ADDIS ABABA UNIVERSITY
GRADUATE STUDIES PROGRAMME
DEPARTMENT OF STATISTICS

MULTIVARIATE TIME SERIES ANALYSIS OF ETHIOPIAN INTERNATIONAL AIR TRAVEL DEMAND

BY

THOMAS SOLOMON

JUNE 2014

ADDIS ABABA

ADDIS ABABA UNIVERSITY
GRADUATE STUDIES PROGRAMME
DEPARTMENT OF STATISTICS

MULTIVARIATE TIME SERIES ANALYSIS OF ETHIOPIAN INTERNATIONAL AIR TRAVEL DEMAND

A Thesis submitted to the School of Graduate Studies of Addis Ababa University in partial fulfillment of the requirements for the Degree of Master of Science in Statistics

By Thomas Solomon

Advisor: Dr. M.K. Sharma (Professor of Statistics)

June 2014

Addis Ababa

ADDIS ABABA UNIVERSITY
GRADUATE STUDIES PROGRAMME
DEPARTMENT OF STATISTICS

This is to certify that the thesis prepared by Thomas Solomon, entitled: ―*Multivariate Time series Analysis of Ethiopian International Air Travel Demand*" and submitted in partial fulfillment of the requirements for the Degree of Master of Science complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Examiner____________________________________ Signature __________Date _________

Examiner____________________________________ Signature _________ Date _________

Advisor_____________________________________ Signature __________Date _________

____

Chair of Department or Graduate Program Coordinator

ABSTRACT

Multivariate Time series Analysis of Ethiopian International Air Travel Demand

Thomas Solomon,

Addis Ababa 2014

Ethiopian Airline is playing a leading role in transforming Addis Ababa and Ethiopia into a world class aviation hub of the African continent not only for trade and business but also for tourism. For the continual of its success in a better way, forecasting air travel demand plays a crucial role for an overall effective planning. An air travel (passengers‘) demand forecast is a scientifically formed opinion about future air traffic. This demand is most preferably measured by Load Factor which relates the proportion of seats purchased to flight distance covered by a given period.

This study utilized a monthly data from January 2009 up to December 2013 to construct multivariate time series model, vector autoregression (VAR) model, and investigation is made on the reaction of a study target variable Load Factor (LF) to Passenger Revenue (PR), Block Hours (BH), and Distance Flown (DF) at international level.

First and foremost all series are seasonally adjusted after they were known to be seasonal through standard tests built in X-12 ARIMA program in Eviews 7 Statistical Software. Post-seasonal adjustment tests also assured that all series are non-seasonal. Stationarity is checked before and after differencing using visual inspection, unit root tests, and variance comparison. Each series are found to be integrated of order one (I (1)). The three information criteria AIC, SC and HQ recommended one lag length. Johansen cointegration test indicated only one long-term equilibrium relationship occurred between the variables. This immediately implied the legitimacy of vector error correction (VEC) model of order one to be fitted than a pure VAR (1) model for the time series data. Exogeneity test also indicated that only LF and PR are endogenous variables. But to determine the short-run bonds between series, impulse response functions and variance decompositions are employed. Granger causality test is also conducted to identify the total possible causal effects among the variables. As one footstep before out-of-sample forecasting, the VECM (1) model has been checked for its accuracy with the aid of RMSE, MAE, MAPE and Theil-U statistics. The summary result of VECM (1) shown that Load Factor (as a measure of travel demand) is Granger caused by all variables in the short-run except Passenger Revenue and significantly explained by all variables in the long-run. At last, forecasting is made for Ethiopian International air travel demand (Load Factor).

KEYWORDS: Load Factor, Vector Autoregression (VAR), Cointegration, Vector Error Correction model (VECM), and forecasting

ACKNOWLEDGMENT

Principally, I would like to pass my utmost thank, that no one deserves as much as him, to the Omnipotent GOD for all things he did in my life. He is my irreplaceable shepherd all in my ways.

My sincere gratitude goes to my advisor Dr. M.K. Sharma, Professor of Statistics in Addis Ababa University, for his constructive advice in all segments of this study.

My gratitude also goes to all members of the Department of Statistics, academic and non-academic, especially to Professor Eshetu W. and Dr. Shibiru T., for all academic support they offered me for the completion of this study.

In addition, I would like to thank the Ethiopian Airline Head Quarter as an organization and some of its members: Mr. Yared Melaku, Mr. Jeba Tamiru, Mrs. Yayeneabeba Kinfemichael, and Mrs. Saba Deressa including their workmates for their active role in helping me during the collection of my data.

I have an endless and heartfelt thank to all my family for their entirely consistent support throughout this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

ADF	**A**ugmented **D**ickey-**F**uller
AIC	**A**kaike **I**nformation **C**riterion
ARIMA	**A**utoregressive **I**ntegrated **M**oving **A**verage
ASK	**A**vailable **S**eat **K**ilometers
BH	**B**lock **H**ours
DF	**D**istance **F**lown
ECM	**E**rror **C**orrection **M**odel
ETB	**E**thiopian **B**irr
FEVD	**F**orecast **E**rror **V**ariance **D**ecomposition
HQ	**H**annan-**Q**uin Information Criteria
I/C	**I**rregular Component to **T**rend-**C**ycle Component Ratio
IRF	**I**mpulse **R**esponse **F**unction
JB	**J**arque – **B**era Statistic
LF	**L**oad **F**actor
LM	**L**agrange **M**ultiplier
MAE	**M**ean **A**bsolute **E**rror
MCD	**M**onthly **C**yclical **D**ominance
MPAE	**M**ean **P**ercentage **A**bsolute **E**rror
MPE	**M**ean **P**ercentage **E**rror
MSE	**M**ean **S**quare **E**rror
OLS	**O**rdinary **L**east **S**quare
PE	**P**ercentage **E**rror

PP	**P**hillips and **P**erron
PR	**P**assenger **R**evenue
QCD	**Q**uarterly **C**yclical **D**ominance
RMSE	**R**oot **M**ean **S**quared **E**rror
RPK	**R**evenue **P**assengers **K**ilometers
SABL	**S**easonal **A**djustment at **B**ell **L**aboratories
SC	**S**chwarz Information **C**riterion
SEASABS	**SEAS**onal Analysis **A**ustralian **B**ureau of **S**tatistics Standard Program
SI	**S**easonal component to **I**rregular Component Ratio
STAMP	**S**tructural **T**ime Series **A**nalyzer, **M**odeler and **P**redictor
STL	**S**easonal-**T**rend Decomposition Procedure
SVAR	**S**tructural **V**ector **A**uto-**R**egressive
TRAMO/SEATS	**T**ime Series **R**egression with **A**RIMA Noise, **M**issing Observations, and **O**utliers/**S**ignal **E**xtraction in **A**RIMA **T**ime **S**eries
USD	**U**nited **S**tates **D**ollar
VAR	**V**ector **A**uto-**R**egressive
VECM	**V**ector **E**rror **C**orrection **M**odel
X-12ARIMA	Filter Based Seasonal Adjustment Software Package with ARIMA Extension

1. INTRODUCTION

1.1 Background of the Study

Air transportation can be defined as a form of transportation of goods and people from one place to another by using air ways. Air transport is one of the world's largest industries, providing many opportunities to the various sectors of the economy, either to the aircraft manufacturers, airline operators, airport authorities or other related sectors. It is also a highly capital intensive industry. Accordingly, air travel demand can be considered as the customers‘ (air passengers‘) degree of response (utilization) to the total transportation services provided by an airline.

Ethiopian Airline (Ethiopian) is the flag carrier of Ethiopia. It is founded in December 21, 1945 and became operational in April 08, 1946. It is under complete government ownership. During the past sixty five plus years, Ethiopian has become one of the continent‘s leading carriers, unrivalled in Africa for efficiency and operational success, turning profits for almost all the years of its existence. It commands a lion‘s share of the pan African network including the daily and double daily east-west flight across the continent. Ethiopian currently serves 70 international and 17 domestic destinations operating the newest and youngest fleets. Consequently, Addis Ababa is now becoming a pivotal link that is connecting four continents.

In the fiscal years from 2007/08 up to 2011/12, the company has recorded a general upward growth in almost all of its performance categories except for its Load Factor which had shown a closely consistent growth from year to year. In 2008/2009 fiscal year, Ethiopian generated revenue of ETB 12.2 billion, 33% higher than that of the previous year (2007/08). The Available Seat Kilometers (ASK), total Block Hours (BK) flown, and Passenger Revenue (PR) during the year were higher than the previous year (2007/08) by 9%, 10%, and 29%, respectively. These increments were mainly due to additional capacity opening of new passenger destinations and cargo operations increased to various destinations. Ethiopian transported 2.8 million passengers in the year 2008/2009, a growth of 12% over the previous year‘s figure, but 6% less than the forecast. In the same year, a Passenger Load Factor of 70% was maintained. In 2009/10 also Ethiopian has registered a totally increased values in all of its categories and profit which are more elevated than the preceding year (Ethiopian Annual Report of 2007/08 – 2009/10).

The Airline‘s level of operation and operating results in the fiscal year 2010/11 were higher than the previous year (2009/10) in all parameters. Capacity availed in terms of Available Seat Kilometers (ASK) and Block Hours during 2010/11 were higher than the preceding year by about 24% and 17%, respectively. This was mainly due to the commencement of new services to Maputo, Bangui, Hangzhou, and Malakal and availing better capacity brand new aircraft – B777-200LR. Schedule Passenger Revenue realized during the year was higher by 49% compared to the preceding year mainly as a result of increase in passenger traffic. Consequently, the net result of the fiscal year 2010/2011 showed a net profit of ETB 1.23 Billion. This is the highest profit among all the years from 2007/08 up to 2011/2012. For 2011/12 fiscal year also, all parameters have increased as compared to all the previous years, except the year 2010/11 which has the maximum of all periods (Ethiopian Annual Report of 2010/11 and 2011/12).

As far as the main goal of an airline is to harmonize its offer with passenger demand, forecasting air travel demand is an integral part of an airline's corporate plan which facilitates designing of appropriate route network, determining capacity and manpower requirements for the corresponding market conditions, and financial projections for the operating capital projects. It also helps to make decisions regarding the development of infrastructure facilities which include the constructions of new hotels, new roads, new airports, new recreational facilities and so forth, thereby ensures the improvement of services to air passengers. Moreover, it assists to reduce the airline company's risk by objectively evaluating the demand side of the air transport business (Boeing Commercial Airplane Company, 1993). According to Riza O. (2010) also, mistakes made in the forecasting process of an air travel demand can be very costly and damaging for local economies. Underestimating demand can lead to increased congestion, delay, and lack of storage facilities, as it happened in Venezuela in 1974.

Most of the time, an air travel demand is measured by load factor, number of passengers and pay load factor. Load factor can be thought as the percentage of actual air seats purchased from the total seats provided by an airline with respect to the amount of distance flown. Pay load factor has similar notion as load factor but it includes cargos movements in addition to passengers. Sometimes, the exact number of air passengers can also be regarded as a measure of air travel demand, irrespective of the total service (seats) supplied by an airline. So, mostly load factor and pay load factor are preferable to measure an air travel demand in a superior quality than the exact number of air passengers. This is mainly because they are associated with distance flown and total service (seats) provided in the aviation.

Various methods are available, ranging from econometric modeling to time series techniques, for representing an air travel demand. Time series approaches are the most common methods for forecasting the traffic demand (Seraj et al., 2001). A seasonal autoregressive moving average model (SARIMA) which can consider seasonal fluctuations to predict an air travel demand is one time series analysis method that can be effective for short term predictions (about 5 years) for domestic or international flights (Inoue et al., 2008) . Multivariate time series, like VAR, forecasting methods can also be used even though they are seldom applied. As noted by Trani et al. (2003), models of air transportation demand can be divided into Macroscopic models and Microscopic models. Macroscopic models are used to estimate the development of air transportation in a certain country or region. Typical indicators are flights, passengers, cargo, etc. Microscopic models estimate air transportation demand between two cities. Typical indicators are the passenger traffic in a specific Origin-Destination (OD) route and the number of passengers in each class when there are various tariffs on a route.

1.2 Statement of the Problem

In recent years, the study on air travel (passengers') demand of Ethiopian Airline has largely been neglected by researchers and academicians. Those few works which are available are mostly performed by air transport bodies and airlines operators and in most cases they are concerned with only descriptive analysis and simple projection method which does not take other factors of air travel demand in account. They did not apply inferential analysis (especially time series forecasting tool like vector autoregressive (VAR) model) despite its necessity for forecasting of the airline's key variables.

On the other side, planning is unarguably a corner stone for management effectiveness of any organization, especially in profit maximizing ones like Airlines. A good planning requires scientifically produced information pertinent to the future, and hence the need for statistical forecasting. Succinctly, ―*Forecasting is concerned with determining what the future 'will' look like"* and ―*Planning is what the future 'should' look like",* (Armstrong et al., 1995). In fact forecasting plays a central role in the managerial activities of major organizations in the form of reducing decision-making risks, allowing managers to anticipate change, broadening the range of options available and providing greater flexibility and certainty of actions. Therefore, the aim of this study is to fill the application gap of multivariate time series analysis (especially VAR) and forecasting for planning and managerial activities in the Ethiopian Airlines.

1.3 Objective of the Study

1.3.1 General Objective

The main objective of this study is to construct a multivariate time series (particularly VAR) model which is to be used to explain and forecast the monthly Ethiopian Airline‘s International Air Travel Demand (Load Factor) in relation to International Passengers Revenue, International Block Hours, and International Distance Flown.

1.3.2 Specific Objectives

Specifically, this study aims:

- To examine the general trend of monthly Ethiopian Airline‘s International Load Factor.
- To investigate the relationship between monthly International Air Travel Demand (Load Factor) and revenue earned form international passengers (Passenger Revenue), total flight time (Block Hours), and total kilometers flown.
- To forecast International Air Travel Demand (Load Factor) by constructing multivariate time series model constructed.

1.4 Significance of the Study

As a sufficient concept is stated about the strong relationship between forecasting, planning, and managerial activities in the ―Statement of the Problem" sub-section, this study has a great inspirational importance for Ethiopian Airline to harmonize forecasting, planning, and management. The findings in this study are organized in a clear way to portray all the possible types of interactions of the international air travel demand (Load Factor) with other variables. So this is the primary advantage that the study gives to Ethiopian Airline‘s mangers to deeply observe the overall condition about the demand and its factors. Then depending on the results, they can compare their simple projections with multivariate time series forecasting methods, and hence, probably change the mode of forecasting to multivariate time series one if it is found to be good. The other significance is that since researches are seldom produced on Ethiopian Airline demand, both in domestic and international aspect, this study can pave the way for others who have an interest of being engaged in forecasting the airline‘s matters, especially by using multivariate time series techniques. This study also provides guidance to individuals who prepare airport activity forecasts as well as to those who review the forecasts.

1.5 Limitation of the Study

This study is limited to some extent due to inexplicitly recorded data on some variables like jet fuel expenditure and jet fuel price, and number of tourists arriving (tourist flow). The variable delaying time (the time difference between proposed flight starting time and actual starting time) is also excluded from this study because of its total absence.

1.6 Organization of the study

This study is divided into five main chapters. The first chapter contains background, problem, objectives, significance, and limitation of the study. Chapter two gives a review of literature on Air Travel Demand and its forecasting tools. Subsequently, chapter three confers the methodology and sources of data used in the study. Chapter four deals with statistical analysis and interpretation of results. Finally, Chapter five presents conclusions of the study. References, Appendix, and Annex are placed at the end.

2. LITERATURE REVIEW

(Review of Air Travel Demand Models)

A review of the available literatures on air travel demand reveal that the construction of models to estimate demand for air-passengers can take many forms, each depending on the objective of the model being developed. For example the models may range from simple extrapolative techniques of univariate time series models (Oberhausen and Koppelman, 1978 used Box-Jenkin approach of Airline model) to multivariate models which are presently becoming more popular. The multivariate type of models themselves may take varying degrees of complexity.

Ippotito (1981) used a cross-sectional model to estimate the origin and destination demand for airlines (US domestic) at each end of the route that incorporated a measure of service. The results of his study confirmed the long held belief that demand is sensitive to flight frequency and availability of ―excess" seats, and that the quantity of seats offered is positively and significantly affected by regulated price. It also confirmed that the price elasticity of demand increases with flight distance.

Kumar and Stephanedes (1988) studied the impact of air travel supply on demand and vice versa between Twin Cities and Chicago. They used a time series analysis as a tool for estimating the impact of air travel supply on demand and vice-versa on non-stop air routes. The models were developed based on monthly data from sales receipts and schedule information over the 1979 to 1983 period. The objective of this study was to develop a simple yet rigorous model that could be used to forecast intercity demand and supply. Several important discoveries were made in the course of their work. It was found that the passenger load factor affects the availability of tickets thus influencing demand; also the service supplied on a particular route by an airline depends mostly on demand.

Another study by Ghobrial (1992) presents an econometric model that estimates the aggregate demand for an airline. The demand is expressed in terms of airline network structure, operating characteristics and firm-specific variables. A number of model formulations with different combinations of explanatory variables are estimated using the two-stage-least-squares procedure. The results suggest that the airline aggregate demand is elastic with respect to yield, and inelastic with respect to network size and hub dominance.

Poore (1993) has conducted a study to test the hypothesis that forecasts of the future demand for air transportation offered by airplane manufacturers and aviation regulators are reasonable and representative of the trends implicit in actual experience. The test compared forecasts issued by Boeing, McDonnell Douglas, Airbus Industry and the International Civil Aviation Organization with actual experience and the results of a baseline model of the demand for revenue passenger kilometers (RPKs). The model is the combination of two equations describing the RPKs demanded by the high- and the low income groups, respectively. While variations in the RPKs demanded by the high-income group are related to changes in income per capita, variations in the RPKs demanded by the low-income segment are related to changes in the population size. The model conforms to the assumptions and conditions for appropriate use of regression analysis.

Seraj et al. (2001) developed several models for the air travel demand with different combinations of fourteen explanatory variables utilizing stepwise regression technique. Among all candidate lines of models, the model least square line with the two variables (i.e., total expenditures and population size) for international air travel demand models in Saudi Arabia; $\mathrm{Y} = -2.2566 + 0.021314\ \mathrm{PS} + 0.39522\ \mathrm{TE}$ was the most appropriate model to represent the demand for international air travel in Saudi Arabia, where PS and TE, with t-statistics 10.975 and 9.437 stand for population size and total expenditure, respectively. The model‘s adjusted R^2 was 0.959. By looking into their results of regression analysis, they found that the population size and total expenditures, collectively and individually, influence international air travel demand of Saudi Arabia in a positive way. They concluded that as the population size goes up by one percentage point, on average, demand for international air travel goes up by 0.021 million passengers, other things kept fixed. Likewise, if the total expenditure goes up by one percentage point, on average, demand for international air travel goes up by 0.395 million passengers, other things kept fixed.

Bahram et al. (2002) employed vector error correction model (VECM) to examine the long-run dynamic relationship between air carrier firms‘ capacity (CAP) measured by the available seat miles and their profits (PRF) for US airline industry. They used quarterly observations on nine US carriers (Alaska, American, America West, Continental, Delta, Northwest, Southwest, Trans World and United) from quarter 3 of 1983 to quarter 3 of 1998. The objective was to investigate whether CAP is related to profitability. If size and profits demonstrate a positive long-run relationship, then there is a market incentive for carrier acquisitions and a move toward an oligopoly market structure. Their result showed that there is a long-run positive relationship between the available seat miles and profits. Granger causality tests also verified that for all carriers the capacity variable Granger causes profits.

Kulendran and Witt (2003) generated one, four and six quarter ahead forecasts of international business passengers to Australia from the following four countries: Japan, New Zealand, the United Kingdom and the United States. They considered various forecasting models: the error correction model (ECM), the structural time series model (STSM), the basic structural model (BSM), and ARIMA model. They concluded that forecasting performance varies with the forecasting horizon and depends on the adequate detection of seasonal unit roots. Consequently, ARIMA and BSM models are the most accurate for short term forecasting (one-quarter ahead) whereas ECM outperforms for medium term forecasting (four and six quarters ahead).

Another study for air travel demand forecasting is done by Grosche et al. (2007). According to their research, there are some variables that can affect the air travel demand, including population, GDP and buying power index. They considered GDP as a representative variable for the level of economic activity.

Tsekeri (2009) has also estimated the short and long-term response of air passengers to change in relative air-sea travel cost components in competitive markets using a dynamic demand model. The model demonstrated the importance of considering the past volumes of air passengers and relative travel cost components to explain current air travel demand.

Constantinos (2013) examined whether or not combining forecasts from autoregressive-integrated-moving average (ARIMA) and seasonal autoregressive-integrated-moving average (SARIMA) models helps to improve forecasting accuracy of Canadian air transportation sector in domestic, transborder (US) and international flights. His study also provided forecasts of air passengers in Canada based the following various time series forecasting models. For domestic flight he fitted ARIMA (3, 1, 2) and SARIMA (0, 1, 1) $(2, 0, 0)_{12}$. His models for transborder flight were also ARIMA (0, 1, 4) and SARIMA (2, 1, 1) $(1, 0, 1)_{12}$. Lastly, ARIMA (0, 1, 4) and SARIMA (2, 1, 1) $(1, 0, 1)_{12}$ were constructed for international flight. According to him, results indicated that all models provide accurate forecasts, with MAPE and RMSPE scores below 10%

on average and residuals did not suffer from autocorrelation problem at 5% significance level for all models, with the exception of SARIMA model for the domestic sector. So from all his findings, he deduced that Canadian government, air transport authorities, and the airlines operating in Canada to use combination techniques (of the above models) to improve their short and medium term forecasts of passenger flows.

