Chapter 4: Stationary Dynamics for Financial Time Series

Financial Econometric Modeling

Stan Hurn, Vance Martin, Peter C.B. Phillips and Jun Yu

Oxford University Press, 2020

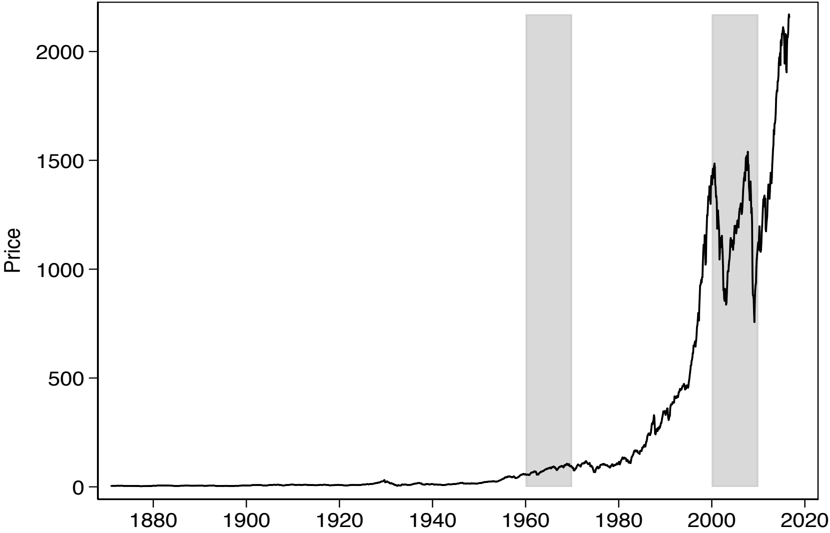
Introduction

- Thus far, the linear regression model has all the variables being measured at the same point in time.
- To allow financial variables to adjust to shocks over time this model is now extended to incorporate dynamic effects.
- In a scalar dynamic model, a single dependent financial variable is explained using its own past history as well as lags of other relevant financial variables.
- Single dependent variable models are often extended to multivariate specifications in which several financial variables are jointly determined and modelled together. Such models are heavily used in central banks, treasuries, international agencies, and the financial industry.

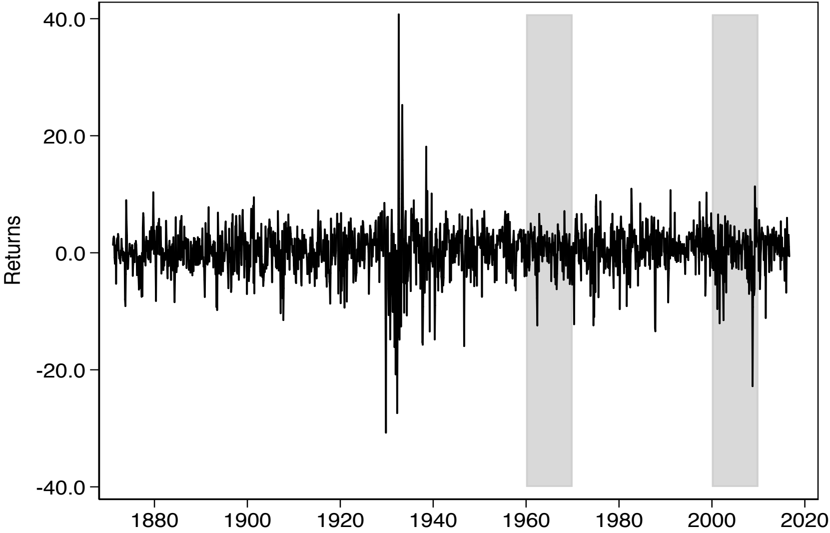
Stationarity

- Standard linear regression requires that the variables involved satisfy a simplifying condition known as stationarity.
- Technically, a time series is said to be covariance stationary, if the mean, variance, and autocovariances all remain invariant to the time periods in which they are calculated.
- Stationarity is important because it allows us to build standard models and use the past behaviour of variables to extrapolate their behaviour in the future.

S&P 500 index


Price
2000
1500
1000
500
0
1880
1900
1920
1940
1960
1980
2000
2020

S&P 500 returns



Single Equation Autoregressive Models

AR model

- The simplest possible dynamic model is an autoregressive model of order 1 or AR(1) model
$$y_t = \phi_0 + \phi_1 y_{t-1} + u_t, \qquad u_t \sim iid\,(0, \sigma_u^2).$$
- The condition $|\phi_1| < 1$ is required for the model to be stationary.
- If the longest lag included being the p^{th} lag, then
$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + u_t, \qquad u_t \sim iid\,(0, \sigma_u^2),$$
in which $\phi_0, \phi_1, \cdots, \phi_p$ are unknown parameters. This is an AR(p) model.

