Written Exam for the M.Sc. in Economics, Winter 2010/2011

ADVANCED MACROECONOMETRICS

Proposed Solution

About the Exam

The examination considers econometric models for the term structure, i.e. models for the relationship between interest rates with different maturities. The purpose of the examination is to assess the students understanding of the CVAR model, their ability to use statistical procedures to make inference on the equilibrium structures and the dynamic adjustment properties, as well as their ability to interpret the results.

All assignments are based on *different* data sets. They all consists of five interest rates collected in the $p = 5$ dimensional data vector,

$$x_t = (R0_t : R1_t : R2_t : R5_t : R10_t)', \tag{0.1}$$

simulated from the following data generating process (DGP):

$$\begin{pmatrix} \Delta R0_t \\ \Delta R1_t \\ \Delta R2_t \\ \Delta R5_t \\ \Delta R10_t \end{pmatrix} = \begin{pmatrix} -0.10 & -0.04 \\ 0.05 & 0.06 \\ 0 & -0.15 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -2 & 0 \\ 1 & 1 \\ 0 & -2 \\ 0 & 1 \end{pmatrix}' \begin{pmatrix} R0_{t-1} \\ R1_{t-1} \\ R2_{t-1} \\ R5_{t-1} \\ R10_{t-1} \end{pmatrix}$$
$$+ \begin{pmatrix} -0.0260 & 0.5050 & 0.0900 & -0.1960 & 0.2300 \\ 0.0030 & 0.6930 & -0.3840 & 0.0590 & 0.1340 \\ 0.0110 & 0.7580 & -0.5390 & 0.1520 & 0.1850 \\ -0.1470 & 0.5260 & -0.4610 & 0.3320 & 0.2580 \\ -0.1450 & 0.2210 & -0.4560 & 0.3640 & 0.1680 \end{pmatrix} \begin{pmatrix} \Delta R0_{t-1} \\ \Delta R1_{t-1} \\ \Delta R2_{t-1} \\ \Delta R5_{t-1} \\ \Delta R10_{t-1} \end{pmatrix}$$
$$+ \begin{pmatrix} 0.0290 & 0.1920 & 0.0110 & -0.0570 & 0.0390 \\ 0.1460 & -0.0720 & -0.0420 & 0.1850 & -0.1400 \\ 0.0640 & 0.1440 & -0.1350 & 0.0930 & -0.1100 \\ -0.0310 & 0.3950 & -0.3720 & 0.2350 & -0.3030 \\ -0.0280 & 0.6010 & -0.6760 & 0.3870 & -0.3780 \end{pmatrix} \begin{pmatrix} \Delta R0_{t-2} \\ \Delta R1_{t-2} \\ \Delta R2_{t-2} \\ \Delta R5_{t-2} \\ \Delta R10_{t-2} \end{pmatrix} + \epsilon_t,$$

with $\epsilon_t \sim N(0, \Omega)$, with

$$\Omega = \frac{1}{1000}\begin{pmatrix} 154.54 & & & & \\ 103.13 & 169.09 & & & \\ 101.63 & 161.47 & 181.38 & & \\ 90.00 & 147.68 & 183.12 & 248.79 & \\ 57.33 & 96.26 & 133.21 & 210.09 & 230.13 \end{pmatrix}.$$

The levels of the variables are set to reflects realizations of interest rates. 306 observations are generated covering monthly data from January 1985 to September 2010.

The students are informed that the regulations of international capital flows were changed in November 2000, and could have affected the equilibrium relationships. All datasets have an outlier at that date, but it is does *not* change the equilibrium structure. In addition, outlying observations are drawn randomly with a probability of 1%, and the typical data set will have approximately 4 outliers, all with a magnitude of 5 standard deviations.

For all data sets it is ensured that, if the correct outliers are modelled with dummy variables, the *trace test* for the cointegration rank will correctly suggest a rank of $r = 2$, and 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`).

1 Background

The students are asked to think about the relationships between pushing and pulling forces and the relationship between the CVAR, including the cointegrating relationships, and the Granger representation.

[1] Based on a baby model

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

the solution should derive the Granger representation. This corresponds to Section 5.4 in Juselius (2006), and, since the initial value is zero, we end with

$$x_t = \beta_\perp \left(\alpha'_\perp \beta_\perp\right)^{-1} \alpha'_\perp \sum_{i=1}^{t} \epsilon_i + \alpha \left(\beta'\alpha\right)^{-1} \sum_{i=0}^{t-1} \left(I_p + \beta'\alpha\right)^i \epsilon_{t-i}. \qquad (1.1)$$

Simply stating the solution is not sufficient here, as a minimum it should be noted that $\beta' x_t$ and $\alpha'_\perp x_t$ are solved separately and combined using the relevant projection identity. The good solution actually does it.

