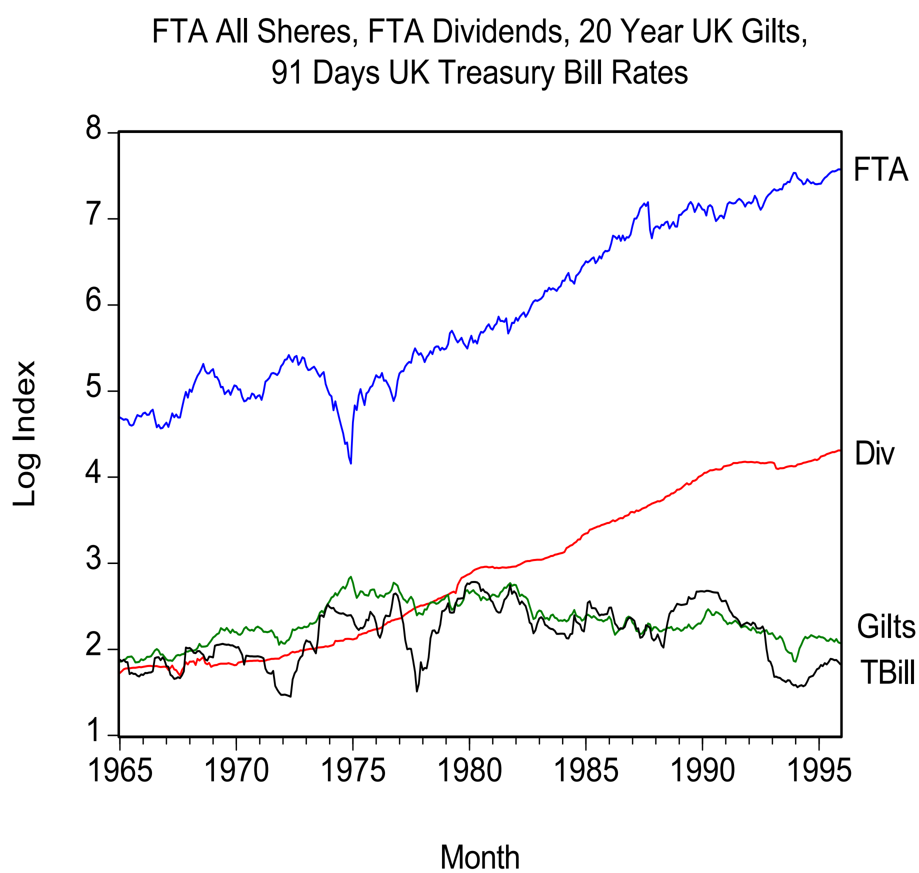
2. Multivariate Time Series

2.1 *Background*

Example. Consider the following monthly observations on FTA All Share index, the associated dividend index and the series of 20 year UK gilts and 91 day Treasury bills from January 1965 to December 1995 (372 months)



Potentially interesting questions:

1. Do some markets have a tendency to lead others?

2. Are there feedbacks between the markets?

3. How about contemporaneous movements?

4. How do impulses (shocks, innovations) transfer from one market to another?

5. How about common factors (disturbances, trend, yield component, risk)?

Most of these questions can be empirically investigated using tools developed in multivariate time series analysis.

Time series models

AR(p)-model

$$\begin{aligned} y_t &= \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t \\ \phi(L) y_t &= \mu + \epsilon_t \end{aligned}$$

where $\epsilon_t \sim \mathrm{WN}(0, \sigma_\epsilon^2)$ (White Noise), i.e.

$$\begin{aligned} E(\epsilon_t) &= 0, \\ E(\epsilon_t \epsilon_s) &= \begin{cases} \sigma_\epsilon^2 & \text{if } t = s \\ 0 & \text{otherwise,} \end{cases} \end{aligned}$$

and $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$ is the lag polynomial of order p with

$$L^k y_t = y_{t-k}$$

being the *L*ag operator ($L^0 y_t = y_t$).