Estimation - AR(1)

- For the AR(1) model the unknown parameters $\{\phi_0, \phi_1, \sigma_u^2\}$ are easily estimated by ordinary least squares.
- The residual sum of squares function for the AR(p) model is

$$S = \sum_{t=2}^{T} u_t^2 = \sum_{t=2}^{T} (y_t - \phi_0 - \phi_1 y_{t-1})^2 ,$$

  where the sample sum of squares begins at $t = 2$ as 1 observation is lost because of the inclusion of 1 lag in the model.
- Estimation proceeds by simple treating the lags of y_t as regressors, so estimation amounts to regressing y_t on a constant and its first lag.
- Once $\widehat{\phi}_0$ and $\widehat{\phi}_1$ are available then

$$\widehat{u}_t = y_t - \widehat{\phi}_0 - \widehat{\phi}_1 y_{t-1}$$

  and the residual variance is

$$\widehat{\sigma}_u^2 = \frac{1}{T-1} \sum_{t=2}^{T} \widehat{u}_t^2 .$$

Estimation - AR(p)

- The unknown parameters $\theta = \{\phi_0, \phi_1, \cdots, \phi_p, \sigma_u^2\}$ of the AR(p) model are estimated by least squares.
- The residual sum of squares function for the AR(1) model is

$$S = \sum_{t=p+1}^{T} u_t^2 = \sum_{t=p+1}^{T} (y_t - \phi_0 - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \cdots - \phi_p y_{t-p})^2 ,$$

  where the sample sum of squares begins at $t = p + 1$ as p observations are lost because of the inclusion of p lags in the model.
- After estimating the parameters the least squares residuals are computed as

$$\widehat{u}_t = y_t - \widehat{\phi}_0 - \widehat{\phi}_1 y_{t-1} - \widehat{\phi}_2 y_{t-2} - \cdots - \widehat{\phi}_p y_{t-p},$$

  which are used to compute the residual variance

$$\widehat{\sigma}_u^2 = \frac{1}{T-p} \sum_{t=p+1}^{T} \widehat{u}_t^2 .$$

Autocorrelation function

- Consider the following sequence of AR models may be estimated equation-by-equation by ordinary least squares

$$\begin{aligned} y_t &= \phi_{10} + \rho_1 y_{t-1} + u_{1t}, \\ y_t &= \phi_{20} + \rho_2 y_{t-2} + u_{2t}, \\ &\vdots \qquad \vdots \qquad \vdots \\ y_t &= \phi_{k0} + \rho_k y_{t-k} + u_{kt}, \end{aligned}$$

  giving the estimated ACF $\{\widehat{\rho}_1, \widehat{\rho}_2, \cdots, \widehat{\rho}_k\}$. The notation adopted for the constant term in the above regressions emphasises that this term differs for each equation.
- For $0 < \phi_1 < 1$, the autocorrelation function of y_t declines exponentially for increasing k so that the effects of previous values on y_t gradually diminish.
- For $-1 < \phi_1 < 0$, the autocorrelation function of y_t alternates in sign as k is even or odd and its modulus declines exponentially.
- Plotting the autocorrelation function is a useful tool to get a feel for the data.

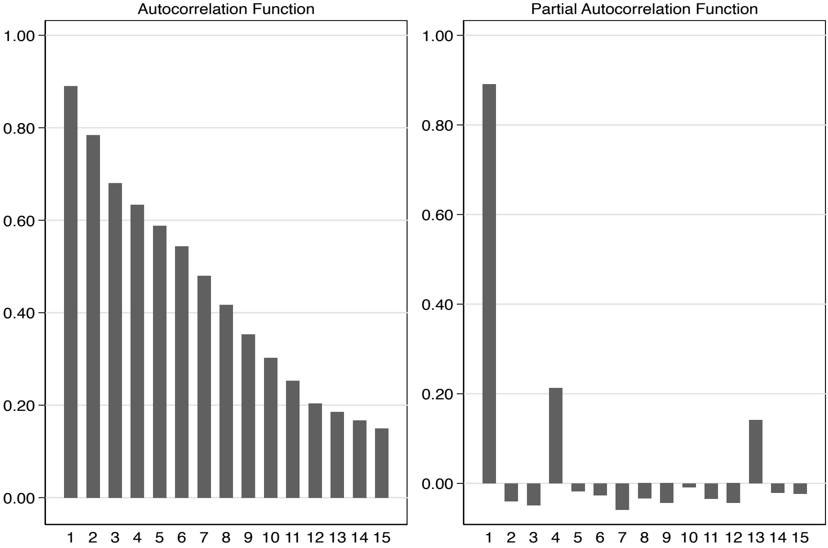
Partial autocorrelation function

- Another measure of the dynamic properties of AR models is the partial autocorrelation function (PACF) at lag k, which measures the relationship between y_t and y_{t-k} but now with the intermediate lags included in the regression model, so that their effects are controlled for.
- To compute the sample PACF the following AR models are estimated equation-by-equation by ordinary least squares