Next, the solution should note that the pulling forces are parametrized as the correction toward equilibrium in the CVAR equations, i.e. the coefficients in α, while the pushing forces are visible in (1.1), particularly the stochastic trends, $\sum \alpha'_{\perp}\epsilon_i$, and their loadings in terms of $\tilde{\beta}_{\perp} = \beta_{\perp}(\alpha'_{\perp}\beta_{\perp})^{-1}$. The good solution may reproduce the familiar graph of attractor and error correction in the 2-dimensional case, see Figure 5.2 in Juselius (2006).

[2] Solutions to the next questions exploits the relationship between the cointegrating relations, β, and the loading to common trend, $\beta_{\perp}$, for the particular example of an interest rate model. For the particular loading, $\beta_{\perp}$, in (1.3), the solution should state that the cointegration space is four dimensional, e.g.

$$\beta = (\beta_{\perp})_{\perp} = \begin{pmatrix} 1 \\ \tau_1 \\ \tau_2 \\ \tau_5 \\ \tau_{10} \end{pmatrix}_{\perp} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ \tau_1^{-1} & 0 & 0 & 0 \\ 0 & \tau_2^{-1} & 0 & 0 \\ 0 & 0 & \tau_5^{-1} & 0 \\ 0 & 0 & 0 & \tau_{10}^{-1} \end{pmatrix}.$$

The solution should note that β is not unique because α and β only enters the likelihood function through $\Pi = \alpha\beta'$, and any choice of $\tilde{\alpha}$ and $\tilde{\beta}$, with $\tilde{\Pi} = \tilde{\alpha}\tilde{\beta}' = \Pi$ are observationally equivalent. The good solution should note that the interest rates cointegrate pair-wise in this scenario.

If $\tau_1 = \tau_2 = \tau_5 = \tau_{10} = 1$ it holds that all interest rate spreads, $R_i - R_j$, are stationary.

[3] In the next scenario, the solution should argue that the proposed β is orthogonal to the loading to the common trend, e.g. by matrix multiplication, and that interest rate spreads are not stationary.

[4] Next consider the scenario with three factors in (1.6). In this case

$$\beta = (\beta_1 : \beta_2) = \begin{pmatrix} 1 & 0 \\ -2 & 0 \\ 1 & 1 \\ 0 & -2 \\ 0 & 1 \end{pmatrix} \quad \text{for} \quad x_t = \begin{pmatrix} R0_t \\ R1_t \\ R2_t \\ R5_t \\ R10_t \end{pmatrix}.$$

If $R10$ is omitted from a statistical analysis, then the cointegration relation β_2 between $R2$, $R5$, and $R10$ can no longer be recovered. The cointegration rank based on the data set $z_t = (R0_t : R1_t : R2_t : R5_t)'$ would be expected to be $r = 1$ and the cointegration vector would be β_1.

The good solution may note that the concept of cointegration is invariant to increases in the information set, such that a cointegration vector in a small system will also prevail in a cointegration analysis of an extended information set.

2 The Statistical Model

[5] The paper should load the data and perform a graphical inspection of the data. It would be natural to look at the levels of the variables, and note that the time series do not look stationary, see Figure 1 (A). To look for indications of the empirical relevance of the theories, the solution could look at spreads between the interest rates as in (B) and note that they do not look stationary. The candidates from the *three-factor* model in (C) appear to be stationary, while the additional cointegrating relation suggested by the *two-factor* model in (D) is much more persistent. Based on the graphical inspection, the three-factor model appears to be the more promising as a candidate for explaining the variation in the data.

Several outliers are visible in the data, but in (C) none of them appear to change equilibrium relationships.