The so called (weak) stationarity condition requires that the roots of the (characteristic) polynomial

$$\phi(L) = 0$$

should lie outside the unit circle, or equivalently the roots of

$$z^p - \phi_1 z^{p-1} - \cdots - \phi_{p-1} z - \phi_p = 0$$

are less than one in absolute value.

Note. Usually the series are centralized such that $\mu = 0$.

MA(q)-model

$$\begin{aligned} y_t &= \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q} \\ &= \mu + \theta(L)\epsilon_t, \end{aligned}$$

where $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$ is again a polynomial in L, this time, of order q, and $\epsilon_t \sim$ WN$(0, \sigma_\epsilon^2)$.

Note. An MA-process is always stationary. But the so called invertibility condition requires that the roots of the characteristic polynomial $\theta(L) = 0$ lie outside the unit circle.

<u>ARMA(p, q)-process</u>

Compiling the two above together yields an ARMA(p, q)-process

$$\phi(L)y_t = \mu + \theta(L)\epsilon_t.$$

<u>ARIMA(p, d, q)-process</u>

A series is called integrated of order d, denoted as $y_t \sim I(d)$, if it becomes stationary after differencing d times. Furthermore, if

$$(1 - L)^d y_t \sim \text{ ARMA}(p, q)$$

we say that $y_t \sim$ ARIMA(p, d, q), where p denotes the order of the AR-lags, q the order of MA-lags, and d the order of differencing.

Example. Univariate time series models for the above (log) series look as follows. All the series prove to be $I(1)$.

Sample: 1965:01 1995:12
Included observations: 372

FTA

<table>
<tr><th>Autocorrelation</th><th>Partial Correlation</th><th></th><th>AC</th><th>PAC</th><th>Q-Stat</th><th>Prob</th></tr>
<tr><td>.|********</td><td>.|********</td><td>1</td><td>0.992</td><td>0.992</td><td>368.98</td><td>0.000</td></tr>
<tr><td>.|********</td><td>.|. |</td><td>2</td><td>0.983</td><td>-0.041</td><td>732.51</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>3</td><td>0.975</td><td>0.021</td><td>1090.9</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>4</td><td>0.966</td><td>-0.025</td><td>1444.0</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>5</td><td>0.957</td><td>-0.024</td><td>1791.5</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>6</td><td>0.949</td><td>0.008</td><td>2133.6</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>7</td><td>0.940</td><td>0.005</td><td>2470.5</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>8</td><td>0.931</td><td>-0.007</td><td>2802.1</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>9</td><td>0.923</td><td>0.023</td><td>3128.8</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>10</td><td>0.915</td><td>-0.011</td><td>3450.6</td><td>0.000</td></tr>
</table>

Dividends

<table>
<tr><th>Autocorrelation</th><th>Partial Correlation</th><th></th><th>AC</th><th>PAC</th><th>Q-Stat</th><th>Prob</th></tr>
<tr><td>.|********</td><td>.|********</td><td>1</td><td>0.994</td><td>0.994</td><td>370.59</td><td>0.000</td></tr>
<tr><td>.|********</td><td>.|. |</td><td>2</td><td>0.988</td><td>-0.003</td><td>737.78</td><td>0.000</td></tr>
<tr><td>.|********</td><td>.|. |</td><td>3</td><td>0.982</td><td>0.002</td><td>1101.6</td><td>0.000</td></tr>
<tr><td>.|********</td><td>.|. |</td><td>4</td><td>0.976</td><td>-0.004</td><td>1462.1</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>5</td><td>0.971</td><td>-0.008</td><td>1819.2</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>6</td><td>0.965</td><td>-0.006</td><td>2172.9</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>7</td><td>0.959</td><td>-0.007</td><td>2523.1</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>8</td><td>0.953</td><td>-0.004</td><td>2869.9</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>9</td><td>0.947</td><td>-0.006</td><td>3213.3</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>10</td><td>0.940</td><td>-0.006</td><td>3553.2</td><td>0.000</td></tr>
</table>

