6 Structural Vector Autoregressive Analysis for Cointegrated Variables*

Helmut Lütkepohl[1]

[1] Department of Economics, European University Institute
helmut.luetkepohl@iue.it

Summary: Vector autoregressive (VAR) models are capable of capturing the dynamic structure of many time series variables. Impulse response functions are typically used to investigate the relationships between the variables included in such models. In this context the relevant impulses or innovations or shocks to be traced out in an impulse response analysis have to be specified by imposing appropriate identifying restrictions. Taking into account the cointegration structure of the variables offers interesting possibilities for imposing identifying restrictions. Therefore VAR models which explicitly take into account the cointegration structure of the variables, so-called vector error correction models, are considered. Specification, estimation and validation of reduced form vector error correction models is briefly outlined and imposing structural short- and long-run restrictions within these models is discussed.

6.1 Introduction

In an influential article, Sims (1980) advocated the use of vector autoregressive (VAR) models for macroeconometric analysis as an alternative to the large simultaneous equations models that were in common use at the time. The latter models often did not account for the rich dynamic structure in time series data of quarterly or monthly frequency. Given that such data became more common in macro economic studies in the 1960s and 1970s, it was plausible to emphasize modelling of the dynamic interactions of the variables of interest. Sims also criticized the way the classical simultaneous equations models were identified and questioned the exogeneity assumptions for some of the variables which often reflect the prefer-

*I thank an anonymous reader for comments on an earlier draft of this paper that helped me to improve the exposition.

ences and prejudices of the model builders and are not necessarily fully backed by theoretical considerations. In contrast, in VAR models all observed variables are typically treated as a priori endogenous. Restrictions are imposed to a large extent by statistical tools rather than by prior beliefs based on controversial theories.

In a VAR analysis, the dynamic interactions between the variables are usually investigated by impulse responses or forecast error variance decompositions. These quantities are not unique, however. To identify those shocks or innovations and the associated impulse responses that reflect the actual ongoings in a given system of variables, usually also requires a priori assumptions which cannot be checked by statistical tools. Therefore *structural* VAR (SVAR) models were developed as a framework for incorporating identifying restrictions for the innovations to be traced out in an impulse response analysis.

In a parallel development it was discovered that the trending properties of the variables under consideration are of major importance for both econometric modelling and the associated statistical analysis. The spurious regression problem pointed out by Granger and Newbold (1974) showed that ignoring stochastic trends can lead to seriously misleading conclusions when modelling relations between time series variables. Consequently, the stochastic trends, unit roots or order of integration of the variables of interest became of major concern to time series econometricians and the concept of cointegration was developed by Granger (1981), Engle and Granger (1987), Johansen (1995) and many others. In this framework, the long-run relations are now often separated from the short-run dynamics. The cointegration or long-run relations are of particular interest because they can sometimes be associated with relations derived from economic theory. It is therefore useful to construct models which explicitly separate the long-run and short-run parts of a stochastic process. Vector error correction or equilibrium correction models (VECMs) offer a convenient framework for this purpose. They also open up the possibility to separate shocks or innovations with permanent and transitory effects. This distinction may be helpful in identifying impulse responses of interest. Therefore these models will be used as the framework in the following exposition.

A variable will be called *integrated of order* d ($I(d)$) if stochastic trends or unit roots can be removed by differencing the variable d times and a stochastic trend still remains after differencing only $d-1$ times. In line with this terminology, a variable without a stochastic trend or unit root is sometimes called $I(0)$. In the following, all variables are assumed to be either $I(0)$ or $I(1)$ to simplify matters. Hence, for a time series variable y_{kt}, it is assumed that the first differences, $\Delta y_{kt} \equiv y_{kt} - y_{k,t-1}$, have no stochastic trend. A set of $I(1)$ variables is called *cointegrated* if a linear combination exists which is $I(0)$. If a system consists of both $I(0)$ and $I(1)$ variables, any linear combination which is $I(0)$ is called a cointegration relation. Admittedly, this terminology is not in the spirit of the original idea of cointegration because it can happen that a linear combination of $I(0)$ variables is called a cointegration relation. In the present context, this terminology is a convenient simplification, however. Therefore it is used here.

Although in practice the variables will usually have nonzero means, polynomial trends or other deterministic components, it will be assumed in the following that deterministic terms are absent. The reason is that deterministic terms do not play