Written Exam for the M.Sc. in Economics, Winter 2014/2015

ADVANCED MACROECONOMETRICS

Remarks on the solution

About the Exam

This project examination deals with an econometric analysis of country C, inspired by the IS-LM model as suggested by the two equations,

$$\begin{aligned} y_t - \gamma_0 t &= \gamma_1 \left(R_{bt} - \pi_t\right) && (1) \\ m_t - \gamma_2 y_t &= \gamma_3 \left(R_{bt} - R_{mt}\right). && (2) \end{aligned}$$

All assignments are based on *different* data sets for the $p = 5$ relevant variables observed quarterly, 1975:1-2012:4. The data series are simulated from a cointegrated VAR(2) as given by,

$$\Delta x_t = \alpha\beta' x_{t-1} + \Gamma_1 \Delta x_{t-1} + \alpha\rho' t + \mu + \epsilon_t,$$

with

$$x_t = \begin{pmatrix} m_t \\ y_t \\ \pi_t \\ R_{mt} \\ R_{bt} \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \quad \text{and} \quad \alpha = \begin{pmatrix} -0.18 & 0 & 2.50 \\ 0 & 0 & 0 \\ 0 & 0.10 & 0 \\ 0 & 0 & -0.10 \\ 0 & -0.08 & 0.15 \end{pmatrix},$$

and $\epsilon_t \sim N(0, \Omega)$. Remaining parameters are calibrated to let x_t behave–more or less–as observed time series. There is a structural break imposed on the series in 1990, where the economy entered a monetary union. In addition, there is a number of outlying observations drawn randomly, and a typical data set will have approximately $3 - 4$ outliers.

For all data sets it is ensured that the lag length can be chosen to $k = 2$ (based on the SW information criteria) and if the correct outliers are modelled with dummy variables, the *trace test* for the cointegration rank will correctly suggest a cointegration rank of $r = 3$. In addition, the true structure of the cointegration space is not rejected by a likelihood ratio (LR) test. It is not important *per se* that the students recover the true DGP, it is more important that they use sound arguments and that they convincingly motivate the choices they make.

There are 5 sections with an unequal number of questions and unequal difficulty. I suggest a tentative weight of 20% for each section. The last Section 5 with extensions to the basic analysis is on the boundary of what they have seen. I will suggest to be a bit flexible here (I have not seen any student solutions yet!).

1 Background and Statistical Model

[1] The solution should interpret (1) and (2) as equilibrium relationships and suggest a cointegration rank of $r = 2$ with

$$x_t = \begin{pmatrix} m_t \\ y_t \\ \pi_t \\ R_{mt} \\ R_{bt} \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 0 & 1 \\ 1 & -\gamma_2 \\ \gamma_1 & 0 \\ 0 & \gamma_3 \\ -\gamma_1 & -\gamma_3 \end{pmatrix},$$

which could also be stated with t included.
The Granger representation should be stated in terms of $\beta_\perp$ and $\alpha_\perp$. As a minimum, it should be mentioned that the model suggests $p-r = 3$ common stochastic trends, and the interpretation should be explained in terms of long-run and short-run impact of shocks, e.g. as impulse-response function coefficients. The good solution may suggest a specific $\beta_\perp$, e.g.

$$\beta_\perp = \begin{pmatrix} -\gamma_1\gamma_2 & -\gamma_3 & \gamma_3 + \gamma_1\gamma_2 \\ -\gamma_1 & 0 & \gamma_1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

but I guess they are not used to find orthogonal complements by hand.

[2] If inflation is stationary, then $r = 3$ with

$$\beta = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -\gamma_2 & 0 \\ 0 & 0 & 1 \\ 0 & \gamma_3 & 0 \\ -\gamma_1 & -\gamma_3 & 0 \end{pmatrix} \quad \text{and} \quad \beta_\perp = \begin{pmatrix} -\gamma_3 & \gamma_3 + \gamma_1\gamma_2 \\ 0 & \gamma_1 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

where $\beta_\perp$ now has a zero row, i.e. a zero loading to the stochastic trends.