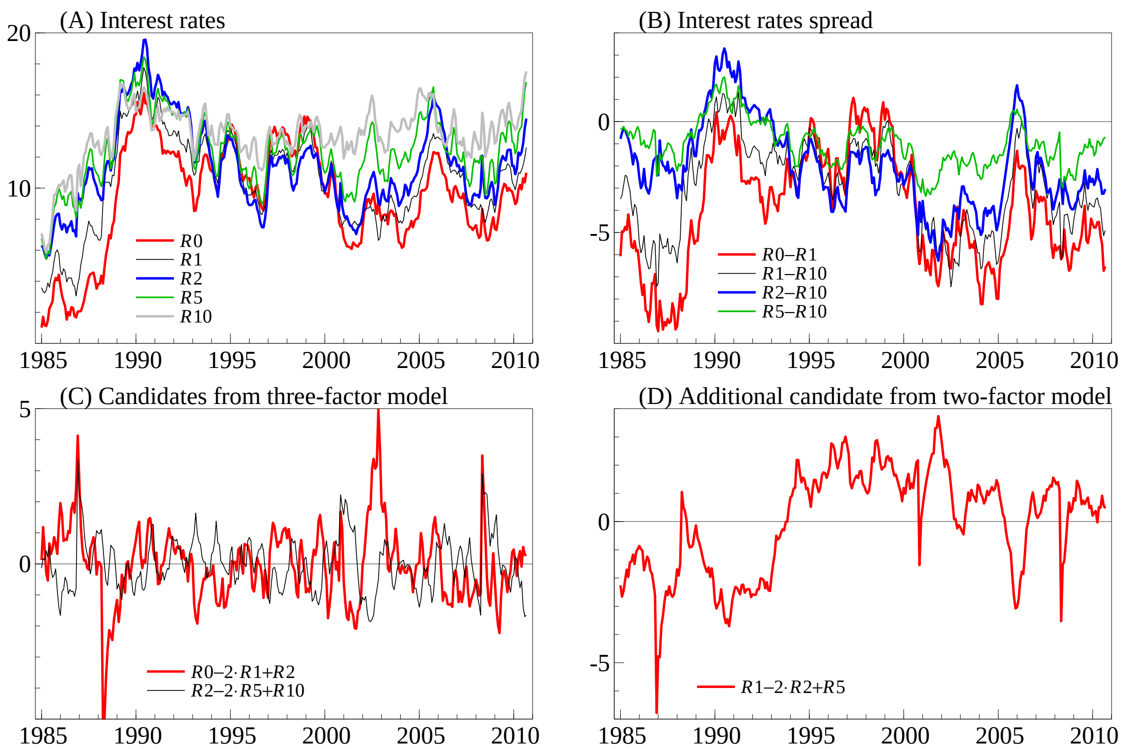

Figure 1: *Data in levels and equilibrium candidates from the factor models.*

[6] Now the paper should state the unrestricted VAR model of order k:

$$x_t = \Pi_1 x_{t-1} + \Pi_2 x_{t-2} + ... + \Pi_k x_{t-k} + \mu_0 + \phi D_t + \epsilon_t,$$

for $t = 1, 2, ..., T$, and initial values, $x_0, ..., x_{-k+1}$, given. We assume that $\epsilon_t \sim Niid(0, \Omega)$, $\Omega > 0$, and that the model has constant parameters.

In terms of *testable* assumptions, we could look at Gaussianity, no-autocorrelation, homoskedasticity, and constant parameters.

Regarding deterministic terms, the model should include a constant, to allow non-zero means in stationary relationships, potential intervention dummies for outliers, and potentially a shift dummy for the liberalization of capital flows 2000 : 11.

Some students may have data sets that trends in sample, and may choose to include a deterministic linear trend. In any case, the choices should be made clear and motivated, and for interest rates it should be noted that a linear trend is probably not a good description for longer samples.

[7] To write the likelihood function we use sequential factorization of the joint density to find

$$f_\theta(x_{-k+1}, ..., x_0, x_1, ..., x_T) = f_\theta(x_{-k+1}, ..., x_0) \prod_{t=1}^{T} f_\theta(x_t \mid x_{t-1}, ..., x_{t-k}),$$

and the likelihood function for $x_1, ..., x_T$ given the k initial observations is the conditional density

$$L(\theta) = \frac{f_\theta(x_{-k+1}, ..., x_0, x_1, ..., x_T)}{f_\theta(x_{-k+1}, ..., x_0)} = \prod_{t=1}^{T} f_\theta(x_t \mid x_{t-1}, ..., x_{t-k}).$$

Under Gaussianity, each term has the form

$$f_\theta(x_t \mid x_{t-1}, ..., x_{t-k}) = (2\pi)^{-\frac{p}{2}} |\Omega|^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}\epsilon_t(\theta)'\Omega^{-1}\epsilon_t(\theta)\right),$$

where $\epsilon_t(\theta) = x_t - \Pi_1 x_{t-1} - \Pi_2 x_{t-2} - ... - \Pi_k x_{t-k} - \mu_0 - \phi D_t$.

[8] Now the solution should estimate the unrestricted VAR and check the assumptions of the model. This includes

- The determination of the number of lags, k, in the model. E.g. based on likelihood ratio tests for the significance of the Π_k or the use of information criteria.
- Testing the estimated residuals for signs of misspecification, cf. question [6].
- Construct and include dummy variables for potential outlier to restore normality of residuals.

  The good solution pay particular attention to the observation 2000:11, where a known shift has taken place. The solution could insert a shift dummy and note that it is not significant.
- The good solution also checks the constancy of parameters by reporting some recursively calculated statistics.

It may be necessary to iterate between the steps to find a good model.

For the present data set, $k = 3$ lags are sufficient to account for the autoregressive nature of the variables. There are 5 large residuals corresponding to observations: 1986:12, 1988:4, 2000:11, 2002:11, and 2008:5. With these dummies there are no

signs of misspecifcation, and the parameters appear constant according to recursive testing. According to $t-$values in Π there are no indications or a level shift in the model. This conclusion is also confirmed by test for long-run exclusion of a level shift for all choices of the cointegration rank.
Most solutions tries to be brief here, but they have to document that they know it is important to have a well-specified model and that they know how to find it.