T-Bill

<table>
<tr><th>Autocorrelation</th><th>Partial Correlation</th><th></th><th>AC</th><th>PAC</th><th>Q-Stat</th><th>Prob</th></tr>
<tr><td>.|********</td><td>.|********</td><td>1</td><td>0.980</td><td>0.980</td><td>360.26</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>**|. |</td><td>2</td><td>0.949</td><td>-0.301</td><td>698.79</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>3</td><td>0.916</td><td>0.020</td><td>1014.9</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>4</td><td>0.883</td><td>-0.005</td><td>1309.5</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>5</td><td>0.849</td><td>-0.041</td><td>1583.0</td><td>0.000</td></tr>
<tr><td>.|****** |</td><td>*|. |</td><td>6</td><td>0.811</td><td>-0.141</td><td>1833.1</td><td>0.000</td></tr>
<tr><td>.|****** |</td><td>.|. |</td><td>7</td><td>0.770</td><td>-0.018</td><td>2059.2</td><td>0.000</td></tr>
<tr><td>.|****** |</td><td>.|. |</td><td>8</td><td>0.730</td><td>0.019</td><td>2263.1</td><td>0.000</td></tr>
<tr><td>.|***** |</td><td>.|. |</td><td>9</td><td>0.694</td><td>0.058</td><td>2447.6</td><td>0.000</td></tr>
<tr><td>.|***** |</td><td>.|. |</td><td>10</td><td>0.660</td><td>-0.013</td><td>2615.0</td><td>0.000</td></tr>
</table>

Gilts

<table>
<tr><th>Autocorrelation</th><th>Partial Correlation</th><th></th><th>AC</th><th>PAC</th><th>Q-Stat</th><th>Prob</th></tr>
<tr><td>.|********</td><td>.|********</td><td>1</td><td>0.984</td><td>0.984</td><td>362.91</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>*|. |</td><td>2</td><td>0.962</td><td>-0.182</td><td>710.80</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>3</td><td>0.941</td><td>0.050</td><td>1044.6</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>4</td><td>0.921</td><td>0.015</td><td>1365.5</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>5</td><td>0.903</td><td>0.031</td><td>1674.8</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>6</td><td>0.885</td><td>-0.038</td><td>1972.4</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>7</td><td>0.866</td><td>-0.001</td><td>2258.4</td><td>0.000</td></tr>
<tr><td>.|*******|</td><td>.|. |</td><td>8</td><td>0.848</td><td>0.019</td><td>2533.6</td><td>0.000</td></tr>
<tr><td>.|****** |</td><td>.|. |</td><td>9</td><td>0.832</td><td>0.005</td><td>2798.6</td><td>0.000</td></tr>
<tr><td>.|****** |</td><td>.|. |</td><td>10</td><td>0.815</td><td>-0.014</td><td>3053.6</td><td>0.000</td></tr>
</table>

Formally, as is seen below, the Dickey-Fuller (DF) unit root tests indicate that the series indeed all are $I(1)$. The test is based on the augmented DF-regression

$$\Delta y_t = \rho y_{t-1} + \alpha + \delta t + \sum_{i=1}^{4} \phi_i \Delta y_{t-i} + \epsilon_t,$$

and the hypothesis to be tested is

$$H_0 : \rho = 0 \text{ vs } H_1 : \rho < 0.$$

Test results:

Series	$\widehat{\rho}$	t-Stat
FTA	-0.030	-2.583
DIV	-0.013	-2.602
R20	-0.013	-1.750
T-BILL	-0.023	-2.403
ΔFTA	-0.938	-8.773
ΔDIV	-0.732	-7.300
ΔR20	-0.786	-8.129
ΔT-BILL	-0.622	-7.095

	ADF critical values	
Level	No trend	Trend
1%	-3.4502	-3.9869
5%	-2.8696	-3.4237
10%	-2.5711	-3.1345

Provided that the series are not *cointegrated* an appropriate modeling approach is VAR for the differences.