[3] If $m_t - y_t$, $R_{bt} - R_{mt}$, and $R_{bt} - \pi_t$ are stationary, then $r = 3$ and

$$\beta = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \quad \text{and} \quad \beta_\perp = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

In this case the LM curve holds for the economy with homogeneity with respect to income, $\gamma_2 = 1$, and the semi-elasticity with respect to opportunity costs, γ_3, determined by the adjustment coefficients in α. The IS curve, on the other hand, does not hold as an equilibrium relationship as y_t is an I(1) process. The good solution may elaborate on the implication if y_t was in fact trend-stationary to make the IS curve hold.

[4] The solution should set up and estimate an empirically relevant model. The model should include the deterministic linear trend suggested by the IS curve, and the possibility of a break in 1990 should be considered. The analysis should state and test the assumptions for the model and include dummies for outlying observations. The good solution uses recursive estimation to test the assumption of constancy of the parameters and also uses this to argue for the presence of a structural break.

2 Estimation and Cointegration Rank

[5] The solution should state the Gaussian log-likelihood function,

$$\log L(\theta, \Omega) = -\frac{Tp}{2}\log(2\pi) - \frac{T}{2}\log|\Omega| - \sum_{t=1}^{T}\epsilon_t(\theta)'\Omega^{-1}\epsilon_t(\theta),$$

where $\epsilon_t(\theta) = x_t - \sum_{i=1}^{k}\Pi_i x_{t-i} - \phi D_t$. Using $\hat{\Omega}(\theta) = T^{-1}\sum_{t=1}^{T}\epsilon_t(\theta)\epsilon_t(\theta)'$, this could be rewritten as the concentrated likelihood,

$$\begin{aligned}
\log L(\theta) &= \log L(\theta, \hat{\Omega}(\theta)) \\
&= -\frac{Tp}{2}\log(2\pi) - \frac{T}{2}\log\left|\hat{\Omega}(\theta)\right| - \frac{1}{2}\sum_{t=1}^{T}\epsilon_t(\theta)'\hat{\Omega}(\theta)^{-1}\epsilon_t(\theta) \\
&= -\frac{Tp}{2}\log(2\pi) - \frac{T}{2}\log\left|\hat{\Omega}(\theta)\right| - \frac{1}{2}\sum_{t=1}^{T}\operatorname{tr}\left(\hat{\Omega}(\theta)^{-1}\epsilon_t(\theta)\epsilon_t(\theta)'\right) \\
&= -\frac{Tp}{2}\log(2\pi) - \frac{T}{2}\log\left|\hat{\Omega}(\theta)\right| - \frac{T}{2}\operatorname{tr}\left(\hat{\Omega}(\theta)^{-1}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t(\theta)\epsilon_t(\theta)'\right) \\
&= -\frac{Tp}{2}\left(1 + \log(2\pi)\right) - \frac{T}{2}\log\left|\hat{\Omega}(\theta)\right|,
\end{aligned}$$

where the first term does not depend on θ.

[6] The solution should explicitly derive the error correction form for the empirical model to obtain, e.g.

$$\Delta x_t = \Pi x_{t-1} + \Gamma_1 \Delta x_{t-1} + \phi D_t + \epsilon_t.$$

The solution should then write the characteristic polynomial for the model,

$$|A(z)| = |(1-z) - \Pi z - \Gamma_1 (1-z) z| = 0,$$

and explain that a unit root, $|A(1)| = 0$, implies that $|A(1)| = |\Pi| = 0$, such that Π has reduced rank.

[7] As a reaction to the internet blog, the solution should explain that the cointegrated VAR analysis does not *assume* the presence of unit roots. The general model is an unrestricted VAR, and unit roots impose testable restrictions on the model. The setup also allows some variables to be stationary, as in question 2 above.

[8] The solution should explain how to calculate the LR statistic for reduced rank based on the ordered eigenvalues, $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq ... \geq \hat{\lambda}_p$, solving the usual eigenvalue problem $\left|S_{11}\lambda - S_{10}S_{00}^{-1}S_{01}\right| = 0$, i.e.

$$\mathrm{LR}(r \mid p) = -T \sum_{i=r+1}^{p} \log(1 - \hat{\lambda}_i).$$

Most solutions have probably identified a structural break in which case the asymptotic distribution should be simulated. The solution should explain how the Brownian motions in the limiting distribution are approximated by random walks, and how the distribution is simulated, e.g. using

$$\mathrm{tr}\left\{\sum_{t=1}^{T} \epsilon_t F_t' \left(\sum_{t=1}^{T} F_t F_t'\right)^{-1} \sum_{t=1}^{T} F_t \epsilon_t'\right\},$$

where $\epsilon_t \sim N(0,1)$, $F_t = (\sum_{i=1}^{t-1} \epsilon_i', t, D_t)'$ corrected for a constant, t is a time trend and D_t is a dummy that breaks after the same proportion of observations as it is the case for the empirical model. Evaluating the statistic for many realizations of random sequences of $\epsilon_1, ..., \epsilon_T$, with some large T, allows a characterization of the distribution, for example in terms of quantiles that can be used as critical values for the test.