3. STATISTICAL DATA AND METHODOLOGY

3.1 Data Source

This study considered a monthly Ethiopian airline‘s data for international flights over the time period January 2009 – December 2013 which are obtained from the Head Quarter Office of Ethiopian Airline located in Bole sub-city, Addis Ababa, Ethiopia.

3.2 Definition and Variables of the Study

The incorporated variables are somewhat technical and need a brief description as follows.

(1) *International Load Factor (LF)*: is the target variable of the study, which describes the percentage of actual air seats purchased out of the total (available) seats provided per month by an airline for international flight. Mathematically, it is expressed as

$$\frac{\text{Number of Passengers}}{\text{Aircraft Configured Seat Number}}$$

(2) *International Passenger Revenue (PR)*: is the monthly aggregate revenue that would be earned from each individual international flight in millions of USD.

(3) *International Block Hours (BH)*: describes monthly flight duration and is a summation of each individual international flight‘s time difference between engine on and engine off. It is measured by thousands of hours.

(4) *International Distance Flown (DF)*: is a monthly distance covered by all international flights. It is cumulative kilometers flown by each international flight. The study measures this variable in millions of kilometers.

3.3 Statistical Methodology

Time series is a set of regular time-ordered observations of quantitative characteristics which can be individual or collective phenomena taken at successive, in most cases equidistant, points of time. It can be divided in to two major parts – univariate and multivariate time series. Univariate time series analysis is one which uses only one endogenous (dependent), one or more other exogenous explanatory time series variables, plus current and past random error terms. Autoregressive integrated moving average (ARIMA) and ARIMA with exogenous variables (ARIMAX) models are specific subsets of univariate models. ARIMA is a model in which a time series is expressed in terms of only past values of itself (the autoregressive component) plus current and lagged values of a ‗white noise' error term (the moving average component), whereas ARIMAX is ARIMA plus one or more other exogenous explanatory time series variables. On the other hand, multivariate time series analysis is an extension of the univariate one and is used when one wants to model and explain the interactions and co-movements among two or more endogenous (dependent) and one or more other exogenous explanatory time series variables. Vector autoregressive (VAR) models are particular lineage of multivariate models in which each time series variable can be treated symmetrically and has an equation explaining its evolution based on its own lags and the lags of other endogenous and/or exogenous model variables. This study begins with seasonality concept and makes its concern on modeling of multivariate time series data. The method used in this study can be divided into two broad sections. The first section is concerned with the Vector Autoregressive (VAR) models for stationary and cointegrated variable(s). In this section model specification and parameter estimation are discussed. The other section deals with Structural Vector Autoregressive (SVAR) Analysis (i.e., Granger Causality, Impulse Response Functions (IRF), and Forecast Error Variance Decomposition (FEVD)).

3.3.1 Seasonality

3.3.1.1 An Overview of Time Series Components

A time series can be decomposed into four main unobserved components:

- *Trend (T)*: indicates the long-term tendency (pattern) in a time series.
- *Cyclical component (C)*: indicates the medium term fluctuation of a time series.
- *Seasonal component (S)*: represents intra-year fluctuations more or less stable year after year with respect to timing, direction and magnitude.
- *Irregular component (I)*: includes unpredictable effects, which are considered as random variables.

3.3.1.2 Seasonal Adjustment

Seasonal Adjustment is the process of estimating and removing the seasonal effects from a time series in order to better reveal certain non-seasonal features. A seasonally adjusted series is the combination of the underlying trend of the series and the irregular factors. Currently, software seasonal adjustment techniques like an X12-ARIMA program are becoming popular and applicable. An X12-ARIMA program contains built-in statistical tests for detecting seasonality and quality measures of seasonal adjustment. This program can be obtained in Eviews 7 and some other statistical software. Detail description on X-12 ARIMA, its tests, and quality control diagnostics can be referred from Annex.

3.3.2. Vector Autoregressive (VAR) Models

The Vector Autoregression (VAR) model, proposed by Sims (1980), is one of the most successful, flexible, and easy to use models for analysis of multivariate time series. It is applied to grasp the mutual influence among multiple time series. VAR model extends the univariate autoregressive (AR) model to dynamic multivariate time series by allowing for more than one evolving variable. It has verified valuable features for describing the dynamic behavior of economic and financial time series and for forecasting. Forecasts from VAR models are quite flexible because they can be made conditional on the potential future paths of specified variables in the model. VAR model often provides superior forecasts to those from univariate time series models and elaborate theory-based simultaneous equations (models). That is, all variables in a VAR model are treated symmetrically in a structural sense; each variable has an equation explaining its evolution based on its own lags and the lags of the other model variables.

Besides to forecasting, the VAR model is also used for structural inference and policy analysis. In structural analysis, certain assumptions about the causal structure of the data under investigation are imposed, and the resulting causal impacts of unexpected shocks or innovations to specified variables on the variables in the model are summarized. These causal impacts are usually summarized with impulse response functions and forecast error variance decompositions.

3.3.3 Stationarity

Time series stationarity is an important point to be described in time series analysis. It is the statistical characteristics of a series such as its mean and auto-covariance over time. If both are constant over time, then the series is said to be a stationary process (i.e. is not a random walk/has no unit root), otherwise, the series is described as being a non-stationary process (i.e. a random walk/has unit root). Sometimes, seasonality can also account for non-stationarity of a series.

A stochastic process (vector) $\{\boldsymbol{Y}_t\}$ can be broadly classified as weak (covariance) or strong (strictly) stationary process.

3.3.3.1 Covariance (Weakly) Stationarity

$\{\boldsymbol{Y}_t\}$ is said to be weakly (covariance) stationary if the first and second moments are time invariant, i.e., stochastic process (vector) $\{\boldsymbol{Y}_t\}$ has

- a mean vector that is independent of time, $E(\boldsymbol{Y}_t) = E(\boldsymbol{Y}_s) = \boldsymbol{\mu} \ \ \forall_{t \neq s}$, and
- covariance matrices $Cov(\boldsymbol{Y}_t, \boldsymbol{Y}_s) = Cov(\boldsymbol{Y}_{t+j}, \boldsymbol{Y}_{s+j}) = \boldsymbol{\Gamma}_{t,s} \ \ \forall_{t,s,j \geq 1}$ which are independent of time. This is, the covariance matrices do not depend on the magnitude of time t and time s, instead depend on the distance between t and s.

3.3.3.2 Strong (Strictly) Stationarity

Strong stationary process is one for which the whole distribution of the variable(s) does (do) not depend on time. Formally, $\{\boldsymbol{Y}_t\}$ is strictly stationary if the probabilistic behavior of every collection of values $\{Y_{t_1}, Y_{t_2}, \dots, Y_{t_i}\}$ is identical to that of the time shifted set $\{Y_{t_{1+j}}, Y_{t_{2+j}}, \dots, Y_{t_{i+j}}\}$. Meaning,

$$P\left[Y_{t_1} \leq c_1, Y_{t_2} \leq c_2, \dots, Y_{t_i} \leq c_i\right] = P\left[Y_{t_{1+j}} \leq c_1, Y_{t_{2+j}} \leq c_2, \dots, Y_{t_{i+j}} \leq c_i\right] for\ j \epsilon [0, \pm 1, \pm 2, \dots].$$

If a time series is strictly stationary, then all of the multivariate distribution functions for subsets of variables must agree with their counterparts in the shifted set for all values of the shift parameter i.

3.3.3.3 Differencing

Differencing techniques are the most commonly employed techniques to transform a time series from a non-stationary to stationary by subtracting each datum (Y_t) in a series from its predecessor (Y_{t-1}). Hence, differencing turns out to be a useful ‗filtering' procedure in the study of non-stationary time series. The set of observations (Y_t's) that correspond to the initial time period (t) when the measurement was taken is described as a series at level.

Using the difference operator Δ, the first difference is defined by

$$\Delta Y_t = Y_t - Y_{t-1} \,. \tag{3.1}$$

(3.1) can also be written using back-ward shift operator B as

$$\Delta Y_t = (1 - B)Y_t \quad , \tag{3.2}$$

where $B^k Y_t = Y_{t-k}$ for $k = 0, 1, 2, \dots$.

In general, for n-order differencing (3.2) will be given as

$$\Delta^n Y_t = (1 - B)^n Y_t \;. \tag{3.3}$$

3.3.3.4 Integration (I (d))

When a time series is not stationary, differencing a series using differencing operations produces other sets of observations such as the first-differenced values, the second-differenced values and so on. If a series is stationary without any differencing it is designated as I (0), or integrated of order zero. On the other hand, a series that has stationary first difference is designated I (1), or integrated of order one. As a whole, if a non-stationary time series has to be differenced *d* times to make it stationary, that time series is said to be integrated of order *d* and denoted as I (*d*) (Gujarati, 2004; Pole et al., 1994; Weigend et al., 1993).

3.3.4 Stationary Vector Autoregressive Model

Let $\boldsymbol{Y_t} = (Y_{1t}, Y_{2t}, \dots, Y_{nt})'$ denote an $(n \times 1)$ vector of time series variables. A VAR model with *p* lags can then be expressed as follows:

$$\boldsymbol{Y}_t = \boldsymbol{c} + \boldsymbol{\Pi}_1 \boldsymbol{Y}_{t-1} + \boldsymbol{\Pi}_2 \boldsymbol{Y}_{t-2} + \cdots + \boldsymbol{\Pi}_p \boldsymbol{Y}_{t-p} + \boldsymbol{\varepsilon}_t, \quad t = 1, \dots, T \,, \tag{3.4}$$

where $\boldsymbol{c}$ denotes an $(n \times 1)$ vector of constants and $\boldsymbol{\Pi}_i$, for $i = 1, 2, \ldots, p$, is an $(n \times n)$ coefficient matrix of autoregressive coefficients. $\boldsymbol{\varepsilon}_t$ is an $(n \times 1)$ unobservable zero mean white noise vector process (serially uncorrelated) with time invariant covariance matrix $\boldsymbol{\Sigma}$. i.e.,

$$E(\boldsymbol{\varepsilon}_t) = \mathbf{0} \qquad \text{and} \qquad Cov(\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_s) = E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s') = \begin{cases} \boldsymbol{\Sigma}, & \forall\, t = S \\ \mathbf{0}, & \forall\, t \neq S \end{cases} \tag{3.5}$$

with $\boldsymbol{\Sigma}$ an $(n \times n)$ symmetric positive definite matrix.

Let c_i denote the i^{th} element of the vector $\boldsymbol{c}$ and let $\Pi_{ij}^{(1)}$ denote the element on i^{th} row, j^{th} column of the matrix $\boldsymbol{\Pi}_1$.Then the first row of the vector system in (3.4) specifies that

$$\begin{aligned} Y_{1t} = c_1 + \Pi_{11}^{(1)} Y_{1,t-1} + \Pi_{12}^{(1)} Y_{2,t-1} + \cdots + \Pi_{1n}^{(1)} Y_{n,t-1} + \Pi_{11}^{(2)} Y_{1,t-2} + \Pi_{12}^{(2)} Y_{2,t-2} + \cdots + \\ \Pi_{1n}^{(2)} Y_{n,t-2} + \cdots + \Pi_{11}^{(p)} Y_{1,t-p} + \Pi_{12}^{(p)} Y_{1,t-p} + \cdots + \Pi_{1n}^{(p)} Y_{n,t-p} + \varepsilon_{1t} \end{aligned} \tag{3.6}$$

$Y_{2t}, \ldots, Y_{nt}$ can also be written in the same manner as Y_{1t} . For instance, specifically, a bivariate VAR (2) model equation has the form

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \Pi_{11}^{(1)} & \Pi_{12}^{(1)} \\ \Pi_{21}^{(1)} & \Pi_{22}^{(1)} \end{pmatrix} \begin{pmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \Pi_{11}^{(2)} & \Pi_{12}^{(2)} \\ \Pi_{21}^{(2)} & \Pi_{22}^{(2)} \end{pmatrix} \begin{pmatrix} Y_{1,t-2} \\ Y_{2,t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}. \tag{3.7}$$

Or explicitly,

$$Y_{1t} = c_1 + \Pi_{11}^{(1)} Y_{1,t-1} + \Pi_{12}^{(1)} Y_{2,t-1} + \Pi_{11}^{(2)} Y_{1,t-2} + \Pi_{12}^{(2)} Y_{2,t-2} + \varepsilon_{1t} \tag{3.8}$$

and

$$Y_{2t} = c_2 + \Pi_{21}^{(1)} Y_{1,t-1} + \Pi_{22}^{(1)} Y_{2,t-1} + \Pi_{21}^{(2)} Y_{1,t-2} + \Pi_{22}^{(2)} Y_{2,t-2} + \varepsilon_{2t} \quad , \tag{3.9}$$

where $Cov(\varepsilon_{1t}, \varepsilon_{2s}) = \sigma_{12}$ for $t = s; 0$ otherwise. Notice that each equation has the same regressors – lagged values of Y_{1t} and Y_{2t}.

Thus, a vector autoregression is a system in which each variable is regressed on a constant and p of its own lags as well as on p lags of each of the other variables in the VAR. Note that each regression has the same explanatory variables.

In lag operator notation, the VAR (p) is written as

$$\boldsymbol{\Pi}(L)\boldsymbol{Y}_t = \boldsymbol{c} + \boldsymbol{\varepsilon}_t \quad , \tag{3.10}$$

where $\boldsymbol{\Pi}(L) = \boldsymbol{I}_n - \boldsymbol{\Pi}_1 L - \cdots - \boldsymbol{\Pi}_p L_p$. The VAR ($p$) is stable if the roots of

$$det[\boldsymbol{\Pi}(L)] = det(\boldsymbol{I}_n - \boldsymbol{\Pi}_1 L - \cdots - \boldsymbol{\Pi}_p L_p) = 0 \qquad (3.11)$$

lie outside the complex unit circle (have modulus greater than one), or, equivalently, if the eigen-values of the companion matrix

$$\boldsymbol{F} = \begin{pmatrix} \boldsymbol{\Pi}_1 & \boldsymbol{\Pi}_2 & \cdots & \boldsymbol{\Pi}_n \\ \boldsymbol{I}_n & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \ddots & \boldsymbol{0} & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_n & \boldsymbol{0} \end{pmatrix}$$

have modulus less than one. Assuming that the process has been initialized in the infinite past, then a stable VAR (p) process is stationary and *ergodic* (i.e., if sample mean, sample autocovariance, and sample autocorrelation converge in probability to their respective population moments) with time invariant means, variances, and autocovariances.

If $\boldsymbol{Y}_t$ in (3.4) is covariance stationary, then the unconditional mean is given by

$$\boldsymbol{\mu} = \left(\boldsymbol{I}_n - \boldsymbol{\Pi}_1 - \cdots - \boldsymbol{\Pi}_p\right)^{-1}\boldsymbol{c}. \qquad (3.12)$$

The mean-adjusted form of the VAR (p) is then

$$\boldsymbol{Y}_t - \boldsymbol{\mu} = \boldsymbol{\Pi}_1(\boldsymbol{Y}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\Pi}_2(\boldsymbol{Y}_{t-2} - \boldsymbol{\mu}) + \cdots + \boldsymbol{\Pi}_p\left(\boldsymbol{Y}_{t-p} - \boldsymbol{\mu}\right) + \boldsymbol{\varepsilon}_t. \qquad (3.13)$$

The basic VAR (p) model may be too restrictive to represent sufficiently the main characteristics of the data. In particular, other deterministic terms (for instance, a linear time trend) and stochastic exogenous variables may be required to represent the data properly. The general form of the VAR (p) model with deterministic terms and exogenous variables is given by

$$\boldsymbol{Y}_t = \boldsymbol{\Pi}_1\boldsymbol{Y}_{t-1} + \boldsymbol{\Pi}_2\boldsymbol{Y}_{t-2} + \cdots + \boldsymbol{\Pi}_p\boldsymbol{Y}_{t-p} + \boldsymbol{\Phi}\boldsymbol{D}_t + \boldsymbol{G}\boldsymbol{X}_t + \boldsymbol{\varepsilon}_t\ , \qquad (3.14)$$

where $\boldsymbol{D}_t$ represents an ($l \times 1$) matrix of deterministic components, $\boldsymbol{X}_t$ represents an ($m \times 1$) matrix of exogenous variables, and $\boldsymbol{\Phi}$ and $\boldsymbol{G}$ are parameter matrices.

3.3.5 Testing Stationarity

Stationarity of a time series is an important phenomenon because it can influence the behavior of the series unless it is detected with standard tests (unit root tests) and properly handled. In non-stationary series, the effect of a shock never dies away and it leads to spurious regressions (i.e., one can regress completely unrelated series then find inflated t-ratio which suggests whether a coefficient of one variable is significant or not to explain the other and high R^2 which indicates how good one term is at predicting another) and forged results of standard tests. In order to ensure the condition of stationarity, a series must to be integrated of order of zero (I (0)).

3.3.5.1 Visual Inspection

The opening stride in the analysis of time series is usually to plot the data and obtain simple descriptive measures of the main property of the series via a visual inspection of the time series plot. This may reveal one or more of the following characteristics: seasonality, trends either in the mean level or the variance of the series, long- term cycles, and so on. If any such patterns are present, then these are signs of non-stationarity.

3.3.5.2 Unit-Root Test

The development of unit root theory, initially proposed by Dickey and Fuller (1979, 1981), has spawned a generation of unit root research. Unit root theory is the cornerstone to the methodology used for testing the stationarity or non-stationarity of a time series. Nowadays, many of the procedures are standard offerings in econometric software packages like Eviews 7, and they have become routine tools for time series analysts. This study applied the most commonly used and unarguably powerful unit root tests: Augmented Dickey- Fuller (ADF) test due to Dickey and Fuller (1979, 1981), and the Phillip-Perron (PP) test due to Phillips (1986) and Phillips and Perron (1988). These test procedures are developed for models with and without intercept terms as well as trend terms. The following discussion outlines the basic features of unit root tests (Hamilton, 1994).

Consider an AR (1) process

$$Y_t = \rho Y_{t-1} + X_t^{'}\delta + \varepsilon_t \quad , \tag{3.15}$$

where X_t are optional exogenous regressors which may consist of constant or a constant and trend, ρ and δ are parameters to be estimated and ε_t is assumed to be white noise.

If $|\rho| \geq 1$, Y is a non-stationary series and the variance of Y increases with time and approaches infinity. On the other hand, if, $|\rho| < 1$, Y is a stationary series. Thus, the hypothesis of (trend) stationarity can be evaluated by testing whether the absolute value of ρ is strictly less than one.

The hypotheses are

H_0: The series are not stationary ($|\rho| = 1$) versus

H_A: The series are stationary($|\rho| < 1$).

3.3.5.2.1 Augmented Dickey-Fuller (ADF) Unit-Root Test

The standard Dickey-Fuller test is conducted by estimating equation (3.15) after subtracting Y_{t-1} from both side of the equation as follows.

$$Y_t - Y_{t-1} = \rho Y_{t-1} - Y_{t-1} + X_t^{'}\delta + \varepsilon_t$$

$$\Rightarrow \Delta Y_t = (\rho - 1)Y_{t-1} + X_t^{'}\delta + \varepsilon_t$$, which in turn can be expressed as

$$\Delta Y_t = \alpha Y_{t-1} + X_t^{'}\delta + \varepsilon_t \quad , \qquad (3.16)$$

where $\alpha = \rho - 1$.

The null and alternative hypothesis can then be stated as

$H_0: \alpha = 0$ against

$H_A: \alpha < 0.$ (3.17)

The test statistic is the conventional t-ratio for α:

$$t_\alpha = \frac{\hat{\alpha}}{(SE(\hat{\alpha}))} \qquad (3.18)$$

where $\hat{\alpha}$ is the estimate of α and $SE(\hat{\alpha})$ is the standard error of $\hat{\alpha}$.

Dickey and Fuller (1979) shown that, under the null hypothesis of a unit root, this statistic does not follow the conventional Student's t-distribution, and they derived asymptotic results and simulated critical values for various tests and sample sizes. MacKinnon (1991, 1996) implemented a much larger set of simulations than those tabulated by Dickey and Fuller. In addition, MacKinnon estimated response surfaces for the simulation results, permitting the calculation of Dickey-Fuller critical values and p-values for arbitrary sample sizes.

The simple Dickey-Fuller unit root test described above is valid only if the series is an AR (1) process. If the series is correlated at higher order lags, the assumption of white noise disturbances $\boldsymbol{\varepsilon}_t$ is violated. The Augmented Dickey-Fuller (ADF) test constructs a parametric correction for higher-order correlation by assuming that the series follows an AR(p) process and adding lagged difference terms of the dependent variable Y to the left-hand side of the test regression:

$$\Delta Y_t = \alpha Y_{t-1} + X_t^{'}\delta + \beta_1 \Delta Y_{t-1} + \beta_2 \Delta Y_{t-2} + \cdots + \beta_p \Delta Y_{t-p} + U_t \qquad (3.19)$$

This augmented specification is then used to test for unit root using the t-ratio in (3.18). An important result obtained by Fuller (1979) is that the asymptotic distribution of the t -ratio for α is independent of the number of lagged first differences included in the ADF regression. Moreover, while the assumption that Y follows an AR process may seem restrictive, Said and Dickey (1984) demonstrate that the ADF test is asymptotically valid in the presence of a moving average component, provided that sufficient lagged difference terms are included in the test regression.