3 The Cointegration Rank

[9] For a model with $k = 3$, the preferred choice for my empirical model, the paper should derive the error correction form of the VAR model, i.e.

$$\Delta x_t = \Pi x_{t-1} + \Gamma_1 \Delta x_{t-1} + \Gamma_2 \Delta x_{t-2} + \mu_0 + \phi D_t + \epsilon_t,$$

as in Section 4.2.4 in Juselius (2006). It is not enough to state the solution.
The characteristic polynomial of the error correction form is

$$A(z) = (1 - z) - \Pi z - \Gamma_1(1 - z)z - \Gamma_2(1 - z)z^2.$$

A unit root implies that

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

so that Π has to be singular. A singular matrix, say $Rank(\Pi) = r$, can be decomposed as

$$\Pi = \alpha\beta',$$

where α and β are $p \times r$ matrices containing the independent columns and rows of Π, respectively.

[10] The eigenvalues of the companion matrix are given by

The Roots of the COMPANION MATRIX // Model: H(5)				
	Real	Imaginary	Modulus	Argument
Root1	0.969	0.007	0.969	0.007
Root2	0.969	−0.007	0.969	−0.007
Root3	0.934	0.000	0.934	0.000
Root4	0.891	−0.000	0.891	−0.000
Root5	0.685	0.164	0.704	0.236
Root6	0.685	−0.164	0.704	−0.236
Root7	0.234	0.433	0.492	1.075
Root8	0.234	−0.433	0.492	−1.075
Root9	−0.048	0.412	0.415	1.688
Root10	−0.048	−0.412	0.415	−1.688
Root11	−0.259	0.169	0.310	2.564
Root12	−0.259	−0.169	0.310	−2.564
Root13	0.160	−0.178	0.239	−0.838
Root14	0.160	0.178	0.239	0.838
Root15	−0.215	0.000	0.215	3.142

We note that all point estimates are inside the unit circle, but that 3-4 roots are close to unity. We conclude that the interest rates are not likely to be stationary.

[11] To avoid trends in the model for interest rates, and to have similarity of the LR test for unit roots, we restrict the constant to be proportional to α, i.e. we write the reduced rank model as

$$H(r) : \Delta x_t = \alpha \begin{pmatrix} \beta \\ \beta_0 \end{pmatrix}' \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} + \Gamma_1 \Delta x_{t-1} + \Gamma_2 \Delta x_{t-2} + \phi D_t + \epsilon_t.$$

To determine the cointegration rank we consider first the LR test for reduced rank of Π. In the present case we have a restricted constant and unrestricted intervention dummies. The latter does not affect the asymptotic distribution, and the critical values reported by CATS are correct. For the present data set we obtain

I(1)-ANALYSIS							
p-r	r	Eig.Value	Trace	Trace*	Frac95	P-Value	P-Value*
5	0	0.486	279.839	268.700	76.813	0.000	0.000
4	1	0.160	76.431	72.897	53.945	0.000	0.000
3	2	0.042	23.171	21.733	35.070	0.521	0.616
2	3	0.023	10.047	9.409	20.164	0.641	0.701
1	4	0.009	2.894	2.649	9.142	0.608	0.653

Two eigenvalues are large, suggesting that the cointegration rank is two in the present case. There is no doubt on the cointegration rank for the present case, and the choice is confirmed by the visual appearance of the candidate cointegrating relationships, and the $t-$ratios for coefficients in α, cf. Juselius (2006), noting however that the critical values are unknown. A recursive trace test may also be used to confirm the choice.

[12] With a cointegration rank of $r = 2$, the three factor model with $p - r = 3$ common trends, appears to be the empirically most relevant. This was also the impression from the graphical inspection.

4 Testing Hypotheses

[13] The paper should estimate the CVAR and comment on the results.

The good paper explains that the unrestricted estimates of the cointegrating relationships are normalized to be conditionally orthogonal, $\hat{\beta}' S_{11} \hat{\beta} = I_r$. This is mathematically convenient, but may not be relevant in terms of economic theory and the unrestricted estimates are often difficult to interpret. The solution should be careful not to attach to much structural interpretation to the unidentified relationships. In the present case we obtain

	β'					
	R0	R1	R2	R5	R10	CONSTANT
Beta(1)	0.144	−0.273	1.000	−1.726	0.821	0.606
Beta(2)	1.000	−1.910	0.950	−0.230	0.344	−2.240

where the true structure is actually visible. The error correction is given by the matrix ($t-$ratios in parentheses)

	α	
	Alpha(1)	Alpha(2)
DR0	−0.050 [−1.179]	−0.114 [−4.572]
DR1	0.076 [1.776]	0.027 [1.089]
DR2	−0.133 [−3.037]	0.007 [0.281]
DR5	0.042 [0.859]	−0.033 [−1.156]
DR10	0.038 [0.807]	−0.021 [−0.760]

indicating that the long-term bond rates do not error correct.