$$
\begin{aligned}
y_t &= \phi_{10} + \phi_{11} y_{t-1} + u_{1t} \\
y_t &= \phi_{20} + \phi_{21} y_{t-1} + \phi_{22} y_{t-2} + u_{2t} \\
y_t &= \phi_{30} + \phi_{31} y_{t-1} + \phi_{32} y_{t-2} + \phi_{33} y_{t-3} + u_{3t} \\
\vdots & \qquad \vdots \qquad\quad \vdots \qquad\quad \vdots \\
y_t &= \phi_{k0} + \phi_{k1} y_{t-1} + \phi_{k2} y_{t-2} + \cdots + \phi_{kk} y_{t-k} + u_{kt},
\end{aligned}
$$

  where the estimated PACF is therefore given by $\{\widehat{\phi}_{11}, \widehat{\phi}_{22}, \cdots, \widehat{\phi}_{kk}\}$.
- The PACF for an AR(p) model is zero for lags greater than p. For example, in the AR(1) model the PACF has a spike at lag 1 and thereafter is $\phi_{kk} = 0, \ \forall\ k > 1$. This is in contrast to the ACF which in general has non-zero values for higher lags, as seen in the simple AR(1) model above.
- Note that by construction the ACF and PACF at lag 1 are equal to each other.

ACF and PACF of S&P 500 dividend returns



- To allow for dynamic effects, regression models may be combined with the ARMA class of models. Some examples are as follows.

$$
\begin{array}{ll}
y_t = \alpha + \beta x_t + u_t, \qquad u_t = \phi_1 u_{t-1} + v_t & \text{[AR disturbance]} \\
y_t = \alpha + \beta x_t + \lambda y_{t-1} + u_t & \text{[Lagged dependent]} \\
y_t = \alpha + \beta x_t + \gamma x_{t-1} + u_t & \text{[Lagged explanatory]} \\
y_t = \alpha + \beta x_t + \lambda y_{t-1} + \gamma x_{t-1} + u_t & \text{[Joint specification]} \\
\qquad\qquad u_t = \phi_1 u_{t-1}. &
\end{array}
$$

- One important reason for including dynamics in a regression model is to correct for potential misspecification problems that arise from incorrectly excluding explanatory variables.

Vector Autoregressive Models

Background

- Thus far we have considered single equations and we have just encountered that autoregressive (AR) model.
- Sometimes it is difficult to delimit a single variable as the dependent variable to be explained in terms of all the other variables - and it is reasonable to suppose that all variables are jointly determined.
- To allow financial variables to adjust to shocks over time this model is now extended to incorporate dynamic effects.
- In a multiple equation system where each variable is dependent on its own lags and the lags or all other variables, then we have a vector autoregressive (VAR) model.
- The assumption of stationarity of all the variables is maintained.

VAR(1) with 2 variables

An example of a two variable VAR(1) model is

$$y_{1t} = \phi_{10} + \phi_{11,1} y_{1t-1} + \phi_{12,1} y_{2t-1} + u_{1t}$$
$$y_{2t} = \phi_{20} + \phi_{21,1} y_{1t-1} + \phi_{22,1} y_{2t-1} + u_{2t}$$

where y_{1t} and y_{2t} are the jointly dependent variables and u_{1t} and u_{2t} are disturbance terms.

VAR(2) with 2 variables

An example of a two variable VAR(2) model is

$$
\begin{aligned}
y_{1t} &= \phi_{10} + \phi_{11,1} y_{1t-1} + \phi_{11,2} y_{1t-2} + \phi_{12,1} y_{2t-1} + \phi_{12,2} y_{2t-2} + u_{1t} \\
y_{2t} &= \phi_{20} + \phi_{21,1} y_{1t-1} + \phi_{21,2} y_{1t-2} + \phi_{22,1} y_{2t-1} + \phi_{22,2} y_{2t-2} + u_{2t}
\end{aligned}
$$

where y_{1t} and y_{2t} are the jointly dependent variables and u_{1t} and u_{2t} are disturbance terms.

VAR(1) with three variables

An example of a three variable VAR(1) model is

$$
\begin{aligned}
y_{1t} &= \phi_{10} + \phi_{11,1} y_{1t-1} + \phi_{12,1} y_{2t-1} + \phi_{13,1} y_{3t-1} + u_{1t} \\
y_{2t} &= \phi_{20} + \phi_{21,1} y_{1t-1} + \phi_{22,1} y_{2t-1} + \phi_{23,1} y_{3t-1} + u_{2t} \\
y_{3t} &= \phi_{30} + \phi_{31,1} y_{1t-1} + \phi_{32,1} y_{2t-1} + \phi_{33,1} y_{3t-1} + u_{3t},
\end{aligned}
$$

where y_{1t}, y_{2t} and y_{3t} are the jointly dependent variables, p is a prescribed lag length which is the same for all equations and u_{1t}, u_{2t} and u_{3t} are disturbance terms.