3.3.5.2.2 Phillips-Perron (PP) Unit-Root Test

Phillips and Perron (1988) propose an alternative (nonparametric) method of controlling serial correlation when testing for a unit root. The PP method estimates the non-augmented DF test equation in (3.17), and modifies the t -ratio of the α coefficient so that serial correlation does not affect the asymptotic distribution of the test statistic. The PP test is based on the statistic

$$\hat{t}_\alpha = t_\alpha \left(\frac{\gamma_0}{f_0}\right)^{1/2} - \frac{T(f_0-\gamma_0)(SE(\hat{\alpha}))}{2f_0^{1/2}s} \quad , \qquad (3.20)$$

where $\hat{\alpha}$ is the OLS estimate of α, t_{α} is the t -ratio of α, $SE(\hat{\alpha})$ is coefficient standard error and S is the standard error of the test regression. In addition, γ_0 is a consistent estimate of the error variance in (3.18) (calculated as $(T-k)\,S^2/T$, where k is the number of regressors). The remaining term, f_0 , is an estimator of the residual spectrum at frequency zero.

3.3.6 Specification of VAR Order

An important aspect of empirical research based on the vector autoregressive (VAR) model is the choice of the lag order, since validity of all inferences in this model depends on the appropriate selection of the lag order and correct model specification. Lag length should not be different from the true VAR length as it will definitely distort all end results. Selecting a very small order may lead to ignoring interesting dynamics of the economic variables. Selecting a very large order leads to inefficiency in estimation, which translates into large coefficients, standard errors and large confidence bands for the impulse response functions (that will be discussed later). In a VAR each variable is regressed against the same number of lags of every variable. In other words, VAR lag specifications are symmetrical. The lag length for the VAR model can be determined using model selection criteria. The general approach is to fit VAR models with orders $m = 0, \ldots, p_{max}$ and choose the value of m which minimizes some model selection criteria (Lutkepohl, 1991). That is, the best model is the one with a lag at which a minimum selection criterion.

The general model selection criteria have the form:

$$IC(p) = ln|\overline{\mathbf{\Sigma}}(p)| + c_T.\varphi(n,p) \ , \tag{3.21}$$

where $\overline{\mathbf{\Sigma}}(p) = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t^{'}$ is the residual covariance matrix without a degrees of freedom correction from a VAR(p) model, c_T is a sequence indexed by the sample size T, and $\varphi(n,p)$ is a penalty function which penalizes large VAR(p) models. In this paper, three most common information criteria are considered to pick out an optimal lag order for the VAR that will be fitted. These are Akaike (AIC), Schwarz (SC), and Hannan-Quinn (HQ):

$$AIC(p) = ln|\overline{\mathbf{\Sigma}}(p)| + \frac{2}{T}pn^2 \tag{3.22}$$

$$SC(p) = ln|\overline{\mathbf{\Sigma}}(p)| + \frac{lnT}{T}pn^2 \tag{3.23}$$

$$HQ(p) = ln|\overline{\boldsymbol{\Sigma}}(p)| + \frac{2\ln(lnT)}{T}pn^2 \quad (3.24)$$

The AIC criterion asymptotically overestimates the order with positive probability, whereas the SC and HQ criteria estimate the order consistently under fairly general conditions if the true order p is less than or equal to p_{max} .

3.3.7 Cointegration Analysis

Engle and Granger (1987) developed the theory that there exists the special case where linear combinations of nonstationary processes are stationary. They defined this linear combination of nonstationary processes as *cointegration* and used the notation CI (*d*, *b*) , where *d* represents the order of integration of the nonstationary processes and *b* represents the number of stationary linear combinations between the nonstationary processes. For instance, consider two I(1) processes, X_{1t} and X_{2t}. If there exists a linear combination of the two processes such that the linear combination is I (0), the two I (1) processes are considered to be CI (1, 1). Broadly, contegrating relationships can be either single or multiple as follows.

3.3.7.1 Single Cointegration Relationship

Let $\boldsymbol{Y}_t = (Y_{1t}, Y_{2t}, \dots, Y_{nt})'$ denote an (n×1) vector of I(1) time series. $\boldsymbol{Y}_t$ is said to be cointegrated if there exists an (n×1) vector $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_n)'$ such that

$$\boldsymbol{\beta}'\boldsymbol{Y}_t = \beta_1 Y_{1t} + \beta_2 Y_{2t} + \cdots + \beta_n Y_{nt} \sim I(0) \quad . \quad (3.25)$$

In words, the nonstationary time series in $\boldsymbol{Y}_t$ are cointegrated if there is a linear combination of them that is stationary or I(0). If some elements of $\boldsymbol{\beta}$ are equal to zero then only the subset of the time series in $\boldsymbol{Y}_t$ with non-zero coefficients is cointegrated. The linear combination $\boldsymbol{\beta}'\boldsymbol{Y}_t$ is often motivated by economic theory and referred to as a *long-run equilibrium relationship*. The intuition is that I (1) time series with a long-run equilibrium relationship cannot drift too far apart from the equilibrium because economic forces will act to restore the equilibrium relationship.

Normalization

The cointegration vector $\boldsymbol{\beta}$ in (3.25) is not unique since for any scalar c the linear combination $c\boldsymbol{\beta}'\boldsymbol{Y}_t = \boldsymbol{\beta}^{*'}\boldsymbol{Y}_t \sim I(0)$. Hence, some normalization assumption is required to uniquely identify $\boldsymbol{\beta}$.

A typical normalization is

$$\boldsymbol{\beta} = (1, -\beta_2, \dots, -\beta_n)' \ .$$

So that the cointegration relationship may be expressed as

$$\boldsymbol{\beta}'\boldsymbol{Y}_t = Y_{1t} - \beta_2 Y_{2t} - \cdots - \beta_n Y_{nt} \sim I(0)$$

or

$$Y_{1t} = \beta_2 Y_{2t} + \cdots + \beta_n Y_{nt} + U_t \ , \tag{3.26}$$

where $U_t \sim I(0)$. In (3.26), the error term U_t is often referred to as the *disequilibrium error* or the *cointegrating residual*. In long-run equilibrium, the disequilibrium error U_t is zero and the long-run equilibrium relationship is

$$Y_{1t} = \beta_2 Y_{2t} + \cdots + \beta_n Y_{nt} \ . \tag{3.27}$$

3.3.7.2 Multiple Cointegration Relationships

If the $(n \times 1)$ vector $\boldsymbol{Y}_t$ is cointegrated, there may be $r, 0 < r < n,$ linearly independent cointegrating vectors. For example, let $n = 3$ and suppose there are $r = 2$ cointegrating vectors $\boldsymbol{\beta}_1 = (\beta_{11}, \beta_{12}, \beta_{13})'$ and $\boldsymbol{\beta}_2 = (\beta_{21}, \beta_{22}, \beta_{23})'$. Then $\boldsymbol{\beta}_1'\boldsymbol{Y}_t = \beta_{11}Y_{1t} + \beta_{12}Y_{2t} + \beta_{13}Y_{3t} \sim I(0)$, $\boldsymbol{\beta}_2'\boldsymbol{Y}_t = \beta_{21}Y_{1t} + \beta_{22}Y_{2t} + \beta_{23}Y_{3t} \sim I(0)$ and the (2×3) matrix

$$\boldsymbol{B}' = \begin{pmatrix} \boldsymbol{\beta}_1' \\ \boldsymbol{\beta}_2' \end{pmatrix} = \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \end{pmatrix} \tag{3.28}$$

forms a basis for the space of cointegrating vectors. The linearly independent vectors $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ in the cointegrating basis $\boldsymbol{B}$ are not unique unless some normalization assumptions are made. Furthermore, any linear combination of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_1$, e.g. $\boldsymbol{\beta}_3 = c_1\boldsymbol{\beta}_1 + c_2\boldsymbol{\beta}_2$ where c_1 and c_1 are constants, is also a cointegrating vector.

As a concluding remark, if there is a set of k integrated variables of order one (I(1)), there may exist up to $k - 1$ independent linear relationships that are I(0). In addition, there can be $r \leq k - 1$ linearly independent cointegrating vectors, which are gathered together into the

$k \times r$ cointegrating matrix. Thus, each element in the r-dimensional vector is I(0), while each element in the k-dimensional vector is I(1) (Engle and Granger, 1987).

3.3.7.3 Testing for Cointegration Using Johansen's Methodology

When there is a significant cointegrating vector, the VAR model should be augmented with an Error Correction term. In other words, pure VAR can be applied only when there is no cointegrating relationship among the variables in the VAR system. Hence, a prerequisite before running any VAR model is to run a cointegration test.

In this study the Johansson procedure is applied. Johansen‘s (1991) procedure considers maximum likelihood for finite-order vector auto regressions (VARs) and is easily calculated for such systems. Johansson‘s procedure allows dealing with models with several endogenous variables. The procedure begins with unrestricted VAR involving potentially non-stationary variables.

The starting point of Johansen‘s procedure (1988, 1991) in determining the number of cointegrating vectors is the VAR representation of $\boldsymbol{Y}_t$. It is assumed a vector autoregressive model of order p and is expressed as follows in (3.29).

$$\boldsymbol{Y}_t = \boldsymbol{A}_1\boldsymbol{Y}_{t-1} + \boldsymbol{A}_2\boldsymbol{Y}_{t-2} + \cdots + \boldsymbol{A}_p\boldsymbol{Y}_{t-p} + \boldsymbol{B}\boldsymbol{X}_t + \boldsymbol{\varepsilon}_t \quad , \tag{3.29}$$

where $\boldsymbol{Y}_t$ is an *n*-vector of non-stationary I(1) variables (i.e., the non-stationary series variables in $\boldsymbol{Y}_t$ are differenced once to achieve stationarity, then Y_t is said to be integrated of order one). This would be then written as $\boldsymbol{Y}_t \sim$ I(1)), $\boldsymbol{X}_t$ is a d-vector of deterministic (other exogenous) variables, and $\boldsymbol{\varepsilon}_t$ is a vector of innovations.

(3.29) can be re-written as:

$$\Delta\boldsymbol{Y}_t = \boldsymbol{\Pi}\boldsymbol{Y}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta\boldsymbol{Y}_{t-i} + \boldsymbol{B}\boldsymbol{X}_t + \boldsymbol{\varepsilon}_t \quad , \tag{3.30}$$

where

$$\boldsymbol{\Pi} = \sum_{i=1}^{p} \boldsymbol{A}_i - \boldsymbol{I}, \quad \boldsymbol{\Gamma}_i = -\sum_{j=i+1}^{p} \boldsymbol{A}_j \quad . \tag{3.31}$$

Granger's representation theorem asserts that if the coefficient matrix $\boldsymbol{\Pi}$ has reduced rank $r < n$, then there exist $n \times r$ matrices α and β each with rank r such that $\boldsymbol{\Pi} = \boldsymbol{\alpha\beta}'$ and $\boldsymbol{\beta}'\boldsymbol{Y_t}$ is I(0), where r is the number of cointegrating relations (the *cointegrating rank*) and each column of it represent the cointegrating vector. The elements of α are known as the *adjustment parameters* in the VEC model. It can be shown that for a given r, the maximum likelihood estimator of $\boldsymbol{\beta}$ defines the combination of $\boldsymbol{Y}_{t-1}$ that yields the r largest canonical correlations of $\Delta\boldsymbol{Y}_t$ with $\boldsymbol{Y}_{t-1}$ after correcting for lagged differences and deterministic variables when present.

Johansen (1988) proposed two tests for estimating the number of cointegrating vectors: the Trace statistic and Maximum Eigenvalue statistic tests. Trace statistics investigate the null hypothesis of r cointegrating relations against the alternative of n cointegrating relations, where n is the number of variables in the system for $r = 0, 1, 2, \dots, n-1$. Define $\hat{\lambda}_i$, $i = 1, 2, \dots, n$ to be a complex modulus of eigenvalues of $\widehat{\boldsymbol{\Pi}}$ and let them be ordered such that $\lambda_1 > \lambda_2 > \cdots > \lambda_n$.

Then the trace statistic is computed as

$$\hat{\lambda}_{trace}(r) = -T \sum_{i=r+1}^{n} \log[1 - \hat{\lambda}_i] \, . \qquad (3.32)$$

The Maximum Eigenvalue statistic tests the null hypothesis of r cointegrating relations against the alternative of $r+1$ cointegrating relations for $r = 0, 1, 2, \dots, n-1$. This test statistic is computed as

$$\hat{\lambda}_{max}(r, r+1) = -\,T\log[1 - \hat{\lambda}_{r+1}] \, , \qquad (3.33)$$

where $\hat{\lambda}_{r+1}$ is the $(r+1)^{th}$ ordered eigenvalue of $\boldsymbol{\Pi}$, and T is the sample size. The critical values tabulated by Johansen and Juselius (1990) will be used for these tests. Neither of these test statistics follows a chi square distribution in general. The asymptotic distributions of the test statistics (3.32) and (3.33) are not normal. Asymptotic critical values for the $\hat{\lambda}_{trace}$ and $\hat{\lambda}_{max}$ statistics have been calculated by Monte Carlo simulation and can be found also in Johansen and Juselius (1990). They are given also by most econometric software packages like Eviews 7. Since the critical values used for the maximum eigenvalue and trace test statistics are based on a pure unit-root assumption, they will no longer be correct when the variables in the system are near- unit-root processes.

3.3.8 Vector Error Correction (VEC) Models

VECM is a natural progression from a VAR model representation, especially when the level series are non-stationary. If cointegration has been detected between series, then it is obviously known that there exist(s) long-term equilibrium relationship(s) between them. So this time VECM will be a suitable model to evaluate the short run properties of the cointegrated series. The main feature of the VECM is its capability to correct for any disequilibrium that may shock the system from time to time. The error correction term (cointegration term) picks up such disequilibrium and guides variables of the system back to equilibrium. That is, this term progressively corrects the deviation from long-run equilibrium through a series of partial short-run adjustments. In case of no cointegration VECM is no longer required and one may directly proceed to Granger causality tests (that will be discussed later) to establish causal links between variables.

If a set of variables are found to have one or more cointegrating vectors, the corresponding error correction representations must be included in the system to evade misspecification and omission of the important constraints. Thus, the VAR in (3.29) should be re-parameterized as a Vector Error Correction Model (VECM) form in (3.30), (Hamilton, 1994). That is,

$$\Delta\boldsymbol{Y}_t = \boldsymbol{\Pi}\boldsymbol{Y}_{t-1} + \sum_{i=1}^{p-1}\boldsymbol{\Gamma}_i\Delta\boldsymbol{Y}_{t-i} + \boldsymbol{B}\boldsymbol{X}_t + \boldsymbol{\varepsilon}_t$$

is known as a *Vector Error Correction Model (VECM)*, where $\boldsymbol{\Pi} = -\boldsymbol{I}_n + \sum_{i=1}^{p}\boldsymbol{A}_i$, $\boldsymbol{I}_n$ is the identity matrix, and $\boldsymbol{\Gamma}_i = -\sum_{j=i+1}^{p}\boldsymbol{A}_j$.

The above specification of VECM contains information on both the short and the long run adjustment to changes in $\boldsymbol{Y}_t$ via estimating $\boldsymbol{\Gamma}$ and $\boldsymbol{\Pi}$, respectively. Matrix $\boldsymbol{\Pi}$ can be decomposed as $\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$, where $\boldsymbol{\alpha}$ is $n \times r$ matrix of speed of adjustments towards the long run equilibrium, and $\boldsymbol{\beta}$ is an $n \times r$ matrix of parameters which determines the cointegrating relationships of long-run coefficients such that $\boldsymbol{\beta}'\boldsymbol{Y}_{t-n}$ represent the multiple cointegration relationships. The columns of $\boldsymbol{\beta}$ are interpreted as long-run equilibrium relationships between variables. Values of $\boldsymbol{\alpha}$ close to zero imply slow convergences and $r,\ 0 \le r \le n,$ is the rank of the matrix $\boldsymbol{\Pi}$ and represents the number of cointegrating vectors in the system which can be determined using the Johansen Maximum Likelihood method.

Generally, the following three conditions should be noticed.

- When $rank(\boldsymbol{\Pi}) = 0$, it means that all of the series are non-stationary but there are no any co integrated variables (no cointegrating relationship) and hence, pure VAR(p) model will be utilized. That is, VECM will not be appropriate in this case.
- When $rank(\boldsymbol{\Pi}) = n$, it suggests that all the n (potentially available) variables are stationary in levels (model in levels), there is no any cointegrating relationship or vector since no unit root at all, and then, pure VAR will be applied. VECM will never be employed in this case also.
- When $rank(\boldsymbol{\Pi}) = r$, for $0 < r < n$, it proposes that there are r cointegrating vectors (relationships), i.e., r stationary linear combinations of the variables. In this case, pure VAR cannot be used and instead VECM will be the right model to be utilized.

3.3.9 Model Checking

It is an obligatory activity to investigate validity and reliability of all inference procedures made by VARs and VECMs before one is going to use these models to forecast future patterns of series. There are several tests for checking forecasting capability (adequacy) of these models.

3.3.9.1 Test of Residual Autocorrelation

The two most popular tests for autocorrelation of residuals are: Breusch-Godfrey LM tests and Portmanteau tests.

3.3.9.1.1 Autocorrelation LM Test

This test was developed by Breusch and Godfrey in 1978. Assume a VAR model for the error $\boldsymbol{u}_t$ given by

$$\boldsymbol{u}_t = \boldsymbol{D}_1\boldsymbol{u}_{t-1} + \cdots + \boldsymbol{D}_h\boldsymbol{u}_{t-h} + \boldsymbol{v}_t \ . \tag{3.34}$$

The quantity $\boldsymbol{v}_t$ denotes a white nose error term. Thus, to test autocorrelation in $\boldsymbol{u}_t$ the following claims should be tested.

$H_0{:}\, \boldsymbol{D}_1 = \cdots = \boldsymbol{D}_h = \boldsymbol{0}$ versus

$H_1{:}\, \boldsymbol{D}_j \neq \boldsymbol{0}$ for at least one j < h. (3.35)

Lagrange multiplier method is utilized to perform the test. The Lagrange Multiplier (LM) test for p^{th} order serial correlation is computed first by estimating an auxiliary regression where the OLS residuals are regressed on the variables in the original model plus p lagged residuals. The test statistic is either T times R^2 from the auxiliary regression or an F test that the coefficients on the lagged residuals are zero. This method is very useful for finding optimal estimates under constraint conditions. Under H_0, the interest is only to estimate the regular VAR model $(\boldsymbol{u}_t = \boldsymbol{v}_t)$. So the constrained case estimates are simple.

The Breusch-Godfrey test statistic, say Q_{BG}^*, is a standard LM test statistic for the null hypothesis $\boldsymbol{\Upsilon} = \boldsymbol{u}_t - \boldsymbol{v}_t = \boldsymbol{0}$:

$$Q_{BG}^* = T\widehat{\boldsymbol{\Upsilon}}'(\widehat{\boldsymbol{\Sigma}}^{\boldsymbol{\Upsilon}\boldsymbol{\Upsilon}})^{-1}\widehat{\boldsymbol{\Upsilon}} \tag{3.36}$$

where, $\widehat{\boldsymbol{\Upsilon}}$ is the generalized least square estimator of $\boldsymbol{\Upsilon}$ and $\widehat{\boldsymbol{\Sigma}}^{\boldsymbol{\Upsilon}\boldsymbol{\Upsilon}}$ is the part of

$$\left[T^{-1}\sum_{t=1}^{T}\begin{pmatrix}\widehat{\boldsymbol{U}}_t \otimes \boldsymbol{I}_n \\ \hat{\boldsymbol{z}}_t \otimes \boldsymbol{I}_n \\ \hat{\boldsymbol{z}}_{1t}\end{pmatrix}\widehat{\boldsymbol{\Omega}}^{-1}\left(\widehat{\boldsymbol{U}}_t' \otimes \boldsymbol{I}_n : \hat{\boldsymbol{z}}_t' \otimes \boldsymbol{I}_n : \hat{\boldsymbol{z}}_{1t}'\right)\right]^{-1} \tag{3.37}$$

corresponding to $\boldsymbol{\Upsilon}$ and $\widehat{\boldsymbol{Z}}_t = (\boldsymbol{1}', \boldsymbol{Y}_t', \dots, \boldsymbol{Y}_{t-p+1}')'$.

Here $\widehat{\boldsymbol{\Omega}} = T^{-1}\sum_{t=1}^{T}\widehat{\boldsymbol{u}}_t\widehat{\boldsymbol{u}}_t'$ is the residual covariance matrix estimator from the restricted auxiliary model. Therefore, under the null hypothesis it follows immediately from (3.36) that as $h \longrightarrow \infty$

$$Q_{BG}^* \xrightarrow{d} \chi^2(hn^2) \ . \tag{3.38}$$

3.3.9.1.2 Portmanteau Autocorrelation Test

Suppose $\boldsymbol{Y}_t = (Y_{1t}, \dots, Y_{nt})'$ is n-dimensional vector of observable time series variables with $r < n$ cointegration relations.The residual auto covariance is

$$\widehat{\boldsymbol{C}}_j = \frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{u}}_t\widehat{\boldsymbol{u}}_{t-j} \quad , \tag{3.39}$$

where, $\widehat{\boldsymbol{u}}_t = \Delta\boldsymbol{Y}_t - \boldsymbol{\Pi}\boldsymbol{Y}_{t-1} - \sum_{i=1}^{p-1}\boldsymbol{\Gamma}_i\Delta\boldsymbol{Y}_{t-i} - \boldsymbol{B}\boldsymbol{X}_t$ is an estimated residual.

