Written Exam for the M.Sc. in Economics, Winter 2012/2013

ADVANCED MACROECONOMETRICS

Proposed Solution

About the Exam

This project examination deals with econometric models for interest rate linkages between countries in a monetary union. Three countries, A, B and C, are *core countries* while three *peripheri countries*, D, E, and F, where closely linked to the union, but only entered in 1997. The students are also informed that the countries entered a financial crisis in 2007, and may anticipate potential breaks in the structures in 1997 and 2007. The data set consists of monthly observations for the *p.a.* interest rates for the six countries covering 1990:1-2012:12.

All assignments are based on *different* data sets. They all consists of six interest rates collected in the $p = 6$ dimensional data vector,

$$x_t = \left(R_t^A : R_t^B : R_t^C : R_t^D : R_t^E : R_t^F\right)',$$

simulated from a cointegrated VAR(2) process

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

with

$$\beta = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad \alpha = \begin{pmatrix} 0 & 0 & 0 & 0 \\ .3 & 0 & 0 & 0 \\ 0 & .2 & 0 & 0 \\ 0 & 0 & -.1 & -.2 \\ 0 & 0 & .2 & 0 \\ 0 & 0 & 0 & .1 \end{pmatrix},$$

and $\epsilon_t \sim N(0, \Omega)$. The remaining parameters are calibrated to let x_t behave–more or less–as observed interest rates.

There are no structural breaks in the data generating process, but quite large (innovational) outliers at 1997:1 and 2007:6. In addition, outlying observations are drawn randomly, and a typical data set will have approximately $4 - 6$ outliers.

For all data sets it is ensured that the lag length can be chosen to $k = 2$ (based on 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 $r = 4$.

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.

The proposed solution below is based on the data for a tentative exam number 1001 (i.e. `Data1001.xls`).

There are 5 sections with an unequal number of questions and difficulties. I will suggest tentative weight of 20% for Section 1 on scenario analysis and statistical modelling, 15% for the short Section 2 on estimation and rank determination, 25% for the longer Section 3 on hypothesis testing, and 25% for Section 4 on identification. The last Section 5 with extensions to the basic analysis is more difficult, and on the boundary of what they have seen. I will suggest to be a bit flexible here (I have not seen the students solutions yet!) and weight with approximately 15%. External examiners may choose to weight differently.

1 Background and Statistical Model

The solution should discuss scenarios and specify an unrestricted VAR model.

[1] First, the solution should explain that a tentative scenario of the form

$$\begin{pmatrix} R_t^A \\ R_t^B \\ R_t^C \\ R_t^D \\ R_t^E \\ R_t^F \end{pmatrix} = \begin{pmatrix} 1 \\ \xi_B \\ \xi_C \\ \xi_D \\ \xi_E \\ \xi_F \end{pmatrix} \left(\sum_{i=1}^t u_{1i} \right) + S_{1t},$$

can be used to motivate a CVAR. In particular, it may be an example a Granger representation with $p - r = 1$ stochastic trend, and hence $r = 5$ cointegrating relationships. Here $\sum_{i=1}^t u_{1i} = \sum_{i=1}^t \alpha'_\perp \epsilon_i$ is the single stochastic trend, the vector of loadings represents $\tilde{\beta}_\perp$, while $S_{1t} = C^*(L)\epsilon_t$ is a linear (stationary) process.

The cointegration space is not unique but can be chosen to highlight pairwise cointegration, e.g.

$$\beta = \begin{pmatrix} -\xi_B & -\xi_C & -\xi_D & -\xi_E & -\xi_F \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

A central bank may suggest that interest spreads are stationary, i.e. full integration between bond markets, so that $\xi_i = 1$, $i = B, C, ..., F$.