Regarding the short-run parameter there are quite a few significant parameters in Γ_1 and Γ_2 and the dummies for outliers are clearly significant.

[14] Estimates of the unrestricted VAR are obtained by solving a generalized eigenvalue problem

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

where $S_{ij} = T^{-1} \sum R_{it} R'_{jt}$ where R_{0t} and R_{1t} are Δx_t and x_{t-1} corrected for the unrestricted terms in the model. The eigenvalue problem has the advantage of having a closed form solution and being numerically robust. In addition the normalization is implicit, and we do not have to impose identifying restrictions from the outset. Finally the procedure estimates the models $H(r)$ for all values of the cointegration rank, r, in one solution, which is very convenient for rank testing.

The main disadvantage is that it has to be modified and changed whenever restrictions are imposed on the parameters.

A more general alternative would be to maximize the likelihood function in question [7] numerically with respect to the parameters in θ. While this is very flexible and allow complicated restrictions on all parameters, it requires that the parameters are identified. Early in the analysis, before we have a full overview of the properties of the variables, it is inconvenient that identification is required.

[15] The paper should explain the Granger representation

$$x_t = \beta_\perp \left(\alpha'_\perp \Gamma \beta_\perp\right)^{-1} \alpha'_\perp \sum_{i=1}^{t} (\epsilon_i + \phi D_i) + C^*(L)(\epsilon_t + \phi D_t) + \tau_0 + \tau_1 t.$$

The paper should note that $C = \tilde{\beta}_\perp \alpha'_\perp$ has reduced rank $p - r$, that the common stochastic trends are given by $\alpha'_\perp \sum_{i=1}^{t} \epsilon_i$ and that they affect the variables with the loading coefficients $\tilde{\beta}_\perp = \beta_\perp (\alpha'_\perp \Gamma \beta_\perp)^{-1}$. The paper should note that the shocks $\alpha'_\perp \epsilon_t$ have permanent effects while the shocks $\alpha' \epsilon_t$ have only transitory effects. The parameters of the Granger representation reads

$\alpha'_\perp$

	R0	R1	R2	R5	R10
CT(1)	−0.267 [−0.863]	0.000 [NA]	0.419 [0.903]	1.000 [NA]	0.000 [NA]
CT(2)	0.269 [0.764]	1.000 [NA]	0.469 [0.887]	0.000 [NA]	0.000 [NA]
CT(3)	−0.162 [−0.573]	0.000 [NA]	0.346 [0.815]	0.000 [NA]	1.000 [NA]

The Loadings to the Common Trends, $\tilde{\beta}_\perp$:

	CT1	CT2	CT3
R0	−1.747 [−2.686]	1.580 [4.386]	0.821 [3.297]
R1	−1.165 [−1.861]	1.276 [3.680]	0.862 [3.599]
R2	−0.258 [−0.373]	0.973 [2.543]	0.554 [2.094]
R5	−0.634 [−1.357]	0.584 [2.258]	0.977 [5.469]
R10	−1.100 [−3.104]	0.190 [0.969]	1.524 [11.243]

and the the long-run impact matrix, C, is given by

The Long-Run Impact Matrix, C

	R0	R1	R2	R5	R10
R0	0.759 [3.370]	1.580 [4.386]	0.294 [0.843]	−1.747 [−2.686]	0.821 [3.297]
R1	0.515 [2.376]	1.276 [3.680]	0.409 [1.219]	−1.165 [−1.861]	0.862 [3.599]
R2	0.241 [1.009]	0.973 [2.543]	0.540 [1.458]	−0.258 [−0.373]	0.554 [2.094]
R5	0.168 [1.040]	0.584 [2.258]	0.347 [1.385]	−0.634 [−1.357]	0.977 [5.469]
R10	0.098 [0.798]	0.190 [0.969]	0.156 [0.820]	−1.100 [−3.104]	1.524 [11.243]

The solution should again explain the idea of the common trends and the loadings. The solution may refer to the theoretical model, and the good solution notes that normalization of $\alpha_\perp$ in CATS makes it difficult to recognize the level, slope and curvature. We could rotate $\tilde{\beta}_\perp$ to get get closer to the theoretical framework, and the students have a small piece of software to do that, but here it is not strictly required.
In the present case, $p-r=3$, we could impose one normalization and two restrictions on each columns of $\beta_\perp$ by rotation to obtain, e.g.

$$\tilde{\beta}_\perp = \begin{pmatrix} 1 & 0 & 0 \\ 1.08502 & 0.34298 & 0.87280 \\ 1 & \frac{1}{2} & 2 \\ 0.96595 & 0.71012 & 1.01944 \\ 1 & 1 & 0 \end{pmatrix},$$

which is close to the theoretical setup in (1.6).

