“Økonomisk Kandidateksamen 2003 II”

Advanced Econometrics (“Videregående Økonometri”)

Project Examination

June 23-25, 2003

The answers given here are based on the following dataset:

$$x_t' = [ESP_t, ITA_t, FRA_t, DEU_t, USA_t]$$

1. Defining the model and the data:

The purpose of this take-home exam is to investigate cointegration and common trends properties of an international portfolio of five daily bond rates of 10 year maturity over the period 1.3.94-25.9.95, spanning approximately 1.5 years.

We consider the following VAR model:

$$\begin{aligned} \Delta x_t &= \Gamma_1 \Delta x_{t-1} + \Pi x_{t-1} + \Phi D_t + \mu_0 + \varepsilon_t, \\ \varepsilon_t &\sim N_{iid}(0, \Omega\ ),\ t = 1, ..., T, \quad x_{-1}, x_0 \text{ given.} \end{aligned} \tag{1}$$

The vector x_t contains the following variables:

$$x_t' = [b1_t, b2_t, b3_t, b4_t, b5_t]$$

measuring daily bond rates (of 10 years maturity) from five different countries as described in your CATS program file (bonds.prg). In addition to the bonds.prg file your exam folder contains the following files: a RATS data file (bonds.rat) and two GiveWin files (bonds.in7 and bonds.bn7). You are free to use any (or both) of the two software packages. The CATS program file (bonds.prg) reads the data series, defines the sample to be analyzed and contains the basic CATS commands for setting up the cointegration model. The GiveWin files contain the data in levels and differences, and the dummy variables.

The vector D_t contains a total of 13 permanent and transitory dummy variables which have been included in the VAR model from the outset. Generally a permanent blip dummy, $DpXX_t$, takes on the value 1 at period XX and the value 0 otherwise, whereas a transitory dummy, $DtrXX_t$, takes on the value 1 at period XX, -1 at period XX+1, and 0 otherwise.

2. Specification:

2.1. Examine the graphs of your variables in levels and differences to find out whether the levels/differences look nonstationary/stationary.

Answer: The levels are drifting off in a nonstationary manner whereas the differences look strongly mean reverting (but a number of outlier observations might violate the normality assumption).

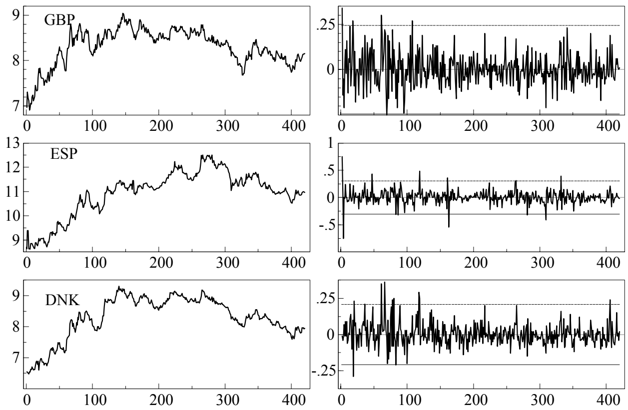

Figure 1:

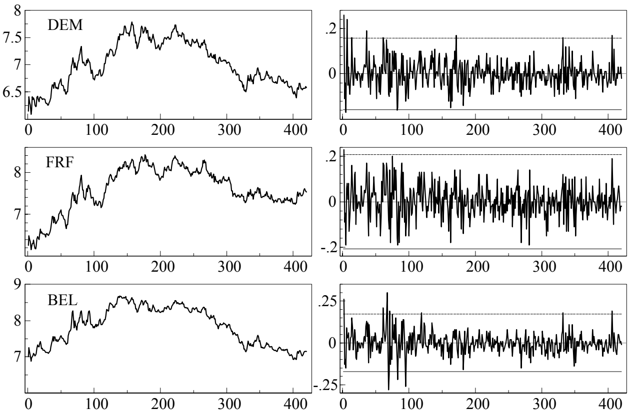

Figure 2:

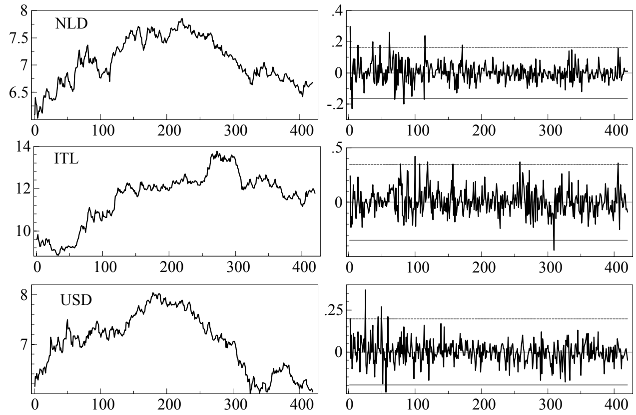

Figure 3:

2.2. Is it likely that some of the variables contain a linear trend? Motivate your answer. Assume that the data do not contain linear trends. How would you specify the constant term μ_0 in the cointegrated VAR model (1)?

