REVIEW OF RESEARCH
IMPACT FACTOR : 5.7631(UIF) UGC APPROVED JOURNAL NO. 48514 ISSN: 2249-894X
VOLUME - 7 | ISSUE - 11 | AUGUST - 2018

METHODOLOGY OF ECONOMETRIC RESEARCH: AN OVERVIEW

Belure Vishal Chandrashekhar
M.A.(Economics), B.ED., SET(2), NET(2).
Assistant Professor & HOD of Economics, Yeshwant Mahavidyalaya, Nanded.

ABSTRACT

Methodology of econometric research is the most scientific and user friendly research technique in the social sciences, especially in economics. Econometrics is the unique combination of Economic theory, Mathematics and Statistical Methods. Most of the recently emerged research in the sphere of economics is done by using econometric methodology. Since its inception most of the economists of this branch of knowledge have received maximum number of Nobel Prizes. Though it is one of the recently developed branches, it is extensively used methodology in empirical economic research. This paper is a modest attempt to explain the step-by-step procedure of econometric research methodology, along with some of its formidable tools and techniques.

KEY WORD: *Econometrics, Research Methodology, Specification, Hypothesis Testing, Interpretation, Forecasting.*

INTRODUCTION:

Econometrics is one of the highly progressive and popular branches of Economics. It is expanding intensively as well as extensively. Though the first Nobel laureate in Economic Science; Ragnar Frisch coined the word 'Econometrics' in 1926, the roots of its origin goes in ancient times. We can also find some of the tenets of this discipline in Kapil Muni's 'Sankhya Darshana'. Different scholars defined econometrics in various ways. Professor Arthur Goldberger defined econometrics as, "A social science in which the tools of economic theory, mathematics and statistical inference are applied to the analysis of economic phenomenon." The joint winner of the first Nobel Prize in Economic Science, Professor Jan *Tinbergen wrote, "Econometrics is the name for a field of science in which mathematical economics* and statistical research are applied in combination." Reviewing above and some other definitions; the researcher can conclude that, '**Econometrics is indisputably a social science.** It is an appropriate conjunction of economic theory, mathematics and statistical methods'. Hence by combining a social science with the logic of certainty and the logic of uncertainty; econometrics becomes a very strong, unique and useful tool for conducting social science research. **When Econometric procedures, techniques and methods are applied to perform research, it is known as Econometric Research.**

ECONOMETRIC RESEARCH METHODOLOGY:

The Methodology of econometric research consist application of specific tools and techniques of econometrics in a systematic way. The step-by-step procedure of econometric research encompasses following stages.

ECONOMETRIC RESEARCH METHODOLOGY:

The Methodology of econometric research consist application of specific tools and techniques of econometrics in a systematic way. The step-by-step procedure of econometric research encompasses following stages.

STAGE I - PROPOSITION OF THE THEORY OR HYPOTHESIS:

The starting point for econometric research is the proposition of theory or hypothesis. **A hypothesis is an empirically testable version of a proposition.** It is a tentative statement about a relationship because when we start a study, we are uncertain as to whether the hypothesis actually holds in the empirical world. By empirical, it means study based upon data. For example, economist Lord J. M. Keynes postulated that there is a direct and proportional relationship between consumption expenditure and national income (GDP) in an economy. Symbolically, Y=f(X). Where Y= consumption expenditure and X= GDP. In econometric research two types of hypotheses are frequently used: **(a) Null hypothesis**- It is also called the hypothesis of no difference. In this hypothesis it is assumed that there is no difference between sample estimates and their population counterparts. The null hypothesis is represented by H_0; **(b) Alternative hypothesis-** it is the opposite side of null hypothesis. This is the hypothesis with difference. The alternative hypothesis is represented by H_a or H_1.

STAGE II - SPECIFICATION OF AN ECONOMETRIC MODEL:

In Econometrics specification of model implies **formulation of maintained hypothesis**. This step involves determination of the type of model i.e. single equation model or simultaneous equation model, the functional form of model i.e. linear or non-linear model, number of dependent and independent variables etc. Let's construct a model based on above stated Keynes' proposition. $Y = \alpha + \beta X + U$, where $\alpha > 0$ and $0 \leq \beta \geq 1$.