[16] The test for long-run exclusion is a subspace restriction of the form

$$\mathcal{H}_0 : \tilde{\beta}^c = H\varphi,$$

where $\tilde{\beta} = \left(\beta' : \beta_0'\right)'$ is the augmented cointegration matrix, H is the $(p+1) \times p$ design matrix and φ contains the $p \cdot r$ free parameters. To test exclusion of $R0_t$ we use

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

and the LR statistic $LR(\mathcal{H}_0 \mid H(r))$ is $\chi^2(r)$ under the null.

The automatic test for exclusion in CATS may be used as long as the set up of the test is explained. For the present data the automatic test produces

TEST OF EXCLUSION								
r	DGF	5% C.V.	R0	R1	R2	R5	R10	CONSTANT
1	1	3.841	9.775 [0.002]	10.506 [0.001]	114.914 [0.000]	149.105 [0.000]	115.775 [0.000]	3.672 [0.055]
2	2	5.991	47.330 [0.000]	50.117 [0.000]	149.734 [0.000]	177.317 [0.000]	121.059 [0.000]	6.658 [0.036]
3	3	7.815	52.357 [0.000]	55.977 [0.000]	149.781 [0.000]	181.330 [0.000]	126.951 [0.000]	11.732 [0.008]
4	4	9.488	54.074 [0.000]	59.958 [0.000]	154.036 [0.000]	183.176 [0.000]	130.897 [0.000]	15.924 [0.003]

For the preferred, $r = 2$, no variables can be excluded. The constant is borderline, but it is not recommended to exclude the constant early in the analysis.

If a variable is excludable it does *not* imply that the variable can be removed from the model altogether. There may still be imporatant short run effects involving the variable, and the particular equation could contain information on the other cointegrating coefficients.

[17] Next the solution should explain that a zero row in α implies that the corresponding variable is weakly exogenous for α, β. this means that α and β can be estimated efficiently in a model for the remaining variables conditional on the weakly exogenous variables.

The solution should explain the idea of conditional models. At least it should be noted that α and β does not appear in the marginal equation for the weakly exogenous variable. If we used only the marginal equations for the endogenous variables to estimate α and β, however, we would loose the contemporaneous effects in the covariance matrix, and efficient estimation requires that we condition, and include the contemporaneous effect of the exogenous variable.

For the present data we get

TEST OF WEAK EXOGENEITY							
r	DGF	5% C.V.	R0	R1	R2	R5	R10
1	1	3.841	1.167 [0.280]	3.042 [0.081]	8.874 [0.003]	0.712 [0.399]	0.629 [0.428]
2	2	5.991	19.525 [0.000]	4.115 [0.128]	8.947 [0.011]	1.887 [0.389]	1.122 [0.571]
3	3	7.815	21.304 [0.000]	5.093 [0.165]	9.353 [0.025]	1.908 [0.592]	2.784 [0.426]
4	4	9.488	21.670 [0.000]	7.129 [0.129]	12.419 [0.014]	5.958 [0.202]	6.457 [0.168]

and for $r = 2$, to two longest maturities are (individually) weakly exogenous, while $R1$ is borderline. A joint test for $R5$ and $R10$ jointly weakly exogenous gives a test statistic of 2.105 or a $p-$ value of 0.716 in the asymptotic $\chi^2(4)$. A joint test of $R1, R5, R10$ weakly exogenous rejects, with a $\chi^2(6)$ statistic of 13.072.

Weak exogeneity of the variables, and the zero rows in α implies the presence of two unit vector in $\alpha_\perp$, and in the particular case of $R5$ and $R10$ weakly exogenous:

$\alpha'_\perp$					
	R0	R1	R2	R5	R10
CT(1)	0.539 [1.588]	1.000 [NA]	0.100 [0.282]	0.000 [NA]	0.000 [NA]
CT(2)	0.000 [NA]	0.000 [NA]	0.000 [NA]	1.000 [NA]	0.000 [NA]
CT(3)	0.000 [NA]	0.000 [NA]	0.000 [NA]	0.000 [NA]	1.000 [NA]

Two common trends are generated by ϵ_{R5} and ϵ_{R10} alone, while the last common trend is the accumulated effect of $\epsilon_{R1} + 0.539\epsilon_{R0} + 0.1\epsilon_{R2}$.

5 Identification

[18] The paper should look at the three factor model and show that it is generically identified. For this case we would have

$$\tilde{\beta} = \begin{pmatrix} 1\cdot\phi_1 & 0 \\ -2\cdot\phi_1 & 0 \\ 1\cdot\phi_1 & 1\cdot\phi_2 \\ 0 & -2\cdot\phi_2 \\ 0 & 1\cdot\phi_2 \end{pmatrix} = \left(\begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} (\phi_1) : \begin{pmatrix} 0 \\ 0 \\ 1 \\ -2 \\ 1 \end{pmatrix} (\phi_2) \right) = (H_1\phi_1 : H_2\phi_2),$$

where ϕ_i is a free parameter to be estimated. Correspondingly,

$$R_1 = H_{1\perp} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad R_2 = H_{2\perp} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Now the structure is identifying if

$$\begin{aligned} Rank\left(R_1'H_2\right) &\geq 1 \\ Rank\left(R_2'H_1\right) &\geq 1. \end{aligned}$$

By direct matrix multiplication this is seen to be true (with equality).

