proof of Theorem 4.1.

If MaxSQF is chosen as the maximizing algorithm, the estimates of β^* and γ achieved by the procedure above is used as starting values. If MaxBFGS or MaxSA is used as

---

[12]See Johansen et al. (2000, Section 4) for a discussion on the problems when this span restriction is not fulfilled.

[13]To use the equation below one needs an estimate of β^* (and not only β). The present version of GRaM use $\hat{\mu} = \sum_{t=1}^T \hat{\beta}^\prime Y_{t-1}$. However, $\hat{\mu} = \sum_{t=1}^T \hat{\beta}^\prime Y_{t-1} - \hat{\phi}_\rho^{m\prime} \left[H_\rho^{m\prime} D_{t-1} \right]_1$ will probably be a more precise estimate.

maximizing algorithm, β^* is estimated again where also the restrictions on ρ are imposed.

4.5 Simulation

If MaxSQP is used as the simulation algorithm the log-likelihood function formulated in Problem 4.2 is maximized, and the first order conditions (30a) - (30c) are used to making the algorithm analytical.[14]

If MasBFGS or MaxSA is used as the simulation algorithm, GRaM uses a modification of Problem 4.2 described in the remarks. Remark 4.2.1 is used if possible (i.e. when there are not so many restrictions on β^* and ρ that the orthogonal complement to the parenthesis (32) is not defined), Remark 4.2.2 otherwise.

4.6 Switching

In Section 4.1 it was shown that the system could be written as

$$\Phi Z_t = \alpha \beta^{*\prime} Y_{t-1}^* + \Phi^* \left(I_{p+1} \otimes \gamma \right) D_t^v + \varepsilon_t, \tag{43}$$

where $Z_t = vec \left(\Delta Y_t, ..., \Delta Y_{t-p+1} \right)$, $D_t^v = vec \left(D_t, ..., D_{t-p} \right)$, $\Phi = \left(I_n, -\Gamma_1, ..., -\Gamma_{p-1} \right)$, and $\Phi^* = \left(I_n, -\Gamma_1^* - \alpha \beta^\prime, -\Gamma_2^*, ..., -\Gamma_p^* \right)$.

Alternatively, the system could be written as

$$Z_{0,t} = \alpha \beta^{*\prime} Z_{1,t} + \Theta Z_{2,t} + \varepsilon, \tag{44}$$

where $Z_{0,t} (\gamma) = \Delta Y_t - \gamma \Delta D_t$, $Z_{1,t} (\gamma) = \left[(Y_{t-1} - \gamma D_{t-1})^\prime, 1 \right]^\prime$, $Z_{2,t} (\gamma) = vec (\Delta Y_{t-1} - \gamma D_{t-1}, ..., \Delta Y_{t-p+1} - \gamma D_{t-p+1})$ and $\Theta = \left(\Gamma_1, ..., \Gamma_{p-1} \right)$.

Furthermore, define

$$
\begin{aligned}
S_{DD} &= T^{-1} \sum D_t^v D_t^{v\prime}, \\
S_{ZD} &= T^{-1} \sum Z_t D_t^{v\prime}, \\
S_{YD} &= T^{-1} \sum Y_{t-1}^* D_t^{v\prime},
\end{aligned}
$$

---

[14]The restrictions (30a) and (30b) are divided into two parts: One part that corresponds to the maximizing problem without restrictions on ρ, and one part based on the restrictions $R_\rho^{m\prime} (\beta^\prime \gamma) = 0$, see description of MaxSQP in Ox (Doornik, 2001).

