Chapter Three

MODELS, METHODOLOGY AND DATA

This chapter presents the models, methodology and data sources for the present study. To start with measures of monetary aggregates in India are briefly presented followed with data sources and variable selection for the study. Methods undertaken for detrending are then narrated along with structural break test methods. Unit root tests under structural breaks are then discussed. This is followed by an approach to causality in time series econometrics. The Toda – Yamamoto method adopted in this study is presented in detail. Finally Vector Autoregression and its mathematical underpinnings are outlined.

3.1 Monetary Aggregates in India

All the money held with public, RBI as well as government is called Total Stock of Money. Money Supply is that part of this Total Stock of Money which is with public. Public refers to the households, firms, local authorities, companies etc. Thus, public money does not include the money held by the government and the money held as CRR with RBI and SLR with themselves by commercial banks. The reason of excluding the above two categories from money supply is that this money held by the Government and RBI is out of circulation. Thus money in circulation is the money supply. This money can be in the following forms.

- Currency Notes and Coins Demand Deposits such as Saving Banks Deposits,
- Other Deposits such as Time Deposits / Term Deposits / Fixed Deposits
- Post Office Saving Accounts
- Cash in Hand (Except SLR) and Deposits of Banks in other Banks / RBI (except CRR)

In other words, this money has two components viz. Currency Component and Deposit Component. Currency Component consist of all the coins and notes in the circulation, while Deposit component is the money of the general public with the banks, which can be withdrawn by them using cheques, withdrawals and ATMs. Deposit can be either Demand Deposit or Time Deposit.

The Reserve bank of India calculates the four concepts of Money supply in India. They are called Monetary Aggregates or Money Stock Measures. They are as follows:

Narrow Money (M1)

At any point of time, the money held with the public has two most liquid components.
Currency Component: This consists of all the coins and notes in the circulation.

Demand Deposit Component: Demand Deposit component is the money of the general public with the banks, which can be withdrawn by them using cheques, withdrawals and ATMs. The above two components i.e. currency component and demand deposit component of the public money is called Narrow Money and is denoted by the RBI as M1. Thus,
M1 = Currency with the public + Demand Deposits of public in Banks. When a third component viz. Post office Savings Deposits is also added to M1, it becomes M2.

M2 = M1 + Post Office Savings deposits.

Broad Money (M_3)

Narrow money is the most liquid part of the money supply because the demand deposits can be withdrawn anytime during the banking hours. Time deposits on the other hand have a fixed maturity period and hence cannot be withdrawn before expiry of this period. When we add the time despots into the narrow money, we get the broad money, which is denoted by M3. M3 = Narrow money + Time Deposits of public with banks. Broad money does not include the interbank deposits such as deposits of banks with RBI or other banks. At the same time, time deposits of public with all banks including the cooperative banks are included in the Broad Money. The major distinction between the M1 and M3 is in the "Treatment of deposits with the banks". Going a little deep, the M3 is the treatment of "Time Deposits" of the public, since demand deposits are available against cheques and ATMs. Adding the Post Office Savings money also into the M3, it becomes M4. Both M2 and M4 which include the Post office Savings with narrow money and broad money respectively are now a days irrelevant. Post Office savings was once a prominent figure when the banks had not expanded in India as we see them today all around. The RBI releases the data at times regarding the money supply in India and Post Office Savings Deposits have not been updated frequently. There is not much change in the money of people deposited with the Post office and RBI did not care to update this money. Further, there was a time when the Reserve Bank used broad money (M3) as the policy target. However, with the weakened relationship between money, output and prices, it replaced M3 as a policy target with a multiple indicators approach. RBI started using the Multiple Indicator Approach since 1998. Currently, Narrow Money (M1) and Broad Money (M3) are relevant indicators of money supply in India. The RBI in all its policy documents, monthly Bulletins and other documents shows these aggregates. The present study takes M_1 as the measure of narrow money supply and M_3 as the measure of broad money supply.

3.2 Data

The present study uses secondary level time series data for the period 1960-2010. The principal data sources are, (1) Reserve Bank of India: Handbook of Statistics on the Indian Economy (various issues), (2) Reserve Bank of India: Report on Currency and Finance (various issues), (3) Reserve Bank of India Bulletin (various issues), and (4) Reserve Bank of India: Annual Report (various issues). The variable selection is detained in the table below.