VAR(p)

An example of a three variable VAR(p) model is

$$y_{1t} = \phi_{10} + \sum_{i=1}^{p} \phi_{11,i} y_{1t-i} + \sum_{i=1}^{p} \phi_{12,i} y_{2t-i} + \sum_{i=1}^{p} \phi_{13,i} y_{3t-i} + u_{1t}$$

$$y_{2t} = \phi_{20} + \sum_{i=1}^{p} \phi_{21,i} y_{1t-i} + \sum_{i=1}^{p} \phi_{22,i} y_{2t-i} + \sum_{i=1}^{p} \phi_{23,i} y_{3t-i} + u_{2t}$$

$$y_{3t} = \phi_{30} + \sum_{i=1}^{p} \phi_{31,i} y_{1t-i} + \sum_{i=1}^{p} \phi_{32,i} y_{2t-i} + \sum_{i=1}^{p} \phi_{33,i} y_{3t-i} + u_{3t},$$

where y_{1t}, y_{2t} and y_{3t} are the jointly dependent variables, p is a prescribed lag length which is the same for all equations and u_{1t}, u_{2t} and u_{3t} are disturbance terms.

Higher dimensional VARs

In matrix notation a VAR with N variables is conveniently represented as

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \cdots + \Phi_p y_{t-p} + u_t,$$

where the parameter matrices are given by

$$\Phi_0 = \begin{bmatrix} \phi_{10} \\ \phi_{20} \\ \vdots \\ \phi_{N0} \end{bmatrix}, \qquad \Phi_i = \begin{bmatrix} \phi_{11,i} & \phi_{12,i} & \cdots & \phi_{1N,i} \\ \phi_{21,i} & \phi_{22,i} & & \phi_{2N,i} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{N1,i} & \phi_{N2,i} & \cdots & \phi_{NN,i} \end{bmatrix}.$$

The disturbances $u_t = \{u_{1t}, u_{2t}, ..., u_{Nt}\}' \sim iid\,(0, \Omega)$ are independent over t with zero mean and covariance matrix

$$\Omega = \mathrm{E}(u_t u_t') = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \cdots & \sigma_N^2 \end{bmatrix}.$$

Estimating the covariance matrix of residuals

To estimate the covariance matrix Ω let $\widehat{u}_t = \{\widehat{u}_{1t}, \widehat{u}_{2t}, \cdots, \widehat{u}_{Nt}\}$ represent the least squares residuals for each equation in the VAR.

An estimate of the covariance matrix is computed using the sample residual moment matrix defined by

$$\widehat{\Omega} = \begin{bmatrix} \widehat{\sigma}_1^2 & \widehat{\sigma}_{12} & \cdots & \widehat{\sigma}_{1N} \\ \widehat{\sigma}_{21} & \widehat{\sigma}_2^2 & & \widehat{\sigma}_{2N} \\ \vdots & & \ddots & \vdots \\ \widehat{\sigma}_{N1} & \widehat{\sigma}_{N2} & \cdots & \widehat{\sigma}_N^2 \end{bmatrix},$$

or more explicitly,

$$\widehat{\Omega} = \frac{1}{T} \begin{bmatrix} \sum_{t=p+1}^{T} \widehat{u}_{1t}^2 & \sum_{t=p+1}^{T} \widehat{u}_{1t}\widehat{u}_{2t} & \cdots & \sum_{t=p+1}^{T} \widehat{u}_{1t}\widehat{u}_{Nt} \\ \sum_{t=p+1}^{T} \widehat{u}_{2t}\widehat{u}_{1t} & \sum_{t=p+1}^{T} \widehat{u}_{2t}^2 & & \sum_{t=p+1}^{T} \widehat{u}_{2t}\widehat{u}_{Nt} \\ \vdots & & \ddots & \vdots \\ \sum_{t=p+1}^{T} \widehat{u}_{Nt}\widehat{u}_{1t} & \sum_{t=p+1}^{T} \widehat{u}_{Nt}\widehat{u}_{2t} & \cdots & \sum_{t=p+1}^{T} \widehat{u}_{Nt}^2 \end{bmatrix}.$$

Popularity

VARs enjoy great popularity in applied research in finance.

(i) Estimation is straightforward, involving the application of ordinary least squares to each equation in the VAR.

(ii) The VAR system provides a convenient framework to forecast financial variables.