2.2 *Vector Autoregression (VAR)*

Suppose we have m time series y_{it}, $i = 1, \ldots, m$, and $t = 1, \ldots, T$ (common length of the time series). Then a vector autoregression model is defined as

$$
\begin{pmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{mt} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix} + \begin{pmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} & \cdots & \phi_{1m}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} & \cdots & \phi_{2m}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{m1}^{(1)} & \phi_{m2}^{(1)} & \cdots & \phi_{mm}^{(1)} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{m,t-1} \end{pmatrix} +
$$

$$
\cdots + \begin{pmatrix} \phi_{11}^{(p)} & \phi_{12}^{(p)} & \cdots & \phi_{1m}^{(p)} \\ \phi_{21}^{(p)} & \phi_{22}^{(p)} & \cdots & \phi_{2m}^{(p)} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{m1}^{(p)} & \phi_{m2}^{(1)} & \cdots & \phi_{mm}^{(p)} \end{pmatrix} \begin{pmatrix} y_{1,t-p} \\ y_{2,t-p} \\ \vdots \\ y_{m,t-p} \end{pmatrix}
$$

$$
+ \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \vdots \\ \epsilon_{mt} \end{pmatrix}.
$$

In matrix notations

$$
\mathbf{y}_t = \mu + \Phi_1 \mathbf{y}_{t-1} + \cdots + \Phi_p \mathbf{y}_{t-p} + \epsilon_t,
$$

which can be further simplified by adopting the matric form of a lag polynomial

$$\Phi(L) = \mathbf{I} - \Phi_1 L - \ldots - \Phi_p L^p.$$

Thus finally we get

$$\Phi(L)\mathbf{y}_t = \epsilon_t.$$

Note that each y_{it} does not only depend on its own history but also on the other series' history (cross dependencies). This gives us several additional tools for analyzing causal as well as feedback effects as we shall see after a while.

A basic assumption in the above model is that the residual vector follow a multivariate white noise, i.e.

$$\begin{aligned} \mathrm{E}(\epsilon_t) &= \mathbf{0} \\ \mathrm{E}(\epsilon_t \epsilon_s') &= \begin{cases} \Sigma_\epsilon & \text{if } t = s \\ \mathbf{0} & \text{if } t \neq s \end{cases}, \end{aligned}$$

which allows for estimation by OLS, because each individual residual series is assumed to be serially uncorrelated with constant variance.

The coefficient matrices must satisfy certain constraints in order that the VAR-model is stationary. They are just analogies with the univariate case, but in matrix terms. It is required that the roots of

$$|\mathbf{I} - \Phi_1 z - \Phi_2 z^2 - \cdots - \Phi_p z^p| = 0$$

lie outside the unit circle. Estimation can be carried out by single equation least squares.

Example. Let us estimate a VAR model for the equity-bond data. First, however, test whether the series are cointegrated. As is seen below, there is no empirical evidence of cointegration (EViews results)

```
Sample(adjusted): 1965:06 1995:12
Included observations: 367 after adjusting end points
Trend assumption: Linear deterministic trend
Series: LFTA LDIV LR20 LTBILL
Lags interval (in first differences): 1 to 4

Unrestricted Cointegration Rank Test
=====================================================================
Hypothesized                 Trace      5 Percent       1 Percent
No. of CE(s)  Eigenvalue  Statistic  Critical Value  Critical Value
=====================================================================
None           0.047131    46.02621       47.21           54.46
At most 1      0.042280    28.30808       29.68           35.65
At most 2      0.032521    12.45356       15.41           20.04
At most 3      0.000872    0.320012        3.76            6.65
=====================================================================
 *(**) denotes rejection of the hypothesis at the 5%(1%) level
 Trace test indicates no cointegration at both 5% and 1% levels
```

VAR(2) Estimates:

Sample(adjusted): 1965:04 1995:12
Included observations: 369 after adjusting end points
Standard errors in ( ) & t-statistics in [ ]