[2] Next the solution should modify the two factor scenario

$$\begin{pmatrix} R_t^A \\ R_t^B \\ R_t^C \\ R_t^D \\ R_t^E \\ R_t^F \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \xi_B & 0 \\ \xi_C & 0 \\ \xi_D & 1 \\ \xi_E & \eta_E \\ \xi_F & \eta_F \end{pmatrix} \begin{pmatrix} \sum_{i=1}^t u_{1i} \\ \sum_{i=1}^t u_{2i} \end{pmatrix} + S_{2t},$$

to allow a third trend driving the spreads of the core countries B and C. This could read

$$\begin{pmatrix} R_t^A \\ R_t^B \\ R_t^C \\ R_t^D \\ R_t^E \\ R_t^F \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \xi_B & 0 & 1 \\ \xi_C & 0 & \kappa_C \\ \xi_D & 1 & 0 \\ \xi_E & \eta_E & 0 \\ \xi_F & \eta_F & 0 \end{pmatrix} \begin{pmatrix} \sum_{i=1}^t u_{1i} \\ \sum_{i=1}^t u_{2i} \\ \sum_{i=1}^t u_{3i} \end{pmatrix} + S_{3t},$$

where the cointegration rank would be $r = 3$.
If $\xi_C = \xi_B$ and $\kappa_C = \kappa_B = 1$ it would hold that $R_t^B - R_t^C \sim I(0)$.
If $\xi_F = \xi_E = \xi_D$ and $\eta_F = \eta_E = \eta_D = 1$, it would hold that $R_t^D - R_t^E \sim I(0)$ and $R_t^D - R_t^F \sim I(0)$.

[3] Now the solution should write the companion form of the VAR, and state that the model is stable if all eigenvalues of the companion matrix are strictly inside the unit circle.

[4] Next, the solution should estimate an unrestricted VAR, include deterministic variables, determine the lag-length, and ensure that the model is well specified.
It is important here that the students explain the steps and motivate the choices they make. A reasonable model could be one with a restricted constant, and the students should be aware of possible level shifts in 1997 and 2007. Some data sets are trending for the current sample, and some students may include a trend. I think this should be accompanied by a sentence stating that this is at most an insample approximation and that a deterministic trend is not really a reasonable model for interest rates. The good solution also performs a recursive estimation to test the assumption of constant parameters.
For the present data set, $k = 2$ lags are sufficient to account for the autoregressive nature of the variables. There are 6 large residuals corresponding to observations: 1997:1, 1997:9, 1998:4, 2001:4, 2001:11, and 2007:6. Two of these correspond to known events, and the importance of a level shift could be tested. For the present data set, the potential level shifts are excludable for (almost) all values of the coin-

tegration rank, and I will not include them in the further analysis:

TEST OF EXCLUSION				
r	DGF	5% C.V.	C(1997:01)	C(2007:06)
1	1	3.841	0.990 [0.320]	2.722 [0.099]
2	2	5.991	4.215 [0.122]	2.832 [0.243]
3	3	7.815	4.221 [0.239]	4.502 [0.212]
4	4	9.488	4.247 [0.374]	7.475 [0.113]
5	5	11.070	13.638 [0.018]	14.988 [0.010]

This is also confirmed from $t-ratios$ on Π. Some students may choose to keep the level shifts and exclude them later, which is also OK, as long as it is discussed and motivated.

[5] Finally the students should inspect and comment on the eigenvalues of the companion matrix.

2 The Cointegration Rank

[6] The solution should explain how to perform ML estimation in the preferred model by solving an eigenvalue problem. This can be reproduced from the book, but here for the preferred model. The solution should state the concentrated regression

$$R_{0t} = \alpha\beta' R_{1t} + \varepsilon_t,$$

where R_{0t} and R_{1t} are OLS residuals of Δx_t and $(x'_{t-1} : 1)'$ on the unrestricted regressors. Next, RRR amounts to solving the eigenvalue problem

$$\left|\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}\right| = 0,$$

where $S_{ij} = T^{-1}\sum_{t=1}^{T} R_{it}R'_{jt}$. That produces p eigenvectors, $\hat{v}_1, ..., \hat{v}_p$ and p corresponding eigenvalues $1 > \hat{\lambda}_1 > ... > \hat{\lambda}_p \geq 0$. The latter can be interpreted as the squared canonical correlations between $\hat{v}'_i R_{it}$ and R_{0t}. The estimate of β is $\hat{\beta} = (\hat{v}_1 : ... : \hat{v}_r)$, and the maximized value of the likelihood function is given by

$$L_{\max}^{-2/T}(H(r)) = |S_{00}| \prod_{i=1}^{r} \left(1 - \hat{\lambda}_i\right).$$

