

ECONOMETRICS MODELLING
Frequently Asked Questions

Prepared by:

Sayed Hossain

Faculty of Management, Multimedia University
63100 Cyberjaya

Website: www.sayedhossain.com

Email: sayed.hossain@yahoo.com

Year 2008.

WHAT IS ECONOMETRICS?

Econometrics is an amalgamation of statistics, mathematics and economic theory. In order to test any hypothesis or to justify any economic theory, econometrics is applied

The task of an econometrician is to produce a good regression model by fulfilling all the regression assumptions. In others words, managing the residual or error term of the regression model is the prime job for an econometrician.

WHAT ARE THE ASSUMPTIONS OF A LINEAR REGRESSION MODEL?

Assumptions are:

- The error term, u follows the normal distribution.
- The variance of error term, u is constant.
- No autocorrelation exists between the error terms u_i and u_j .
- No exact collinearity exists between the two explanatory variables.
- The error term u has a zero mean.
- X_i is uncorrelated with the disturbance term u , where $i= 1,2,3,\dots,n$.

WHAT IS A HYPOTHEIS?

Hypothesis is set to test population phenomena, not sample phenomena.

DISCUSS THE FEATURES OF A GOOD ECONOMETRICS MODEL.

Characteristics are:

- An econometric model should be specified on the basis of economic theory, past experience and personal intuition.
- The sample data must be fitted to the model reasonably well, which means that a high R^2 value is desirable and majority of the independent variables should be significant.
- A single equation model (that we are dealing now) must meet all the assumptions of Classical Linear Regression Model (CLRM):
- The model is free from autocorrelation, heteroscedasticity and multicollinearity problems
- The error term u follows normal distribution.
- Coefficients that have been estimated must have the expected signs.
- A model must have a strong forecasting power.
- There must be a proper functional form between dependent and independent variables. In other words, the relationship between dependent and independent variables must be specified (linear or non-linear) as per actual relationship.

WHAT ARE THE SIX ASSUMPTIONS RELATED TO ECONOMETRICS MODELLING ?

Heteroscedasticity

H_0 : Homoscedasticity (the variance of u is constant)

H_1 : Heteroscedasticity (the variance of u is unequal)

Autocorrelation

H_0 : not autocorrelated (no correlation between error term, u_i and u_j)

H_1 : autocorrelated (correlation between error term, u_i and u_j)

Normality test

H_0 : Normal distribution (the error term u follows normal distribution)

H_1 : Not normal distribution (the error term u follows normal distribution)

Time Series Hypothesis

H_0 : Unit Root, Random Walk, non- stationary

H_1 : Stationary or white noise

Wald Statistics

Null H_0 . Lagged values of coefficients in each equation are zero

Alt H_1 : Not zero

WHO DEVELOPED THE COINTEGRATION CONCEPT?

The concept of cointegration was developed by Engle and Granger in 1987.

WHY TIME SERIES DATA MUST BE STATIONARY?

Stationary time series data is necessary to have a valid *t statistics* and *F statistics*.

WHAT DOES COINTEGRATION TELL?

Cointegration tells us about the presence of long run relation among two or more variables.

WHETHER VARIABLES NEED TO BE INTEGRATED IN THE SAME ORDER TO BE COINTEGRATED?

When we go for running cointegration analysis, we assume that all variables are non-stationary. Secondly they are all integrated of the same order. Even, if the variables are not integrated in the same order, we still can continue with cointegration analysis. We call this situation Multicointegration.

WHAT ARE THE TOOLS AVAILABLE FOR IDENTIFYING LONG RUN RELATIONSHIP OR COINTEGRATION?

There are indeed two tools to identify whether there exist a long run relationship among variables. They are:

1. Engel-Granger's residual based test
2. Johansen-Juselius (JJ) test.

Since *Engel-Granger's Test* suffers from some shortcomings, *Johansen-Juselius (JJ) test* is preferred for cointegration analysis.

WHICH ONE IS PREFERRED, MAXIMUM EIGENVALUE TEST OR TRACE STATISTICS ?