The Portmanteau test for residual autocorrelation checks the null hypothesis that all residual auto-covariances are zero. That is,

$$H_0: E\left(\boldsymbol{u}_t \boldsymbol{u}_{t-i}^{'}\right) = 0 \text{ for i=1,2,…} \quad . \tag{3.40}$$

Here, the $\boldsymbol{u}_t$ ‘s in (3.40) are residuals to be estimated by $\widehat{\boldsymbol{u}}_t$.

The null hypothesis in (3.40) is tested against the alternative that at least one auto covariance is different from zero, i.e., autocorrelation is nonzero. The test statistic is based on the residual auto covariances and has the form

$$Q_p = T \sum_{j=1}^{h} tr\left(\widehat{\boldsymbol{C}}_j^{'} \widehat{\boldsymbol{\Omega}}^{-1} \widehat{\boldsymbol{C}}_j \widehat{\boldsymbol{\Omega}}^{-1}\right), \tag{3.41}$$

where

$$\widehat{\boldsymbol{C}}_j = \frac{1}{T} \sum_{t=1}^{T} \widehat{\boldsymbol{u}}_t \widehat{\boldsymbol{u}}_{t-j} \quad \text{and} \tag{3.42}$$

$$\widehat{\boldsymbol{\Omega}} = \frac{1}{T} \sum_{t=1}^{T} \widehat{\boldsymbol{u}}_t \widehat{\boldsymbol{u}}_t^{'} \quad . \tag{3.43}$$

The approximate distribution of the test statistic in (3.41) is the chi-squared distribution with $n^2(h-p)$ degrees of freedom in large samples (T) if $\frac{h}{T} \longrightarrow 0$.

A related statistic with potentially superior small sample properties is the adjusted Portmanteau statistic

$$Q_p^* = T^2 \sum_{j=1}^{h} \frac{1}{T-j} tr\left(\widehat{\boldsymbol{C}}_j^{'} \widehat{\boldsymbol{C}}_0^{-1} \widehat{\boldsymbol{C}}_j \widehat{\boldsymbol{C}}_0^{-1}\right) \tag{3.44}$$

Nevertheless, the asymptotic properties of Q_p^* are the same as those of Q_p.

3.3.9.2 Normality of the Residuals

The other point that should be carefully checked is the normality of residuals from VAR and VECM. Although a number of normality tests are available, the most widely used test for normality of regression disturbances is due to Jarque and Bera (1981). Especially in multivariate time series analysis, Lütkepohl (1991) also suggested Jarque-Bera test to examine the multivariate normality of residuals. It tests skewness (3rd moment) and kurtosis (4th moment) properties of the residuals $\boldsymbol{u}_t = (u_{1t}, \dots, u_{nt})'$ against those of a multivariate normal distribution of the appropriate dimension. To construct the Jarque-Bera test consider a mean adjusted n-dimensional stationary, stable VAR (p) process

$$\boldsymbol{Y}_t - \boldsymbol{\mu} = \boldsymbol{A}_1(\boldsymbol{Y}_{t-1} - \boldsymbol{\mu}) + \cdots + \boldsymbol{A}_p\left(\boldsymbol{Y}_{t-p} - \boldsymbol{\mu}\right) + \boldsymbol{u}_t, \tag{3.45}$$

where $\boldsymbol{u}_t$ is a zero mean white noise with nonsingular covariance matrix $\boldsymbol{\Sigma}_u$ and $\boldsymbol{A}_1, \dots, \boldsymbol{A}_p$ are coefficient matrices. Then define

$$\hat{u}_t = (\boldsymbol{Y}_t - \overline{\boldsymbol{Y}}) - \widehat{\boldsymbol{A}}_1(\boldsymbol{Y}_{t-1} - \overline{\boldsymbol{Y}}) - \cdots - \widehat{\boldsymbol{A}}_p\left(\boldsymbol{Y}_{t-p} - \overline{\boldsymbol{Y}}\right), \quad t = 1, \dots, T, \tag{3.46}$$

where $\widehat{\boldsymbol{A}}_1, \dots, \widehat{\boldsymbol{A}}_p$ are estimators of the coefficients $\boldsymbol{A}_1, \dots, \boldsymbol{A}_p$ in (3.45) based on a sample $\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_T$. Let $\widehat{\boldsymbol{P}}$ be a matrix satisfying $\widehat{\boldsymbol{P}}\widehat{\boldsymbol{P}}' = \widehat{\boldsymbol{\Sigma}}_u$ such that $plim\left(\widehat{\boldsymbol{P}} - \boldsymbol{P}\right) = 0$,

where $$\widehat{\boldsymbol{\Sigma}}_u = \frac{1}{T - np - 1} \sum_{t=1}^{T} \widehat{\boldsymbol{u}}_t \widehat{\boldsymbol{u}}_t' .$$

Furthermore, define the standardized residuals and their sample moments as

$$\widehat{\boldsymbol{w}}_t = (\widehat{w}_{1t}, \dots, \widehat{w}_{nt})' = \widehat{\boldsymbol{P}}^{-1}\widehat{\boldsymbol{u}}_t \tag{3.47}$$

$$\widehat{\boldsymbol{b}}_1 = (\widehat{b}_{11}, \dots, \widehat{b}_{n1})' \quad \text{with} \quad \widehat{b}_{j1} = \frac{1}{T}\sum_{t=1}^{T} \widehat{w}_{jt}^3, \quad j = 1, \dots, n, \tag{3.48}$$

and $$\widehat{\boldsymbol{b}}_2 = (\widehat{b}_{12}, \dots, \widehat{b}_{n2})' \quad \text{with} \quad \widehat{b}_{j2} = \frac{1}{T}\sum_{t=1}^{T} \widehat{w}_{jt}^4, \quad j = 1, \dots, n. \tag{3.49}$$

Consequently, the statistic $\hat{\lambda}_s = T\widehat{\boldsymbol{b}}_1'\widehat{\boldsymbol{b}}_1/6$ can be used to test

$$H_0: E\begin{bmatrix} u_{1t}^3 \\ \vdots \\ u_{nt}^3 \end{bmatrix} = \boldsymbol{0} \quad \text{against} \quad H_A: E\begin{bmatrix} u_{1t}^3 \\ \vdots \\ u_{nt}^3 \end{bmatrix} \neq \boldsymbol{0} \quad \text{(Skewness)} \tag{3.50}$$

and $\hat{\lambda}_k = T(\widehat{\boldsymbol{b}}_2 - \boldsymbol{3}_n)'(\widehat{\boldsymbol{b}}_2 - \boldsymbol{3}_n)/24$ may be used to test

$$H_0: E\begin{bmatrix} u_{1t}^4 \\ \vdots \\ u_{nt}^4 \end{bmatrix} = \boldsymbol{3}_n \quad \text{against} \quad H_A: E\begin{bmatrix} u_{1t}^4 \\ \vdots \\ u_{nt}^4 \end{bmatrix} \neq \boldsymbol{3}_n \quad \text{(Kurtosis).} \tag{3.51}$$

Furthermore, $\hat{\lambda}_{sk} = \hat{\lambda}_s + \hat{\lambda}_k$ may be used for a joint test of the null hypotheses in (3.50) and (3.51).

Asymptotically,

$$\hat{\lambda}_s = T\hat{\boldsymbol{b}}_1'\hat{\boldsymbol{b}}_1/6 \xrightarrow{d} \chi^2(n)\ , \tag{3.52}$$

$$\hat{\lambda}_k = T(\hat{\boldsymbol{b}}_2 - \boldsymbol{3}_n)'(\hat{\boldsymbol{b}}_2 - \boldsymbol{3}_n)/24 \xrightarrow{d} \chi^2(n)\ , \tag{3.53}$$

and

$$\hat{\lambda}_{sk} = \hat{\lambda}_s + \hat{\lambda}_k \xrightarrow{d} \chi^2(2n). \tag{3.54}$$

Or

$$\sqrt{T}\begin{bmatrix} \hat{\boldsymbol{b}}_1 \\ \hat{\boldsymbol{b}}_2 - \boldsymbol{3}_n \end{bmatrix} \xrightarrow{d} N\left(0, \begin{bmatrix} 6\boldsymbol{I}_n & 0 \\ 0 & 24\boldsymbol{I}_n \end{bmatrix}\right). \tag{3.55}$$

3.3.10 Forecasting

The ultimate goal of estimating VAR and VECM is forecasting. Since multivariate time series analysis is an extension of univariate one, the way of forecasting is almost similar in both cases.

Consider first the problem of forecasting future values of $\boldsymbol{Y}_T$ when the parameters $\boldsymbol{\Pi}$ of the VAR (p) process are assumed to be known and there are no deterministic terms or exogenous variables. The best linear predictor, in terms of minimum mean squared error (MSE), of $\boldsymbol{Y}_{T+1}$ or 1-step forecast based on information available at time T is

$$\boldsymbol{Y}_{T+1|T} = \boldsymbol{c} + \boldsymbol{\Pi}_1\boldsymbol{Y}_T + \cdots + \boldsymbol{\Pi}_p\boldsymbol{Y}_{T-p+1} \tag{3.56}$$

for $T \geq p$.

Forecasts for longer horizons h (h-step forecasts) can be obtained using the chain-rule of forecasting as

$$\boldsymbol{Y}_{T+h|T} = \boldsymbol{c} + \boldsymbol{\Pi}_1\boldsymbol{Y}_{T+h-1|T} + \cdots + \boldsymbol{\Pi}_p\boldsymbol{Y}_{T+h-p|T}\ , \tag{3.57}$$

where, $\boldsymbol{Y}_{T+j|T} = \boldsymbol{Y}_{T+j}$ for $j \leq 0$. The h-step forecast errors may be expressed as

$$\boldsymbol{Y}_{T+h} - \boldsymbol{Y}_{T+h|T} = \sum_{s=0}^{h-1} \boldsymbol{\Psi}_s \boldsymbol{\varepsilon}_{T+h-s} \quad , \tag{3.58}$$

where, the matrices $\boldsymbol{\Psi_s}$ are determined by recursive substitution,

$$\boldsymbol{\Psi}_s = \sum_{j=1}^{p-1} \boldsymbol{\Psi}_{s-j} \boldsymbol{\Pi}_j \tag{3.59}$$

with $\boldsymbol{\Psi}_0 = \boldsymbol{I}_n$ and $\boldsymbol{\Pi_j} = \mathbf{0}_{n\times n}$ for $j > p$. The forecasts are unbiased since all of the forecast errors have expectation zero, and the MSE matrix for $\boldsymbol{Y}_{T+h|T}$ is

$$\boldsymbol{\Sigma}(\text{h}) = \text{MSE}\left(\boldsymbol{Y}_{T+h} - \boldsymbol{Y}_{T+h|T}\right)$$

$$\boldsymbol{\Sigma}(\text{h}) = \text{MSE}\left(\sum_{s=0}^{h-1} \boldsymbol{\Psi}_s \boldsymbol{\varepsilon}_{T+h-s}\right)$$

$$= \sum_{s=0}^{h-1} \boldsymbol{\Psi}_s \boldsymbol{\Sigma} \boldsymbol{\Psi}_s' \quad . \tag{3.60}$$

Now consider forecasting $\boldsymbol{Y}_{T+h}$ when the parameters of the VAR (p) process are estimated using multivariate least squares. The best linear predictor of $\boldsymbol{Y}_{T+h}$ is now

$$\widehat{\boldsymbol{Y}}_{T+h|T} = \widehat{\boldsymbol{\Pi}}_1 \widehat{\boldsymbol{Y}}_{T+h-1|T} + \cdots + \widehat{\boldsymbol{\Pi}}_p \widehat{\boldsymbol{Y}}_{T+h-p|T} \quad , \tag{3.61}$$

where, $\widehat{\Pi}_j$ are the estimated parameter matrices. The h-step forecast error is given by

$$\boldsymbol{Y}_{T+h} - \widehat{\boldsymbol{Y}}_{T+h|T} = \sum_{s=0}^{h-1} \boldsymbol{\Psi}_s \boldsymbol{\varepsilon}_{T+h-s} + (\boldsymbol{Y}_{T+h|T} - \widehat{\boldsymbol{Y}}_{T+h|T}) \tag{3.62}$$

and the term $\boldsymbol{Y}_{T+h|T} - \widehat{\boldsymbol{Y}}_{T+h|T}$ captures the part of the forecast error due to estimating the parameters of the VAR. The MSE matrix of the h-step forecast is then,

$$\widehat{\boldsymbol{\Sigma}}(h) = \boldsymbol{\Sigma}(h) + MSE\left(\boldsymbol{Y}_{T+h|T} - \widehat{\boldsymbol{Y}}_{T+h|T}\right). \tag{3.63}$$

In practice, the second term $\text{MSE}\left(\boldsymbol{Y}_{T+h|T} - \widehat{\boldsymbol{Y}}_{T+h|T}\right)$ is often ignored and $\widehat{\boldsymbol{\Sigma}}(h)$ is computed using (3.60) as:

$$\widehat{\boldsymbol{\Sigma}}(h) = \sum_{s=0}^{h-1} \widehat{\boldsymbol{\Psi}}_s \widehat{\boldsymbol{\Sigma}} \widehat{\boldsymbol{\Psi}}_s' \tag{3.64}$$

with $\widehat{\boldsymbol{\Psi}}_s = \sum_{j=1}^{s} \widehat{\boldsymbol{\Psi}}_{s-j} \widehat{\boldsymbol{\Pi}}_j$. Lütkepohl (1991) gives an approximation to $MSE\left(\boldsymbol{Y}_{T+h|T} - \widehat{\boldsymbol{Y}}_{T+h|T}\right)$ which may be interpreted as a finite sample correction to (3.64).

3.3.11 Measures of Forecasting Accuracy

Accuracy measurement of time series models is considered as a mandatory step before one has to go for forecasting. It would be unreliable work unless any finding is provided with its respective accuracy measurements. By accuracy it broadly means that how well the model fitted the data on hand. So what expected always is the forecasted values should adhere themselves to the actual values. There are many ways to evaluate the forecasting performance of a VEC model, ranging from mean error (ME) measures to Theil's U statistic measure that will be discussed in this subsection.

If Y_{jt}, j=1, 2, . . ., n (number of variables), is the actual observation on the j^th^ variable at period t and F_{jt} is the forecast for jth variable at time t (Y_{jt}), then the residual of the j^th^ variable at time t is defined as:

$$\hat{\varepsilon}_{jt} = Y_{jt} - F_{jt} \ . \tag{3.65}$$

Usually F_{jt} is calculated using data, $Y_{j1}, Y_{j2}, \ldots, Y_{jt-1}$. It is a one step forecast because it is forecasting one period ahead of the last observation used in the calculation. Therefore, $\hat{\varepsilon}_{jt}$ is described as a one step forecast error. It is the difference between the observation Y_{jt} and forecast made using all observations excluding Y_{jt}.

If there are observations and forecasts for T time periods, then there will be T error terms for each variable j, and the following standard statistical measures can be defined as

$$Mean\ Error(ME) = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_{jt} \quad , \tag{3.66}$$

$$Mean\ Absolute\ Error(MAE) = \frac{1}{T}\sum_{t=1}^{T}\left|\hat{\varepsilon}_{jt}\right| \quad \text{, and} \tag{3.67}$$

$$Mean\ Squared\ Error(MAE) = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_{jt}^{2} \ . \tag{3.68}$$

To make comparisons a little bit work is needed with relative or percentage error measures. First define a relative or percentage error as

$$PE_t = \left(\frac{Y_{jt} - F_{jt}}{Y_{jt}}\right) \times 100\% \ . \tag{3.69}$$

Then the following two relative measures in (3.70) and (3.71) are frequently used.

$$Mean\ Percentage\ Error(MPE) = \frac{1}{T}\sum_{t=1}^{T} PE_t \quad (3.70)$$

$$Mean\ Percentage\ Absolute\ Error(MPAE) = \frac{1}{T}\sum_{t=1}^{T} |PE_t| \quad (3.71)$$

(3.69) can be used to compute the percentage error for any time period and averaged as in (3.70) to give the mean percentage error. However, as with the ME, the MPE is likely to be small since positive and negative PE's tend to offset one another. Hence, the MPAE in (3.71) is defined using absolute values of PE.

Alternatively, Theil's U statistic can be used as a measure of forecasting accuracy. Like MPAE statistic, high values suggest poor performance in the forecast. Theil's U can be estimated as

$$U = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Y_{jt} - F_{jt}\right)^2}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n} F_{jt}^2} + \sqrt{\frac{1}{n}\sum_{t=1}^{n} Y_{jt}^2}} \quad (3.72)$$

The scaling of U is such that it will always lie between 0 and 1. If U = 0, $Y_{jt} = F_{jt}$ for all forecasts and there is a perfect fit. In contrary, if U = 1, the predictive performance is not good.

3.3.12 Structural Vector Autoregressive (SVAR) Analysis

Because VAR models represent the correlations among a set of variables, they are often used to analyze certain aspects of the relationships between the variables of interest. In the following, three ways to interpret a VAR model will be discussed. They are all closely related and they are all beset with problems that will be pointed out subsequently.

3.3.12.1. Granger Causality Test

Granger (1969) has defined a concept of causality which, under suitable conditions, is fairly easy to deal with in the context of VAR models. Therefore, it has become quite popular in recent years. The idea is that a cause cannot come after the effect. Thus, if a variable *x* affects a variable *z*, the former should help improving the predictions of the latter variable.

To formalize the above idea, the following intuitive notion of a variable's forecasting ability is constructed by Granger (1969). If a variable (or group of variables) Y_1 is found to be helpful for predicting another variable (or group of variables) Y_2 ,then Y_1 is said to Granger-cause Y_2. Formally, Y_1 fails to Granger-cause Y_2 if for all s > 0, the MSE of a forecast of $Y_{2,t+s}$ based on $(Y_{2,t}, Y_{2,t-1}, \ldots)$ is the same as the MSE of a forecast of $Y_{2,t+s}$ based on $(Y_{2,t}, Y_{2,t-1}, \ldots)$ and $(Y_{1,t}, Y_{1,t-1}, \ldots)$. Clearly, the notion of Granger causality does not imply true causality. It only implies forecasting ability. If Y_1 causes Y_2 and Y_2 also causes Y_1, the process (Y_{1t}, Y_{2t}) is called a *feedback system*.

For instance, in a bivariate VAR (p) model for $\boldsymbol{Y}_t = (Y_{1t}, Y_{2t})'$, Y_2 fails to Granger-cause Y_1 if all of the p VAR coefficient matrices $\boldsymbol{\Pi}_1, \ldots, \boldsymbol{\Pi}_p$ are lower triangular. That is, the VAR (p) model has the form

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \Pi_{11}^{(1)} & 0 \\ \Pi_{21}^{(1)} & \Pi_{22}^{(1)} \end{pmatrix} \begin{pmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{pmatrix} + \cdots + \begin{pmatrix} \Pi_{11}^{(p)} & 0 \\ \Pi_{21}^{(p)} & \Pi_{22}^{(p)} \end{pmatrix} \begin{pmatrix} Y_{1,t-p} \\ Y_{2,t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}. \quad (3.73)$$

So that all of the coefficients on lagged values of Y_2 is zero in (3.73) for Y_1.Similarly, Y_1 fails to Granger-cause Y_2 if all of the coefficients on lagged values of Y_1 are zero in the equation for Y_2.

The p linear coefficient restrictions implied by Granger non-causality may be tested using the Wald statistic which is asymptotically distributed as χ^2 with p (number of lags in the VAR) degree of freedom. A large value of Wald statistic is an evidence against the null hypothesis of non-causality. Notice that if Y_2 fails to Granger-cause Y_1 and Y_1 fails to Granger-cause Y_2, then the VAR coefficient matrices $\boldsymbol{\Pi}_1, \boldsymbol{\Pi}_2, \ldots, \boldsymbol{\Pi}_p$ are diagonal. Testing for Granger non-causality in general n variable (multivariate) VAR (p) models follows the same logic used for bivariate models.

3.3.12.2. Impulse Response Functions

Impulse response function is an important tool in a VAR system in revealing the direction and magnitude at which one variable (especially the target variable) reacts to the change (shock) applied on the other exogenous variables in the system.