[7] Next the solution should explain how the LR statistic can be calculated from the eigenvalues. This follows directly, as

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

The asymptotic distribution for a true null hypothesis is given by

$$\operatorname{tr}\left\{\int_0^1 dBF(u)'\left(\int_0^1 F(u)F(u)'du\right)^{-1}\int_0^1 F(u)dB'\right\},$$

where $B(u)$ is a Brownian motion in $[0:1]$ and $F(u)$ is the vector $(B(u)':1)'$. This can be simulated by replacing integrals with sums, replacing $B(u)$ with a random walk $\sum_{i=1}^{t-1}\epsilon_i$, such that the stochastic increment dB becomes ϵ_t, i.e.

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

where $F_t=(\sum_{i=1}^{t-1}\epsilon_i':1)'$. Evaluating the statistic for many realizations of random sequences of $\epsilon_1,...,\epsilon_T$, with T large, allows a characterization of the distribution, for example in terms of quantiles that can be used as critical values for the test. The explanation may be less detailed, but the students should have an idea of what is going on.

[8] In the present case, the rank determination produces the following table

I(1)-ANALYSIS							
p-r	r	Eig.Value	Trace	Trace*	Frac95	P-Value	P-Value*
6	0	0.422	377.048	364.690	103.679	0.000	0.000
5	1	0.346	226.818	211.750	76.813	0.000	0.000
4	2	0.188	110.402	103.577	53.945	0.000	0.000
3	3	0.147	53.476	50.242	35.070	0.000	0.000
2	4	0.029	9.882	9.181	20.164	0.656	0.721
1	5	0.006	1.742	1.633	9.142	0.821	0.840

where the choice $r=4$ and hence the scenario with $p-r=2$ stochastic trends seems reasonable. Other informal indicators in the model (graphs of $\beta' x_t$, eigenvalues, strength of error-correction) confirm this choice.

3 Hypotheses Testing

[9] Now the solution should impose the reduced rank, $\Pi=\alpha\beta'$, and test for *long-run exclusion.* This is a zero row in β and hence a sub-space restriction that cannot be obtained as a normalization of the cointegration space. The degrees of freedom is

given by the number of columns in β, i.e. r. In the present case,

TEST OF EXCLUSION									
r	DGF	5% C.V.	RA	RB	RC	RD	RE	RF	1
1	1	3.841	14.548 [0.000]	0.198 [0.656]	30.445 [0.000]	2.842 [0.092]	3.733 [0.053]	0.765 [0.382]	8.301 [0.004]
2	2	5.991	23.668 [0.000]	1.448 [0.485]	32.611 [0.000]	61.445 [0.000]	34.987 [0.000]	29.043 [0.000]	13.615 [0.001]
3	3	7.815	36.867 [0.000]	9.368 [0.025]	33.307 [0.000]	74.272 [0.000]	35.772 [0.000]	31.771 [0.000]	18.818 [0.000]
4	4	9.488	72.211 [0.000]	41.783 [0.000]	68.034 [0.000]	109.411 [0.000]	69.687 [0.000]	66.982 [0.000]	22.992 [0.000]
5	5	11.070	77.626 [0.000]	48.004 [0.000]	74.170 [0.000]	115.634 [0.000]	73.402 [0.000]	72.835 [0.000]	29.282 [0.000]

For the preferred $r = 4$ no variables are excludable. The good solution may use the information to explain that the strongest relationships involve certain variables.