To determine the cointegration rank, other indicative sources of information may also be considered, e.g. number of characteristic roots close to unity, strength of error correction, visual appearance of linear combinations, or recursive trace tests.

3 Hypotheses Testing

[9] Based on the obtained cointegration rank, the solution should tests if $R_{bt} - R_{mt}$, $R_{bt} - \pi_t$, or π_t are stationary with or without the included deterministic components. The solution should explain how to formulate and test the hypotheses and how to calculate the degrees of freedom. For $r = 3$, stationarity of π_t would correspond to $\beta = (e_3, \varphi_2, \varphi_3)$ where $e_3 = (0, 0, 1, 0, 0)'$ and φ_2, φ_3 unrestricted. This imposes $4 - (r - 1) = 2$ overidentifying restrictions and the LR statistic is asymptotically $\chi^2(2)$ under the null.

[10] Next, the solution should test the hypothesis that one of the stochastic trends affects

only money and income and in the same magnitude, i.e. corresponding to

$$\beta_{\perp} = \begin{pmatrix} 1 & * \\ 1 & * \\ 0 & * \\ 0 & * \\ 0 & * \end{pmatrix},$$

where $*$ denotes an unrestricted coefficient. The solution should explain that a known vector in $\beta_{\perp}$ is implemented as a subspace restriction on β of the form

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \varphi,$$

with $\varphi \in \mathbb{R}^{4\times 3}$ unrestricted. This imposes 3 restrictions and the LR statistic is $\chi^2(3)$.

[11] Next, the solution should test the hypothesis that one of the stochastic trends is determined by cumulated shocks to income, i.e. $CT_{1t} = \sum_{i=1}^{t} \epsilon_{y,i}$. This corresponds to

$$\alpha_{\perp} = \begin{pmatrix} 0 & * \\ 1 & 0 \\ 0 & * \\ 0 & * \\ 0 & * \end{pmatrix} \quad \text{and} \quad \alpha = \begin{pmatrix} * & * & * \\ 0 & 0 & 0 \\ * & * & * \\ * & * & * \\ * & * & * \end{pmatrix},$$

corresponding to weak exogeneity of income. The statistic is $\chi^2(3)$.
Likewise, $CT_{1t} = \sum_{i=1}^{t} (\epsilon_{R_m,i} - \epsilon_{R_b,i})$, corresponds to

$$\alpha_{\perp} = \begin{pmatrix} 0 & * \\ 0 & * \\ 0 & * \\ 1 & * \\ -1 & * \end{pmatrix} \quad \text{and} \quad \alpha = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \\ a & b & c \\ a & b & c \end{pmatrix},$$

which is again $\chi^2(3)$.

[12] Finally, the solution should test that each of the shocks have only transitory effects. For m_t that corresponds to

$$\alpha_{\perp} = \begin{pmatrix} 0 & 0 \\ * & * \\ * & * \\ * & * \\ * & * \end{pmatrix} \quad \text{and} \quad \alpha = \begin{pmatrix} a & * & * \\ 0 & * & * \\ 0 & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix},$$

for some coefficient a, which imposes two overidenifying restrictions.
The solution should note that there can be at most $r = 3$ shocks with only transitory effects, as we need $p - r = 2$ common stochastic trends.

4 Identification

[13] The solution should explain that an identifying structure could be obtained by imposing individual restrictions of the form,

$$\beta^c = (\beta_1^c, \beta_2^c, ..., \beta_r^c) = (H_1\varphi_1, H_2\varphi_2, ..., H_r\varphi_r),$$

where H_j is a known matrix and φ_j is a vector with parameters to be estimated, $j = 1, 2, ..., r$. Generic identification is checked by considering the usual rank conditions. For all i and $k = 1, \ldots, r-1$ and any set of indices $1 \leq i_1 < \cdots < i_k \leq r$ not containing i it holds that

$$R(i.i_1, \ldots, i_k) = rank\left(R_i'\left[H_{i_1} \cdots H_{i_k}\right]\right) \geq k$$

where $R_i = H_{i\perp}$ for all i. The explanation may be less formal, but has to stress that the restrictions imposed on relation i cannot be satisfied by any linear combination of other relations.