(and similarly for S_{ZZ} and $S_{YZ} = S'_{ZY}$),

$$M_{ij} = T^{-1} \sum_{t=1} TZ_{it}Z'_{jt} \quad, i,j = 0,1,2,$$

$$S_{ij} = M_{ij} - M_{i2}M_{22}^{-1}M_{2j} \quad, i,j = 0,1,$$

and implicitly define the matrix M with dimension $q(p+1)^2 n \times qn$ by[15]

$$vec\left(I_{p+1} \otimes \gamma\right) = Mvec\gamma'.$$

Theorem 4.3 *The conditional maximum likelihood estimators for β, α, $\Theta \equiv (\Gamma_1, ..., \Gamma_{p-1})$, γ and Ω under the restrictions* (27a), (28a) *and* (29a) *are given by*

$$\widehat{\beta^*} = H_\beta \cdot \widehat{\phi_\beta} \tag{45}$$

$$\widehat{\alpha} = S_{01}\beta^* \left(\beta^{*\prime}S_{11}\beta^*\right)^{-1} \tag{46}$$

$$\widehat{\Theta} = M_{02}\left(M_{22}\right)^{-1} - \alpha\beta^{*\prime}M_{12}\left(M_{22}\right)^{-1} \tag{47}$$

$$\begin{aligned}\widehat{vec\gamma'} &= H_\gamma^v\left[H_\gamma'M'\left(S_{DD} \otimes \Phi^{*\prime}\Omega^{-1}\Phi^*\right)MH_\gamma\right]^{-1} \\ &\times \left[H_\gamma'M'vec\left(\Phi^{*\prime}\Omega^{-1}\Phi S_{ZD} - \Phi^{*\prime}\Omega^{-1}\alpha\beta^{*\prime}S_{YD}\right)\right]\end{aligned} \tag{48}$$

$$\widehat{\Omega} = S_{00} - \alpha\beta^{*\prime}S_{10} - S_{01}\beta^*\alpha' + \alpha\beta^{*\prime}S_{11}\beta^*\alpha\prime \tag{49}$$

where $\widehat{\phi_\beta}$ is the solution of the eigenvalue problem in Section 4.3, and

$$H_\beta = \left(R_\beta^m\right)_\perp \quad and \quad H_\gamma = \left(R_\gamma^v, \left(\beta' \otimes R_\rho^m\right)\right)_\perp \tag{I}$$

or

$$H_\beta = \left(R_\beta^m, \left(J'\gamma R_\rho^m\right)\right)_\perp \quad and \quad H_\gamma = \left(R_\gamma^v\right)_\perp \tag{II}$$

or

$$H_\beta = \left(R_\beta^m, \left(J'\gamma R_\rho^m\right)\right)_\perp \quad and \quad H_\gamma = \left(R_\gamma^v, \left(\beta' \otimes R_\rho^m\right)\right)_\perp \tag{III}$$

with $J = (I_n, 0_{n\times 1})$.

Proof. The equations (45), (46) and (49) are shown in Section 4.3; for (47) see Johansen (1995, Section 6.1) and (48) see Theorem 4.1 on page 26. ∎

---

[15]See footnote 9 on page 26 for the explicit form of M.

From (I) - (III) it can be seen that the restrictions on ρ are imposed on either β^* or γ, or on both.

The advantage of using (III) in an algorithm is that is secures that the likelihood value increases for each iteration. However, it might reduce the parameter space the algorithm can search.

The advantage of using (I) or (II) in an algorithm is that it seeks a broader parameter space than if (III) is used. However, it might lead to a decrease in the likelihood value.

GRaM use the following algorithm:

Algorithm 4.1 *Estimators for β, α, $\Theta \equiv (\Gamma_1, ..., \Gamma_{p-1})$, γ and Ω under the restrictions* (27a), (28a) *and* (29a) *may be obtained by the following iterative procedure, starting from a set of initial values* $\{\beta, \alpha, \Theta, \gamma, \Omega\}$:

$$
\begin{aligned}
\widehat{\beta^*}_i &= \beta\left(\hat{\alpha}_{i-1}, \hat{\gamma}_{i-1}, \hat{\Omega}_{i-1}; H_\beta\right) \\
\widehat{\alpha'}_i &= \alpha\left(\hat{\beta}_i, \hat{\gamma}_{i-1}, \hat{\Omega}_{i-1}\right) \\
\widehat{\Theta}_i &= \Theta\left(\hat{\beta}_i, \hat{\gamma}_{i-1}, \hat{\Omega}_{i-1}\right) \\
\widehat{vec\gamma'}_i &= \gamma\left(\hat{\beta}_i, \hat{\alpha}_i, \hat{\Theta}_i, \hat{\Omega}_{i-1}; H_\gamma\right) \\
\widehat{\Omega}_i &= \Omega\left(\hat{\beta}_i, \hat{\alpha}_i, \hat{\gamma}_i\right) \\
i &= 1, 2, ...
\end{aligned}
$$