[10] Next, the solution should test for *stationarity* of some interest rate spreads, $R_t^i - R_t^j$. If some level shifts are maintained, they could be included in the relationships. The solution should explain that a certain (augmented) cointegration vector,

$$\begin{pmatrix} \beta_1^c \\ \rho_1 \end{pmatrix} = (1 : -1 : 0 : 0 : 0 : 0 : \rho_1)',$$

for some constant ρ_1, involves 5 restrictions and one normalization, and hence $5 - (r-1) = 2$ degrees of freedom. For the present case, some examples could read

TEST OF STATIONARY SPREADS									
RA	RB	RC	RD	RE	RF	1	LR	df	$p-val$
1	-1	0	0	0	0	0.281	0.696	2	0.706
1	0	-1	0	0	0	0.582	0.308	2	0.857
1	0	0	-1	0	0	-4.142	42.082	2	0.000
1	0	0	0	-1	0	2.525	41.371	2	0.000
1	0	0	0	0	-1	1.955	36.917	2	0.000
0	0	0	1	-1	0	0.054	5.204	2	0.074
0	0	0	0	1	-1	0.369	1.597	2	0.450

In each case the error correction should be discussed, i.e. which variable corrects the disequilibrium. In the present case, the structure seems to be in line with the scenario with two factors and equal loadings.

[11] Next, the hypothesis of weak exogeneity should be tested. It should be explained that a zero row in α produces a unit vector in $\alpha_\perp$, and hence that unexpected shocks to a weakly exogenous variables constitute innovations to one of the stochastic trends.

In the present case, the following is obtained:

TEST OF WEAK EXOGENEITY								
r	DGF	5% C.V.	RA	RB	RC	RD	RE	RF
1	1	3.841	0.997 [0.318]	9.962 [0.002]	28.423 [0.000]	0.248 [0.619]	6.388 [0.011]	0.356 [0.551]
2	2	5.991	1.484 [0.476]	14.406 [0.001]	75.208 [0.000]	16.178 [0.000]	8.451 [0.015]	0.979 [0.613]
3	3	7.815	1.917 [0.590]	23.385 [0.000]	75.971 [0.000]	16.489 [0.001]	8.567 [0.036]	3.026 [0.388]
4	4	9.488	3.016 [0.555]	41.630 [0.000]	75.996 [0.000]	22.177 [0.000]	22.556 [0.000]	3.328 [0.505]
5	5	11.070	3.024 [0.696]	41.677 [0.000]	76.369 [0.000]	26.833 [0.000]	27.253 [0.000]	7.969 [0.158]

where for $r = 4$, RA and RF appear weakly exogenous. The good solution may test weak exogenity of several variables jointly. In the present case the stochastic trends appear to be generated by ϵ_{RA} and ϵ_{RF}.

[12] The opposite hypothesis that shocks have only transitory effects can be tested by unit vectors in α. Again it should be explained that for $r = 4$ the restriction involves 5 zeros in α, but that $r - 1 = 3$ can be obtained as a normalization and that the degrees of fredom is 2. In the present case

TEST OF UNIT VECTOR IN ALPHA								
r	DGF	5% C.V.	RA	RB	RC	RD	RE	RF
1	5.000	11.070	99.545 [0.000]	95.048 [0.000]	36.033 [0.000]	36.880 [0.000]	75.175 [0.000]	100.249 [0.000]
2	4.000	9.488	84.114 [0.000]	62.029 [0.000]	33.374 [0.000]	7.437 [0.115]	48.528 [0.000]	67.188 [0.000]
3	3.000	7.815	40.985 [0.000]	5.183 [0.159]	7.740 [0.052]	3.861 [0.277]	12.615 [0.006]	20.943 [0.000]
4	2.000	5.991	39.493 [0.000]	2.475 [0.290]	1.221 [0.543]	2.188 [0.335]	1.277 [0.528]	11.278 [0.004]
5	1.000	3.841	5.168 [0.023]	0.002 [0.963]	1.110 [0.292]	0.012 [0.914]	1.249 [0.264]	0.283 [0.595]

and the findings are totally in line with the results above.