Any covariance stationary VAR (p) process has a Wald representation of the form

$$\boldsymbol{Y}_t = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t + \boldsymbol{\Psi}_1 \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\Psi}_2 \boldsymbol{\varepsilon}_{t-2} + \cdots \quad , \quad (3.74)$$

where, the $(n \times n)$ moving average matrices $\boldsymbol{\Psi}_s$ are determined recursively using (3.59). It is tempting to interpret the $(i,j)^{th}$ element, $\boldsymbol{\Psi}_{ij}^s$, of the matrix $\boldsymbol{\Psi}_s$ as the dynamic multiplier or impulse response

$$\frac{\partial Y_{i,t+s}}{\partial \varepsilon_{j,t}} = \frac{\partial Y_{i,t}}{\partial \varepsilon_{j,t-s}} = \Psi_{ij}^s,\ i,j = 1,2,\dots,n\ . \tag{3.75}$$

However, this interpretation is only possible if $Var(\boldsymbol{\varepsilon}_t) = \boldsymbol{\Sigma}$ is a diagonal matrix so that the elements of $\boldsymbol{\varepsilon}_t$ are uncorrelated. One way to make the errors uncorrelated is to follow Sims(1980) and estimate the triangular structural VAR (p) model

$$\left.\begin{aligned}
Y_{1t} &= c_1 + \boldsymbol{\gamma}_{11}'\boldsymbol{Y}_{t-1} + \cdots + \boldsymbol{\gamma}_{1p}'\boldsymbol{Y}_{t-p} + \eta_{1t} \\
Y_{2t} &= c_2 + \beta_{21}Y_{1t} + \boldsymbol{\gamma}_{21}'\boldsymbol{Y}_{t-1} + \cdots + \boldsymbol{\gamma}_{2p}'\boldsymbol{Y}_{t-p} + \eta_{2t} \\
Y_{3t} &= c_3 + \beta_{31}Y_{1t} + \beta_{32}Y_{2t} + \boldsymbol{\gamma}_{31}'\boldsymbol{Y}_{t-1} + \cdots + \boldsymbol{\gamma}_{3p}'\boldsymbol{Y}_{t-p} + \eta_{3t} \\
&\vdots \\
Y_{nt} &= c_n + \beta_{n1}Y_{1t} + \cdots + \beta_{n,n-1}Y_{n-1,t} + \boldsymbol{\gamma}_{n1}'\boldsymbol{Y}_{t-1} + \cdots + \boldsymbol{\gamma}_{np}'\boldsymbol{Y}_{t-p} + \eta_{nt}
\end{aligned}\right\} \tag{3.76}$$

In matrix form, the triangular structural VAR (p) model is

$$\boldsymbol{B}\boldsymbol{Y}_t = \boldsymbol{c} + \boldsymbol{\Gamma}_1\boldsymbol{Y}_{t-1} + \boldsymbol{\Gamma}_2\boldsymbol{Y}_{t-2} + \cdots + \boldsymbol{\Gamma}_p\boldsymbol{Y}_{t-p} + \boldsymbol{\eta}_t\ , \tag{3.77}$$

where,

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ -\beta_{21} & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\beta_{n1} & -\beta_{n2} & \dots & 1 \end{pmatrix} \tag{3.78}$$

is a lower triangular matrix with 1's along the diagonal and $\boldsymbol{c}$ is vector of constants. The algebra of least squares will ensure that the estimated covariance matrix of the error vector $\boldsymbol{\eta}_t$ is diagonal. The uncorrelated/orthogonal errors $\boldsymbol{\eta}_t$ are referred to as *structural errors*. The triangular structural model (3.76) imposes the recursive causal ordering

$$Y_1 \longrightarrow Y_2 \longrightarrow \cdots \longrightarrow Y_n\ . \tag{3.79}$$

The ordering (3.79) means that the contemporaneous values of the variables to the left of the arrow affect the contemporaneous values of the variables to the right of the arrow but not vice versa. These contemporaneous effects are captured by the coefficients β_{ij} in (3.76).

For instance, the ordering $Y_1 \longrightarrow Y_2 \longrightarrow Y_3$ imposes the restrictions: Y_{1t} affects Y_{2t} and Y_{3t} but Y_{2t} and Y_{3t} do not affect Y_{1t}; Y_{2t} affects Y_{3t} but Y_{3t} does not affect Y_{2t}.

For a VAR (p) with n variables there are $n!$ possible recursive causal orderings. Which ordering to use in practice depends on the context and whether prior theory can be used to justify a particular ordering. Results from alternative orderings can always be compared to determine the sensitivity of results to the imposed ordering.

Once a recursive ordering has been established, the Wald representation of $\boldsymbol{Y}_t$ based on the orthogonal errors $\boldsymbol{\eta}_t$ is given by

$$\boldsymbol{Y}_t = \boldsymbol{\mu} + \boldsymbol{\Theta}_0\boldsymbol{\eta}_t + \boldsymbol{\Theta}_1\boldsymbol{\eta}_{t-1} + \boldsymbol{\Theta}_2\boldsymbol{\eta}_{t-2} + \cdots \quad , \tag{3.80}$$

where, $\boldsymbol{\Theta}_0 = \boldsymbol{B}^{-1}$ is a lower triangular matrix. The impulse responses to the orthogonal shocks $\eta_{j,t}$ are

$$\frac{\partial Y_{i,t+s}}{\partial \eta_{j,t}} = \frac{\partial Y_{i,t}}{\partial \eta_{j,t-s}} = \theta_{ij}^{s}, \; i,j = 1,2,\ldots,n; s > 0 \quad , \tag{3.81}$$

where, θ_{ij}^{s} is the $(i,j)^{th}$ element of $\boldsymbol{\Theta}_s$. A plot of θ_{ij}^{s} against s is called the *orthogonal impulse response function (IRF)* of Y_i with respect to η_j. With n variables there are n^2 possible impulse response functions.

In practice, the orthogonal IRF (3.81) based on the triangular VAR (p) (3.76) may be computed directly from the parameters of the non triangular VAR (p) (3.4) as follows. First, decompose the residual covariance matrix $\boldsymbol{\Sigma}$ as

$$\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}' \quad , \tag{3.82}$$

where, $\boldsymbol{A}$ is an invertible lower triangular matrix with 1's along the diagonal and $\boldsymbol{D}$ is a diagonal matrix with positive diagonal elements. Next, define the structural errors as

$$\boldsymbol{\eta}_t = \boldsymbol{A}^{-1}\boldsymbol{\varepsilon}_t \; . \tag{3.83}$$

These structural errors are orthogonal by construction since

$$Var(\boldsymbol{\eta}_t) = \boldsymbol{A}^{-1}\boldsymbol{\Sigma}\boldsymbol{A}^{-1'} = \boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{D}\boldsymbol{A}'\boldsymbol{A}^{-1'} = \boldsymbol{D} \; . \tag{3.84}$$

Finally, re-express the Wald representation (3.74) as

$$\boldsymbol{Y}_t = \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{\varepsilon}_t + \boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{\varepsilon}_{t-2} + \cdots$$

$$= \boldsymbol{\mu} + \boldsymbol{\Theta}_0\boldsymbol{\eta}_t + \boldsymbol{\Theta}_1\boldsymbol{\eta}_{t-1} + \boldsymbol{\Theta}_2\boldsymbol{\eta}_{t-2} + \cdots \quad ,$$

where, $\boldsymbol{\Theta}_j = \boldsymbol{\Psi}_j\boldsymbol{A}$. Notice that the structural $\boldsymbol{B}$ matrix in (3.78) is equal to $\boldsymbol{A}^{-1}$.

3.3.12.3. Forecast Error Variance Decompositions

The forecast error variance decomposition (FEVD) answers the question: what portion of the variance of the forecast error in predicting $Y_{i,T+h}$ is due to the structural shock η_j? Using the orthogonal shocks $\boldsymbol{\eta}_T$ the h-step ahead forecast error vector, with known VAR coefficients, may be expressed as

$$\boldsymbol{Y}_{T+h} - \boldsymbol{Y}_{T+h|T} = \sum_{s=0}^{h-1} \boldsymbol{\Theta}_s\boldsymbol{\eta}_{T+h-s} \quad . \tag{3.85}$$

For a particular variable $Y_{i,T+h}$, this forecast error has the form

$$Y_{i,T+h} - Y_{i,T+h|T} = \sum_{s=0}^{h-1} \theta_{i1}^{s}\eta_{1,T+h-s} + \cdots + \sum_{s=0}^{h-1} \theta_{in}^{s}\eta_{n,T+h-s} \quad . \tag{3.86}$$

Since the structural errors are orthogonal, the variance of the h-step forecast error is

$$Var(Y_{i,T+h} - Y_{i,T+h|T}) = \sigma_{\eta_1}^2 \sum_{s=0}^{h-1}(\theta_{i1}^{s})^2 + \cdots + \sigma_{\eta_n}^2 \sum_{s=0}^{h-1}(\theta_{in}^{s})^2 \quad , \tag{3.87}$$

where $\sigma_{\eta_n}^2 = Var(\eta_{jT})$. The portion of $Var(Y_{i,T+h} - Y_{i,T+h|T})$ due to shock η_j is then

$$FEVD_{i,j}(h) = \frac{\sigma_{\eta_j}^2 \sum_{s=0}^{h-1}\left(\theta_{ij}^{s}\right)^2}{\sigma_{\eta_1}^2 \sum_{s=0}^{h-1}\left(\theta_{i1}^{s}\right)^2 + \cdots + \sigma_{\eta_n}^2 \sum_{s=0}^{h-1}\left(\theta_{in}^{s}\right)^2} \quad , i,j = 1,2,\ldots,n. \tag{3.88}$$

In a VAR with n variables there will be n^2 $FEVD_{i,j}(h)$ values. It must be kept in mind that the $FEVD$ in (3.88) depends on the recursive causal ordering used to identify the structural shocks $\boldsymbol{\eta}_T$ and is not unique. Different causal orderings will produce different $FEVD$ values.

4. STATISTICAL RESULTS AND DISCUSSIONS

This chapter provides the statistical analyses, results and discussions of seasonal and nonstationary multivariate time series using VAR model that incorporates co-integration relationships. Specification, estimation and inference in VAR model will be performed by using Eviews7 statistical software. The discussion covered all the necessary descriptions of original and transformed data as well interpretations of analyses results. These analyses included deseasonalization of the seasonal variables, pre and post tests for both seasonal adjustment and stationarity, optimal lag length selection, test for optimality of the lag selected (lag exclusion test), co-integration test, VEC model estimation, structural analysis (i.e., Granger-causality tests, the computation of impulse response functions, and forecast error variance decompositions). Finally, the chapter concludes with adequacy checking of the VEC models and forecasting from the models.

4.1 Descriptive Analysis and Time Plot

The preliminary works to begin the analysis are computation of some descriptive statistics and inspection of time plots for the original data.

Table 4.1: Summary of Descriptive Statistics for all Original Series

Series	Mean	Minimum	Maximum	Std. Dev.	CV	Jarque-Bera	Probability
Load Factor	71.19722	64.35880	78.9056	3.782256	0.05312	1.139294	0.565725*
Block Hours	14.20452	7.995810	19.3971	3.031984	0.21350	2.527538	0.282587*
Distance Flown	9.667997	5.398781	13.4377	2.125080	0.21981	2.469979	0.290838*
Passenger Revenue	77.494972	41.566976	122.0000	20.445475	0.26383	2.907045	0.233746*

*** P-values < 0.05: Statistically significant**

Table 4.1 above shows the mean, minimum, maximum, standard deviation, coefficient of variation (CV), Jarque-Bera statistics, and the corresponding probability values (P-values). From the Table, the Load Factor and Passenger Revenue are the lowly and highly scattered series in the study with a CV of 5.312 % and 26.383%, respectively. Block Hours and Distance Flown series are almost equally dispersed. Regarding normality, all the P-values of the Jarque-Bera statistic in the Table are greater than 5% suggesting that initially all the series are normally distributed.

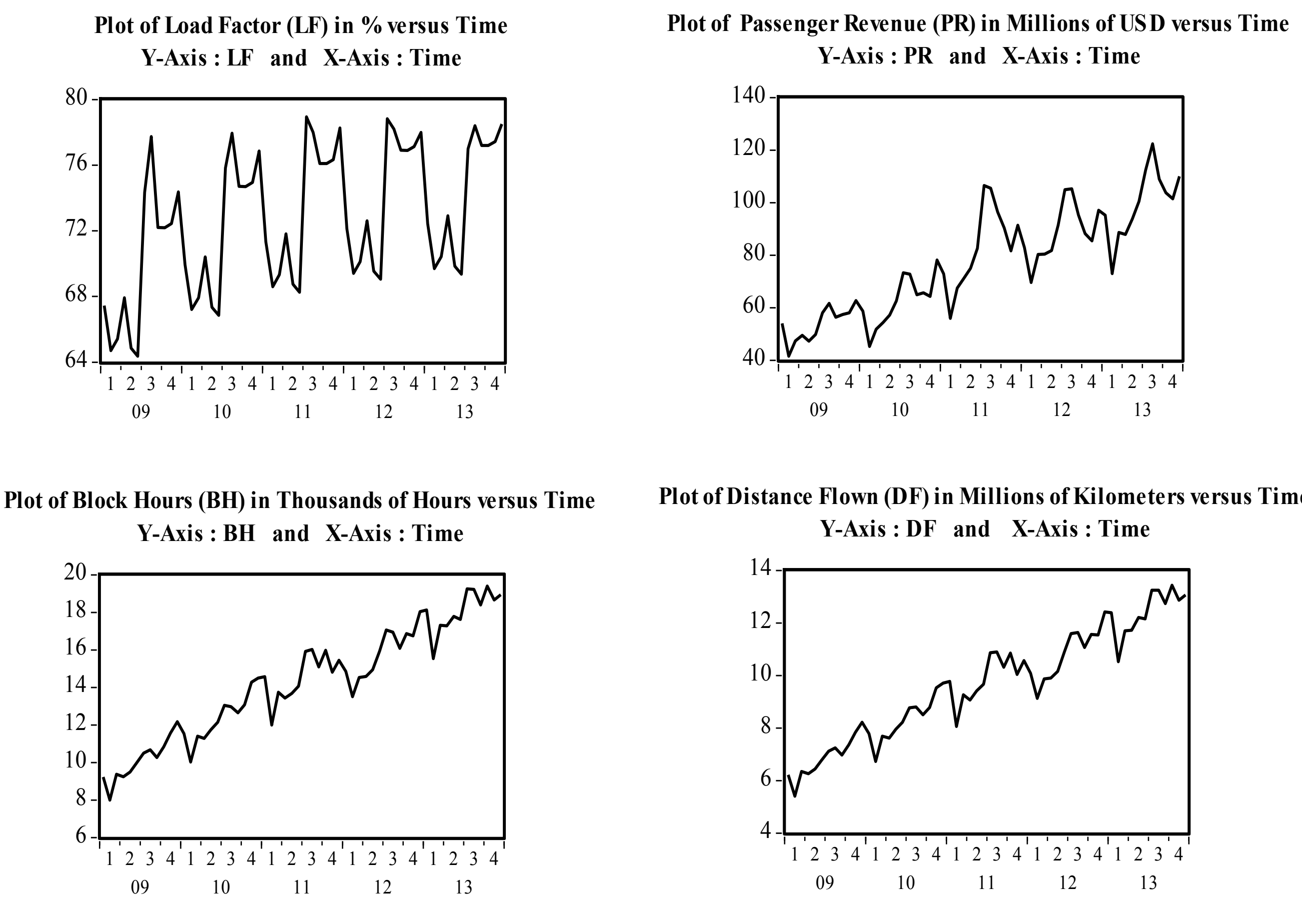
On its part, Table 4.2 below presents the correlation matrix for the four series and all are significantly inter-correlated as the figures tend to one in magnitude. Accordingly, Load Factor of the aviation is highly correlated with Passenger Revenue, Block Hours, and Distance Flown.

Table 4.2: Correlation Matrix of the Variables

Variables	Load Factor	Block Hours	Distance Flown	Passenger Revenue
Load Factor	1.000000	0.876008	0.873992	0.899532
Block Hours	0.876008	1.000000	0.999878	0.991071
Distance Flown	0.873992	0.999878	1.000000	0.990727
Passenger Revenue	0.899532	0.991071	0.990727	1.000000

In most practices, time plots of a certain data are very helpful to detect and extract useful insight pre- information about the data on hand. Thus, inspection should be made on the time plots of the original series. Figure 4.1 below provides the individual plots of original series against time period covered by the study.

Figure 4.1: Time Plots of each Original Series



From Figure 4.1, it can be seen that there is a clear seasonality and a general upward trend in all the series. This implies that all the data are not stationary. But it should be strongly noticed that only graph inspections are not enough to certainly conclude the series are seasonal and patterned. There are standard tests for both seasonality and stationarity which has been discussed previously in methodology and will be applied in analysis of the data as follow.

4.2 Seasonality Test

Before a series is seasonally adjusted, it should be shown that the series is seasonal. As stated in the previous chapter, X-12 ARIMA seasonal adjustment method will be employed to formally test the presence of seasonality for all the series. When using X-12 ARIMA for seasonal adjustment, two diagnostics commonly used to determine seasonality are M7, a diagnostic developed at Statistics Canada for X-11-ARIMA, and the F-tests for seasonality of the series and residuals. Then the variable(s) in which seasonality is observed will be seasonally adjusted.

Table 4.3: F-tests for Seasonality and Adjustment Quality Diagnostics of Original LF

Test for the presence of seasonality assuming stability:				
	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between months	1514.9564	11	137.72331	181.173**
Residuals	36.4885	48	0.76018	
Total	1551.4449	59		
****Seasonality present at the 0.1 percent level**				
Nonparametric Test for the Presence of Seasonality Assuming Stability:				
	Kruskal-Wallis Statistic	Degrees of Freedom	Probability Level	
	56.7443	11	0.000%	
Seasonality present at the one percent level				
Moving Seasonality Test:				
	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between Years	2.7531	4	0.688283	0.898
Error	33.7353	44	0.766712	
No evidence of moving seasonality at the five percent level				
COMBINED TEST FOR THE PRESENCE OF IDENTIFIABLE SEASONALITY: IDENTIFIABLE SEASONALITY PRESENT				
Test for the presence of residual seasonality:				
No evidence of residual seasonality in the entire series at the 1 percent level. F = 0.01				
No evidence of residual seasonality in the last 3 years at the 1 percent level. F = 1.64				
No evidence of residual seasonality in the last 3 years at the 5 percent level.				
M1 = 0.532, M2 = 0.429, M3 = 0.856, M4 = 0.563, M5 = 0.433, M6 = 0.098, M7 = 0.390				
Q = 0.45				

Table 4.3 presents the full F-tests for seasonality of the original Load Factor (LF). The *combined test for the presence of identifiable seasonality* indicates that LF series has a seasonal pattern that can be identified by X-12 ARIMA. The M7 diagnostic (0.390 < 1) also strengthens the identifiability of the LF series with the aid of X-12 ARIMA. From the Table, the F- tests assert that seasonality exists in the monthly original LF series at 0.1% level of significance. But the good is that, before seasonal adjustment, the seasonality never passed to the years with a confidence of 95% and also the residuals of the series are free from seasonality at 1% significance level. In addition to M7, all the M-statistics are shown to be less than one, and hence the Q-statistic (0.45) produced from them is also less than one. This condition assures that the seasonal adjustment performed on LF is acceptable.

The remaining pre adjustment tests for Passenger Revenue (PR), Block Hours (BH), and Distance Flown (DF) can be seen from Table A1 (a– c) in Appendix. As of these Tables, similar tests are conducted for the remaining three series and based on the *combined tests* and M7 statistic (less than 1 for each variable) seasonal effect of all variables can be identified by X-12 ARIMA. The results of the F-tests indicate as seasonality existed in the monthly original series of PR, BH, and DF as well at 0.1% significance level. But for these three dataset, as that of the LF series, seasonality is not inherited by the yearly data from the monthly ones. At 1% level of significance, the residuals also do not exhibit any seasonality for all these variables as it is the case for LF series. These all Tables indicate that, in all the original dataset of this study, there is no seasonality across the years (moving seasonality) and residuals (residual seasonality) at 5% and 1% significance levels, respectively. Furthermore, all the M-statistics and the Q-statistics for each series do not exceed one, and hence all the seasonal adjustment procedures made on all series are acceptable.

4.3 Post- Seasonal Adjustment Features

A seasonally adjusted series should not have any estimable seasonal effects. A lack of residual seasonality is the most fundamental requirement of a good quality seasonal adjustment. This is to mean that no one should expect seasonality in either the seasonally adjusted series or irregular component. Some of the most important diagnostics in X-12 are the diagnostics to detect residual seasonality. Also, the level of the seasonally adjusted series should not have bias, meaning that the local levels of the original series and the seasonally adjusted series should be similar. X-12

ARIMA has procedures that give us unbiased estimates of the seasonal component, so there are no diagnostics present to check for bias.

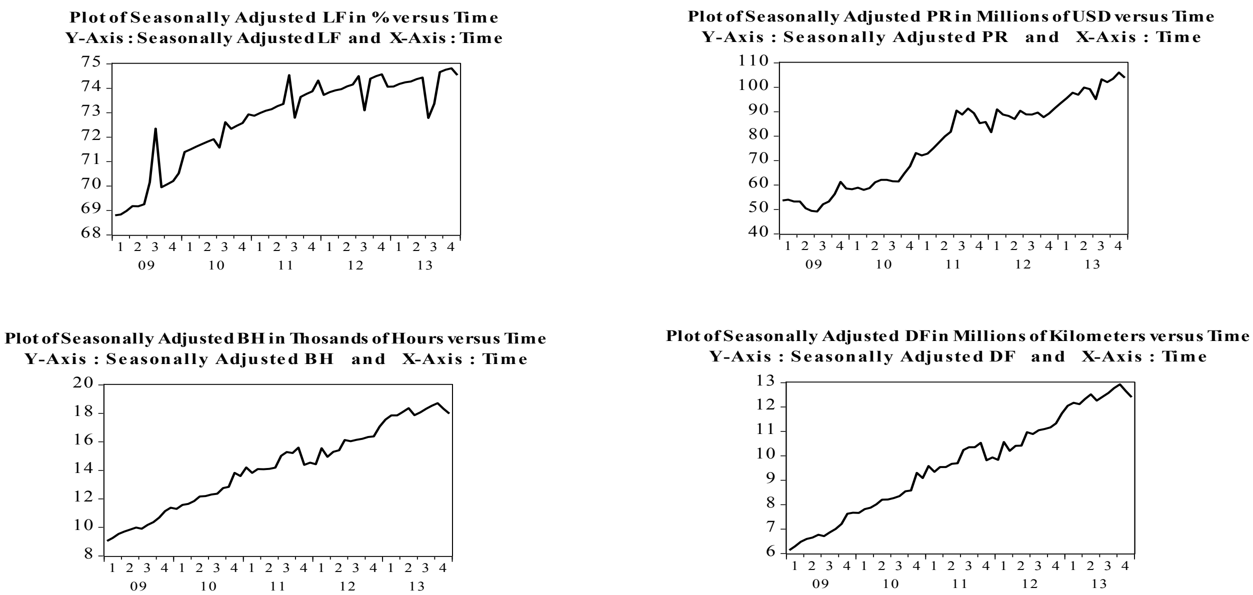
From Table 4.4 below, at 1 % level of risk, the *test for residual seasonality* (at 1% significance level) shows that there is no any estimable seasonal effect left in the seasonally adjusted series of LF and its irregular component as it is also indicated by the F- tests at 0.1% and 1% (for *Kruskal-Wallis test*) significance level. The *combined test for the presence of seasonality* together with M7 diagnostic (a value of 3 which is greater than 1) is also assuring that no more seasonal adjustment will be necessary at 1 % significance level. Similar deductions can be drawn from the *tests of residual seasonality* (with 99% confidence) and F-tests (with 99.9% confidence) in Tables A2 (a - c) in Appendix which exhibit that seasonality is completely removed from PR, BH, and DF series. Moreover, the *combined tests for seasonality* and M7 (greater than 1 for each variable) guarantee that no more seasonal adjustment is required for each variable.