This equation is the Simple Linear Regression Model(SLRM), because there are only two variables in the above model and both variables are in the linear form. **The term regression is 'a statistical method of measuring(estimating) the extent to which variations in one variable are associated with variations in other variables.**' The regression analysis is considered as the bread-and-butter tool of econometrics. The above regression line is also called the Single equation model. The variable X and Y represent income(GDP) and consumption expenditure in the economy respectively. Parameters α & β are intercept and slope coefficient respectively. In economic theory they represent the autonomous consumption level and Marginal Propensity to Consume–MPC. The last term **'U' is the unique feature** of an econometric model. It is called the '**error term'** or **'disturbance term'**. It represents the random or stochastic part of the above stated relationship between variable X and Y. It is also called **unexplained variation** of the regression model. The remaining part of the right-hand side i.e. $\alpha + \beta X$ is called the **explained variation** part of the regression. Specification of model is a very crucial part of the econometric research. It needs great skills, theoretical knowledge and common sense on the part of a researcher.

STAGE III - COLLECTION OF DATA:

Data is the plural word for datum. The observations gathered from an experiment, survey or observational study is called data. There are many types and forms of data, like population and sample data, primary and secondary data, quantitative and qualitative/nominal data, experimental and non-experimental data etc. There are three types of data mainly used in the econometric research: (a) Time series data; (b) Cross- section data; (c) Panel data.

Time series data are a series of observations taken over a period of time usually at equal intervals, for example- data on Personal Consumption Expenditure-PCE and GDP of India from the year 1950-51 to 2017-18. **Cross-section** data are observations across individuals for a single time period, for example- cross-countries data of New Human Development Report-NHDI in the year 2017. **Panel data** is also known as longitudinal data. It is a special type of pooled data which combine time series and cross-section data in a

very specific way. Panel data includes observations on the same variables from the same cross-sectional sample from two or more different time periods. For example, if the researcher surveyed 100 students when they graduated from a college and then administered the same questionnaire to the same students after two years, he would have created a panel data set.

STAGE IV - ESTIMATION OF ECONOMETRIC MODEL:

After the econometric model has been specified and the data set is collected, the researcher must proceed with the estimation stage. **To estimate an econometric model** means to obtain as nearly as possible numerical values of the population parameters by using sample data set. The estimation of an Econometric model is a purely technical process which requires the knowledge of various estimation methods with their pros and cons. Broadly speaking there are two types of estimation methods in econometrics: **(a) Single equation methods–** Ordinary Least Squares-OLS Method, Generalized Least Squares-GLS Method etc. for example- in our Keynesian regression model we can obtain the OLS estimate of β as $\hat{\beta} = \sum xy/x^2$ and $\hat{\alpha} = (\sum Y/n) - \hat{\beta}(\sum X/n)$; **(b) Simultaneous equation methods**- Maximum Likelihood-ML Method, Three Stage Least Squares-3SLS Method etc. are applicable to estimate all parameters in a system of simultaneous equations at the same time. In case of the estimation of simultaneous equation model, researcher has to deal with the problem of identification first. There are two types of estimators; point estimator and interval estimators.

Point estimator is a single valued output whereas interval estimator consist a range of two output values. Now-a-days n numbers of user friendly software are available to perform any kind of estimation, like SPSS, GRETL, EVIEWS, LIMDEP, R, STATA etc. The researcher can also use MS Excel program to perform some estimation procedures.

STAGE V - TESTING OF HYPOTHESIS AND INTERPRETATION OF RESULTS:

After estimation of regression model, the next step is testing or evaluation of the results. In this stage it is examined that whether sample estimates are in accordance with underlying theory. Are they statistically significant? Is there any statistically significant proximity between sample estimates and population counterparts? In econometric research there are three types of criteria used in hypothesis testing/ tests of significance: (a) The **theoretical or *a priory criteria*** are set by economic postulations. They mainly refer to the sign and size of the parameters. They should be defined in the first stage of research; (b**)** The **statistical criteria or first order tests** are some routinely used tests in statistics, like the coefficient of determination(r^2) test, standard error(σ) test, Z/t/F tests etc. ; (c) The **econometric criteria or second order tests** are mainly invented in econometric research. They are Multicollinearity, Heteroskedasticity, Autocorrelation etc.

There are numerous tests to perform the task of hypothesis testing. These tests are classified into two groups: (a) Parametric tests; (b) Non-parametric tests. **Parametric tests** are based on the assumption that the given data set is normally distributed. For example- Student's 't' test which is used to test the statistical significance of an individual sample estimate, the 'F' test is used to test the overall significance of the all regression estimates. **Non-parametric tests** are those tests which do not rely on the assumption of normal distribution. These tests are free from distribution and parameter i.e. there is no need to estimate parameters. For example- One sample sign test, Chi- Square test, Run's test etc. Non-parametric tests can be used even for nominal and ordinal data. These tests can also be applied to very small samples. In our Keynesian model, the formula to obtain the calculated value of 't' test for $\hat{\beta}$ is $t^* = (\hat{\beta} - \beta_0)/ SE(\hat{\beta})$. Where β_0 0 - null hypothesis and SE ($\hat{\beta}$) means standard error of $\hat{\beta}$. In the beginning of the hypothesis testing step the researcher has to fix the level of significance i.e. α if necessary. It may be 1 percent, 5 percent or otherwise. In parametric tests researcher has to compare calculated values with predetermined critical or table values. If the calculated value is less than the critical value; the null hypothesis is accepted and vice versa. In hypothesis testing the H_1 sometimes called a directional hypothesis, may be one tailed or two tailed. If the values identified by H_1 are all rest on only one side of the H_0, then it is a **one-tailed hypothesis(test)**. In one-tailed test critical region or the region of rejection is either on the extreme left or on the extreme right of the