Use (I). *However, if an iteration leads to a decreased likelihood, use the estimates from the previous iteration and use* (III) *in this iteration. In the next iteration, use* (I), *unless this leads to a decreased likelihood. (And so forth...) Continue until convergence.*

4.7 Standard errors

GRaM computes standard errors. The standard errors for γ are based on the conditional variance matrix for γ. With given estimates of $\alpha, \beta^*, \Gamma_1, ..., \Gamma_{p-1}$ and Ω, the conditional variance matrix for γ under the restriction (27a) becomes

$$
V\left(\widehat{vec\gamma'}\right) = T^{-1} H \left[H'M'\left(S_{DD} \otimes \Phi^{*\prime}\Omega^{-1}\Phi^*\right) MH \right]^{-1} H',
$$

where S_{DD} is defined elsewhere. The standard error for $vec\left(\gamma'\right)$ are the square root of the diagonal elements in this matrix.

The conditional variance matrix for β^* is shown in Doornik (1995) to be given by

$$V\left(\widehat{vec}\beta^*\right) = T^{-1}\left(I_r \otimes H_\beta^m\right)\left[\left(I_r \otimes H_\beta^m\right)'\left(\alpha'\Omega^{-1}\alpha \otimes S_{11}\right)\left(I_r \otimes H_\beta^m\right)\right]^{-1}\left(I_r \otimes H_\beta^m\right)',$$

(with S_{11} defined elsewhere) and the standard error for $vec\beta^*$ are given by the square root of the diagonal elements in this matrix.

Since $\rho = \beta'\gamma$, the variance of ρ depends on the variances of both β and γ. It is therefore simpler to calculate the variance of ρ conditioned of β or γ. Note that $\rho' = \overline{H_\rho}H_\rho'\rho' + \overline{R_\rho}R_\rho'\rho'$ where $\overline{H} = H\left(H'H\right)^{-1}$ with $H_\rho = \left(R_\rho\right)_\perp$. Since $R_\rho'\rho' = 0$ and $\rho = \beta'\gamma$, we have $\rho' = \overline{H_\rho}H_\rho'\gamma'\beta$. Vectorizing this yields $vec\rho' = \left(I_r \otimes \overline{H_\rho}H_\rho'\gamma'\right)vec\beta$. Then the 'variance' of ρ becomes

$$V\left(\widehat{vec}\rho'\right) = \left(I_r \otimes \overline{H_\rho}H_\rho'\gamma'\right)V\left(vec\beta\right)\left(I_r \otimes \gamma'\overline{H_\rho}H_\rho'\right)$$

if we condition on γ. The standard error for $vec\left(\rho'\right)$ are the square root of the diagonal elements in this matrix.

Note that the standard errors of β^* and ρ are only computed if appropriate normalization restrictions are imposed on the cointegration space.

4.8   Distribution for the likelihood tests

When GRaM tests restrictions, it applies the χ^2-distribution. For restrictions on β it has been shown in e.g. Boswijk (1995) or Boswijk and Doornik (2003) that this is the correct distribution. Boswijk and Doornik (2003) also shows that restrictions on β^* is χ^2 in the H_c case, (i.e D_t is empty, or at least does not include a trend). In the H_l case μ is not included in the cointegration vectors when estimating cointegrated VAR models in its normal form (1), and therefore similar results are not available.

For restrictions on γ Johansen et al. (2000) shows that (at least a subset of) restrictions of the form $R'\gamma'$ are χ^2-distributed.