Table 4.4: F-tests for Seasonality of Load Factor Series after Adjustment

Test for the presence of seasonality assuming stability:				
	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between months	0.0234	11	0.00213	0.004
Residuals	26.5476	48	0.55307	
Total	26.5710	59		
No evidence of stable seasonality at the 0.1 percent level				
Nonparametric Test for the Presence of Seasonality Assuming Stability:				
	Kruskal-Wallis Statistic	Degrees of Freedom	Probability Level	
	1.7659	11	99.916%	
No evidence of seasonality at the one percent level				
Moving Seasonality Test:				
	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between Years	0.7577	4	0.189416	1.214
Error	6.8632	44	0.155982	
No evidence of moving seasonality at the five percent level				
COMBINED TEST FOR THE PRESENCE OF IDENTIFIABLE SEASONALITY: IDENTIFIABLE SEASONALITY NOT PRESENT				
Test for the presence of residual seasonality:				
No evidence of residual seasonality in the entire series at the 1 percent level. F = 0.02				
No evidence of residual seasonality in the last 3 years at the 1 percent level. F = 1.38				
No evidence of residual seasonality in the last 3 years at the 5 percent level.				
M7 = 3.000				

In addition to the standard tests above, the time plots of each seasonally adjusted series are shown below in Figure 4.2. The plots confirm the fact attained from the above seasonality tests in that the peaks can be seen appropriately suppressed and the troughs are smoothed.

Figure 4.2: Post – Seasonal Adjustment Time Plots for each Series



4.4 Stationarity Test for Individual Series

In practice, using the non-stationary time series in VAR modeling is problematic with regard to statistical inference since the standard statistical tests used for inference are based on the condition that all of the series used must be stationary.

Thus, inspections and standard testes should be conducted on each variable for the presence of unit root(s) and in so doing the order of integration of each series is determined.

4.4.1 Visual Inspection

Time plot examination is among the first round works in time series analysis as it provides a hint and an easy outlook of the series' important properties. As shown in Figure 4.2 above, all the series are seasonally adjusted but they are still with an increasing pattern in line with time increment. This means that all the series are non-stationary. But the time plots should not be the only instruments to detect stationarity of the series. Rather the clue obtained from the time plots ought to be authenticated by standard tests for stationarity (unit root tests). This study utilizes the two most commonly applied unit root tests: Augmented Dickey-Fuller test and a Phillips and Perron test.

4.4.2 Unit Root Test

Unit root tests are confirmatory strive for stationarity detection. Augmented Dickey-Fuller test and a Phillips and Perron test are employed to test stationarity and determine the maximum order of integration of each series. The claims of these tests will be as follows.

H_0: The series is non stationary versus

H_1: The series is stationary

Table 4.5 and 4.6 below provide the outcome of ADF and PP tests when there is intercept without trend and when both intercept and trend exist at level and first difference, respectively, for each series. The critical values used for the tests are the McKinnon (1991) critical values. As asserted from Table 4.5, at level, one cannot have enough evidence to reject the null hypothesis that states the series contain a unit root (is non-stationary) for all the time series as the P-Values for each series are greater than 5% level of significance (α).

Table 4.5: Unit Root Test Results (At Level)

Series	With Intercept				With Intercept and Trend			
	Test Statistic		Prob.*		Test Statistic		Prob.*	
	ADF	PP	ADF	PP	ADF	PP	ADF	PP
LF	-1.44	-2.16	0.56	0.22	-2.93	-3.33	0.16	0.07
PR	1.34	1.26	0.99	0.99	-1.72	-1.91	0.74	0.65
BH	1.59	1.81	0.99	0.99	-2.19	-2.04	0.49	0.57
DF	1.94	2.11	0.99	0.99	-2.11	-1.93	0.53	0.64
Critical Value (5%)	-2.88				-3.44			

*MacKinnon (1996) one-sided p-values

Once it is confirmed that all the series are non-stationary, the next step is to go for differencing so as to make the data stationary. It is also at this instance that the orders of integration for the four non-stationary series of this study are determined. The order of integration is the number of unit roots that should be contained in the series until stationarity is achieved. Then exactly the same tests are performed to the first differences. Table 4.6 below implies, at first difference, the P-Values are less than 1% significance level for all the series. Meaning, all the first differences are stationary at 99 % confidence level.

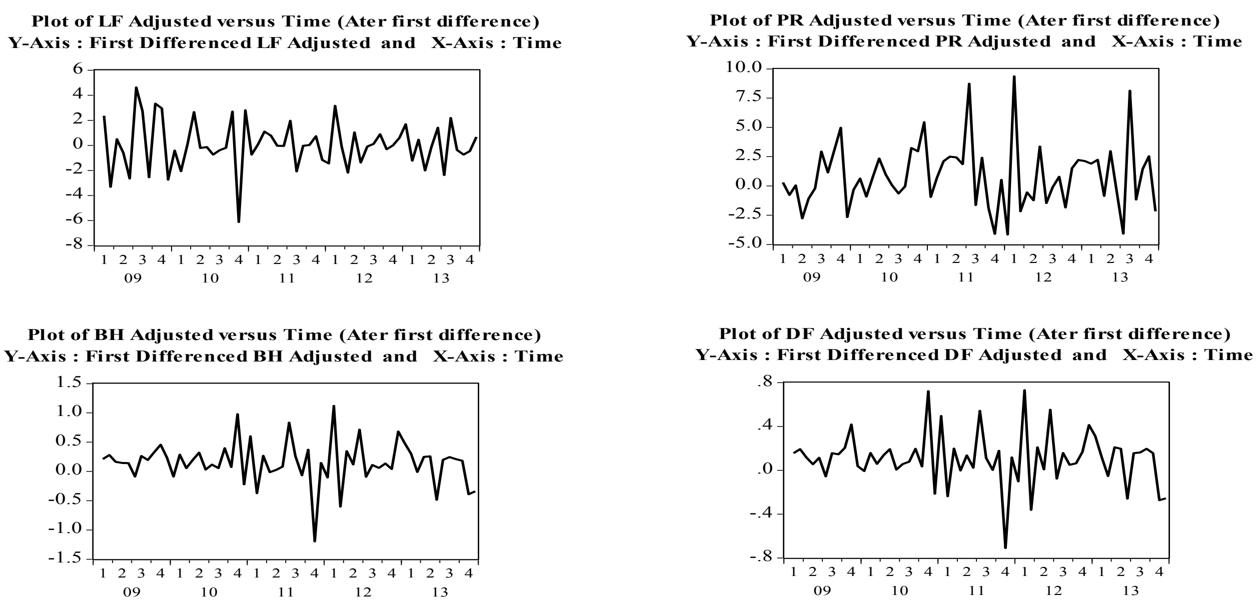
Table 4.6: Unit Root Test Results (After First Difference)

Series	With Intercept				With Intercept and Trend			
	Test Statistic		Prob.*		Test Statistic		Prob.*	
	ADF	PP	ADF	PP	ADF	PP	ADF	PP
LF	-13.07	-19.33	0.00	0.00	-13.03	-19.26	0.00	0.00
PR	-15.08	-15.01	0.00	0.00	-15.29	-15.31	0.00	0.00
BH	-15.12	-15.29	0.00	0.00	-15.45	-16.54	0.00	0.00
DF	-15.47	-15.42	0.00	0.00	-15.94	-16.83	0.00	0.00
Critical Value (5%)	-2.88				-3.44			

*MacKinnon (1996) one-sided p-values

Consequently, based on the ADF and PP test results, it can be concluded that all series are non stationary at level and stationary at first difference.

Figure 4.3: Time Plots of Seasonally Adjusted Series after First Difference



Furthermore, time plots for each seasonally adjusted and first differenced series are presented above in Figure 4.3. In the figure it can be clearly seen that there is no seasonality and upward or downward pattern with time, meaning, all the series are non-seasonal and stationary. Therefore, based on all the above methods of stationarity detection, the four time series are non-stationary originally and stationary after first differences. This implies all the series are integrated of order one (I(1)).

4.5 VAR Model Specification

4.5.1 Specification of VAR Order

A critical element in the specification of VAR models is the determination of the lag length of the VAR, since all inferences in the VAR model depend on the correct model specification. The importance of lag length determination is demonstrated by Braun and Mittnik (1993) who have shown that estimates of a VAR whose lag length differs from the true lag length are inconsistent as are the impulse response functions and variance decompositions derived from the estimated VAR. For example, Johansen (1991) and Gonzalo (1994) pointed out that VAR order selection can affect proper inference about cointegrating vectors and rank. Lütkepohl (1993) indicated that over-fitting (selecting a higher order lag length than the true lag length) causes an increase in the mean-square forecast errors of the VAR and that under-fitting the lag length often generates autocorrelated errors. Hafer and Sheehan (1989) found that the accuracy of forecasts from VAR models varies substantially for alternative lag lengths.

In this study, determination of optimal lag order for the VAR/VEC model is performed using the Akaike information criterion (AIC), Schwarz information criterion (SC), and Hannan-Quin(HQ) information criterion. In each criterion, the lag with a minimum criterion value is selected as an optimum lag length for the model. The results are shown in Table 4.7 below.

Table 4.7: VAR Lag Order Selection Results

Lag	AIC	SC	HQ
0	1.477747	11.74201	11.88164
1	0.004127*	5.859841*	6.557956*
2	0.004457	5.929077	7.185684
3	0.005027	6.030597	7.845695
4	0.005274	6.042776	8.416366
5	0.005334	5.994265	8.926347
6	0.006327	6.072719	9.563293
7	0.007247	6.072470	10.12154
8	0.009315	6.128377	10.73593

* indicates lag order selected by the criterion

From Table 4.7, the AIC, SC and HQ test suggest the appropriate lag length for the VAR model to be one (1) since the minimum AIC, SC and HQ values occur at lag one. Thus, it should be assumed that VAR (1) is the best for the data among all contender models.

4.5.2 Lag Exclusion Test

This test carries out confirmation for suitability of each lag selected by the above three criteria for the VAR. For each lag, the Chi square χ^2 (Wald) statistics of all variables are reported separately and jointly in Table 4.8 below.

Table 4.8: VAR Lag Exclusion Wald Tests

Chi-squared test statistics for lag exclusion:
Numbers in [ ] are p-values

	LF	PR	BH	DF	Joint
Lag 1	10.67937 [0.030414]	26.99636 [1.99e-05]	25.87778 [3.35e-05]	22.87876 [0.000134]	80.94704 [1.12e-10]
Lag 2	3.497042 [0.478328]	4.116885 [0.390418]	5.497757 [0.239927]	5.484011 [0.241139]	23.65859 [0.097231]
Lag 3	4.213082 [0.377936]	3.106088 [0.540232]	4.182982 [0.381808]	5.092473 [0.277940]	26.19202 [0.051386]
Lag 4	2.000285 [0.735706]	8.778823 [0.066872]	8.536231 [0.073797]	9.098116 [0.058693]	23.54798 [0.099852]
df	4	4	4	4	16

As it can be seen from Table 4.8 above, only the first lag is significant for LF at 5% significance level and for the remaining variables and for the joint at 1% significance level. Therefore, provided that VAR models usually need the same lag length for all the series, the Wald exclusion test assures that VAR (1) is found optimal for the data set and hence could be adopted.

4.5.3 Cointegration Analysis

So far it is noticed that the variables in the study are integrated of the same order one. This concrete fact paves the way for the necessity of cointegration test to see the existence of a linear combination(s) of the four variables which is (are) stationary (I (0)). This is what economically termed as *long term relationships* among the variables. The idea behind cointegration analysis is that, although variables may tend to trend up and down over time, groups of variables may drift together. To determine the number of cointegrating relationships, the Johansen (1995) approach of cointegration test is applied. The two tests for cointegration are the trace test and the maximum eigenvalue statistics. Here, these two test statistics are compared to special critical values to determine the number of cointegrating vector(s) in the model. The maximum eigenvalue and trace tests proceed sequentially from the first hypothesis – no cointegration – to an increasing number of co-integrating vectors.

From the results of Johansen cointegration test presented in Table 4.9 below, it can be observed that the trace or estimated LR statistic (63.35851) exceeds the respective critical value (47.85613) with P-value (0.0009). The maximum eigenvalue test also supports the same thing as the trace test. This implies that the null hypothesis of no cointegration relations is rejected at the 5% significance level in favour of the alternative one which states that there exists one cointegration relation. Therefore, the rank of cointegration matrix is equal to one, meaning, there is only one cointegrating equation in the system.

Table 4.9: Johansen Cointegration Test Results (By Assumption: Linear Deterministic Trend)

Hypothesised Number of Cointegration Equation(s)	Eigenvalue	Trace Test			Maximum Eigenvalue Test		
		Statistic	Critical Value (5%)	Prob.**	Statistic	Critical Value (5%)	Prob.**
None *	0.222015	63.35851	47.85613	0.0009	38.66131	27.58434	0.0013
At most 1	0.099832	24.69720	29.79707	0.1726	16.19672	21.13162	0.2136
Atmost 2	0.051055	8.500472	15.49471	0.4135	8.070276	14.26460	0.3716
Atmost 3	0.002790	0.430196	3.841466	0.5119	0.430196	3.841466	0.5119

Normalized cointegrating coefficients (standard error in parentheses)

LF	PR	BH	DF
1.000000	-0.540468	-0.927617	0.944683
	(0.03997)	(0.09354)	(0.09852)
	[-13.5210]	[-9.91697]	[9.58913]

* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Consequently, the cointegrating vector is given by

$$\beta = (1, -0.540468, -0.927617, 0.944683).$$

The values correspond to the cointegrating coefficients of LF (normalized to one), PR, BH, and DF, respectively.

As far as the main purpose of cointegration analysis is to get a stationary series from two or more nonstationary series, the resulting stationary series is written as a linear combination of the nonstationary series under study. Accordingly, if this stationary series is designated by S_t, then using the results obtained from Table 4.9 above S_t is given by

$$S_t = LF_t - 0.540468PR_t - 0.927617BH_t + 0.944683DF_t \quad (4.1)$$

(4.1) above enlightens that S_t is stationary in spite of the fact that all the four series are non-stationary.

4.6 Model Estimation

After deduction is made that variables in the VAR model appeared to be cointegrated, the immediate stride is to estimate the short run behavior and the adjustment to the long run models, which is represented by VECM. The VEC model has the following structure:

$$\Delta Y_t = \mu + \sum_{i=1}^{p} \Gamma_i \Delta Y_{t-i} + \alpha B X_{t-1} + \varepsilon_t \quad , \quad (4.2)$$

where BX_t is the error correction term given by $\beta^{'} Y_t$ and β is the cointegrating vector. The responses of LF, PR, BH and DF to short-term output movements are captured by the Γ_i coefficient matrices. The α coefficient vector reveals the speed of adjustments to the equilibrium, which measures the deviation from the long-run relationship between LF, PR, BH, and DF. Coefficient estimates of the VEC model are presented in Table 4.10 below. The Table is built from two components. The first part puts together the coefficients of the cointegration vector (long-term equilibrium equation) with their respective standard errors and t- statistics, which are derived by normalizing the Load Factor (LF). From this result, all of the variables are significant at the conventional significance levels, and it can be deduced that there exist long-run causal relationships among LF, PR, BH, and DF. This long-run equilibrium model is:

$$LF_t = 47.27715 + 0.540468PR_t + 0.927617BH_t - 0.944683DF_t \, . \quad (4.3)$$

(4.3) above indicates that, in the long run, a one million dollar increase in the monthly Passenger Revenue accounts for an average increase of about 0.54 % in the monthly Load Factor.

Likewise, a one thousand hours flight time increase per month will result in an average increase by around 0.93 % in the monthly Load Factor of the Ethiopian Aviation, in the long run. In contrast, a one million kilometer increase in the monthly flight distance, on average in the long run, will come up with a decrease of about 0.94 % in the load factor per month.

Table 4.10: Vector Error Correction Estimates

Vector Error Correction Estimates
Standard errors in ( ) & t-statistics in [ ]

Cointegrating Eq:	CointEq1			
LF(-1)	1.000000			
PR(-1)	-0.540468 (0.03997) [-13.5210]			
BH(-1)	-0.927617 (0.09354) [-9.91697]			
DF(-1)	0.944683 (0.09852) [9.58913]			
C	-47.27715			
Error Correction:	**D(LF)**	**D(PR)**	**D(BH)**	**D(DF)**
CointEq1	-0.588086 (0.18673) [-3.14934]	0.662929 (0.20171) [3.28654]	0.087086 (0.26189) [0.33253]	0.607007 (0.18972) [3.19943]
D(LF(-1))	-0.652171 (0.21241) [-3.07039]	0.867706 (0.38976) [2.22626]	0.384553 (0.45889) [0.83801]	0.022715 (0.02388) [0.95109]
D(PR(-1))	-0.057621 (0.09491) [-0.60711]	-0.278937 (0.16772) [-1.66314]	0.071514 (0.41041) [0.17425]	0.005741 (0.01512) [0.37975]
D(BH(-1))	0.757550 (0.187992) [4.02969]	0.971061 (0.36552) [2.65664]	-0.920145 (0.56427) [-1.63068]	-0.129807 (0.45777) [-0.28357]
D(DF(-1))	0.938681 (0.32081) [2.92601]	0.952187 (0.34654) [2.74773]	0.920495 (0.67738) [1.35890]	-0.114545 (0.67165) [-0.17054]
C	-0.596773 (0.31018) [-1.92396]	0.062378 (0.26496) [0.23543]	0.101714 (0.03383) [3.00644]	0.123822 (0.03513) [3.52441]
R-squared	0.651586	0.680120	0.123152	0.168651
Adj. R-squared	0.619367	0.651933	0.102185	0.131674
Sum sq. resids	34.72280	423.5239	7.712675	3.441194
F-statistic	4.571581	4.109382	2.112504	2.190930
Log likelihood	-68.72787	-143.7643	-23.59193	0.619467
Akaike AIC	3.853418	4.992143	0.986398	0.179351
Schwarz SC	4.062853	5.201577	1.195832	0.388786

The second part of Table 4.10 contains the coefficients of the error correction terms (cointEq1) for the cointegration vector. These coefficients are called the *adjustment coefficients*. They measure the short-run adjustments of the deviations of the endogenous variables from their long- run values. These first row coefficients identify the fraction of the long-term gap that is closed by each endogenous variable in each period (months). In another saying, these figures provide information on the short run disequilibria percentage adjustment of each endogenous variable within one period of time (month in this case).

But before going to construct the individual VEC models, each variable should be checked whether they are endogenous (can be separately modeled as dependent variable) or exogenous (can only be included as independent variable in other variable's VECM with its lagged values). This can be done through the following exogeneity test.

Granger Causality/ Block Exogeneity Wald Test

This test detects whether the lags of one variable can Granger-cause any other variables in the VEC system. The null hypothesis is that all lags of one variable can be excluded from each VECM. Table 4.11 below presents an exogeneity test when each series are treated as dependent (endogenous) variable. From the Table, $\Delta\mathrm{LF}$ (with a joint P-value of 0.0023 < 5% significance level) and $\Delta\mathrm{PR}$ (with a joint P-value of 0.0020 < 5% significance level) are endogenous variables and each of them can have an independent VECM. But a P-value of 0.1078 less than 5% significance level indicates that lagged difference values of PR can be excluded from VECM of LF. Thus, $\Delta\mathrm{PR}_{t-1}$ does not Granger-cause $\Delta\mathrm{LF}$. In contrary, $\Delta\mathrm{BH}$ and $\Delta\mathrm{DF}$ are exogenous variables with joint P-values of 0.7503 and 0.6156, respectively, greater than 5% level of significance. That is, $\Delta\mathrm{BH}$ and $\Delta\mathrm{DF}$ cannot have an explicitly written VEC models.

Thus from the two parts of VEC estimates in Table 4.10, the following two VECMs can be straightforwardly estimated only for endogenous variables (LF and PR) by introducing the error correction term as another independent variable in the restricted VAR model.