Answer: It is highly unlikely that bonds rate would contain a linear trends, as it would imply that part of the bond yield would be predictable which is against the assumption of rational behavior in the financial market. The constant term should be specified so that the projection onto the nonstationary space, $\beta_\perp$, is restricted to zero, i.e. $\mu_0 = \alpha(\beta'\alpha)^{-1}\beta'\mu_0$. In practise, this can be achieved by restricting the constant to be in the cointegration relations.

2.3. The 15 dummy variables have been selected based on a VAR model containing nine different bond rates. Since your model contains only a subset of these nine variables, some of the presently included dummies may not be significant, others might be highly significant. You should now consider the possibility that there has been a shift in the mean of the cointegration relations on the day defined by the most significant permanent dummy variable. Re-specify your model so that you can test this hypothesis and explain the underlying logic of your re-specification. When you re-estimate your model leave out all dummy variables with an absolute t-ratio less than 3.5.

Answer: A permanent blip in the Δx_t correspond to a level shift in the x_t. Thus, blip in the differences might correspond to a mean shift in the cointegration relations. (Furthermore, by allowing for this possibility we achieve 'similarly' in the rank test.). Therefore, a big permanent outlier blip in the differences, $Dpxx_t = 1 \,\text{at}\, xx,\ 0, \text{otherwise}$, might signal an equilibrium mean shift and we should test for this possibility by including a shift dummy $Dsxx_t = 1$ for $t = xx, ..., 0\,\text{otherwise}$. My comment: Unfortunately, there was a big permanent outlier at observation 4, which in some cases turned out to be the largest. A mean shift in this case would probably not be detectable and it would be more

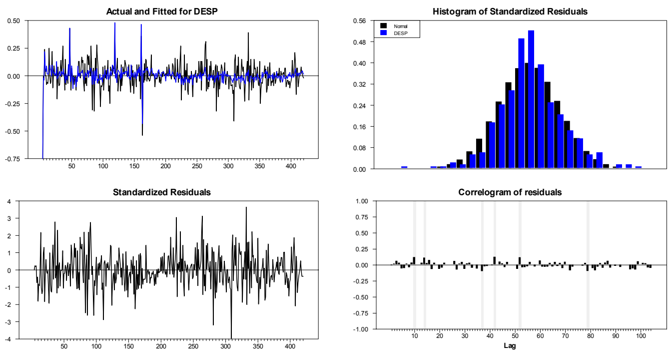

Figure 4: One Example of the residual graphs.

relevant to focus on the next big one.

2.4. Examine the graphs of the VAR residuals to find out whether the multivariate normality assumption underlying the re-specified VAR model seems reasonable. Examine also whether the residuals look homoscedastic and uncorrelated. Compare this visual inspection with the formal misspecification tests of residual normality, residual ARCH and residual autocorrelation. Is your VAR model an adequate description of the observed data? Motivate! Would you trust the subsequently estimated results based on this model. Motivate!

Answer: There are clearly outlier observations, but the homoscedasticity assumption is not too bad. Violation of normality because of skewness can bias the estimates, whereas excess kurtosis seems to be less serious. Cointegration estimates seem to be quite robust against ARCH according to the results by Rahbek + coauthors.

2.5. Determine the rank r of $\Pi = \alpha\beta'$ based on the trace test.

2.5.a Which asymptotic tables are most appropriate in this case? Are they likely to be good approximations? Discuss!

Answer: Table 2 is the closest asymptotic table, but is not correct because of the shift dummy. The asymptotic table + a χ^2 would be roughly OK.

2.5.b Check the results of the trace test with other useful information in your estimated model. Is your choice of r consistent with this information? Motivate your answer!

Answer: Check the roots of the model, the t-ratios of the r:th β vector, the graphs of the r:th cointegration relation possibly the recursive graphs of the trace tests (should grow linearly if $\lambda_r > 0$).

3. Testing:

Independently of your choice of r in question 2.4 continue with $r = 2$. My comment: This turned out to be a problem in some cases as the rank was quite clearly 1. However, to illustrate long-run identification it was important to assume 2.

3. 1. Test the hypothesis that there is no mean shift in the cointegration relations on the day of the most significant dummy variable (as selected in question 2.2). Test also the hypothesis that any one of the bond rates can be excludable from the cointegration relations (Hint: use the option proc=tsprop in CATS).

Answer: This should be straigtforward.