normal curve. On the contrary, if the H_1 values lie on both directions of H_0, then it will be a **two-tailed hypothesis(test)**. In two-tailed test the critical area is kept on both sides of the normal distribution. Testing of hypothesis is a very sensitive task. Sometimes it leads to the type I or type II error. **Type I error** occurs when the researcher rejects the H_0 when it is true. Type **II error** occurs when the researcher accepts H_0 when it is false. The probabilities of type I and type II errors are denoted by **α & β** respectively. The term 1- **β** is the measure of the **power of test**. In conducting a hypothesis test, a **P-value** i.e. probability value helps the researcher to determine the significance of his results. If the value of $P \leq 0.05$, the researcher rejects H_0. And if the value of $P > 0.05$ he accepts H_0. P value is called the exact or observed level of significance or exact probability of committing a type I error. In other words, the P value is the lowest significance level at which a H_0 can be rejected.

Interpretation may be regarded as the next to last step of econometric research. Interpretation of research refers to the process of assigning meaning to the obtained estimates of given regression line. Interpretation of results is also important for taking policy decisions. In above consumption- income regression hypothesis, let our estimated regression line is $\hat{Y}= 50 + 0.6\ X$. It means that the autonomous consumption level i.e. $\hat{\alpha}=50$ and the marginal propensity consume-MPC i.e. $\hat{\beta} =0.6$.

STAGE VI - FORECASTING OR PREDICTION:

It is the final stage of an econometric research. Once the model in consideration successfully passes through hypothesis testing and interpretation procedure, it can be used for forecasting or prediction of economic phenomenon. **Forecasting** is the prediction of values of a variable based on known past values of that variables or other related variables. In most forecasts the values being predicted are for future time periods, but cross-sectional forecasting of values for not in the sample set are also common. There are five important methods of econometric forecasting: (a) Exponential smoothing methods; (b) Single equation regression method; (c) Simultaneous equation regression methods; (d) Box- Jenkins ARIMA method; (e) Vector Autoregression- VAR method.

Exponential smoothing methods are some supplementary forecasting methods. They provide forecasts using weighted averages of past values of the data and forecast errors. They are commonly used in inventory control systems where many items are to be forecasted and low cost is a primary concern. There are several exponential smoothing methods of forecasting, like Single Exponential Smoothing-SES method, Holt's linear method, Holt-Winter's method etc.

In single and simultaneous equation regression methods once researcher successfully obtains numerical estimates of all the parameters, he can insert any value of independent or exogenous variables to obtain the estimates or forecast values of dependent or endogenous variables. For example, in the above stated estimated regression line $\hat{Y}= 50 + 0.6\ X$, if researcher assumes the value of independent variable X= 100, the estimated or forecasted value of dependent variable will be $\hat{Y}=110$. The difference between actual and estimated or forecasted values of a variable is called residual.

Box- Jenkins Autoregressive Integrated Moving Average-ARIMA methodology is one of the most famous forecasting methods in recent econometric research. This method is developed by G. P. E. Box and G. M. Jenkins in 1978. It is basically a short-term forecasting method based on time series information of single variable. ARIMA methodology requires following four steps to be performed.

(a) Identification of ARIMA (p, d, q) Model: Here the letter p stands for the value of lag length of AR component, d is the value of differencing of data to make it stationary if required. And q is the value of MA component. Determining values of p, d and q need specific skills, experience and some knowledge of stationarity tests, AIC and BIC criteria etc. For example the ARIMA (2, 1, 2) model will be written as,

$$Y_t = \theta + \phi_1 \Delta Y_{t-1} + \phi_2 \Delta Y_{t-2} + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2}$$

where Y_t = a dependent variable, θ= an intercept term, ε_t = error term, ϕ's = parameters of AR components, Δ's= difference operators and ψ's = parameters of MA components.

(b) Estimation of ARIMA (p, d, q) Model: In this step the researcher has to obtain the estimated values of the chosen ARIMA (p, d, q) model by using OLS method. But sometimes it requires some non-linear methods of estimation like ML method.