	DFTA	DDIV	DR20	DTBILL
DFTA(-1)	0.102018	-0.005389	-0.140021	-0.085696
	(0.05407)	(0.01280)	(0.02838)	(0.05338)
	[1.88670]	[-0.42107]	[-4.93432]	[-1.60541]
DFTA(-2)	-0.170209	0.012231	0.014714	0.057226
	(0.05564)	(0.01317)	(0.02920)	(0.05493)
	[-3.05895]	[0.92869]	[0.50389]	[1.04180]
DDIV(-1)	-0.113741	0.035924	0.197934	0.280619
	(0.22212)	(0.05257)	(0.11657)	(0.21927)
	[-0.51208]	[0.68333]	[1.69804]	[1.27978]
DDIV(-2)	0.065178	0.103395	0.057329	0.165089
	(0.22282)	(0.05274)	(0.11693)	(0.21996)
	[0.29252]	[1.96055]	[0.49026]	[0.75053]
DR20(-1)	-0.359070	-0.003130	0.282760	0.373164
	(0.11469)	(0.02714)	(0.06019)	(0.11322)
	[-3.13084]	[-0.11530]	[4.69797]	[3.29596]
DR20(-2)	0.051323	-0.012058	-0.131182	-0.071333
	(0.11295)	(0.02673)	(0.05928)	(0.11151)
	[0.45437]	[-0.45102]	[-2.21300]	[-0.63972]
DTBILL(-1)	0.068239	0.005752	-0.033665	0.232456
	(0.06014)	(0.01423)	(0.03156)	(0.05937)
	[1.13472]	[0.40412]	[-1.06672]	[3.91561]
DTBILL(-2)	-0.050220	0.023590	0.034734	-0.015863
	(0.05902)	(0.01397)	(0.03098)	(0.05827)
	[-0.85082]	[1.68858]	[1.12132]	[-0.27224]
C	0.892389	0.587148	-0.033749	-0.317976
	(0.38128)	(0.09024)	(0.20010)	(0.37640)
	[2.34049]	[6.50626]	[-0.16867]	[-0.84479]

Continues ...

	DFTA	DDIV	DR20	DTBILL
R-squared	0.057426	0.028885	0.156741	0.153126
Adj. R-squared	0.036480	0.007305	0.138002	0.134306
Sum sq. resids	13032.44	730.0689	3589.278	12700.62
S.E. equation	6.016746	1.424068	3.157565	5.939655
F-statistic	2.741619	1.338486	8.364390	8.136583
Log likelihood	-1181.220	-649.4805	-943.3092	-1176.462
Akaike AIC	6.451058	3.569000	5.161567	6.425267
Schwarz SC	6.546443	3.664385	5.256953	6.520652
Mean dependent	0.788687	0.688433	0.052983	-0.013968
S.D. dependent	6.129588	1.429298	3.400942	6.383798

Determinant Residual Covariance	18711.41
Log Likelihood (d.f. adjusted)	-3909.259
Akaike Information Criteria	21.38352
Schwarz Criteria	21.76506

As is seen the number of estimated parameters grows rapidly very large.

Defining the order of a VAR-model

In the first step it is assumed that all the series in the VAR model have equal lag lengths. To determine the number of lags that should be included, multivariate extensions of criterion functions like SC and AIC can be utilized in the same manner as in the univariate case.

The likelihood ratio (LR) test can also be used in determining the order of a VAR. The test is generally of the form

$$\text{LR} = T(\log|\hat{\Sigma}_k| - \log|\hat{\Sigma}_m|)$$

where $\hat{\Sigma}_k$ denotes the maximum likelihood estimate of the residual covariance matrix of VAR(k) and $\hat{\Sigma}_m$ the estimate of VAR(m) ($m > k$) residual covariance matrix. If VAR(k) (the shorter model) is the true one, then

$$\text{LR} \sim \chi^2_{df},$$

where the degrees of freedom, df, equals the difference of in the number of estimated parameters between the two models.

In a p variate VAR(k)-model each series has $m - k$ lags less than those in VAR(m). Thus the difference in each equation is $p(m - k)$, so that in total $df = p^2(m - k)$.

Note that often, when T small a modified LR

$$\text{LR}^* = (T - pm)(\log|\hat{\Sigma}_k| - \log|\hat{\Sigma}_m|)$$

is used to correct for small sample bias.