3.2. Test the hypothesis that any one of the bond rates is weakly exogenous for the long-run parameters (Hint: use the option proc=tsprop in CATS). If you find more than one weakly exogenous variable test the joint hypothesis of weak exogeneity. Interpret the results. Does weak exogeneity of a variable imply that it could not have been affected by any of the cointegration relations under any circumstances? (Hint: use the estimated residual covariance matrix as an illustration.) Motivate your answer!

Answer: testing should be straightforward. Weak exogeneity is not invariant under linear transformations. Hence, conditioning on current changes of variables might very well change weak exogeneity found in the reduced form.

3.3. Assume that one of the bond rates is both long-run excludable and weakly exogenous in your model. Would it be a good idea to re-specify your model, and in that case how? Motivate your answer!

Answer: Even if a variable is found to be long-run excludable and weakly exogenous, it is still possible that it is very significant in the Γ_1 matrix and, thus, have explanatory power. But, if the column of Γ_1 corresponding to such a variable had no significant, or only borderline significant coefficients then it would be better to throw it out.

3.4. Test the hypothesis of long-run price homogeneity between the bond rates in both of your cointegration relations (assuming $r = 2$). Specify the design matrices H and R. How many binding restrictions are there in this test? Answer: $r \times 1 = 2$.

How many free parameters? Answer: $(p-1) \times 2$. Note, we have not required identification. Motivate!

Assume that you were able to accept the homogeneity hypothesis. How many common trends would this imply between the included bond rates? Answer: $p - r = 3$ common trends which influence the bond rates in exactly the same way. However, if one of the bond rates was found to be long-run excludable, its cumulated residuals would be a stochastic trend which would not have an influence on the other bond rates.

4. Identification of the long-run structure

4. 1. Estimate a just-identified structure for the two cointegration relations. Is the structure empirically identified? Motivate your answer!

Answer: I would expect students to impose one identifying zero restriction on each equation and to normalize on a nonzero coefficient, i.e. on a variable which corresponds to a nonzero α coefficient. Empirical identification would then require that the estimated coefficient of the variable which was set to zero in the other cointegration relation should be statistically significant. My comment: Some students can have experienced problems here, if the choice of $r = 2$ was clearly wrong. In this case it might be difficult to find an empirically just identified structure.

4. 2. If possible, impose further restrictions on the cointegration relations. Are they over-identifying? Motivate your answer! (Avoid calculating the rank conditions using the matrices **H** and **R.** It will take too long time to do it!)

Answer: If any of the estimated coefficients of the just-identified structure is insignificant, then one should be able to impose such restrictions without loosing cointegration (but, it should of course be tested!). However, a zero coefficient on the same variable in both cointegration relations is a testable restriction but not an overidentifying restriction. The same is true if one imposes a homogeneity restriction on both relations.

5. The identification of the short-run structure

5. 1. Estimate a parsimonious short-run adjustment structure (i.e. with few and significant coefficients) for the five bond rates by keeping the two cointegration relations at their estimated values (Hint: use the option 'calculator' in GiveWin to create the estimated cointegration variables). At this stage you can disregard the possibility of simultaneous effects. My comment: I think there is a bug in the PcGive calculator, which happens when you first make a mistake and then overwrite the incorrectly calculated relation with the revised one. Except for this it should be straightforward to do this part and it serves essentially the purpose of estimating a parsimonious forecasting model. Otherwise more or less a check that students have learned how to use PcFiml.

5. 2. Re-estimate the short-run structure by allowing for simultaneous effects between the bond rates, so that the off-diagonal elements of the standardized residual covariance matrix (the residual correlation matrix) become as small as possible. Comment on the results.

Answer: This exercise has probably been quite difficult. The reason is that there is a lot of simultaneous effects in this data set and it is not obvious that all equations can be both generically and empirically identified. I myself was not able to find an identified system accounting for most of the current correlations. There is always the possibility to estimate the system by OLS, but the likelihood function does not really have a meaning and the estimated coefficients will suffer from simultaneity bias (the last point has not been discussed during my course so students should not be penalized even if they do not mention this)

5. 3. Re-estimate the system as a partial system by conditioning on the US bond rate and interpret the results. Discuss whether this is a good idea or not. For example: Are the empirical results robust to this change of model

specification? Do we loose any important information on how the bond rates are being generated? Do we gain anything?

Answer: US bond rate is not generally weakly or strongly exogenous in this system. Thus, we loose some information regarding the feed-back from Europe to USA. In most cases this information is probably not very important and is likely to be compensated by more robust estimates. If it is the case that the US bond market is the main driving force behind the Euro bond market then conditioning on the US bond rate will probably be a good idea. One should note, however, that the different time zones may play an important role when deciding whether conditioning or not. This is more generally so for modelling the current effects.