(iii) The model provides a basis for performing so-called causality tests between financial variables.

(iv) Theoretical models in finance can be tested through the imposition of restrictions on the VAR parameters.

Lag Length Selection

- If the lag length is too short, there is a risk that aspects of the dynamic mechanism are excluded from the model. If the lag structure is too long then there are redundant lags which can reduce the precision of the parameter estimates, thereby raising the standard errors and yielding t statistics that may be biased downwards.
- In choosing the lag structure of a VAR, care must be exercised in relation to the sample size as degrees of freedom quickly diminish for even moderate lag lengths. For each integer increase in the lag length, an additional matrix of coefficients must be estimated. In a K dimensional system this means an additional K^2 coefficient parameters are needed for each extra lag.
- For these reasons, an important practical consideration in constructing and estimating a VAR(p) model is the choice of the lag order p.

Choosing lag length

- A common data-driven approach to selecting lag order is to use information criteria.
- The three most commonly used information criteria (IC) for selecting a parsimonious time series model are

$$AIC = \log|\widehat{\Omega}| + \frac{2pK^2}{T}$$

$$HIC = \log|\widehat{\Omega}| + \frac{2\log(\log(T))}{T}pK^2$$

$$SIC = \log|\widehat{\Omega}| + \frac{\log(T)}{T}pK^2.$$

  In these expressions, $\widehat{\Omega}$ is an estimate of the covariance matrix.
- In the scalar case, the determinant of the estimated covariance matrix, $|\widehat{\Omega}|$, is replaced by the estimated residual variance, $\widehat{\sigma}_u^2$.

Selection procedure

Choosing an IC optimal lag order using any of the above criteria requires the following steps.

Step 1: Choose a maximum number of lags, $p_{\max}$, for the VAR model. This choice may be informed by the ACFs and PACFs of the data, the frequency with which the data are observed and the sample size.

Step 2: Estimate the model sequentially for all lags up to and including $p_{\max}$. For each regression, compute the relevant information criterion, holding the sample size fixed.

Step 3: Choose the specification corresponding to the minimum values of the information criterion. In some cases there will be disagreement between different information criteria on the choice of lag length. The final decision is then a matter of individual judgement.

Interpreting VARs

Interpreting VARs

- In a VAR model, all lagged variables are assumed to contribute information in determining the behaviour of each dependent variable.
- But in most empirical applications of VARs there are often large numbers of estimated coefficients which are statistically insignificant.
- A question of considerable importance in empirical work is whether the coefficients of all the lagged values of a particular explanatory variable in a given equation are zero or not.
- This question bears on whether the information content of the past values of one variable influences the behaviour of another variable in the system.
- This notion has a close connection with that of causal influence in the sense that predictions might be improved by measuring and including such influences.

Granger causality

- Have another look at

$$y_{1t} = \phi_{10} + \phi_{11,1} y_{1t-1} + \phi_{12,1} y_{2t-1} + \phi_{13,1} y_{3t-1} + u_{1t}$$
$$y_{2t} = \phi_{20} + \phi_{21,1} y_{1t-1} + \phi_{22,1} y_{2t-1} + \phi_{23,1} y_{3t-1} + u_{2t}$$
$$y_{3t} = \phi_{30} + \phi_{31,1} y_{1t-1} + \phi_{32,1} y_{2t-1} + \phi_{33,1} y_{3t-1} + u_{3t},$$

- The information content of variable y_2 on variable y_1 might be tested by considering restriction

$$\phi_{12,1} = 0.$$

  This restriction can be tested jointly using a χ^2 test with 1 degrees of freedom.
- If y_{2t} plays a role in predicting future values of y_{1t}, then y_{2t} is said to cause y_{1t} in Granger's sense (Named after Professor Sir Clive Granger, who won the Nobel Prize in Economics in 2003)
- It is important to remember that Granger causality is based on the presence (or absence) of predictability and does not of itself signify causal influence in a philosophical sense.
- Evidence of Granger causality and the lack of Granger causality from y_{2t} to y_{1t}, are denoted, respectively, as

$$y_{2t} \rightarrow y_{1t} \qquad y_{2t} \nrightarrow y_{1t}\,.$$

More generally ...