While doing *Johansen-Juselius Test*, if there comes up a different result between *trace statistic* and *maximum eigenvalue test*, *maximum eigenvalue* result is preferred (Banerjee et al 1993)

WHAT ARE THE FINDINGS OF ENGEL AND GRANGER ?

Engel and Granger (1997) identify that if cointegration exists between two variables in the long run, then there must be either unidirectional or bi-directional Granger Causality between these two variables.

Engel and Granger also identifies that the cointegrated variables must have an ECM (Error correction model) representation

HOW TO CHOOSE OPTIMAL LAG LENGTH WHILE GOING FOR COINTEGRATION ANALYSIS?

There requires an optimum lag length while going for cointegration analysis using Johansen-Juselius test. In this regards there are few tests such as *Akaike Information Criterion*, *Schwarz Bayesian Criterion*, *Sim's modified likelihood ratio test* etc to identify the optimal lag length. It is up to us to decide which method we would choose.

WHETHER EVIEWS SOFTWARE CAN SUPPORT TO IDENTIFY OPTIMAL LAG LENGTH?

Yes. In this regard, EVIEWS 5.1 software supports us to find out the optimal lag length using *Akaike Information Criterion*.

HOW TO KNOW THE NUMBERS OF COINTEGRATED VECTORS?

Cointegration analysis is done to find out the number of cointegrated vector necessary to run VAR model.

WHAT IS A VECTOR AUTOREGRESSIVE MODEL (VAR MODEL) ?

The Vector Autoregressive Model (VAR) describes about two or more variables operating in a system where the dependent variables are found as lagged ones on the right hand side of the equation.

Variables used in VAR are all assumed to be endogenous. But if a VAR is furnished with some exogenous variables together with endogenous, we can call it as VARMAX model

WHAT IS THE PURPOSE OF APPOINTING VAR ?

VAR is a technique that facilitates to capture both the dynamic and interdependent relationships of the said variables

WHEN SHALL WE APPOINT VECM MODEL?

Once the variables included in the VAR model found to be cointegrated, we will use Vector Error Correction Model (VECM), not VAR. Indeed VECM is special type of restricted VAR, is introduced to correct a disequilibrium that may shock the whole system.

WHAT ARE THE OBJECTIVES OF VECM

In VECM model, the dynamics of both short run and long run adjustment will be made. VECM will also allow us to find out the causal factors that affect our variables. But if we found that variables are not cointegrated, we will restrict our analysis on VAR only.

WHETHER OLS ASSUMPTIONS ARE NECESSARY TO FULFILL IN VECM ?

Yes, we will satisfy all the OLS assumption of our causal VECM model (that is residuals are homoscedastic, no serial correlation, normality of the residuals, appropriate sign of the variables, not high multicollinearity that is VIF should be less than 4 etc)

HOW TO REMOVE HETREROSCEDASTICITY FROM THE VECM MODEL OR MODEL IN GENERAL?

Whenever we are on the way to make any simple econometrics or VECM model, we will convert all the variables except dummy variable in natural log form in order remove heteroscedasticity problem from the model initially.

WHY WALD STATISTICS WILL BE APPOINTED IN VECM?

Wald statistic will be used to find out the causal affect on our dependent variable.

SHOULD WE USE SEASONAL ADJUSTED DATA IN VECM MODEL?

Since there exists huge criticism over seasonal adjusted data, we will not use any seasonal adjusted data.

HOW TO SELECT LAG LENGTH USING EIEWS 5.1?

Steps are as follows:

- From Eviews 5.1 >> choose *View/Lag Structure/Lag Length Criteria, max lags 12*
- Afterwards, we will find the model selection criterion such as log likelihood, LR, FPE, AIC SC, HQ. Note that estimation period of the VARs in this option is the same for all the models considered and therefore depends on the max. number of lags we want to consider. Once we have decided the optimal number of lags, the estimation period changes according to the selected lag number.

HOW TO COMIT DIAGNOSTIC CHECKING USING EIEWS ?