Example Let $m = 12$ then in the equity-bond data different VAR models yield the following results. Below are EViews results.

```
VAR Lag Order Selection Criteria
Endogenous variables: DFTA DDIV DR20 DTBILL
Exogenous variables: C
Sample: 1965:01 1995:12
Included observations: 359
==========================================================
Lag   LogL      LR     FPE       AIC      SC       HQ
----------------------------------------------------------
 0  -3860.59    NA    26324.1   21.530   21.573   21.547
 1  -3810.15  99.473  21728.6*  21.338*  21.554*  21.424*
 2  -3796.62  26.385  22030.2   21.352   21.741   21.506
 3  -3786.22  20.052  22729.9   21.383   21.945   21.606
 4  -3783.57  5.0395  24489.4   21.467   22.193   21.750
 5  -3775.66  14.887  25625.4   21.502   22.411   21.864
 6  -3762.32  24.831  26016.8   21.517   22.598   21.947
 7  -3753.94  15.400  27159.4   21.560   22.814   22.059
 8  -3739.07  27.018* 27348.2   21.566   22.994   22.134
 9  -3731.30  13.933  28656.4   21.612   23.213   22.248
10  -3722.40  15.774  29843.7   21.651   23.425   22.357
11  -3715.54  12.004  31443.1   21.702   23.649   22.476
12  -3707.28  14.257  32880.6   21.745   23.865   22.588
==========================================================
 * indicates lag order selected by the criterion
 LR: sequential modified LR test statistic
     (each test at 5% level)
 FPE: Final prediction error
 AIC: Akaike information criterion
 SC: Schwarz information criterion
 HQ: Hannan-Quinn information criterion
```

Criterion function minima are all at VAR(1) (SC or BIC just borderline). LR-tests suggest VAR(8) based upon the first significant LR$(m, m-1)$ statistics starting from $m = 12$.

Model Diagnostics

Even though the information criteria were seen to favour VAR(1), we decided to fit a VAR(2) model to the data, because the VAR(1) residuals do not pass the white noise test, as we shall see below.

To investigate whether the VAR residuals are white noise, the hypothesis to be tested is

$$H_0 : \Upsilon_1 = \cdots = \Upsilon_h = 0,$$

where $\Upsilon_k = (\gamma_{ij}(k))$ denotes the matrix of the k'th cross autocovariances of the residuals series ϵ_i and ϵ_j:

$$\gamma_{ij}(k) = E(\epsilon_{i,t-k} \cdot \epsilon_{j,t}),$$

whose diagonal elements reduce to the usual autocovariances γ_k. Note, however, that cross-autocovariances, unlike univariate autocovariances, are not symmetric in k, that is $\gamma_{i,j}(k) \neq \gamma_{i,j}(-k)$, because the covariance between residual series i and residual series j k steps ahead is in general not the same as the covariance between residual series i and the residual series j k steps before. Stationarity ensures, however, that $\Upsilon_k = \Upsilon'_{-k}$ (exercise).

In order to test $H_0 : \Upsilon_1 = \cdots = \Upsilon_h = 0$, we may use the (Portmanteau) Q-statistics[††]

$$Q_h = T \sum_{k=1}^{h} \text{tr}(\widehat{\Upsilon}_k' \widehat{\Upsilon}_0^{-1} \widehat{\Upsilon}_k \widehat{\Upsilon}_0^{-1})$$

where

$$\widehat{\Upsilon}_k = (\widehat{\gamma}_{ij}(k)) \text{ with } \widehat{\gamma}_{ij}(k) = \frac{1}{T-k} \sum_{t-k}^{T} \widehat{\epsilon}_{t-k,i} \widehat{\epsilon}_{t,j}$$

are the estimated (residual)cross autocovariances and $\widehat{\Upsilon}_0$ the contemporaneous covariances of the residuals. Alternatively (especially in small samples) a modified statistic is used

$$Q_h^* = T^2 \sum_{k=1}^{h} (T-k)^{-1} \text{tr}(\widehat{\Upsilon}_k' \widehat{\Upsilon}_0^{-1} \widehat{\Upsilon}_k \widehat{\Upsilon}_0^{-1}).$$

The statistics are asymptotically χ^2 distributed with $p^2(h-k)$ degrees of freedom. Note that in computer printouts h is running from $1, 2, \ldots h^*$ with h^* specified by the user.