[13] Finally, the solution should test that one of the stochastic trends is composed by the average of shocks to the new member countries, D, E, and F:

$$ST_t = \sum_{i=1}^{t} (\epsilon_{D,t} + \epsilon_{E,t} + \epsilon_{F,t}).$$

This amounts to a known vector $(0 : 0 : 0 : 1 : 1 : 1)'$ in $\alpha_\perp$and hence a subspace restriction on α of the form

$$\alpha = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ -a_1 - b_1 & -a_2 - b_2 & -a_3 - b_3 & -a_4 - b_4 \end{pmatrix}$$

where $*$ indicates an unrestricted coefficient. This gives a LR with 4 degrees of freedom. In the present case we get

α				
	Alpha(1)	Alpha(2)	Alpha(3)	Alpha(4)
DRA	0.018 [1.486]	0.068 [1.068]	−0.006 [−0.765]	−0.002 [−0.052]
DRB	−0.044 [−3.310]	−0.129 [−1.844]	0.040 [4.246]	0.058 [1.351]
DRC	−0.107 [−8.198]	−0.188 [−2.713]	0.003 [0.311]	−0.040 [−0.941]
DRD	0.027 [1.650]	−0.309 [−3.582]	−0.009 [−0.750]	0.004 [0.071]
DRE	−0.031 [−1.901]	0.182 [2.098]	−0.008 [−0.661]	0.087 [1.655]
DRF	0.004 [0.263]	0.127 [1.471]	0.016 [1.414]	−0.091 [−1.734]

which is not rejected with a statistic of 6.323 [0.176].

4 Identification

[14] Now the solution should begin with a just identifying structure for β, state the relevant design matrices, and formally check the rank conditions for identification. The solution should at least state the idea of identification, and the relevant rank conditions. In practice CATS can be used to calculate the ranks (quite a large number for $r = 4!$).

In the present case, one may start with e.g.

$$\beta = \begin{pmatrix} 1 & 1 & * & * \\ * & 0 & 0 & 0 \\ 0 & * & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & * & 0 \\ * & * & 0 & * \end{pmatrix}$$

or something similar, which produces

β'							
	RA	RB	RC	RD	RE	RF	1
Beta(1)	1.000 [NA]	−0.967 [−24.163]	0.000 [NA]	0.000 [NA]	0.000 [NA]	−0.038 [−0.908]	0.383 [3.321]
Beta(2)	1.000 [NA]	0.000 [NA]	−1.018 [−25.985]	0.000 [NA]	0.000 [NA]	0.025 [0.601]	0.505 [4.393]
Beta(3)	−0.028 [−0.781]	0.000 [NA]	0.000 [NA]	1.000 [NA]	−0.988 [−25.015]	0.000 [NA]	0.096 [0.968]
Beta(4)	−0.037 [−1.164]	0.000 [NA]	0.000 [NA]	1.000 [NA]	0.000 [NA]	−0.960 [−28.232]	0.315 [3.176]

This uses the following design matrices:

H'_1						
1	0	0	0	0	0	0
0	1	0	0	0	0	0
0	0	0	0	0	1	0
0	0	0	0	0	0	1

H'_2						
1	0	0	0	0	0	0
0	0	1	0	0	0	0
0	0	0	0	0	1	0
0	0	0	0	0	0	1

H'_4						
0	0	0	1	0	0	0
0	0	0	0	0	1	0
0	0	0	0	0	0	1
1	0	0	0	0	0	0

H'_3						
0	0	0	1	0	0	0
0	0	0	0	1	0	0
1	0	0	0	0	0	0
0	0	0	0	0	0	1

and the rank conditions state that the set of restrictions imposed by $H_1, \ldots H_r$ are formally identifying if 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

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

where $R_i = H_{i\perp}$ for all i. In the present case the conditions are summarized in the following table:

Rank Conditions					
R(i.j)		R(i.jk)		R(i.jkl)	
(1.2):	1	(1.23):	3	(1.234):	3
(1.3):	2	(1.24):	2		
(1.4):	1	(1.34):	2		
(2.1):	1	(2.13):	3	(2.134):	3
(2.3):	2	(2.14):	2		
(2.4):	1	(2.34):	2		
(3.1):	2	(3.12):	3	(3.124):	3
(3.2):	2	(3.14):	2		
(3.4):	1	(3.24):	2		
(4.1):	1	(4.12):	2	(4.123):	3
(4.2):	1	(4.13):	2		
(4.3):	1	(4.23):	2		

This confirms the generic identification.