[14] Next the solution should impose identifying restrictions. It should explain how to proceed and give an economic interpretation of the long-run relationships and the equilibrium adjustment with reference to the IS-LM framework. The solution may use the automatic approach as implemented in CATSmining, but has to explain the approach taken.

[15] Next it should carefully interpret the Granger representation, i.e. explaining the pushing forces in the system in terms of the stochastic trends and their impact.

[16] If the equation for the short interest rate, R_{mt}, corresponds to the rule-based monetary policy of a central bank, then the residual,

$$\epsilon_{R_m,t} = \Delta R_{mt} - E\left(\Delta R_{mt} \mid x_{t-1}, ..., x_{t-k}\right),$$

measures *unexpected* monetary policy shocks. To test if inflation is not *controllable* by the central bank in the long run, we should test that unexpected monetary policy shocks have no long-term effect on inflation. This can be formulated as an hypothesis on the long-run impact matrix of the Granger representation, i.e. that element 3,4 in $C = \beta_\perp\left(\alpha_\perp'(I - \Gamma_1)\beta_\perp\right)^{-1}\alpha_\perp'$ is zero. The ong-run impact matrix, C, is reported by CATS together with standard errors or t-values for Wald testing.
The solution should also explain that this hypothesis is difficult to test using a likelihood ratio test, because a restriction on a single element in C imposes complicated nonlinear restrictions on the parameters in α and β. Hence, this requires restricted numerical optimization of the restricted likelihood function.

5 Extensions

[17] (Great Moderation) Now, a heteroskedastic model is considered,

$$x_t = \sum_{i=1}^{k} \Pi_i x_{t-i} + \phi D_t + \epsilon_t, \quad \epsilon_t \sim \begin{cases} N(0, \Omega_1) & \text{if} \quad t < 1989:4 \\ N(0, \Omega_2) & \text{if} \quad t \geq 1990:1, \end{cases},$$

where Ω_2 implies smaller variances of individual variables than Ω_1. The likelihood function is now

$$\begin{aligned} \log L(\theta, \Omega) \quad = \quad & -\frac{Tp}{2}\log(2\pi) \\ & -\frac{T_1}{2}\log|\Omega_1| - \sum_{t=1}^{T_1} \epsilon_t(\theta)'\Omega_1^{-1}\epsilon_t(\theta) \\ & -\frac{T-T_1}{2}\log|\Omega_2| - \sum_{t=T_1+1}^{T} \epsilon_t(\theta)'\Omega_2^{-1}\epsilon_t(\theta), \end{aligned}$$

with T_1 corresponding to the observation $1989:4$. The solution should state the maximum likelihood estimators

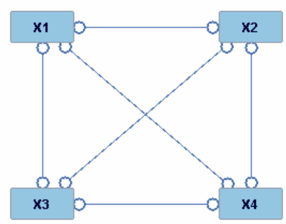
$$\hat{\Omega}_1(\theta) = \frac{1}{T_1}\sum_{t=1}^{T_1} \epsilon_t(\theta)\epsilon_t(\theta)' \quad \text{and} \quad \hat{\Omega}_2(\theta) = \frac{1}{T-T_1}\sum_{t=T_1+1}^{T} \epsilon_t(\theta)\epsilon_t(\theta)',$$

and the very good solution would derive this taking derivatives of the likelihood function. This would correspond to a concentrated likelihood function

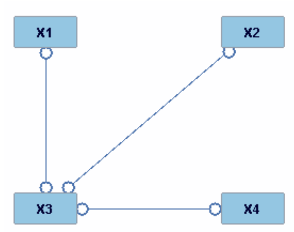
$$\log L(\theta) = -\frac{Tp}{2}(1+\log(2\pi)) - \frac{T_1}{2}\log\left|\hat{\Omega}_1(\theta)\right| - \frac{T-T_1}{2}\log\left|\hat{\Omega}_2(\theta)\right|.$$

There is no closed form estimators for θ.
The impulse response function in the heteroskedastic model (as well as the ordering etc.) depends on Ω, and now the results are different before and after 1990. An impulse response analysis using $\hat{\Omega}_2$ and $\hat{\theta}$ would be standard and the interpretation would be the effect of shocks in the second regime, after the moderation.