Table 3.1. Variable–wise time periods of Annual Time Series Data for the Study		
Variables	Data Period	No. Of Time Points
Constant Price GDP or Real GDP (**RGDP** with base year as 2004-05)	1960 – 2010	63
Real GDP Growth Rate (**RGDPGR**)	1960 – 2010	62 (1st data point excluded)
Broad Money (**BM**)	1960 – 2010	60
Broad Money Growth Rate (**BMGR**)	1960 – 2010	59 (1st data point excluded)
Narrow Money (**NM**)	1960 – 2010	60
Narrow Money Growth Rate (**NMGR**)	1960 – 2010	59 (1st data point excluded)
Revenue Expenditure (**REVEXP**)	1960 – 2010	63
Capital Expenditure (**CAPEXP**)	1960 – 2010	63
Govt. Expenditure (**G**)	1960 – 2010	63
Revenue Deficit (**REVDEF**)	1972-2010	42
Fiscal Deficit (**GFD**)	1970-2010	44
Whole-sale Price Index (**WPIAC**) (base year 1952-53)	1960 – 2010	62
WPI Inflation (**INFLA**)	1960 – 2010	61 (1st data point excluded)
Consumer Price Index with base year 1960-61 (**CPI**)	1960-2010	54
Bank Rate (**BR**)	1961-2010	53
Cash Reserve Ratio (**CRR**)	1972-2010	42
Statutory Liquidity Ratio (**SLR**)	1961-2010	53

Source: Compiled from RBI: Handbook of Statistics on the Indian Economy, 2010.

Although the objectives of the study pertain to the period 1960-2010, in case of some variables data was obtained for a longer period while for a few others the data series was available over a shorter period.

3.3 Exponential Detrending

Long run macroeconomic data is most likely to have a trend – linear or non-linear. A glance at the time series line plots for each variable during 1961-2010 (not presented) reveals strong non-linear trends in all three variables. Both parabolic and exponential curves are fitted to each variable and the goodness of fit statistics are compared. The results are strongly suggestive of exponential trends in each of the variables. Accordingly, the exponentially detrended series on each variable are preferred for analysis. The detrended data is generated using the following steps. First, the natural logarithm of the variable is regressed linearly on a constant and time, i.e., the linear regression $ln(y_t) = ln(\alpha) + \beta.t + error,$ is run where y_t is the variable to be detrended. This is a log-linear form of the exponential growth (or smoothing) function $y_t = \alpha.\exp(\beta.t)$. Second, the parameters α and β are estimated using OLS and predicted $ln(y_t)$ series is generated. Third, anti-log of predicted $ln(y_t)$ is generated, which is predicted y_t in non-logarithmic form. Finally $e_t = y_t - \hat{y_t}$ is the residual from the exponential smoothing (or curve fitting) in non-logarithmic form and is thus the part of y_t that is free from any exponential trend (where $\hat{y_t}$ is predicted y_t in non-logarithmic form). Hence, e_t is exponentially detrended y_t. This method is applied to detrend both variables – fiscal deficit and broad money supply.

Standard tests for stationarity may be misleading for non-linearly trended data (for instance quadratic or exponential, both of which are rising at a rising rate over time) as because standard tests of stationarity such as Augmented Dickey-Fuller and

Philips-Perron tests include linear trend terms only (i.e., some 'constant' times 'time'). For an exponentially growing variable, stationarity may not be attained even at second difference, although for de-trended series it may be attained either at level (if trend stationary) or at first difference. Moreover, the autocorrelation function (ACF) helps us to select the lag lengths p (order of AR) and q (order of MA) and the ACF of the residuals is an important diagnostic tool. Unfortunately ACF as used in linear models may be misleading for non-linear models. The reason is that autocorrelation coefficients measure the degree of linear association between Y_t and Y_{t-i} (Y is the time series variable in question). As such ACF may fail to detect important non-linear relationships in the data. It is thus desirable to work with detrended data.

3.4 Testing Stationarity in the Presence of Structural Breaks

In the long run macroeconomic variables are expected to experience structural breaks, some of which may be the result of macroeconomic policy shifts, regime changes, or random shocks (droughts, warfare, socio-political instability and violence, etc.) at the domestic level or due to similar factors at the international level. The present study applies the Bai-Perron (1998 and 2003) multiple unknown structural break point test to original as well as the detrended series and compares the periods of break for each of the variables. Instead of going into the mathematical details, the method of break date determination as performed using EVIEWS 9 is as follows. First the time series variable in question is regressed (using OLS) on a constant only allowing for serial correlation that varies across break dates (regimes) through the use of HAC covariance estimation. Three break dates are considered along with a trimming percentage of 20, which implies around 12 observations per regime (as the period