[15] Next the solution should state the Granger representation

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

where $S_t = C^*(L)\epsilon_t$ in a stationary process, and comment in detail on the results.

To see how it relates to the scenarios, it may be necessary to renormalize the matrix $\tilde{\beta}_\perp$. In the present case the Granger representation could read

$$\begin{pmatrix} R_t^A \\ R_t^B \\ R_t^C \\ R_t^D \\ R_t^E \\ R_t^F \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0.9946 & -0.0415 \\ 1.0065 & 0.0256 \\ 0.9973 & 1 \\ 0.9810 & 1.0126 \\ 1 & 1.04171 \end{pmatrix} \begin{pmatrix} \sum_{i=1}^t u_{1i} \\ \sum_{i=1}^t u_{2i} \end{pmatrix} + S_t,$$

with $\sum_{i=1}^t u_i = \sum_{i=1}^t \alpha_\perp' \epsilon_i$, with $\alpha_\perp$ correspondingly rotated. This just identified structure is quite close to the suggested scenario.

[16] Now the structure should be simplified by removing insignificant parameters. The choices should be motivated from significance and from relevance in relation to the theoretical setup.
The final model should be stated. Both in terms of the reduced form estimates (β and α) and in terms of the Granger representation.
If the student has chosen a different specification of dummies, it may not be possible to find the true structure. It is more important that firm arguments for the choices are given, and that all relevant restrictions are imposed.

[17] A recursive estimation of the preferred identified structure should be performed, and some selected diagnostic output presented. It should be explained that one possible and simple remedy in the case of instability could be to allow for a level shift, i.e. assuming that the instability is related only to the equilibrium levels.

[18] Finally, the students could mention additional weaknesses and uncertainties. An example could be that they are uncertain on the inclusion of a level shift or the value of the cointegration rank.

5 Extensions

These extensions have little to do with the former analysis and are slightly more advanced.

[19] Considering the simple VAR model

$$W_t = \Pi_1 W_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \Omega),$$

the solution could e.g. explain that the causal interpretation of the symmetric covariance, Ω, as contemporaneous effects is problematic, because many 'structural' representations are equivalent to the same reduced form equation and likelihood function.
One simple suggested solution is to insist on a causal chain, i.e. to have a lower triangular matrix C, such that

$$CW_t = C\Pi_1 W_{t-1} + u_t, \quad u_t = C\epsilon_t \sim N(0, I).$$

The triangular matrix C is unique given the ordering of the variables (up to the sign of the diagonal), and is called the Choleski decomposition of Ω^{-1}. This may also be explained as sequential conditioning, and the parameters in the Choleski lower triangular matrix can be estimated by OLS. Given the assumption of a causal chain, and given the ordering of the variables, the model suggests a unique set of orthogonal shocks, u_t, and a corresponding unique set of impulse-response functions. It remains controversial, however, because the uniqueness depends on the ordering, and different orderings will give different impulse response functions.
Next, we are informed that:

$$\begin{array}{lcclcc}
\operatorname{corr}(x_t, y_t) & = & \underset{(0.15)}{0.49} & \qquad \operatorname{corr}(x_t, y_t \mid z_t) & = & \underset{(0.17)}{0.53} \\
\operatorname{corr}(y_t, z_t) & = & \underset{(0.20)}{0.65} & \qquad \operatorname{corr}(y_t, z_t \mid x_t) & = & \underset{(0.23)}{0.70} \\
\operatorname{corr}(x_t, z_t) & = & \underset{(0.11)}{-0.03} & \qquad \operatorname{corr}(x_t, z_t \mid y_t) & = & \underset{(0.13)}{0.27}
\end{array}$$