It is also possible to test for Granger causality in the reverse direction by performing a joint test of the lags of y_{1t} in the y_{2t} equation. Combining both sets of causality results can yield a range of statistical causal patterns:

Unidirectional: (from y_{1t} to y_{2t})	$y_{1t} \rightarrow y_{2t}$ $y_{2t} \nrightarrow y_{1t}$
Unidirectional: (from y_{2t} to y_{1t})	$y_{2t} \rightarrow y_{1t}$ $y_{1t} \nrightarrow y_{2t}$
Bidirectional: (feedback)	$y_{2t} \rightarrow y_{1t}$ $y_{1t} \rightarrow y_{2t}$
Independence:	$y_{2t} \nrightarrow y_{1t}$ $y_{1t} \nrightarrow y_{2t}$

Impulse response analysis

- Granger causality testing is one method of identifying the system dynamics of a VAR that enhances understanding of variable interactions over time.
- An alternative but related approach focuses on impulse responses by tracking the transmission effects of shocks to the system on the dependent variables. This approach to examining system dynamics is called impulse response analysis.
- A potential candidate for the shocks in a VAR system is the vector of disturbances $u_t = \{u_{1t}, u_{2t}, ..., u_{Nt}\}$, which represents contributions to the dependent variable that are not predicted from past information. The primary problem in the direct use of the fitted disturbances in studying impulse responses is that these terms are correlated, which complicates the interpretation of the shocks u_t with respect to the underlying economic and financial forces.
- One solution that aids interpretation is to transform the VAR into a new system in which the disturbances in the equations are uncorrelated so that the effects of these uncorrelated shocks on the system variables can be traced over time, thereby enabling determination of the responses to impulses associated with the individual shocks.

Illustration

- Consider the bivariate VAR model of equity returns and dividend returns

$$re_t = \phi_{10} + \sum_{i=1}^{6} \phi_{11,i} re_{t-i} + \sum_{i=1}^{6} \phi_{12,i} rd_{t-i} + u_{1t},$$
$$rd_t = \phi_{20} + \sum_{i=1}^{6} \phi_{21,i} re_{t-i} + \sum_{i=1}^{6} \phi_{22,i} rd_{t-i} + u_{2t}.$$

- Let the relationship between the two VAR disturbances u_{1t} and u_{2t}, be represented by the linear regression equation

$$u_{2t} = \rho u_{1t} + v_{2t},$$

where ρ is a parameter capturing the correlation between u_{1t} and u_{2t}.

- Note that v_{2t} is a new disturbance term which, from the properties of the linear regression model, is uncorrelated with u_{1t}. This equation is a structural relationship between the two shocks u_{1t} and u_{2t} in which the residual v_{2t} is that part of the impulse u_{2t} which is uncorrelated with u_{1t} .

Illustration

- Substituting for u_{1t} and u_{2t} and rearranging yields

$$rd_t = (\phi_{20} - \rho\phi_{10}) + \rho\, re_t + \sum_{i=1}^{6}(\phi_{21,i} - \rho\phi_{11,i})re_{t-i} + \sum_{i=1}^{6}(\phi_{22,i} - \rho\phi_{12,i})rd_{t-i} + v_{2t},$$

$$= \beta_{20} + \rho\, re_t + \sum_{i=1}^{6}\beta_{21,i}re_{t-i} + \sum_{i=1}^{6}\beta_{22,i}rd_{t-i} + v_{2t},$$

  in which $\beta_{20} = \phi_{20} - \rho\phi_{10}$, $\beta_{21,i} = \phi_{21,i} - \rho\phi_{11,i}$ and $\beta_{22,i} = \phi_{22,i} - \rho\phi_{12,i}$.
- The difference between the original VAR equation for rd_t and this one is the inclusion of equity returns at time t, re_t, as an explanatory variable.
- Moreover, since v_{2t} is independent of u_{1t}, the VAR equation for re_t and this equation for rd_t contain disturbances that are now uncorrelated with each other.
- The VAR is now called a structural VAR (SVAR)

$$re_t = \beta_{10} + \sum_{i=1}^{6}\beta_{11,i}re_{t-i} + \sum_{i=1}^{6}\beta_{12,i}rd_{t-i} + v_{1t},$$

$$rd_t = \beta_{20} + \rho\, re_t + \sum_{i=1}^{6}\beta_{21,i}re_{t-i} + \sum_{i=1}^{6}\beta_{22,i}rd_{t-i} + v_{2t},$$

  where $\beta_{10} = \phi_{10}$, $\beta_{11,i} = \phi_{11,i}$, $\beta_{12,i} = \phi_{12,i}$ and $v_{1t} = u_{1t}$.

Parameter estimates

Parameter estimates of a bivariate SVAR(6) model for the United States monthly equity and dividend returns for the period January 1871 to September 2016.