1970-2010 implies 45 observations). Since the errors are assumed to be serially correlated, quadratic spectral kernel based HAC covariance estimation is specified using pre-whitened residuals. The kernel bandwidth is determined automatically using the Andrews AR(1) method. The default **method** setting in EVIEWS 9 **(sequential L+1 breaks vs. L)** instructs the software to perform sequential testing of *l+1* versus *l* breaks using the methods outlined in Bai (1997) and Bai and Perron (1998). The error distribution is allowed to differ across breaks to allow for heterogeneity. This test employs the same HAC covariance settings as used in the original equation but assumes regime specific error distributions. The break dates along with the respective F-statistic values are presented in the results empirical section. Stationarity related issues are discussed next.

Perhaps the most widely used unit root test to examine the stationarity of a time series (order of its integration) is the Augmented Dickey-Fuller test (ADF test) which makes use of equation (3.1). This generalised form includes both trend and intercept in the model.

$$\Delta y_t = a_0 + \gamma . y_{t-1} + a_1 . t + \sum_{i=1}^{p} \beta_i . \Delta y_{t-i} + \varepsilon_t \qquad (3.3.1)$$

Equation (3.1) tests the null hypothesis of a unit root against a trend stationary alternative. The optimum number of lagged Δy_t terms (introduced to tackle serial correlations in the errors) may be determined by the optimum value of some information criterion such as Schwartz's Information Criterion (SIC). Phillips and Perron (1988) proposed a nonparametric method of controlling serial correlation while testing for unit root. They estimate the unaugmented Dickey-Fuller test equation [Equation (3.3.1) without the term

($\sum_{i=1}^{p} \beta_i . \Delta\, y_{t-i}$) on the right hand side], and modifies the t-ratio of the γ coefficient so that serial correlation does not affect the asymptotic distribution of the test statistic.

Kwiatkowski, Phillips, Schmidt and Shin (1992) propose a test of the null hypothesis that the observed series is stationary around a deterministic trend. The series is expressed as the sum of deterministic trend, random walk and stationary error and the test is the LM test of the null hypothesis that the random walk has zero variance. The asymptotic distribution of the statistic is derived under the null and under the alternative that the series is difference stationary. KPSS test is quite contrary to the ADF and PP tests which consider the null hypothesis of unit root (i.e. a non-stationary series) as opposed to the former (KPSS) which considers a null hypothesis of stationary series.

The ADF and other traditional stationarity tests do not normally include a structural break term. But one can insert structural break dummies (say, seasonal dummies, for example) in equation (3.3.1) that may include both slope and intercept dummies. The point of break may be exogenously determined (approximately) by a visual scrutiny of the time series line plots. Importantly, the ADF test fails to perform well in the presence of structural breaks especially when the breaks are ignored. In such situations unit root tests with structural breaks are more suitable [see Perron (1989); Zivot and Andrews (1992)]. Perron (1989) demonstrated, assuming an exogenously fixed break date, that the power to reject the null hypothesis of unit root decreases (given that the alternative hypothesis of stationarity is actually true) when the structural break is ignored.

Zivot and Andrews (1992) suggest an improvement over the Perron (1989) test where they presume that the exact break point is unknown and endogenise the break date determination. A data dependent algorithm is used to proxy Perron's subjective procedure to determine the break points endogenously. Following Perron's characterization of the form of structural break, they adopt the following three models to test for unit roots.

$$\Delta y_t = a_0 + \gamma . y_{t-1} + a_1 . t + \delta . DU_t + \sum_{i=1}^{p} \beta_i . \Delta y_{t-i} + \varepsilon_t \qquad \text{(Model A)}$$

$$\Delta y_t = a_0 + \gamma . y_{t-1} + a_1 . t + \theta . DT_t + \sum_{i=1}^{p} \beta_i . \Delta y_{t-i} + \varepsilon_t \qquad \text{(Model B)}$$

$$\Delta y_t = a_0 + \gamma . y_{t-1} + a_1 . t + \theta . DU_t + \delta . DT_t + \sum_{i=1}^{p} \beta_i . \Delta y_{t-i} + \varepsilon_t \quad \text{(Model C)}$$

Here DU_t captures mean shift occurring at each possible break-date (TB) while DT_t is corresponding trend shift variable. Formally the values assigned to DU_t and DT_t may be summarised as follows. $DU_t = 1\ for\ t > TB\ ,\ and\ = 0\ otherwise.$ On the other hand $DT_t = t - TB\ for\ t > TB\ ,\ and = 0\ otherwise.$