[19] Now we impose a just identifying structure inspired from (1.7). For a particular choice we get

β'						
	R0	R1	R2	R5	R10	CONSTANT
Beta(1)	1.000 [NA]	−1.910 [−18.213]	0.833 [7.896]	0.000 [NA]	0.239 [2.016]	−2.366 [−2.067]
Beta(2)	0.000 [NA]	0.002 [0.059]	1.000 [NA]	−1.961 [−31.209]	0.893 [15.877]	1.076 [2.758]

,

which is quite close to the theoretical candidates, and adjustment given by

α		
	Alpha(1)	Alpha(2)
DR0	−0.121 [−4.721]	−0.057 [−1.532]
DR1	0.038 [1.481]	0.070 [1.855]
DR2	−0.012 [−0.450]	−0.116 [−3.006]
DR5	−0.027 [−0.918]	0.033 [0.767]
DR10	−0.015 [−0.546]	0.031 [0.745]

Imposing over-identifying restrictions reveals the theoretical structure

β'						
	R0	R1	R2	R5	R10	CONSTANT
Beta(1)	1.000 $[NA]$	−2.000 $[NA]$	1.000 $[NA]$	0.000 $[NA]$	0.000 $[NA]$	−0.100 [−0.749]
Beta(2)	0.000 $[NA]$	0.000 $[NA]$	1.000 $[NA]$	−2.000 $[NA]$	1.000 $[NA]$	0.143 [3.326]

,

and the equilibrium adjustment

α		
	Alpha(1)	Alpha(2)
DR0	−0.113 [−4.615]	−0.044 [−1.186]
DR1	0.031 [1.256]	0.067 [1.790]
DR2	−0.017 [−0.667]	−0.118 [−3.063]
DR5	−0.025 [−0.866]	0.030 [0.689]
DR10	−0.002 [−0.075]	0.022 [0.528]

This reduction is valid with a test statistic of 7.232 corresponding to a $p-$value of 0.300 in the asymptotic $\chi^2(6)$. Here we could still accept the two zero rows in α, but the students are not specifically asked to impose that. Nevertheless the results are

β'						
	R0	R1	R2	R5	R10	CONSTANT
Beta(1)	1.000 $[NA]$	−2.000 $[NA]$	1.000 $[NA]$	0.000 $[NA]$	0.000 $[NA]$	−0.105 [−0.756]
Beta(2)	0.000 $[NA]$	0.000 $[NA]$	1.000 $[NA]$	−2.000 $[NA]$	1.000 $[NA]$	0.143 [3.292]

and

α		
	Alpha(1)	Alpha(2)
DR0	−0.099 [−4.024]	−0.056 [−1.482]
DR1	0.053 [2.150]	0.048 [1.277]
DR2	0.008 [0.319]	−0.141 [−3.711]
DR5	0.000 [0.000]	0.000 [0.000]
DR10	0.000 [0.000]	0.000 [0.000]

[20] The granger representation for this case is given by

$\alpha'_\perp$					
	R0	R1	R2	R5	R10
CT(1)	0.000 [NA]	0.000 [NA]	0.000 [NA]	1.000 [NA]	0.000 [NA]
CT(2)	0.000 [NA]	0.000 [NA]	0.000 [NA]	0.000 [NA]	1.000 [NA]
CT(3)	0.542 [1.496]	1.000 [NA]	0.125 [0.337]	0.000 [NA]	0.000 [NA]

and

The Loadings to the Common Trends, $\tilde{\beta}_\perp$:			
	CT1	CT2	CT3
R0	−1.680 [−1.582]	0.981 [2.476]	1.524 [3.915]
R1	−0.387 [−0.347]	0.575 [1.381]	1.164 [2.847]
R2	0.906 [0.660]	0.168 [0.328]	0.804 [1.597]
R5	−0.006 [−0.006]	0.848 [2.501]	0.505 [1.517]
R10	−0.918 [−1.480]	1.527 [6.600]	0.206 [0.906]

and the students should discuss this (with or without weak exogeneity). Again it is not easy to recognize $\beta_\perp$, but the students should realize that we may rotate to get, exactly,

$$\tilde{\beta}_\perp = \begin{pmatrix} 1 & 0 & 0 \\ 1 & \frac{1}{4} & 1 \\ 1 & \frac{1}{2} & 2 \\ 1 & \frac{3}{4} & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \alpha_\perp = \begin{pmatrix} 0.82601 & -0.71436 & -0.01653 \\ 1.52400 & -1.31800 & -0.03050 \\ 0.19050 & -0.16475 & -0.00381 \\ -1.68000 & 0.76200 & 1.10250 \\ 0.98100 & 0.54600 & -0.54300 \end{pmatrix}$$

where the permanent shocks are now more complicated to interpret; this just reflects that the level, slope and curvature are *not* simple accumulated shocks to simple equations.