VEC Model of Load Factor:

$$\begin{aligned}\Delta\mathrm{LF} = &-0.59[\mathrm{LF}_{t-1} - 0.54\mathrm{PR}_{t-1} - 0.93\mathrm{BH}_{t-1} + 0.94\mathrm{DF}_{t-1} - 47.3] - 0.65\Delta\mathrm{LF}_{t-1} \\ &- 0.06\Delta\mathrm{PR}_{t-1} + 0.76\Delta\mathrm{BH}_{t-1} + 0.94\Delta\mathrm{DF}_{t-1} - 0.59 \qquad (4.4)\end{aligned}$$

VEC Model of Passenger Revenue:

$$\Delta PR = 0.66[LF_{t-1} - 0.54PR_{t-1} - 0.93BH_{t-1} + 0.94DF_{t-1} - 47.3] + 0.87\Delta LF_{t-1} - 0.28\Delta PR_{t-1} + 0.97\Delta BH_{t-1} - 0.95\Delta DF_{t-1} + 0.062 \quad (4.5)$$

Where, ‗Δ' denotes first difference (D), the value in the closed bracket is the error correction term and the coefficients of error correction term are called *adjustment coefficients*.

Table 4.11: VEC Granger Causality/Block Exogeneity Wald Tests

Dependent variable: D(LF)			
Excluded	Chi-sq	df	Prob.
D(PR)	2.585663	1	0.1078
D(BH)	12.80246	1	0.0003
D(DF)	12.38657	1	0.0004
All	14.48618	3	0.0023
Dependent variable: D(PR)			
Excluded	Chi-sq	df	Prob.
D(LF)	6.208201	1	0.0127
D(BH)	8.623484	1	0.0033
D(DF)	8.236895	1	0.0041
All	9.579177	3	0.0020
Dependent variable: D(BH)			
Excluded	Chi-sq	df	Prob.
D(LF)	0.790258	1	0.3740
D(PR)	0.001542	1	0.9687
D(DF)	0.057172	1	0.8110
All	1.211434	3	0.7503
Dependent variable: D(DF)			
Excluded	Chi-sq	df	Prob.
D(LF)	1.043092	1	0.3071
D(PR)	4.76E-06	1	0.9983
D(BH)	0.007355	1	0.9317
All	1.797007	3	0.6156

Therefore, from (4.4) and (4.5) above it can be realized that, each month, 59 % and 66 % of the long term gaps are closed by LF and PR, respectively. That is, 59% and 66% of the short run disequilibria in LF and PR are adjusted within one month, respectively. On other hand, the long term BH and DF gaps are closed by about 8.7% and 61% in each month as it can be referred from Table 4.10, respectively. It is also possible to say that 8.7% and 61% of the short run disequilibria in BH and DF are adjusted within a single month. These results imply that BH and PR have the shortest and longest speed, respectively, to get back to the equilibrium after a shock. LF and DF share almost equal speed to achieve equilibrium after a shock. Additionally, LF is significantly affected by BH, DF and its own lagged values in the short-run. On its part, PR is significantly determined by lagged values of all the variables, except its own lagged values, in the short run. However, BH and DF are insignificantly affected by all of the variables in the short run.

4.7 Model Checking

Subsequent to model development, it is necessary to verify whether the fitted model is suitable. All the time it is after model validity examination that forecasting will be made.

4.7.1 Test of Residual Autocorrelation

Table 4.12 below presents the results of the Portmanteau Q-statistic and Lagrange Multiplier (LM) test for the whole VEC model residual serial correlation.

Table 4.12: Test of Residual Autocorrelation

Lag	Q-Stat		Adj Q-Stat		LM-Stat	
	Value	Prob.	Value	Prob.	Value	Prob.
1	2.432637	NA*	2.473868	NA*	14.38020	0.5704
2	18.54117	0.9117	19.13787	0.8938	19.07306	0.2649
3	42.37984	0.5412	44.23121	0.4619	22.35275	0.1322
4	54.94693	0.6604	57.69594	0.5604	13.66433	0.6237
5	70.40964	0.6593	74.56435	0.5251	16.42390	0.4238
6	87.10019	0.6249	93.10941	0.4481	17.37960	0.3615
7	101.1300	0.6672	108.9922	0.4552	15.29466	0.5032
8	108.5863	0.8364	117.5957	0.6448	8.167734	0.9437
9	121.3304	0.8706	132.5887	0.6594	15.63678	0.4786
10	132.5038	0.9139	145.9968	0.7057	13.76859	0.6159
11	147.6776	0.9103	164.5769	0.6443	23.66651	0.0970
12	164.7226	0.8886	185.8831	0.5300	21.16612	0.1722

***The test is valid only for lags larger than the VAR lag order.**

The tests in Table 4.12 above are used to test for the overall significance of the residual autocorrelations up to lag 12. Both tests imply that residuals do not suffer from autocorrelation problem up to lag 12 as all *p*-values go beyond the 5% level of risk.

4.7.2 Testing Normality

Multivariate version of the Jarque-Bera tests is used to test the normality of the residuals. It compares the 3rd and 4th moments (skewness and kurtosis) to those from a normal distribution. The test has null hypothesis indicating that the error term in the model has skewness and kurtosis corresponding to a normal distribution.

Table 4.13: Normality Test

Component	Skewness		Kurtosis		Jarque-Bera Statistic	
	Value	Prob.	Value	Prob.	Value	Prob.
1	-0.512526	0.1051	4.331375	0.0353	7.058224	0.0293
2	0.409035	0.1958	3.384622	0.5431	2.042931	0.3601
3	-0.616863	0.0511	3.445014	0.4817	4.300290	0.1165
4	-0.083268	0.7923	2.474528	0.4061	0.759639	0.6840
Joint		0.0854		0.2001		0.077

The results in Table 4.13 show that there is no evidence to reject the null hypothesis of normality for the whole VEC model residuals. So from the Jarque-Bera test, it can be deduced that the residuals fulfill the normally assumption. In addition, the individual VEC models are examined whether their residuals are normally distributed or not. As it can be referred from Figure A5 (a) and (b), which contain Jarque-Bera statistic with their respective P-values (both less than 5%), the residuals for LF and PR VEC models follow normal distribution.

4.8 Structural Analysis

4.8.1 Granger Causality Test

Granger causality test is considered a useful technique for determining whether one time series is good for forecasting the other. Table 4.14 below presents results from the pair wise Granger causality tests at 5% significance level. The result shows that at 95% confidence level, Block Hours (BH) and Distance Flown (DF) Granger cause the Load Factor (LF) but the converses do not hold. Passenger Revenue (PR) does not Granger cause LF. That is, only the change in PR does not account for the change in LF. Beside, PR is Granger caused by all the variables but the reverses fail. That is, the changes in all variables will result in the change in PR.

Table 4.14: Pairwise Granger Causality Tests

Null Hypothesis:	Obs	F-Statistic	Prob.
PR does not Granger Cause LF	59	0.22910	0.6341
LF does not Granger Cause PR		7.32725	0.0090
BH does not Granger Cause LF	59	11.6683	0.0008
LF does not Granger Cause BH		2.98006	0.0863
DF does not Granger Cause LF	59	11.4995	0.0009
LF does not Granger Cause DF		2.50532	0.1155
BH does not Granger Cause PR	59	13.5908	0.0003
PR does not Granger Cause BH		0.00248	0.9604
DF does not Granger Cause PR	59	13.8178	0.0003
PR does not Granger Cause DF		0.00616	0.9375
DF does not Granger Cause BH	59	1.74584	0.1884
BH does not Granger Cause DF		0.05641	0.8126

4.8.2 Impulse-Response Functions

Impulse responses trace out the reaction of the variables in the VAR to shocks of each variable. Therefore, for each variable a unit shock is applied to the error and the effects upon the VAR system over time are noted. Thus, if there are m variables in a system, a total of m^2 impulse responses could be generated. A standard Cholesky decomposition is used in order to identify the short run effects of shocks on the levels of the endogenous variables in the VECM.

The x-axis in Figure A1 (a) and (b) in the Appendix part provides the time horizon or the duration of the shock whilst the y-axis gives the direction and intensity of the impulse or the percent variation in the dependent variable away from its base line level. In our case there are 8 potential impulse response functions. The outcomes and combined graphs of these IRF functions are given in Table A5 ((a) and (b)) and Figure A1 ((a) and (b)) of Appendix with the Cholesky ordering of LF, PR, BH, and DF, respectively.

Figure A1 (a) shows the responses of LF, PR, BH, and DF with respect to one standard deviation innovation in LF. The result indicates LF innovations have a positive impact on PR and BH. For PR, it displays a slow rising trend until it reaches 0.072 intensity value and it calms down at around 2 month time horizon. Similarly the shocks of LF show a positive effect upon BH with a fast intensifying up to an impulse intensity of 0.23, then declines rapidly to an intensity of 0.092 and finally stabilizes at 3 month time horizon. But the one standard shocks in LF have totally a negative impact on DF. This impact initially exhibits a slight decrease up to a -0.2 and becomes constant after 4 month time horizon.

Figure A1 (b) shows that the effects of a one standard deviation shock in PR on the remaining variables. From the figure, the shocks have a positive response for LF. That is, it reveals a sluggish diminishing pattern up to 1.11 level of intensity and moves upward moderately until it becomes steady at around 3. The shocks have also a parallel effect in BH but an opposite response in DF.

4.8.3 Forecast Error Variance Decomposition

Variance decompositions offer a slightly different method for examining VAR system dynamics. The decomposition used to understand the proportion of the fluctuation in a series explained by its own shocks versus shocks from other variables. In general one expects a variable to explain almost all its forecast error variance at short horizons and smaller proportions at longer horizons. The decomposition results of the models of endogenous variables (LF and PR) are presented in Table A6 ((a) and (b)) and plotted in Figure A2 ((a) and (b)). These two results provide the forecast error percentage in each variable that could be attributed to innovations of the other variables, for different time periods. The Cholesky ordering employed is LF, PR, BH, and DF.

The variance decomposition analysis result of Figure A2 (a) shows that, at the first horizon, variation of LF is explained only by its own shock. In the second month 97.83 % of the variability in the LF fluctuations is explained by its own innovations and the remaining 2.17% is explained by BH (1.71%), DF (0.29%), and PR (0.17%). Even up to the tenth month, much of the variability of LF (85.93%) is explained by its own shock and the rest portion is occupied by DF (8.74%), BH (4.01%), and PR (1.32%). It can also be observed that, after ten months, the variability of LF determined by DF has shown an increment to 8.74% and the LF shock revealed a total of 14.07% decrement. However, the percentages of BH and PR to LF variability explanation seem to never increase beyond 5% and 2%, respectively, even after large amount of duration.

In a similar fashion, Figure A2 (b) displays that, in the first month, 81 .25% of the variability of PR is explained by its own shock and 18.75% is determined by LF. After ten months the variability of PR explained by its shocks and LF attained 70.05% and 26.85%, respectively.

4.9 Forecasting

Forecasting is a deep-seated ambition of time series analysis or developing a time series model. The previous discussions validate that vector error correction model of order one is the paramount model to suitably describe the series of this study. This section conducts an examination on the forecasting accuracy of the fitted model and then makes a forecast for January 2014 to December 2014. Meaning, one year ahead forecast is made and can be seen from Table 4.16.

4.9.1 Evaluation of Accuracy

The mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE) and Theil U statistics are used to assess the forecasting performance. The RMSE and MAE statistics are scale-dependent measures, but allow a comparison between the actual and forecast values. The Theil-U statistics is independent of the scale of the variables and is constructed to lie between zero and one, zero indicating a perfect fit. In evaluating the performance of the forecasting models, the lower the RMSE, MAE, MAPE and Theil-U statistic, the better the forecasting accuracy.

In practice, to evaluate the out-of-sample forecasting ability of the model it is recommended to divide the available data set into two subsamples. The first subsample of the data is used to build (estimate) the VEC model, and the second subsample (that contains end time series values of whole period covered which are not used to estimate the model) is used to evaluate the forecasting performance of the model. Usually the two subsamples of data are referred as estimation and forecasting subsamples, respectively. Therefore, the data from January 2009 up to December 2012 are utilized to estimate the VEC model of order one and then the forecast is performed for the time period January 2013 – December 2013. Table 4.15 reports the forecasting accuracy statistics of the estimated model. The result indicates that all estimated models are good enough to describe the series. For each endogenous series (LF and PR) RMSE, MAE, MAPE values are less than 5% and Theil-U statistics are close to zero, which indicate the differences between the actual values and the predicted values are small. That is, the predictive powers of the models are healthier and suitable for n-step ahead forecasting.

Table 4.15: Forecasting Accuracy Statistics

Forecast sample: January 2013 to December 2013		
	Variables	
Accuracy Measures	LF	PR
Root Mean Squared Error	1.090529	4.369759
Mean Absolute Error	0.891087	3.970237
Mean Absolute percent error	1.249139	3.931896
Theil Inequality Coefficient	0.007653	0.022276

4.9.2. Out-of-Sample and In-Sample Forecasting Analysis

Out-of-sample forecasted values for the series under study, using the vector error correction model, are presented in Table 4.16 below.

Table 4.16: Forecasts from the VECM (1) Models

Months		LF	PR
January	**2014**	71.39764	92.39132
February	**2014**	71.31635	93.70771
March	**2014**	71.47449	94.64862
April	**2014**	71.53258	93.66481
May	**2014**	71.78363	94.04010
June	**2014**	71.53519	95.94809
July	**2014**	71.22674	98.67355
August	**2014**	71.18153	99.72597
September	**2014**	70.76529	102.4194
October	**2014**	70.95727	103.9212
November	**2014**	70.94412	104.6392
December	**2014**	70.77500	105.7876

Besides, Figures A3 and A4 in Appendix reveal the forecasts (in-sample) of LF and PR (from their individual VEC models) graphed together with the corresponding actual series and residuals in a comparable mode for the time phase form January 2009 up to December 2013. It can be easily observed from the figures that the forecast series are much closer to the actual series.

5. CONCLUSION

Ethiopian Airline is the fastest growing Airline service provider in the world. One pointer to know the extent of service capacity expansion is Load Factor which is measured in percentage. A diminishing pattern in this quantity can be a best indicator of the Airline‘s poor expansion in accommodating a large number of passengers. In converse, a considerable rise in the growth of Load Factor points out an admirable performance of the aviation.

This study examines the long-run dynamic relationship between Load Factor, Passenger Revenue, Block Hours, and Distance Flown of Ethiopian Airline. The ultimate objective is to investigate whether Load Factor is related to the remaining variables in the long and short run by making use of records on each series from January 2009 – December 2013. VEC (Vector Error Correction) of order one is utilized.

Initially, all the series were identified to be seasonal and non-stationary using F-tests and M7 diagnostics for seasonality, and then employing Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) unit root tests for stationarity. Thus, all the series are firstly adjusted for seasonality then differenced once to make them non- seasonal and stationary series. Using AIC, SC, and HQ lag order selection criteria, the appropriate lag was found to be one and optimality test (lag exclusion test) of lag length also approved the selected lag order. On the other hand, Johansen cointegration test suggested the existence of only one cointegrating vector at 95% confidence level and it has been clearly identified that VECM(1) is the best among all competitor models to fit the data. But before constructing the individual VEC models exogeneity test was employed then LF and PR were found to be endogenous, whereas BH and DF were identified to be exogenous variables. From the VEC models it can be realized that LF is significantly determined by all the variables in the long-run, while it is considerably affected by BH, DF and its own lagged values in the short-run. Specifically, in the long-run, a one million dollar increase in the monthly Passenger Revenue and a one thousand hours flight time increase per month, account for an average increase of about 0.54 % and 0.93 % in the monthly Load Factor. In contrary, a one million kilometer increase in the monthly flight distance contributes a decrease of about 0.94 % for the monthly Load Factor of the Ethiopian International Aviation, in the long run. Additionally, 58.81% of the short run disequilibrium in LF is adjusted within one month. Jarque-Bera verified that residuals are normally distributed, whilst Portmanteau Q-statistic tests confirmed that residuals do not exhibit autocorrelation. Granger causality test also provided information at 5 % significance level that BH and DF can significantly Granger cause the Load Factor, but PR has insignificant causal effect on LF. The converses did not hold for any of the variables.

Impulse response functions (IRF) and forecast error variance decomposition (FEVD) were also done. The IRF outcomes show that a one standard deviation innovation in LF accounts for a positive but slight impact on PR and BH as well as a negative effect on DF. A one standard deviation shock in PR is found to have a positive consequence on LF and also the same effect as LF did on DF and BH.

A support was also given to the IRF end results by FEVD. From the FEVD it is understood that although most proportion of the variation in LF is explained by its own innovations, DF has shown up to share this variability explanation with 8.74 % after 10 months. Some percent of PR is also explained by LF next to its own shocks.

REFERENCES

ANDERSON, O.D. (1976): ―*Time Series Analysis: The Box-Jenkins Approach*", Butterworth, London.

ANDERSON, J.E. and KRAUS, M. (1980): ―*Quality of Service and the Demand for Air Travel*", The Review of Economics and Statistics, pp. 533-540.

ANDREW, S. (1999): ―*Seasonal Adjustment: Comparison of Philosophies*", Australian Bureau of Statistics, Working Paper No. 99/2, (December).

ARMSTRONG, J.S. and FILDES, R. (1995): ―*Correspondence On the Solution of Error Measures for Comparisons Among Forecasting Methods*", Journal of Forecasting, Vol. 14, pp. 67-71.

ASSOCIATION OF EUROPEAN AIRLINES (1998): ―*Medium-Term Forecast of European Scheduled Passenger Traffic, 1994-1998*".

AUSTRALIAN BUREAU OF STATISTICS (2005): ―*An Introductory Course on Time Series Analysis*", (January).

BAHRAM, A., RICHARD, D.G., and KAMBIZ, R., (2002): ―*The Long –Run Dynamic Equilibrium between Air Carrier Capacity and Profitability*"**.**

BARNABÁS, F and ZOLTÁN, M.J. (2002): ―*Manual to Hungarian Economic Statistics*", (December).

BARON, R. (1973): ―*Analysis of Seasonality and Trends in Statistical Series – Methodology and Applications in Israel*", Technical Publication No. 39, Publication conducted in 2011, Jerusalem: Central Bureau of Statistics, Israel.

BELL, W.R. and HILLMER, S.C. (1984): ―*Issues Involved with the Seasonal Adjustment of Economic Time Series*", Journal of Business and Economic Statistics, 2, 291-320.

BERA, A.K. and JARQUE, C.M. (1981): ―*An Efficient Large-Sample Test for Normality of Observations and Regression Residuals*", Australian National University, Working Paper in Economics and Econometrics 40.

BOEING COMMERCIAL AIRPLANE COMPANY (1987): ―*Air Freight, Air Travel, and Strategic Planning*", London.

BOEING COMMERCIAL AIRPLANE COMPANY (1993): ―*Proceedings of Airline Evaluation Seminar, Airline Evaluation Seminar*", London.

BOX, G.E.P. and JENKINS, G.M. (1976): ―*Time Series Analysis Forecasting and Control*", San Francisco, Holden-Day, (2nd Ed).

BRAUN, P. A. AND S. MITTNIK (1993): ―*Misspecifications in Vector Autoregressions and Their Effects on Impulse Responses and Variance Decompositions*”, Journal of Econometrics 59 (October), 319 – 41.

Canner, W.H.P. (1986): ―*The Air Transport Industry*”, Publishing Limited, Brown, Son & Ferguson, (1st Ed.).

CLEVELAND, W. S., DEVLIN, S. J. and TERPENNING, I. J. (1982): ―*The SABL Seasonal and Calendar Adjustment Procedures*”, In O. D. Anderson, editor: ―*Time Series Analysis: Theory and Practice 1*”, pp. 539-564, North- Holland, Amsterdam.

CLEVELAND, R.B., CLEVELAND, W.S., MCRAE, J.E., and TERPENNING, I. (1990): ―*STL: A Seasonal-Trend Decomposition Procedure Based on Loess*”, Journal of Official Statistics, Vol. 6, No. 1, 3 – 73.

CONSTANTINOS, B. (2013): ―*Forecasting Air Passenger Traffic Flows in Canada: An Evaluation of Time Series Models and Combination Methods*”, Master‘s Thesis, Laval University, Canada.

DAGUM, E. B. (1980): ―*The X11-ARIMA Seasonal Adjustment Method*”, Statistics Canada, Ottawa.

DAGUM, E. B. (1988): ―*The X-11-ARIMA/88 Seasonal Adjustment Method – Foundations and User's Manual*”, Statistics Canada, Ottawa.

DICKEY, D.A., and W.A. FULLER (1979): ―*Distribution of the Estimators for Autoregressive Time Series with a Unit-root*”, Journal of the American Statistical Association, 74, 427-431.

DICKEY, D.A., and W.A. FULLER (1981): ―*Likelihood Ratio Statistics for Autoregressive Time Series with a Unit-root*”, Econometrica, 49, 1057-1072.

ENGLE, R.F. AND GRANGER, C.W. (1987): ―*Cointegration and Error Correction: Representation, Estimation and Testing*”, Econometrica, 55,251-276.

ETHIOPIAN AIRLINE (2008): ―*Ethiopian Annual Report of 2007-2008:* http://www.ethiopianairlines.com/en/corporate/annualreport.aspx”.

ETHIOPIAN AIRLINE (2009): ―*Ethiopian Annual Report of 2008-2009:* http://www.ethiopianairlines.com/en/corporate/annualreport.aspx”.