The null hypothesis in all three models is that $\gamma = 0$, which implies that $\{y_t\}$ has a unit root with drift without any structural break. The alternative hypothesis if $\gamma < 0$, implies that the series is a trend-stationary with a single break occurring at some unknown time point. Zivot and Andrews regard every point as a potential break-date (*TB*) and run a regression for every possible break-date sequentially. From all possible break-points (*TB*), the procedure selects as its choice of break-date (*TB*) the date which minimizes the one-sided t-statistic for testing $\gamma = 0$ against $\gamma < 0$ [or $\gamma = (\varphi - 1) < 0$]. According to Zivot and Andrews, the presence of the end points cause the asymptotic distribution of the statistics to diverges towards infinity. Therefore, some region must be chosen such that the end points of the sample are not included. More recently, Sen (2003) showed that if one uses model A and if the break occurs according to model C then there would be a sizeable loss in power of the test. However, if break is characterized according to model A, but model C is used then the loss in power is negligible, suggesting the superiority of model C over model A. While Zivot and Andrews (1992) and Perron (1997) determined the point of break 'endogenously' from the data, Lumsdaine and Papell (1997) suggested an improvement over the Zivot and Andrews (1992) model by incorporating a couple of structural breaks. However, such endogenous tests have been subject to criticism for their treatment of breaks under the null hypothesis. If the breaks are absent under the null hypothesis of unit root these tests may suggest evidence of stationarity with breaks (Lee and Strazicich, 2003). Lee and Strazicich (2003) on the other hand propose a two break minimum Lagrange Multiplier (LM) unit root test in which the alternative hypothesis unambiguously implies that the series is trend stationary.

3.5 Toda – Yamamoto Modified Granger Causality under VAR Environment

A simple definition of Granger Causality, in the case of two time-series variables, X and Y is as follows. "x is said to Granger-cause y if y can be better predicted using the histories of both x and y than it can by using the history of y alone." The absence of Granger causality can be tested by estimating the following VAR model (equations.3.5.1 and 3.5.2).

$$y_t = \alpha + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{i=1}^{p} \beta_i x_{t-i} + u_{1t} \quad (3.5.1)$$

$$x_t = \beta + \sum_{i=1}^{p} \lambda_i y_{t-i} + \sum_{i=1}^{p} \delta_i x_{t-i} + u_{2t} \quad (3.5.2)$$

For the present study y_t represents detrended real WPI for India and x_t represents broad money supply or G. X does not Granger cause Y is tested by H_{01}: $\beta_1 = \beta_2 = \cdots = \beta_p = 0$ against the alternative that $\beta_1 \neq \beta_2 \neq \cdots \neq \beta_p \neq 0$. On the other hand Y does not Granger cause X is tested by H_{02}: $\lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$ against the alternative the $\lambda_1 \neq \lambda_2 \neq \cdots \neq \lambda_p \neq 0$. In each case rejection of null hypothesis implies the presence of Granger causality. The modified Wald test for testing Granger causality as proposed by Toda and Yamamoto (1995) avoids the problems associated with the usual Granger causality testing (which ignores non-stationarity and co-integrations between series while testing for causality). If the Wald test is being used to test linear restrictions on the parameters of a VAR model, and the data are non-stationary (which is most likely), then the Wald test *statistic* does not follow its usual asymptotic chi-square distribution under the null hypothesis (Toda and Yamamoto, 1995).

The approach to modified Granger causality as adopted in this study is outlined as follows. **First**, each time series variable is tested for stationarity (or for its order of integration) using standard tests such as ADF, PP and KPSS. The maximum order of integration (*m*) for the group of time-series is determined. Structural breaks if any are identified and a structural break dummy variable is created. **Second**, a VAR model is set up in level, regardless of the orders of integration of the various time-series. None of the variables are differenced.

