Appendix A. Vector Autoregression and Vector Error Correction Models

The structural approach to time series modeling uses economic theory to model the relationship among the variables of interest. Unfortunately, economic theory is often not rich enough to provide a dynamic specification that identifies all of these relationships (Engle and Granger, 1987). Furthermore, estimation and inference are complicated by the fact that endogenous variables may appear on both the left and right sides of equations. These problems lead to alternative, non-structural approaches to modeling the relationship among several variables. This section briefly describes the estimation and analysis of vector autoregression (VAR) and the vector error correction (VEC) models. We also describe tools for testing the presence of cointegrating relationships among several non-stationary variables.

The vector autoregression (VAR) is commonly used for forecasting systems of interrelated time series and for analyzing the dynamic impact of random disturbances on the system of variables. The VAR approach sidesteps the need for structural modeling by treating every endogenous variable in the system as a function of the lagged values of all of the endogenous variables in the system. The vector autoregression (VAR) is commonly used for forecasting systems of interrelated time series and for analyzing the dynamic impact of random disturbances on the system of variables. The VAR approach sidesteps the need for structural modeling by treating every endogenous variable in the system as a function of the lagged values of all of the endogenous variables in the system.

The mathematical representation of a VAR is:

$$y_t = A_1 y_{t-1} + ... + A_p y_{t-p} - Bx_t + \varepsilon_t, \quad (1)$$

where y_t is a vector of endogenous variables, x_t is a vector of exogenous variables, $A_1,\ldots,A_p$ and B are matrices of coefficients to be estimated, and ε_t is a vector of innovations that may be contemporaneously correlated but are uncorrelated with their own lagged values and uncorrelated with all of the right-hand side variables. Since only lagged values of the endogenous variables appear on the right-hand side of the equations, simultaneity is not an issue and OLS yields consistent estimates. Moreover, even though the innovations ε_t may be contemporaneously correlated; OLS is efficient and equivalent to GLS since all equations have identical regressors.

Multivariate simultaneous equations models are used extensively for macro-econometric analysis when Sims (1980) advocated VAR models as alternatives. At that time longer and more frequently observed financial time series called for models which described the dynamic structure of the variables. VAR models possess the following purposes, first, VAR models typically treats all variables as a priori endogenous, and thereby VAR models accounts for Sims' critique that the exogeneity assumptions for some of the variables in simultaneous equations models are ad hoc and often not backed by fully developed theories. Restrictions, including exogeneity of some of the variables, may be imposed on VAR models based on statistical procedures. Further, VAR models are natural tools for forecasting. Their setup is such that current values of a set of variables are partly explained by past values of the variables involved. They can also be used for economic analysis, however, because they describe the joint generation mechanism of the variables involved. Structural VAR analysis attempts to investigate structural economic hypotheses with the help of VAR models. Impulse response analysis, forecast error variance decompositions, historical decompositions and the analysis of forecast scenarios are the tools which have been proposed for disentangling the relations between the variables in a VAR model.

Traditionally VAR models are designed for stationary variables without time trends. Trending behavior can be captured by including deterministic polynomial terms. In the 1980s the discovery of the importance of stochastic trends in economic variables and the development of the concept of cointegration by Granger (1981), Engle and Granger (1987), Johansen (1995) and others have shown that stochastic trends can also be captured by VAR models. If there are trends in some of the variables it may be desirable to separate the long-run relations from the short-run dynamics of the generation process of a set of variables.

A vector error correction (VEC) model is a restricted VAR designed for use with nonstationary series that are known to be cointegrated. Vector error correction (VEC) models offer a convenient framework for separating long-run and short-run components of the data generation process. You may test for cointegration using an estimated VAR object, Equation object estimated using nonstationary regression methods. The VEC has cointegration relations built into the specification so that it restricts the long run behavior of the endogenous variables to converge to their cointegrating relationships while allowing for short-run adjustment dynamics. The cointegration term is known as the error correction term since the deviation from long-run equilibrium is corrected gradually through a series of partial short-run adjustments. To take the simplest possible example, consider a two variable system with one cointegrating equation and no lagged difference terms. The cointegrating equation is:

$$y_{2,t} = \beta y_{1,t}, \tag{2}$$

and,

$$\begin{aligned} \Delta y_{1,t} &= \alpha_1(y_{2,t-1} - \beta y_{1,t-1}) + \varepsilon_{1,t} \\ \Delta y_{2,t} &= \alpha_2(y_{2,t-1} - \beta y_{1,t-1}) + \varepsilon_{2,t} \end{aligned}. \tag{3}$$

In this simple model, the only right-hand side variable is the error correction term. In long run equilibrium, this term is zero. However, if y_1 and y_2 deviate from the long run equilibrium, the error correction term will be nonzero and each variable adjusts to partially restore the equilibrium relation. The coefficient measures the speed of adjustment of the i^{th} endogenous variable towards the equilibrium. A sufficient (and considerably stronger) condition is that $y_{i,t\text{-}1}$ and $x_{i,t}$ be predetermined; that is, they should satisfy the sequential moment restriction:

$$E[\varepsilon_{i,t}^{y} \mid x_{i,t}, y_{i,t-1}, x_{i,t-1}, \ldots, y_{i,1}, x_{i,1}] = 0\,. \tag{4}$$

If we are willing to assume that $\varepsilon_{i,t}$ is serially uncorrelated, the estimate of the pooled least squares is a consistent estimator for all models presented above. The assumption that $\varepsilon_{i,t}$ is serially uncorrelated is, however, restrictive, especially for models including a smaller number of lags. We choose $\varepsilon_{i,t}$ following a moving average process for the presence of serially correlated errors. Because the presence of moving-average errors would introduce bias in the least-squares estimator, estimations of the more general models proceed differently. For VAR model, the optimal lags of AIC is 5 lag length. We use SIC Max-lag 2 for VAR and VEC specification in terms of both above discussion and our data characteristic and frequency.