[††]See e.g. Lütkepohl, Helmut (1993). *Introduction to Multiple Time Series*, 2nd Ed., Ch. 4.4

The table below contains the Q-statistics of the VAR(1) residuals fitted to financial market data from January 1965 to December 1995.

```
VAR(1) Residual Portmanteau Tests for Autocorrelations
H0: no residual autocorrelations up to lag h
Sample: 1966:02 1995:12
Included observations: 359
================================================
Lags   Q-Stat    Prob.   Adj Q-Stat  Prob.    df
------------------------------------------------
 1    1.847020   NA*      1.852179   NA*      NA*
 2    27.66930   0.0346   27.81912   0.0332   16
 3    44.05285   0.0761   44.34073   0.0721   32
 4    53.46222   0.2725   53.85613   0.2603   48
 5    72.35623   0.2215   73.01700   0.2059   64
 6    96.87555   0.0964   97.95308   0.0843   80
 7    110.2442   0.1518   111.5876   0.1320   96
 8    137.0931   0.0538   139.0485   0.0424  112
 9    152.9130   0.0659   155.2751   0.0507  128
10    168.4887   0.0797   171.2972   0.0599  144
11    179.3347   0.1407   182.4860   0.1076  160
12    189.0256   0.2379   192.5120   0.1869  176
================================================
*The test is valid only for lags larger than the
VAR lag order. df is degrees of freedom for
(approximate) chi-square distribution
```

There is still left some autocorrelation in the VAR(1) residuals. Let us next check the residuals of the VAR(2) model.

```
VAR Residual Portmanteau Tests for Autocorrelations
H0: no residual autocorrelations up to lag h
Sample: 1965:01 1995:12
Included observations: 369
==============================================
Lags   Q-Stat    Prob.   Adj Q-Stat Prob.  df
----------------------------------------------
 1    0.438464    NA*    0.439655   NA*   NA*
 2    1.623778    NA*    1.631428   NA*   NA*
 3    17.13353   0.3770  17.26832  0.3684  16
 4    27.07272   0.7143  27.31642  0.7027  32
 5    44.01332   0.6369  44.48973  0.6175  48
 6    66.24485   0.3994  67.08872  0.3717  64
 7    80.51861   0.4627  81.63849  0.4281  80
 8    104.3903   0.2622  106.0392  0.2271  96
 9    121.8202   0.2476  123.9049  0.2081 112
10    136.8909   0.2794  139.3953  0.2316 128
11    147.3028   0.4081  150.1271  0.3463 144
12    157.4354   0.5425  160.6003  0.4718 160
==============================================
*The test is valid only for lags larger than
the VAR lag order.
df is degrees of freedom for (approximate)
chi-square distribution
```

Now the residuals pass the white noise test. On the basis of these residual analyses we can select VAR(2) as the specification for further analysis. Mills (1999) finds VAR(6) as the most appropriate one. Note that there ordinary differences (opposed to log-differences) are analyzed. Here, however, log transformations are preferred.

Vector ARMA (VARMA)

Similarly as is done in the univariate case one can extend the VAR model to the vector ARMA model

$$\mathbf{y}_t = \mu + \sum_{i=1}^{p} \Phi_i \mathbf{y}_{t-i} + \epsilon_t + \sum_{j=1}^{q} \Theta_j \epsilon_{t-j}$$

or

$$\Phi(L)\mathbf{y}_t = \mu + \Theta(L)\epsilon_t,$$

where $\mathbf{y}_t$, μ, and ϵ_t are $m \times 1$ vectors, and Φ_i's and Θ_j's are $m \times m$ matrices, and

$$\begin{aligned}\Phi(L) &= \mathbf{I} - \Phi_1 L - \ldots - \Phi_p L^p \\ \Theta(L) &= \mathbf{I} + \Theta_1 L + \ldots + \Theta_q L^q.\end{aligned}$$

Provided that $\Theta(L)$ is invertible, we always can write the VARMA(p, q)-model as a VAR(∞) model with $\Pi(L) = \Theta^{-1}(L)\Phi(L)$. The presence of a vector MA component, however, implies that we can no longer find parameter estimates by ordinary least squares. We do not pursue our analysis to this direction.