Third, the optimum lag length for each variable in the VAR, say *p*, is determined using AIC, SIC, HQ, or other usual statistics. Care is taken so that there is no serial correlation in the residuals. The length p may be increased slightly until autocorrelation issues are resolved. Normality of the VAR residuals is highly desirable. **Fourth**, if both the time-series have the same order of integration, then Johansen Co-integration test is applied to test for co-integration (based on the selected VAR model). It provides some cross-check on the validity of the Causality results. **Fifth**, the favoured VAR model is constructed and additional *m* lags of each variable are inserted into each equation. In EVIEWS 9 these new *m* variables are to be treated as exogenous to the VAR system. The structural break dummy is also added (not shown) as an exogenous variable. It is thus ensured that the additional *m* lags and the structural break dummy would not be dropped while testing for Granger non-causality (via the Wald tests). The new VAR is presented in equations (3.5.3) and (3.5.4).

$$y_t = \alpha + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=p+1}^{p+m} \alpha_j y_{t-j} + \sum_{i=1}^{p} \beta_i x_{t-i} + \sum_{j=p+1}^{p+m} \beta_j x_{t-j} + u_{1t}$$

[3.5.3]

$$x_t = \beta + \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=p+1}^{p+m} \alpha_j y_{t-j} + \sum_{i=1}^{p} \beta_i x_{t-i} + \sum_{j=p+1}^{p+m} \beta_j x_{t-j} + u_{2t}$$

[3.5.4]

Finally, the hypothesis that the coefficients of only the first *p* lagged values of *x* are restricted to zero in the first equation, is tested using the standard Wald test (to test H_{01}: *x* does not Granger cause *y*). Analogously, a similar procedure is followed to test that *y* does not Granger cause *x*. The Wald statistic under the null hypothesis will be asymptotically distributed as chi-square with *p* degrees of freedom. Importantly enough, if two or more time-series are co-integrated, then there must be Granger causality between them (either uni-directional or both ways). The converse however is not true. Thus causality may be present without co-integration. The following section presents empirical results of the study along with discussions.

According to Zapata and Rambaldi (1997) the advantage of using the Toda-Yamamoto procedure is that in order to test Granger causality in the VAR framework, it is not necessary to pre-test the variables for the integration and co-integration properties, provided the maximal order of integration of the process does not exceed the true lag length of the VAR model. According to Toda and Yamamoto (1995), Toda-Yamamoto procedure however does not substitute the conventional unit roots and co-integration properties pre-testing in time series analysis. They are considered as complimentary to each other.

3.6 Vector Auto-regression (VAR)

For a set of n time series variables $y_t = (y_{1t}, y_{2t,}...,y_{nt})'$, a VAR model of order p (VAR(p)) can be written as, $y_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + u_t$,

where, the A_i 's are (nxn) coefficient matrices and $u_t = (u_{1t}, u_{2t},...,u_{nt})'$ is an unobservable i.i.d. zero mean error term. Consider a two-variable VAR (1) with k = 2.

(3.6.1) $y_t = b_{10} - b_{12} z_t + c_{11} y_{t-1} + c_{12} z_{t-1} + \varepsilon_{yt}$

(3.6.2) $z_t = b_{20} - b_{21} y_t + c_{21} y_{t-1} + c_{22} z_{t-1} + \varepsilon_{zt}$

with $\varepsilon_{it} \sim i.i.d(0, \sigma^2_{\varepsilon i})$ and $\text{cov}(\varepsilon_y, \varepsilon_z) = 0$

In matrix form: (3.5.3) $\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$

More simply, (3.5.4) $BX_t = \Gamma_0 + \Gamma_1 X_{t-1} + \varepsilon_t$

which is a **structural VAR (SVAR) or the Primitive System.**

To normalize the LHS vector, we need to multiply the equation by inverse B:

$B^{-1} B X_t = B^{-1} \Gamma_0 + B^{-1} \Gamma_1 X_{t-1} + B^{-1} \varepsilon_t$, thus,

(3.6.5) $X_t = A_0 + A_1 X_{t-1} + e_t$

which is a VAR in standard form (unstructured VAR or UVAR).

Alternately,

(3.6.6) $\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$

These error terms are composites of the structural innovations from the primitive system.

$e_t = B^{-1}\varepsilon_t$ where $B^{-1} = \frac{1}{|B|} B^a = \frac{1}{|B|}(B^*)^T = \frac{1}{(1-b_{21}b_{12})}\begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$

B^* =cofactor of B and $(B^*)^T$ =transpose.