6 Extensions

[21] The granger representation is given by

$$x_t = C\sum_{i=1}^{t}\epsilon_i + C_0^*\epsilon_t + C_1^*\epsilon_{t-1} + C_2^*\epsilon_{t-2} + ... + A,$$

where A depends on initial values, $C = \beta_\perp\left(\alpha_\perp'\Gamma\beta_\perp\right)^{-1}\alpha_\perp'$, and C_i^* are convergent. The idea of the structural MA form is to find a representation

$$\begin{aligned} x_t &= CB^{-1}\sum_{i=1}^{t}u_i + C_0^*B^{-1}u_t + C_1^*B^{-1}u_{t-1} + C_2^*B^{-1}u_{t-2} + ... + A \\ &= R\sum_{i=1}^{t}u_i + R_0^*u_t + R_1^*u_{t-1} + R_2^*u_{t-2} + ... + A, \end{aligned}$$

with structural shocks $u_t = B\epsilon_t$. Typically, we require $E\left[u_t u_t'\right] = I_p$, such that the structural shocks are orthogonal, at it is straightforward to interpret the impulse response functions.

Now the solution could explain that the orthogonalization of shocks involves at Choleski decomposition, and that the zeros in the lower triangular Choleski factor can be places in the long-run impact R by a suitable rotation. This material is quite detailed, and the notation is not easy, but the answer should have an idea about the reason for doing it and an idea about the implementation.

For our case of $p = 5$ and $r = 2$, the true loading to the interpretable factors (level, slope and curvature) is

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1/4 & 1 \\ 1 & 2/4 & 2 \\ 1 & 3/4 & 1 \\ 1 & 1 & 0 \end{pmatrix},$$

and a natural identification for the structural long-run impact matrix could be

$$R = \begin{pmatrix} 0 & 0 & * & 0 & 0 \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & 0 \end{pmatrix},$$

where $*$ is an unrestricted coefficient, and $E\left[u_t u_t'\right] = I_p$. For the current data the estimated long-run impact (based on the over-identified β and α unrestricted is

$$\hat{R} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0.899 & 0.339 & 0.500 \\ 0 & 0 & 0.798 & 0.678 & 1 \\ 0 & 0 & 0.523 & 0.839 & 0.500 \\ 0 & 0 & 0.248 & 1 & 0 \end{pmatrix}.$$

The reason that the impact of the first factor does not resemble a level factor is that maintained assumption $E[u_t u_t'] = I_p$ is not valid for the 'economic' shocks. In this case the structural model is difficult to interpret and difficult to reconcile with the theoretical results. This question is probably difficult.

[22] This may also be a difficult question that asks them to build a model for a new case, and we do not require very much. In the homoskedastic case, $\epsilon_t | x_{t-1}, ..., x_{t-k} \sim N(0, \Omega)$, then the likelihood function is given by

$$L(\theta) = \prod_{t=1}^{T} f_\theta(x_t \mid x_{t-1}, ..., x_{t-k}).$$

where each term is given by

$$f_\theta(x_t \mid x_{t-1}, ..., x_{t-k}) = (2\pi)^{-\frac{p}{2}} |\Omega|^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}\epsilon_t(\theta)'\Omega^{-1}\epsilon_t(\theta)\right),$$

and

$$\epsilon_t(\theta) = x_t - \sum_{i=1}^{k} \Pi_i x_{t-i} - \mu_0 - \phi D_t.$$

The parameters are Ω and $\theta = \{\Pi_1, ..., \Pi_k, \mu_0, \phi\}$.

In the case of multivariate ARCH, i.e. if $\epsilon_t | x_{t-1}, ..., x_{t-k} \sim N(0, \Omega_t)$, with conditional covariance matrix given by

$$\Omega_t = \Omega + A\epsilon_{t-1}\epsilon_{t-1}' A',$$

the likelihood function is the same with each term replaced with

$$\begin{aligned} f_\theta(x_t \quad &| \quad x_{t-1}, ..., x_{t-k}) \\ &= (2\pi)^{-\frac{p}{2}} |\Omega_t|^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}\epsilon_t(\theta)'\Omega_t^{-1}\epsilon_t(\theta)\right) \\ &= (2\pi)^{-\frac{p}{2}} \left|\Omega + A\epsilon(\theta)_{t-1}\epsilon(\theta)_{t-1}' A'\right|^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}\epsilon_t(\theta)'\left(\Omega + A\epsilon(\theta)_{t-1}\epsilon(\theta)_{t-1}' A'\right)^{-1}\epsilon_t(\theta)\right), \end{aligned}$$

where the parameters are now Ω, A and $\theta = \{\Pi_1, ..., \Pi_k, \mu_0, \phi\}$. There is no closed form solution, but the likelihood function can be maximized using a standard numerical procedure.