Residual analyses are to be carried out for diagnostic checking. To become best model or diagnostic checking, following investigations are necessary

Fitted Model

R- square value should be

Adjusted R square value is necessary when we would like to compare among models furring.

Normality test :

Null Hypothesis H_0 : Residuals (u) are normally distributed

Alternative H_1 : not normal

Suggestion: Histogram-Normality test (Perform Jarque-Bera Statistic)

Heteroscedasticity test

Null hypothesis H_0 : Variance of residuals (u) are constant: Homoscedasticity

Alternative H_1 : Variance of residuals (u) are unequal : Heteroscedasticity

Suggestion : Serial correlation LM test: ARCH LM test (Autoregressive Conditional Heteroscedasticity test)

Autocorrelation

Null hypothesis H_0 : no correlation between error terms u_i and u_j

Alternative H_1 : correlated between u_i and u_j

Suggestion : Breusch-Godfrey serial correlation LM test : BG test

Stability test: Stability of the dependent variable

We will use CUSUM test (Recursive residual of cumulative sums test).

This test looks at the cumulative sums of the residuals.

Decision: All the residuals are stable if the cumulative sums are located within the two standard deviation band. If the cumulative sums are outside the band meaning that parameter (here dependent variable) concerned is not stable.

Here cumulative sums is plotted against the time

HOW TO PERFORM GRANGER CAUSALITY TEST ?

Granger causality test can be performed using Wald statistics.

Null hypothesis H_0 . Lagged values of coefficients in each equation are zero

Alt H_1 : Not zero

Granger causality is done to see the short run causality running from independent variable to dependent variable

It is found that test statistics for granger test should follow chi-square distribution instead of F distribution. So we would follow Chi-square result.

WHAT ARE THE EXTENSIONS OF COINTEGRATION ANALYSIS ?

Cointegration is done to see whether there exists a long run equilibrium relationship among all the variables.

If we get one or more than one cointegrated vector in the model, we say that there exists a long run relationship among the variables.

WHAT ARE THE TEN STEPS NECESSARY TO BUILD A VAR MODEL ?

- Convert all the variables involved in natural log form to avoid heteroscedasticity problem.
- Choose optimal lag for unit root testing.
- Checking stationarity.
- If the variables are non-stationary, convert it into stationary to estimate.
- Choose optimal lag length for system model
- Go for cointegration analysis using non-stationary data
- If the variables are cointegrated, employ VECM model. If not, estimate VAR model.
- Appoint Wald statistics for estimating short run causality running from independent variable to dependent variable.
- Diagnostic checking is necessary (R square, hetro, auto, normality etc)
- Interpret the result of error correction term and exogeneity.

WHAT IS THE DYNAMIC BEHAVIOUS OF VAR MODEL?

The dynamic behaviour of VAR model can be analysed using following two approaches:

1. Impulse response function
2. Variance decomposition

1. Impulse response function (IRF)

- In IRF, we bring a shock in residuals or innovations of each equation in the system VAR model that would affect on all endogenous variables.
- Normally we introduce one period standard deviation shock to one of the endogenous variable
- We assume that the model is in a stable equilibrium and errors or innovations should not be correlated.
- So we need a common ordering in IRF as well as in Variance decomposition.

Ordering of the variables:

- Variables that are not caused by any other variables in the system will be placed first in the list of ordering.
- The rest of the variables are ordered according to the correlation among these variables
- The target variable will be put at the last in the ordering. In our case the target variable in GDP.

2. Variance decomposition

- Variance decomposition breaks down the variance of the forecast error for each variable into components.
- In other words, each variable is explained as a linear combinations of its own current innovations and lagged innovations of all the variables in the system.
- Like IRF, ordering is also important.

Prepared by:

Sayed Hossain
Faculty of Management, Multimedia University
63100 Cyberjaya
Website: www.sayedhossain.com
Email: sayed.hossain@yahoo.com

Please comment at my guestbook at: www.sayedhossain.com

NOTE

Use the information of this website on your own risk. This website shall not be responsible for any loss or expense suffered in connection with the use of this website.