Thus (3.5.7) $\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \frac{1}{(1-b_{21}b_{12})}\begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}\begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$

Or $e_{1t} = \frac{\varepsilon_{yt} - b_{12}\varepsilon_{zt}}{\Delta}$ where $\Delta = 1 - b_{21}b_{12}$

$e_{2t} = \frac{-b_{21}\varepsilon_{yt} + \varepsilon_{zt}}{\Delta}$

$\varepsilon's$ are white noise, thus *e*'s are $(0, \sigma_i^2)$:

$E(e_{it}) = 0$

$Var(e_{1t}) = E(e_{1t}^2) = \frac{E(\varepsilon_{yt}^2 + b_{12}^2\varepsilon_{zt}^2)}{\Delta^2} = \frac{\sigma_y^2 + b_{12}^2\sigma_z^2}{\Delta^2}$ is time independent, and the same is true for $Var(e_{2t})$. But covariances are not zero. This can be seen from,

Covar $(e_{1t}, e_{2t}) = E(e_{1t}e_{2t}) = \frac{E[(\varepsilon_{yt} - b_{12}\varepsilon_{zt})(\varepsilon_{zt} - b_{21}\varepsilon_{yt})]}{\Delta^2} = \frac{-(b_{12}\sigma_z^2 + b_{21}\sigma_y^2)}{\Delta^2} \neq 0$.

So the shocks in a standard VAR are correlated. The only way to remove the correlation and make the co-var = 0 is if we assume that the contemporaneous effects are zero: $b_{12} = b_{21} = 0$.

The var/covar matrix of the VAR shocks are represented as,

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}.$$

Identification

We can estimate (3.5.6) with OLS, since the RHS consists of predetermined variables and the error terms are white noise. The errors are serially uncorrelated but correlated across equations. Although SUR could be used in these cases, here we do not need it since all the RHS variables are identical, so there is no efficiency gain in using SUR over OLS. But we cannot use OLS to estimate the SVAR because of contemporaneous effects, which are correlated with the $\varepsilon 's$ (structural innovations).

To see how a structural innovation ε_{it} affects the dependent variables in our original model.

Sims (1980) suggested using a recursive system. For this we need to restrict some of the parameters in the VAR. Assume y is contemporaneously affected by z but not vice-versa. Thus we assume that $b_{21} = 0$. In other words, y is affected by both structural innovations of y and z, while z is affected only by its own structural innovation. This is a triangular decomposition also called Cholesky decomposition. Then we have 9 parameter estimates and 9 unknown structural parameters, and SVAR is exactly identified.

Now the SVAR system becomes:

$$(3.6.8)\quad \begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}\begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$B^{-1} = \frac{1}{(1-b_{21}b_{12})}\begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} = \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix}.$$

Hence the VAR system in standard form can be written as,

$$(3.6.8')\quad \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} - b_{12}b_{20} \\ b_{20} \end{bmatrix} + \begin{bmatrix} (c_{11} - b_{12}c_{21}) & (c_{12} - b_{12}c_{22}) \\ c_{21} & c_{22} \end{bmatrix}\begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} - b_{12}\varepsilon_{zt} \\ \varepsilon_{zt} \end{bmatrix}$$

If we match the coefficients in (8') with the estimates in (3.5.6) $\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$, we can extract the coefficients of the SVAR:

$a_{10} = b_{10} - b_{12}b_{20}$ $\quad a_{20} = b_{20}$ $\quad e_1 = \varepsilon_y - b_{12}e_z$

$a_{11} = c_{11} - b_{12}c_{21}$ $\quad a_{21} = c_{21}$ $\quad e_2 = e_z$

$a_{12} = c_{12} - b_{12}c_{22}$ $\quad a_{22} = c_{22}$ $\quad Cov_{12} = \dfrac{-(b_{12}\sigma_z^2 + b_{21}\sigma_y^2)}{\Delta^2} = -b_{12}\sigma_z^2$

Impulse response functions

We want to trace out the time path of the effect of structural shocks on the dependent variables of the model. For this, we first need to transform the VAR into a VMA representation.

Rewrite the UVAR more compactly.

$$(3.5.9)\; X_t = A_0 + A_1 X_{t-1} + e_t \Rightarrow X_t = \frac{A_0}{I - A_1 L} + \frac{e_t}{I - A_1 L}$$

First, consider the first component on the RHS:

$$\frac{A_0}{I-A_1}=(I-A_1)^{-1}A_0=\frac{(I-A_1)^a A_0}{|I-A_1|}=\frac{\begin{bmatrix}1-a_{11} & -a_{12}\\ -a_{21} & 1-a_{22}\end{bmatrix}A_0}{\begin{vmatrix}1-a_{11} & -a_{12}\\ -a_{21} & 1-a_{22}\end{vmatrix}}=\frac{\begin{bmatrix}1-a_{22} & a_{21}\\ a_{12} & 1-a_{22}\end{bmatrix}\begin{bmatrix}a_{10}\\ a_{20}\end{bmatrix}}{(1-a_{11})(1-a_{22})-a_{21}a_{12}}$$

$$=\frac{1}{\Delta}\begin{bmatrix}(1-a_{22})a_{10}+a_{21}a_{20}\\ a_{12}a_{10}+(1-a_{22})a_{20}\end{bmatrix}=\begin{bmatrix}\bar{y}\\ \bar{z}\end{bmatrix}$$