Restrictions on ρ can be reformulated into restrictions on β or γ; so if restrictions on β and γ can be tested based on a χ^2-distribution, restrictions on ρ can be tested based on the same distribution as well.

In order to apply the χ^2 test one needs to know the degrees of freedom in the test, i.e. how many independent restrictions that are imposed. The number of restrictions on γ is given by $rank\left(R_\gamma^v\right)$.

If there are no restrictions on β^* there are $(n+1-r)r$ independent elements i β^*. If restrictions are imposed, Boswijk (1995) shows that the number of independent ele-

ments in β^* is given by

$$rank\left[\left(I_r \otimes \beta_\perp^{*\prime}\right) H_\beta\right].$$

If both restrictions on β and γ are considered;

$$H_\beta^v = \left(R_\beta^v, \left(I_r \otimes J'\gamma\right) R_\rho^v\right)_\perp.$$

The number of independent restrictions imposed on β^* and ρ is therefore given by $(n+1-r)r - rank\left[\left(I_r \otimes \beta_\perp^{*\prime}\right) H_\beta\right]$.

The degrees of freedom used in the χ^2 test is the sum of restrictions on γ and the restrictions on β^* and ρ, i.e. the degrees of freedom are given by

$$(n+1-r)r - rank\left[\left(I_r \otimes \beta_\perp^{*\prime}\right) H_\beta\right] + rank\left(R_\gamma^v\right).$$

References

Anderson, T. W. (1951), 'Estimating linear restrictions on regression coefficients for multivariate normal distributions', *Annals of Mathematical Statistics* **22**, 327–351.

Boswijk, H. P. (1995), 'Identifibility of cointegrated systems', *Discussion Paper ti 7-95-078, Tinberger Institute, University of Amsterdam* .

Boswijk, H. P. and J. A. Doornik (2003), Identifying, estimating and testing restricted cointegrated systems: An overview. http://www1.fee.uva.nl/ke/boswijk/BoswijkDoornik.pdf.

Doornik, J. A. (1995), Testing general restrictions on the cointegrating space. Nuffield College, Oxford OX1 INF, UK.

Doornik, J. A. (2001), *Object-Oriented Matrix Programming using Ox.*, London: Timberlake Consultants Press.

Doornik, J. A. and D. F. Hendry (2001), *Modelling Dynamic Systems Using PcGive, Volume II*, Timberlake Consumltats ltd. London.

Goffe, W. L., G. D. Ferrier and J. Rogers (1994), 'Global optimalization of statistical functions with simulated annealing', *Journal of Econometrics* **60**, 65–99.

Hendry, D. F. and G. E. Mizon (1998), 'Exogeneity, causality, and co-breaking in economic policy analysis of a small econometric model of money in the UK', *Empirical Economics* **23**, 267–294.

Hungnes, H. (2002), 'Restricting growth rates in cointegrated VAR models', *Revised version of Discussion Papers 309, Statistics Norway* . (Downloadable at http://folk.ssb.no/hhu).

Hungnes, H. (2005), 'Structural breaks in cointegrated VAR models', *unp.* .

Johansen, S. (1995), *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford University Press.

Johansen, S., R. Mosconi and B. Nielsen (2000), 'Cointegration analysis in the presence of structural breaks in the deterministic trend', *Econometrics Journal* **3**, 216–249.

Lütkepohl, H. and J. Wolters (1998), 'A money demand system for german M3', *Empirical Economics* **23**, 371–386.

Lütkepohl, H. (1996), *Handbook of Matrices*, New York: John Wiley & Sons.

Magnus, J. R. and H. Neudecker (1988), *Matrix Differential Calculus wih Applications in Statistics and Econometrics*, New York: John Wiley & Sons.

Saikkonen, P. and H. Lütkepohl (2000), 'Testing for the cointegrating rank of a VAR process with structural shifts', *Journal of Buisness & Economic Statistics* **18**, 451–464.