Lag	Equity Returns		Dividend Returns	
	re	*rd*	*re*	*rd*
0			-0.006	
			(0.003)	
1	0.297	-0.049	0.003	0.910
	(0.024)	(0.188)	(0.003)	(0.024)
2	-0.070	0.520	0.007	0.019
	(0.025)	(0.255)	(0.003)	(0.032)
3	-0.029	-0.248	0.008	-0.254
	(0.025)	(0.251)	(0.003)	(0.032)
4	0.030	0.318	0.002	0.226
	(0.025)	(0.251)	(0.003)	(0.032)
5	0.052	-0.231	0.012	0.014
	(0.025)	(0.255)	(0.003)	(0.032)
6	-0.005	-0.341	0.014	-0.025
	(0.024)	(0.186)	(0.003)	(0.024)
Constant	0.264		0.019	
	(0.100)		(0.012)	
T	1742		1742	
RSS	25945.84		419.0077	

Equity Shock

- Impulse responses show the time-forms of system variable responses to incoming structural shocks.
- Impulse responses impart interpretable information about the internal dynamics within a VAR system that govern the transmission effects of shocks.
- Let the size of the shock be one standard deviation of v_{1t} given by

$$\Delta re_t = \sqrt{\frac{25945.84}{1742}} = 3.8593.$$

- Dividend returns change by

$$\Delta rd_t = \widehat{\rho} \times 3.8593 = -0.0061 \times 3.8593 = -0.0237\,.$$

One period effects

- Consider

$$re_{t+1} = \beta_{10} + \sum_{i=1}^{6} \beta_{11,i} re_{t+1-i} + \sum_{i=1}^{6} \beta_{12,i} rd_{t+1-i} + v_{1t+1},$$

$$rd_{t+1} = \beta_{20} + \rho\, re_{t+1} + \sum_{i=1}^{6} \beta_{21,i} re_{t+1-i} + \sum_{i=1}^{6} \beta_{22,i} rd_{t+1-i} + v_{2t+1}\,.$$

- Expected change in equities at $t+1$ comes from

$$\mathrm{E}_t(\Delta re_{t+1}) = \beta_{11,1} \Delta re_t + \beta_{12,1} \Delta rd_t\,.$$

Thus the impulse response of equities at time $t+1$ to an equity shock is estimated as

$$\mathrm{E}_t(\Delta re_{t+1}) = 0.2974 \times 3.8593 - 0.0489 \times (-0.0237) = 1.1491\,.$$

- Expected change in dividends at $t+1$ comes from

$$\mathrm{E}_t(\Delta rd_{t+1}) = \rho \Delta\, re_{t+1} + \beta_{21,1} \Delta re_t + \beta_{22,1} \Delta rd_t,$$

which is estimated as

$$\mathrm{E}_t(\Delta rd_{t+1}) = -0.0061 \times 1.1491 + 0.0028 \times 3.8593 + 0.9096 \times (-0.0237) = -0.0179.$$

Dividend Shock

- Let the size of the shock be one standard deviation of v_{2t} which is

$$\Delta rd_t = \sqrt{\frac{419.0077}{1742}} = 0.4904.$$

- The effect of a dividend shock on equities at time t is

$$\Delta re_t = 0.0000\,.$$

- The effects of the dividend shock at time t is

$$\mathrm{E}_t(\Delta re_{t+1}) = \beta_{11,1}\Delta re_t + \beta_{12,1}\Delta rd_t,$$

  which is estimated as

$$\mathrm{E}_t(\Delta re_{t+1}) = 0.2974 \times 0.00 + (-0.0489) \times 0.4904 = -0.0240.$$

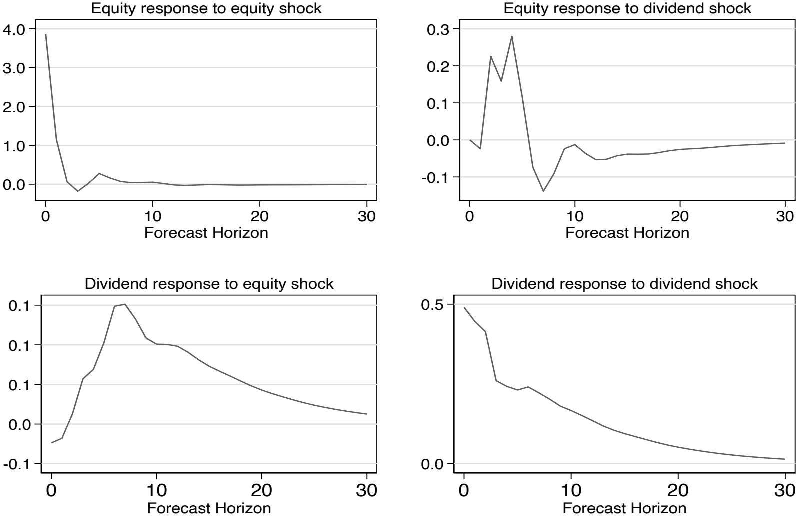
- The effect of the dividend shock on dividends the next month is

$$\mathrm{E}_t(\Delta rd_{t+1}) = \rho\,\Delta re_{t+1} + \beta_{21,1}\Delta re_t + \beta_{22,1}\Delta rd_t\,,$$

  which is estimated as

$$\mathrm{E}_t(\Delta rd_{t+1}) = 0.2974 \times (-0.0241) + 0.0028 \times 0.00 + 0.9097 \times 0.4904 = 0.4463\,.$$