Stability requires that the roots of $I-A_1L$ lie outside the unit circle. We will assume that it is the case. Then, we can write the second component as:

$$\frac{e_t}{I-A_1L}=\sum_{i=0}^{\infty}A_1^i e_{t-i}=\sum_{i=0}^{\infty}\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}^i\begin{bmatrix}e_{1,t-i}\\ e_{2,t-i}\end{bmatrix}$$

We can thus write the VAR as a VMA with the standard VAR's error terms.

$$\text{(3.6.10)}\qquad \begin{bmatrix}y_t\\ z_t\end{bmatrix}=\begin{bmatrix}\bar{y}\\ \bar{z}\end{bmatrix}+\sum_{i=0}^{\infty}\underbrace{\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}^i}_{A^i}\begin{bmatrix}e_{1,t-i}\\ e_{2,t-i}\end{bmatrix}$$

But these are composite errors consisting of the structural innovations. We must thus replace the e's with the $\varepsilon's$ from (7) $e_t=\frac{1}{|\ |}\begin{bmatrix}1 & -b_{12}\\ -b_{21} & 1\end{bmatrix}\varepsilon_t$

$$\text{(3.6.10a)}\qquad \begin{bmatrix}y_t\\ z_t\end{bmatrix}=\begin{bmatrix}\bar{y}\\ \bar{z}\end{bmatrix}+\sum_{i=0}^{\infty}\underbrace{\frac{A^i}{1-b_{12}b_{21}}\begin{bmatrix}1 & -b_{12}\\ -b_{21} & 1\end{bmatrix}}_{\Phi_i}\begin{bmatrix}\varepsilon_{y,t-i}\\ \varepsilon_{z,t-i}\end{bmatrix}\qquad =\begin{bmatrix}\bar{y}\\ \bar{z}\end{bmatrix}+\sum_{i=0}^{\infty}\begin{bmatrix}\Phi_{11}^{(i)} & \Phi_{12}^{(i)}\\ \Phi_{21}^{(i)} & \Phi_{22}^{(i)}\end{bmatrix}^i\begin{bmatrix}\varepsilon_{y,t-i}\\ \varepsilon_{z,t-i}\end{bmatrix}$$

$$=\bar{X}+\sum_{i=0}^{\infty}\Phi_i\varepsilon_{t-i}\,.$$

Impact multipliers

They trace the impact effect of a one unit change in a structural innovation. Ex: find the impact effect of $\varepsilon_{z,t}$ on y_t and z_t :

$$\frac{dy_t}{d\varepsilon_{z,t}} = \Phi_{12}(0) \qquad \frac{dz_t}{d\varepsilon_{z,t}} = \Phi_{22}(0)$$

Lets trace the effect one period ahead on y_{t+1} and z_{t+1}

$$\frac{dy_{t+1}}{d\varepsilon_{z,t}} = \Phi_{12}(1) \qquad \frac{dz_{t+1}}{d\varepsilon_{z,t}} = \Phi_{22}(1)$$

Note that this is the same effect on y_t and z_t of a structural innovation one period ago:

$$\frac{dy_t}{d\varepsilon_{z,t-1}} = \Phi_{12}(1) \qquad \frac{dz_t}{d\varepsilon_{z,t-1}} = \Phi_{22}(1)$$

Impulse response functions are the plots of the effect of $\varepsilon_{z,t}$ on current and all future y and z. IRs show how $\{y_t\}$ or $\{z_t\}$ react to different shocks. Impulse response function of y to a one unit change in the shock to z may be expressed as,

$= \Phi_{12}(0)$, $\Phi_{12}(1)$, $\Phi_{12}(2)$, ...

Cumulated effect is the sum over IR functions: $\sum_{i=0}^{n} \Phi_{12}(i)$.

Long-run cumulated effect: $\lim_{n\to\infty} \sum_{i=0}^{n} \Phi_{12}(i)$

In practice we cannot calculate these effects since the SVAR is under-identified. So we must impose additional restrictions on the VAR to identify the impulse responses.