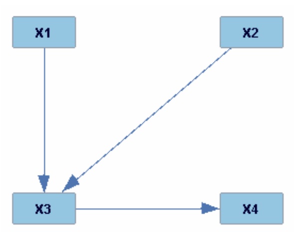
[18] (Inference on Contemporaneous Causal Structures) Using graph theory, we would start with a fully saturated skeleton:



Since $\text{Corr}(\epsilon_{1t}, \epsilon_{2t}) = 0.0543$, $\text{Corr}(\epsilon_{1t}, \epsilon_{4t} \mid \epsilon_{3t}) = 0.0216$, and $\text{Corr}(\epsilon_{2t}, \epsilon_{4t} \mid \epsilon_{3t}) = 0.0046$ are insignificant, we delete the edges $x_1 - x_2$, $x_1 - x_4$, and $x_2 - x_4$, to get an undirected graph:



We observe that $\mathrm{Corr}(\epsilon_{1t}, \epsilon_{2t}) = 0.0543$ is insignificant while $\mathrm{Corr}(\epsilon_{1t}, \epsilon_{2t} \mid \epsilon_{3t}) = -0.1925$ is strongly significant suggesting to orient the triple $x_1 - x_3 - x_2$ as an unshielded collider, $x_1 \rightarrow x_3 \leftarrow x_2$. No additional colliders are found, as $\mathrm{Corr}(\epsilon_{1t}, \epsilon_{4t} \mid \epsilon_{3t}) = 0.0213$ and $\mathrm{Corr}(\epsilon_{2t}, \epsilon_{4t} \mid \epsilon_{3t}) = 0.0026$ are insignificant. This suggests to orient $x_3 \rightarrow x_4$ to avoid a collider which is not supported by the conditional correlation structure. The class of observationally equivalent structures in this case consists of one member, given by the directed graph:



An orthogonal impulse response analysis could be based on a lower-triangular Choleski decomposition of the the residuals covariance matrix, and would depend on the ordering of the variables. The derived contemporaneous causal structure could guide the ordering of the variables, e.g. as $x_t = (x_1, x_2, x_3, x_4)'$ or $x_t = (x_2, x_1, x_3, x_4)'$.

[19] (Measurement Errors) Now we discuss a model with measurement errors:

$$\begin{aligned} \Delta x_t &= \alpha \beta' x_{t-1} + \epsilon_t \\ y_t &= x_t + w_t, \end{aligned}$$

where only y_t is observed, and $\epsilon_t \sim N(0, \Omega)$, $w_t \sim N(0, \Sigma_w)$.
The solution should derive an equation to show the behavior of y_t. We note that $x_t = y_t - w_t$, and

$$\begin{aligned} \Delta (y_t - w_t) &= \alpha \beta' (y_{t-1} - w_{t-1}) + \epsilon_t \\ \Delta y_t &= \alpha \beta' y_{t-1} + \epsilon_t + w_t - \left(I + \alpha \beta'\right) w_{t-1}, \end{aligned}$$

which is an vector ARMA-type process.
To discuss cointegration, one could start with the Granger representation for x_t :

$$x_t = \beta_\perp \left(\alpha'_\perp \beta_\perp\right)^{-1} \alpha'_\perp \sum_{i=1}^{t} \epsilon_i + C^*(L) \epsilon_t + A,$$

where $C^*(L)$ is a convergent matrix polynomial and A depends on initial values. Therefore, we have for y_t,

$$y_t = x_t + w_t = \beta_\perp \left(\alpha'_\perp \beta_\perp\right)^{-1} \alpha'_\perp \sum_{i=1}^{t} \epsilon_i + C^*(L)\epsilon_t + A + w_t.$$

We note that y_t is I(1) due to $\sum_{i=1}^{t} \epsilon_i$, and that

$$\beta' y_t = \beta' C^*(L)\epsilon_t + \beta' w_t,$$

is still stationary such that β is the cointegration matrix of rank r.
If the measurement error was an I(1) process, $w_t = w_{t-1} + \xi_t$, then cointegration would be lost because $\beta' w_t$, and hence $\beta' y_t$ is now an I(1) process.