(c) Diagnostic Testing of ARIMA (p, d, q) Model: In this step the researcher uses some sophisticated statistical tests for checking the significance of estimated AR and MA parameters. These tests include Partial Autocorrelation and Autocorrelation Functions (PAC and ACF) tests, Box- Pierce Q statistic etc.

(d) Forecasting with ARIMA (p, d, q) Model: After successfully passing through the above three stages, the selected ARIMA (p, d, q) model is finally used for prediction purpose. The forecasting step is as simple as single equation method. The forecast values obtained by this method are more reliable than earlier methods.

Now-a-days well-documented software programs are developed for conducting ARIMA forecasting. Many statistical software companies are offering ARIMA add-ons to perform the task of forecasting.

The VAR i.e. Vector Auto-regression Modeling technique is developed by Christopher Sims in 1980. It is one of the types of simultaneous equation model in which the value of a variable is expressed as a linear function of the past or lagged values of that variable and all other variables included in the model. Since the lagged values of each variable present on the right hand-side of the equation and there is a vector of two (or more) variables appears on the left-hand side of these equations, they are called VAR models. VAR models do not generate the problem of simultaneity bias. They can be estimated directly by applying OLS method for each equation independently. Once researcher obtains the estimated VAR model; it is very easy to use this model for forecasting purpose. Like ARIMA modeling the VAR modeling is also superior and widely used forecasting methods in econometric research.

LIMITATIONS:

Econometric research methodology is not free from criticism. It is argued that, econometrics mainly deals with quantitative data. It throws very little light on qualitative data. Econometrics is based on statistics and statistics is based on certain assumptions. If assumptions are not satisfied, the study based on them generates errors. Econometrics is regarded as a positive science in the sense that it does not help us in value judgments. Econometric methods are complex, tedious and time consuming. An econometric researcher has to be a competent economist, mathematician and statistician by training. It is difficult for others to use this methodology for research.

CONCLUSION:

Econometrics is the unification of Economics, Mathematics and Statistics. The main objective of econometrics is to perform research. Econometric research methodology is one of the most powerful and widely used scientific research methodologies in social sciences. The scope and areas of application of econometric research methods are expanding constantly. Econometric methodology is used in describing economic reality, testing hypotheses about economic theory, policy and forecasting future values of the variables. It is used to conduct research in Economics, Political Science, Behavioral Sciences, Management Sciences etc.

REFERENCES:

1) Black J. (2005). *A Dictionary of Economics.* New Delhi, India. Oxford University Press.
2) Christopher C. & James N. (2017). *The Concise Dictionary of Mathematics.* Daryaganj, New Delhi, India. Oxford University Press.
3) Franses P. H. (2002). *A Concise Introduction to Econometrics: An intuitive guide.* Cambridge, UK. Cambridge University Press.
4) Gujarati D. N. (2004). *Basic Econometrics.* New York, US. McGraw-Hill Higher Education.
5) Gupta S. P. (2008). *Statistical Methods.* Deryagang, New Delhi. Sultan Chand & Sons.
6) Koutsoyiannis A. (2001). *Theory of Econometrics.* New York. US. Palgrave Publishers Ltd.

7) Makridakis S., Wheelwright S. C. & Hyndman R. J. (2005). *Forecasting: Methods and applications.* Singapore. John Wiley & Sons, Inc.
8) Neuman W. L. (2015). *Social Research Methods: Qualitative and quantitative approaches.* Noida-UP, India. Dorling Kindersley India Pvt. Ltd.
9) Panneerselvam R. (2016). *Research Methodology.* Delhi, India. PHI Learning Private Ltd.
10) Ramanathan R. (2002). *Introductory Econometrics with Applications.* Singapore. Thomson Asia Pvt. Ltd.
11) Singh S. P., Parashar A. K. & Singh H. P. (1999). *Econometrics and Mathematical Economics.* New Delhi, India. S. Chand and Company Ltd.
12) Stock J. H. & Watson M. W. (2018). *Introduction to Econometrics.* Noida, UP. Pearson India Education Services Pvt. Ltd.
13) Studenmund A. H. (2018). *Using Econometrics: A practical guide.* Noida, UP. Pearson India Education Services Pvt. Ltd.
14) Upton G. & Cook I. (2014). A *Dictionary of Statistics.* New Delhi, India. Oxford University, Press.
15) *https://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1969/.*

Belure Vishal Chandrashekhar
M.A.(Economics), B. ED., SET(2), NET(2)
Assistant Professor & HOD of Economics,Yeshwant Mahavidyalaya, Nanded.