If we use the Cholesky decomposition and assume that y does not have a contemporaneous effect on z, then $b_{12} = 0$. Thus the error structure becomes lower triangular:

(3.6.11) $$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

The ε_y shock doesn't affect z directly but it affects it indirectly through its lagged effect in VAR.

Granger Causality: If the z shock affects e_1, e_2 and the y shock doesn't affect e2 but it affects e_1, then z is causally prior to y.

3.7 Causality in Time Series Econometrics – Granger and Sims

Perhaps the most influential explicit approach to causality in economics is due to Clive W. J. Granger (1969). Granger causality is an inferential approach, in that it is data based without direct reference to background economic theory; and it is a process approach, in that it was developed to apply to dynamic time-series models. Granger causality is an example of the modern probabilistic approach to causality, which is a natural successor to Hume (e.g., Suppes 1970). Where Hume required constant conjunction of cause and effect, probabilistic approaches are content to identify cause with a factor that raises the probability of the effect: A causes B if $P(B|A) > P(B)$, where the vertical "|" indicates "conditional on". The asymmetry of causality is secured by requiring the cause (A) to occur before the effect (B). Granger's (1980) definition is more explicit about temporal dynamics than is the generic probabilistic account, and it is cast in terms of the incremental predictability of one variable conditional on another:

X_t Granger-causes Y_{t+1} if $P(Y_{t+1}|$ all information dated t and earlier$) \neq P(Y_{t+1}|$ all information dated t and earlier omitting information about X$)$.

More generally, since the future cannot predict the past, if variable *X* (Granger) causes variable *Y*, then changes in *X* should *precede* changes in *Y*. Therefore, in a regression of *Y* on other variables (including its own past values) if we include past or lagged values of *X* and it significantly improves the prediction of *Y*, then we can say that *X* (Granger) causes *Y*. A similar definition applies if *Y* (Granger) causes *X*.

Christopher Sims (1972) famously used Granger-causality to demonstrate the causal priority of money over nominal income. Later, as part of a generalized critique of structural econometric models, Sims (1980) advocated vector auto-regressions (VARs) – a theoretical time-series regression, but generally including more variables with lagged values of each appearing in each equation. In the VAR context, Granger-causality generalizes to the multivariate case. Sims (1980) advocated VARs as a reaction to the manner in which the Cowles Commission program, which identified structural models through a priori theory, had been implemented. From a causal perspective, it was closely related to Granger's analysis. Starting with VAR such as equations (3.6.1) and (3.6).2, Sims wished to work out how various "shocks" would affect the variables of the system. This is complicated by the fact that the error terms in (3.6.1) and (3.6.2), which might be taken to represent the shocks, are not in general independent, so that a shock to one is a shock to both, depending on how correlated they are. Sims's initial solution was to impose an arbitrary orthogonalization of the shocks (a Choleski decomposition). This amounts to imposing a recursive order on X_t and Y_t , such that the covariance matrix of the error terms is diagonal (i.e., ε_{yt} and εz_t are uncorrelated). A shock to X can then be represented by a realization of ε_{yt} and a shock to Y by a realization of ε_{2t}. Initially, Sims treated the choice of recursive order as a matter of indifference. Criticizing the VAR program from the point of view of structural models, Leamer (1985) (in Causality in Economics and Econometrics), and Cooley and LeRoy (1985) pointed out that the substantive results (e.g., impulse response functions and innovation accounts) depend on which recursive order is chosen.

A simple physical example makes it clear what is happening. Suppose that X measures the direction of the rudder on a ship and Y the direction of the ship. The ship is pummelled by heavy seas. If the helmsman is able to steer on a straight course, effectively moving the rudder to exactly cancel the shocks from the waves, the direction of the rudder (in ignorance of the true values of the shocks) will not predict the course of the ship. The rudder would be structurally effective in causing the ship to turn, but it would not Granger-cause the ship's course.

Sims (1982, 1986) advocated structural vector Autoregressions (SVARs). SVARs can be identified through the contemporaneous causal order only. Ironically, since the initial impulse behind the VAR program was to avoid theoretically tenuous identifying assumptions, the choice of restrictions on contemporaneous variables used to transform the VAR into the SVAR are typically only weakly supported by economic theory. Nevertheless, the move from the VAR to the SVAR is a move from an inferential to an a priori approach. It is also a move from a fully non-structural, process approach to a partially structural approach, since the structure of the contemporaneous variables, though not of the lagged variables, is fully specified (Hoover, 2006). The present study hinges on the Toda – Yamamoto version of the Granger causality which makes use of the usual unstructured VAR as opposed to Sim's SVAR.