Impulse responses



Cholesky Decomposition

- The calculation of impulse responses is presented within a regression framework which involves estimating the SVAR by least squares and then using recursive substitution methods to compute the effects of the structural shocks on the variables in the model.
- A more convenient approach in the multivariate case is to perform the calculations using matrices. Formally this is achieved by defining a lower triangular matrix L with the property that the estimated VAR covariance matrix $\widehat{\Omega}$ in is decomposed as

$$\widehat{\Omega} = LL'.$$

- This decomposition of a matrix is commonly referred to as the Cholesky decomposition. As the matrix being decomposed in this case is the covariance matrix, L can be referred to as the standard deviation matrix, or even square-root matrix, as multiplying L by the transposition of itself recovers the covariance matrix.
- Impulse responses based on the Cholesky decomposition are also the same estimates obtained using the recursive substitution method applied earlier to the estimated SVAR model.

Variance decomposition

- To gain explicit quantitative insight on the relative importance of various structural shocks on the variables in the system a variance decomposition can also be performed.
- The forecast variances for each variable over alternative forecast horizons are decomposed into the separate relative effects of each structural shock with the results expressed as a percentage of the overall movement.
- In the case of the SVAR model of equities and dividends, the approach is to express the forecast error variances of equities and dividends in terms of the structural shocks v_1 and v_2.

Forecast Error Variance at $T+1$ for Equity

- The forecast error variance of equities at $T+1$ is defined as

$$\mathrm{var}(e_{1T+1}) = \mathrm{E}_T[(e_{1T+1} - \mathrm{E}_T(e_{1T+1}))^2] = \mathrm{E}_T(e^2_{1T+1}) = \mathrm{E}_T(v^2_{1T+1}),$$

which uses the property that $\mathrm{E}_T(e_{1T+1}) = 0$ and uses the fact that the one-step-ahead forecast error for equities simply equals the equity structural shock

$$e_{1T+1} = v_{1T+1}\,.$$

- This result shows that the equities forecast error variance at $T+1$ is totally determined by its own shocks. This result immediately follows from the triangular ordering of the SVAR model.
- The forecast error variance of equities at $T+1$ is estimated as

$$\widehat{\mathrm{var}}(e_{1T+1}) = 3.8593^2 = 14.8940.$$

and 100% of this forecast error variance of 14.8940 at $T+1$ is the result of own shocks to equities and nothing comes from dividend shocks.

Forecast Error Variance at $T+1$ for Dividends

- After a little manipulation it can be shown that

$$e_{2T+1} = \rho(re_{T+1} - \mathrm{E}_T(re_{T+1})) + v_{2T+1} = \rho v_{1T+1} + v_{2T+1}.$$

  The forecast error at $T+1$ for dividends is a function of both structural shocks unlike equities which is simply a function of its own shock v_{1T+1}.

- The forecast error variance of dividends at $T+1$ is

$$\begin{aligned}\mathrm{var}(e_{2T+1}) &= \mathrm{E}_T[(\rho v_{1T+1} + v_{2T+1})^2] \\ &= \mathrm{E}_T(\rho^2 v_{1T+1}^2 + v_{2T+1}^2 + 2\rho v_{1T+1} v_{2T+1}) \\ &= \rho^2 \mathrm{E}_T(v_{1T+1}^2) + \mathrm{E}_T(v_{2T+1}^2).\end{aligned}$$

- The dividend forecast error variance at $T+1$ is estimated as

$$\widehat{\mathrm{var}}(e_{2T+1}) = (-0.0061)^2 \times 3.8593^2 + 0.49044^2 = 0.0010 + 0.24053 = 0.2411.$$

- This expression shows that for the dividends forecast error variance of 0.2411, $0.24053/0.2411 = 0.99766$, or 99.766% is due to own shocks and the remaining is due to equity shocks.

Variance decomposition

Variance decomposition expressed in percentages computed from a VAR(6) model estimated using monthly data on United States equity returns, re_t, and dividend returns, rd_t, over the period January 1871 to September 2016.

Period	Decomposition of *re*		Decomposition of *rd*	
	re	*rd*	*re*	*rd*
1	100.000	0.000	0.234	99.766
2	99.996	0.004	0.201	99.799
3	99.684	0.316	0.172	99.828
4	99.531	0.469	0.631	99.369
5	99.057	0.943	1.217	98.783
10	98.785	1.215	8.822	91.178
15	98.732	1.267	11.522	88.478
20	98.694	1.306	12.571	87.429
25	98.680	1.320	12.911	87.089
30	98.675	1.325	13.012	86.988