We note that the unconditional correlation $\operatorname{corr}(x_t, z_t)$ is insignificant, while $\operatorname{corr}(x_t, z_t \mid y_t)$ is borderline significant. The insight from graph-theory is that this is consistent with the causal structure

$$z_t \rightarrow y_t \leftarrow x_t,$$

called an unshielded collider, while it is inconsistent with the causal chains

$$z_t \rightarrow y_t \rightarrow x_t \quad \text{and} \quad z_t \leftarrow y_t \leftarrow x_t,$$

and inconsistent with the common cause

$$z_t \leftarrow y_t \rightarrow x_t.$$

So (under the assumption of the graph-theoretic approach) the only member of the observationally equivalent class is the unshielded collider. This suggests an ordering that allow contemporaneous effects from z_t and x_t to y_t.

[20] Students are given the information that a univariate LR test for $\pi = 0$ in the model

$$\Delta y_t = \delta + \pi y_{t-1} + \epsilon_t,$$

has the property that as $T \rightarrow \infty$,

$$LR(\pi = 0) \overset{D}{\rightarrow} \begin{cases} \chi^2(1) & \text{if} \quad \delta \neq 0 \\ DF^2 & \text{if} \quad \delta = 0 \end{cases}$$

The solution should explain that

(i) This is problematic because the statistic has two different distributions, depending on the true value of the unknown parameter δ. The investigator therefore has to choose critical values based on some (ad hoc) argument.

(*ii*) This is an example of a non-similar test (or a non-pivotal test statistic). We would strongly prefer to work with similar tests. This suggest an alternative way of formulating the hypothesis, namely as $H_0^* : \pi = \delta = 0$, where the statistic, $LR(\pi = \delta = 0)$ is always distributed according to a DF-type distribution in the limit.

(*iii*) This is exactly the models we work with in the CVAR case, where leading deterministic terms are included in the cointegration space, e.g.

$$\Delta x_t = \alpha \begin{pmatrix} \beta \\ \rho \end{pmatrix}' \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} + \Gamma_1 \Delta x_{t-1} + \epsilon_t.$$

The reduced rank hypothesis is tested on the augmented matrix

$$\Pi^* = \alpha \begin{pmatrix} \beta \\ \rho \end{pmatrix}' \quad \text{rather than} \quad \Pi = \alpha\beta'.$$

Rank tests in the model with an unrestricted constant are non-similar and the asumptotic distributions depend on the presence of a drift.

[21] This is probably a difficult question, as the model class has not been discussed during lectures. The solution should use the insight that given $\beta = \bar{\beta}$, the CVAR model is a linear regression,

$$\Delta x_t = \alpha\bar{\beta}' x_{t-1} + \Gamma_1 \Delta x_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \Omega),$$

and OLS estimates of the remaining parameters, i.e. $\theta = \{\alpha, \Gamma_1, \Omega\}$, coincide with ML.

For known threshold parameter, $\delta = \bar{\delta}$, this also applies to the threshold case

$$\begin{aligned} \Delta x_t &= R_t \cdot \alpha_1 \bar{\beta}' x_{t-1} + (1 - R_t) \cdot \alpha_2 \bar{\beta}' x_{t-1} + \Gamma_1 \Delta x_{t-1} + \epsilon_t \\ &= \alpha_1 \cdot \left\{ R_t \bar{\beta}' x_{t-1} \right\} + \alpha_2 \cdot \left\{ (1 - R_t) \bar{\beta}' x_{t-1} \right\} + \Gamma_1 \Delta x_{t-1} + \epsilon_t, \end{aligned}$$

$\epsilon_t \sim N(0, \Omega)$, where R_t is a function of x_{t-1}, $\bar{\beta}$, and $\bar{\delta}$. The regressors are therefore directly observable, and ML estimates of the parameters $\theta = \{\alpha_1, \alpha_2, \Gamma_1, \Omega\}$ can be found using OLS.

If we also seek to estimate δ, notice that the likelihood function becomes non-differentiable. As a consequence, standard algorithms for maximizing the function (based on derivatives) fail, and standard arguments for deriving asymptotics (based on expansions) also fail. In terms of estimation, a simple grid search would work in this case, but the asymptotic analysis is non-standard.