ETHIOPIAN AIRLINE (2010): ―*Ethiopian Annual Report of 2009-2010:* http://www.ethiopianairlines.com/en/corporate/annualreport.aspx”.

ETHIOPIAN AIRLINE (2011): ―*Ethiopian Annual Report of 2010-2011:* http://www.ethiopianairlines.com/en/corporate/annualreport.aspx”.

ETHIOPIAN AIRLINE (2012): ―*Ethiopian Annual Report of 2011-2012:* http://www.ethiopianairlines.com/en/corporate/annualreport.aspx”.

FINDLEY, D. F., MONSELL C. M., BELL, W. R., OTTO, M. C., AND CHEN, B. (1998): ―*New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program*", Journal of Business and Economic Statistics, 16.

FRIEDSTORM, L. AND THUNE-LARSEN (1989): ―*A Econometric Air Travel Demand Model For the Entire Conventional Domestic Network: The Case of Norway*", Transportation Research B, vol. 2313 No.3, pp. 213-223.

GANG X., SHOUYANG, W., AND KIN, K.L. (2013): ―*Air Passenger Forecasting by Using a Hybrid Seasonal Decomposition and Least Squares Support Vector Regression Approach*", China.

GARY, F. (2008): ―*Uncovering Trends in Seasonal Transportation Data: New Tool Analyzes Airline Revenue Passenger-Mile Trend Data, Adjusting RPM Data To Remove Seasonality*", U.S. Department of Transportation Research and Innovative Technology Administration, (May).

GHOBRIAL, A. (1992): ―*Aggregate Demand Model for Domestic Airlines*", Journal of Advanced Transportation, Alberta 26 (2), 3 – 14.

GHYSELS AND OSBORN (2001): ―*The Econometric Analysis of Seasonal Data*".

GÓMEZ, V. AND MARAVALL, A. (1997): ―*Programs TRAMO and SEATS; Instructions for the User*", Working Paper 9628. Servicio de Estudios. Banco de España.

GONZALO, J. (1994): ―*Five Alternative Methods of Estimating Long-Run Equilibrium*", Journal of Econometrics, 60:203–233.

GRANGER, C.W.J. (1969): ―*Investigating Causal Relations by Econometric models and Cross Spectral Methods*", Econometrica, 4, 112-120.

GROSCHE, T., ROTHLAUF, F., AND HEINZL, A. (2007): ―*Gravity Models for Airline Passenger Volume Estimation*", Journal of Air Transport Management, 13(4), 175–183.

GUJARATI, D.N. (2004): ―*Basic Econometrics*", McGraw-Hill Inc., New York, (4^{th} Ed.).

HAFER, R. W. AND R. G. SHEEHAN (1989): ―*The Sensitivity of VAR Forecasts to Alternative Lag Structures*", International Journal of Forecasting 5, 399-408.

HAMILTON, J.D. (1994): ―*Time Series Analysis*", Princeton University Press, Princeton.

HARVEY, A. (1990): ―*Forecasting, structural time series models, and the Kalman filter*", Published by Cambridge University Press.

HIGGINSON, J. (1976): ―*An F-test for the Presence of Residual Seasonality in Table D11 of the X-11 Seasonal Adjustment Program*", Research paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.

INOUE, G. AND NIU, K. (2008): ―*Analysis of International Air Passenger Flow in Asia and Study of Improvement of Air Travel Demand Predictions*", Airport Department.

IPPOLITO, R.A. (1981): ―*Estimating Airline Demand with Quality of Service Variables*", Journal Of Transport Economic and Policy (January), pp. 7-15.

JOHANSEN, S. (1988): ―*Statistical Analysis of Cointegration Vectors*", Journal of Economic Dynamics and Control, 12, 231-254.

JOHANSEN, S. AND K., JUSELIUS (1990): ―*Maximum Likelihood Estimation and Inference on Cointegration—with Applications to the Demand for Money*", Oxford Bulletin of Economics and Statistics, 52, 169–210.

JOHANSEN, S. (1991): ―*Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models*", Econometrica, 59, 1551-1580.

KAJAL, L. AND WENXIONG, Y., HERMAN S., AND PEG, Y. (2003): ―*Monthly Output Index for the U.S. Transportation Sector*".

KOOPMAN, S. J., HARVEY, A. C., DOORNIK, J. A .AND SHEPHARD N. (1995): ―*Stamp 5.0 (Structural Time Series Analyzer, Modeler and Predictor*", International Thompson Publishing, London.

KULENDRAN, N. AND WITT, S. (2003): ―*Forecasting the Demand for International Business Tourism*", Journal of Travel Research, 41(3):265–271.

KUMAR, V., AND STEPHANEDES (1988): ―*Dynamic Forecasting of Demand and Supply In Nonstop Air Routes*", Transportation Research Record, Vol. 1158, pp. 14-20.

LOTHIAN, J., AND MORRY, M. (1978): ―*A Set of Quality Control Statistics for the X-11-ARIMA Seasonal Adjustment Method*", Research Paper, Seasonal Adjustment and Time Series Staff, Statistics Canada.

LÜTKEPOHL, H. (1991): ―*Introduction to Multiple Time Series Analysis*", Springer-Verlag, Berlin.

LÜTKEPOHL, H. (1993): ―*Introduction to Multiple Time Series Analysis*", Berlin:Springer-Verlag, (Second Ed.), Chapter 4.

MACKINNON, J.G. (1991): ―*Critical Values for Cointegration Tests*". In R. F. Engle and C. W.J. Granger (eds): ―*Long-run Economic Relationships: Readings in Cointegration*", Ch. 13, pp. 267–76. Oxford: Oxford University Press.

MACKINNON, J.G. (1996): ―*Numerical Distribution Functions for Unit-root and Cointegration Tests*", Journal of Applied Econometrics, 11, 601–18.

MARCEL, V.V., ROBERTO, W., AND PIM, O. (2011): ―*Method series: Seasonal Adjustment, Statistical Methods*", Statistics Netherlands, The Hague, Heerlen.

MCDONALD-JOHNSON, K., MONSELL, B., FESCINA, R., FELDPAUSCH, R., HOOD, C., and WROBLEWSKI, M. (2006): ―*Seasonal Adjustment Diagnostics: Census Bureau Guideline*", U.S. Census Bureau, U.S. Department of Commerce, Washington, D.C.

OBERHAUSEN, P.J. and KOPPELMAN, F.S. (1978): ―*Time-Series Analysis of Intercity Air Travel*", Research Board, National Research Council, Washington, D.C., U.S.A.

OTT, J. (1990): ―*Air Transport: Airport Official Blast Carrier Marketing Tactics for Connecting Flights*", Aviation Week & Space Technology, (December 17), pp. 33-43.

PETER, L. (2005): ―*Seasonal Adjustment*", edited by Head of section Christian Harhoff, Statistics Denmark, (January).

PHILLIPS, P.C.B. (1986): ―*Understanding Spurious Regressions in Econometrics". Journal of Econometrics*", 33, 311-340.

PHILLIPS, P. C. B. and PERRON, P. (1988): ―*Testing for a Unit-root in Time Series Regression*", Biometrika, 75, 335–346.

POLE, A.M. and WEST, P.H (1994): ―*Applied Bayesian Forecasting and Time series Analysis*", Chapman-hall., New York.

POORE, J.W. (1993): ―*Forecasting the Demand for Air Transportation Services*", Journal of Transportation Engineering, Virginia 19 (5), 22-34.

RIZA, O.Y.(2011): ―*Air Passenger Demand Forecasting for Planned Airports, Case study: Zafer and Or-Gi Airports In Turkey*", Master's Thesis , Department of Civil Engineering, Middle East Technical University, (January).

SAID and DICKEY (1984): ―*Testing for unit roots in autoregressive-moving average models of unknown order*", East Carolina University, U.S.A.

SERAJ, Y.A., ABDULLAH, O.B., SAJJAD, M.J. (2001): ―*An Econometric Analysis of International Air Travel Demand in Saudi Arabia*", Journal of Air Transport Management, Jeddah, Saudi Arabia, 7, 143 – 148.

SHISKIN, J., YOUNG, A. H. AND MUSGRAVE, J.C. (1967): ―*The X11 Variant of the Census Method II Seasonal Adjustment Program*", Technical Paper 15, Bureau of the Census, U.S. Department of Commerce, Washington, D.C.

SIMS, C.A. (1980): ―*Macroeconomics and Reality*", Econometrica, 48, 1 – 48.

SIMS, C. A., STOCK, J. H. & WATSON, M. W. (1990): ―*Inference in Linear Time Series Models with Some Unit Roots*", Econometrica, 58(1), 113-144.

SINGAPORE DEPARTMENT OF STATISTICS (2006): ―*Seasonal Adjustment of Economic Time Series*", (September).

TRANI, A. AND TEODOROVIC, D. (2003): ―*Demand in Air Transportation Systems*", Department of Civil and Environmental Engineering, Virginia Tech., (June).

TSEKERIS, T. (2009): ―*Dynamic Analysis of Air Travel Demand in Competitive Island Markets*", Journal of Air Transport Management, 15(6): 267 – 273.

WEIGEND, A., and N., GERSHENFELD (1993): ―*Time Series Prediction-Forecasting the Future and Understanding the Past*", Addison-Wesley, Reading, Massachussets.

WORLD BANK (2012): http://www.worldbank.org, (March).

APPENDIX

Table A1: Seasonality F-Tests and Adjustment Quality Diagnostics of Original Series

Table A1 (a): F-tests for Seasonality and Adjustment Quality Diagnostics of Original PR

Test for the presence of seasonality assuming stability:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between months	7302.5129	11	663.86481	130.619**
Residuals	243.9578	48	5.08246	
Total	7546.4707	59		

****Seasonality present at the 0.1 per cent level**

Nonparametric Test for the Presence of Seasonality Assuming Stability:

Kruskal-Wallis Statistic	Degrees of Freedom	Probability Level
56.3639	11	0.000%

Seasonality present at the one percent level

Moving Seasonality Test:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between Years	24.8880	4	6.221996	1.542
Error	177.5488	44	4.035201	

No evidence of moving seasonality at the five percent level

COMBINED TEST FOR THE PRESENCE OF IDENTIFIABLE SEASONALITY:

IDENTIFIABLE SEASONALITY PRESENT

Test for the presence of residual seasonality:

No evidence of residual seasonality in the entire series at the 1 percent level. F = 0.09

No evidence of residual seasonality in the last 3 years at the 1 percent level. F = 0.56

No evidence of residual seasonality in the last 3 years at the 5 percent level.

M1 = 0.099, M2 = 0.113, M3 = 0.000, M4 = 0.201, M5 = 0.225, M6 = 0.521, M7 = 0.211
Q = 0.19

Table A1 (b): F-tests for Seasonality and Adjustment Quality Diagnostics of Original BH

Test for the presence of seasonality assuming stability:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between months	1594.6490	11	144.96809	54.912**
Residuals	126.7196	48	2.63999	
Total	1721.3686	59		

****Seasonality present at the 0.1 percent level**

Nonparametric Test for the Presence of Seasonality Assuming Stability:

Kruskal-Wallis Statistic	Degrees of Freedom	Probability Level
53.4551	11	0.000%

Seasonality present at the one percent level

Moving Seasonality Test:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between Years	9.5947	4	2.398663	0.948
Error	111.3712	44	2.531163	

No evidence of moving seasonality at the five percent level

COMBINED TEST FOR THE PRESENCE OF IDENTIFIABLE SEASONALITY:

IDENTIFIABLE SEASONALITY PRESENT

Test for the presence of residual seasonality:

No evidence of residual seasonality in the entire series at the 1 percent level. F = 0.23

No evidence of residual seasonality in the last 3 years at the 1 percent level. F = 0.80

No evidence of residual seasonality in the last 3 years at the 5 percent level.

M1 = 0.182, M2 = 0.194, M3 = 0.000, M4 = 0.161, M5 = 0.147, M6 = 0.056, M7 = 0.299
Q = 0.18

Table A1 (c): F-tests for Seasonality and Adjustment Quality Diagnostics of Original DF

Test for the presence of seasonality assuming stability:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between months	1690.9310	11	153.72100	59.340**
Residuals	124.3450	48	2.59052	
Total	1815.2760	59		

****Seasonality present at the 0.1 percent level**

Nonparametric Test for the Presence of Seasonality Assuming Stability:

Kruskal-Wallis Statistic	Degrees of Freedom	Probability Level
53.2098	11	0.000%

Seasonality present at the one percent level

Moving Seasonality Test:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between Years	7.9251	4	1.981283	0.863
Error	101.0621	44	2.296867	

No evidence of moving seasonality at the five percent level

COMBINED TEST FOR THE PRESENCE OF IDENTIFIABLE SEASONALITY:

IDENTIFIABLE SEASONALITY PRESENT

Test for the presence of residual seasonality:

No evidence of residual seasonality in the entire series at the 1 percent level. F = 0.15

No evidence of residual seasonality in the last 3 years at the 1 percent level. F = 0.67

No evidence of residual seasonality in the last 3 years at the 5 percent level.

M1 = 0.200, M2 = 0.214, M3 = 0.000, M4 = 0.282, M5 = 0.144, M6 = 0.114, M7 = 0.284
Q = 0.20

Table A2: Post- Seasonal Adjustment Tests

Table A2 (a): F-tests for Seasonality of Passenger Revenue Series after Adjustment

Test for the presence of seasonality assuming stability:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between months	24.0917	11	2.19016	0.514
Residuals	204.7018	48	4.26462	
Total	228.7936	59		

No evidence of stable seasonality at the 0.1 percent level

Nonparametric Test for the Presence of Seasonality Assuming Stability:

Kruskal-Wallis Statistic	Degrees of Freedom	Probability Level
3.7502	11	97.666%

No evidence of seasonality at the one percent level

Moving Seasonality Test:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between Years	1.2852	4	0.321289	0.135
Error	104.3351	44	2.371253	

No evidence of moving seasonality at the five percent level

COMBINED TEST FOR THE PRESENCE OF IDENTIFIABLE SEASONALITY:

IDENTIFIABLE SEASONALITY NOT PRESENT

Test for the presence of residual seasonality:

No evidence of residual seasonality in the entire series at the 1 percent level. F = 0.10

No evidence of residual seasonality in the last 3 years at the 1 percent level. F = 0.49

No evidence of residual seasonality in the last 3 years at the 5 percent level.

M7 = 2.189

Table A2 (b): F-tests for Seasonality of Block Hours Series after Adjustment

Test for the presence of seasonality assuming stability:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between months	7.9035	11	0.71850	0.316
Residuals	109.2417	48	2.27587	
Total	117.1452	59		

No evidence of stable seasonality at the 0.1 percent level

Nonparametric Test for the Presence of Seasonality Assuming Stability:

Kruskal-Wallis Statistic	Degrees of Freedom	Probability Level
1.4590	11	99.967%

No evidence of seasonality at the one percent level

Moving Seasonality Test:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between Years	3.3667	4	0.841686	0.825
Error	44.8811	44	1.020025	

No evidence of moving seasonality at the five percent level

COMBINED TEST FOR THE PRESENCE OF IDENTIFIABLE SEASONALITY:

IDENTIFIABLE SEASONALITY NOT PRESENT

Test for the presence of residual seasonality:

No evidence of residual seasonality in the entire series at the 1 percent level. F = 0.19

No evidence of residual seasonality in the last 3 years at the 1 percent level. F = 0.65

No evidence of residual seasonality in the last 3 years at the 5 percent level.

M7 = 3.000

Table A2 (c): F-tests for Seasonality of Distance Flown series after Adjustment

Test for the presence of seasonality assuming stability:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between months	8.0585	11	0.73259	0.311
Residuals	113.0569	48	2.35535	
Total	121.1154	59		

No evidence of stable seasonality at the 0.1 percent level

Nonparametric Test for the Presence of Seasonality Assuming Stability:

Kruskal-Wallis Statistic	Degrees of Freedom	Probability Level
1.1508	11	99.990%

No evidence of seasonality at the one percent level

Moving Seasonality Test:

	Sum of Squares	Degrees of Freedom	Mean Square	F-value
Between Years	1.0636	4	0.265898	0.244
Error	47.9685	44	1.090194	

No evidence of moving seasonality at the five percent level

COMBINED TEST FOR THE PRESENCE OF IDENTIFIABLE SEASONALITY:

IDENTIFIABLE SEASONALITY NOT PRESENT

Test for the presence of residual seasonality:

No evidence of residual seasonality in the entire series at the 1 percent level. F = 0.13

No evidence of residual seasonality in the last 3 years at the 1 percent level. F = 0.59

No evidence of residual seasonality in the last 3 years at the 5 percent level.

M7 = 2.980

Table A3: Least Squares Estimator of LF

Dependent Variable: D(LF)
Method: Least Squares
D(LF) = C(1)*(LF(-1) - 0.540467863445*PR(-1) - 0.92761687861*BH(-1) + 0.94468277047*DF(-1) - 47.2771483121) + C(2)*D(LF(-1)) + C(3) *D(PR(-1)) + C(4)*D(BH(-1)) + C(5)*D(DF(-1)) + C(6)

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.588086	0.186726	-3.149337	0.0103
C(2)	-0.652171	0.212407	-3.070391	0.0118
C(3)	-0.057621	0.094909	-0.607114	0.5463
C(4)	0.757550	0.187992	4.029694	0.0024
C(5)	0.938681	0.320805	2.926013	0.0151
C(6)	-0.596773	0.310176	-1.923956	0.0897

R-squared	0.651586	Mean dependent var	0.062676
Adjusted R-squared	0.619367	S.D. dependent var	1.808796
S.E. of regression	1.584788	Akaike info criterion	3.853418
Sum squared resid	34.72280	Schwarz criterion	4.062853
Log likelihood	-68.72787	Hannan-Quinn criter.	3.935339
F-statistic	4.571581	Durbin-Watson stat	2.128233
Prob(F-statistic)	0.001505		

Table A4: Least Squares Estimator of PR

Dependent Variable: D(PR)
Method: Least Squares
D(PR) = C(7)*(LF(-1) - 0.540467863445*PR(-1) - 0.92761687861*BH(-1) + 0.94468277047*DF(-1) - 47.2771483121) + C(8)*D(LF(-1)) + C(9) *D(PR(-1)) + C(10)*D(BH(-1)) + C(11)*D(DF(-1)) + C(12)

	Coefficient	Std. Error	t-Statistic	Prob.
C(7)	0.662929	0.201709	3.286538	0.0082
C(8)	0.867706	0.389759	2.226260	0.0302
C(9)	-0.278937	0.167718	-1.663137	0.1021
C(10)	0.971061	0.365517	2.656636	0.0240
C(11)	0.952187	0.346535	2.747729	0.0206
C(12)	0.062378	0.264959	0.235426	0.8148

R-squared	0.680120	Mean dependent var	0.817632
Adjusted R-squared	0.651933	S.D. dependent var	2.813492
S.E. of regression	2.800542	Akaike info criterion	4.992143
Sum squared resid	423.5239	Schwarz criterion	5.201577
Log likelihood	-143.7643	Hannan-Quinn criter.	5.074064
F-statistic	4.109382	Durbin-Watson stat	2.013538
Prob(F-statistic)	0.001343		

Table A5: Impulse Response Results (Cholesky Ordering: LF PR BH DF)

Table A5 (a): Response of LF

Period	LF	PR	BH	DF
1	1.584788	0.000000	0.000000	0.000000
2	0.707372	0.072319	0.229310	-0.095068
3	0.570479	0.060804	0.091592	-0.144731
4	0.351148	0.082558	0.141682	-0.203685
5	0.306381	0.079909	0.124629	-0.213003
6	0.253537	0.084614	0.128889	-0.226599
7	0.235466	0.084852	0.127147	-0.231316
8	0.222716	0.085738	0.127576	-0.234435
9	0.216955	0.085959	0.127282	-0.235927
10	0.213540	0.086162	0.127334	-0.236781

Table A5 (b): Response of PR

Period	LF	PR	BH	DF
1	1.212795	2.524314	0.000000	0.000000
2	1.113426	2.150071	0.508846	-0.420855
3	1.445829	2.163550	0.443438	-0.145376
4	1.378483	2.171589	0.399866	-0.204049
5	1.393663	2.169909	0.429614	-0.199768
6	1.412428	2.167921	0.419615	-0.191979
7	1.410077	2.168827	0.421929	-0.193672
8	1.414102	2.168293	0.421591	-0.192521
9	1.414713	2.168361	0.421675	-0.192317
10	1.415383	2.168301	0.421639	-0.192181

Table A6: Variance Decomposition Results (Cholesky Ordering: LF PR BH DF)

Table A6 (a): Variance Decomposition of LF

Period	S.E.	LF	PR	BH	DF
1	1.584788	100.0000	0.000000	0.000000	0.000000
2	1.754645	97.82866	0.169873	1.707917	0.293555
3	1.853985	97.09407	0.259716	1.773860	0.872353
4	1.904978	95.36334	0.433818	2.233319	1.969521
5	1.946818	93.78513	0.583850	2.548172	3.082853
6	1.982296	92.09393	0.745336	2.880534	4.280198
7	2.015395	90.45889	0.898312	3.184708	5.458090
8	2.046950	88.87523	1.046269	3.475714	6.602788
9	2.077577	87.36473	1.186834	3.749328	7.699108
10	2.107517	85.92673	1.320495	4.008602	8.744176

Table A6 (b): Variance Decomposition of PR

Period	S.E.	LF	PR	BH	DF
1	2.800542	18.75386	81.24614	0.000000	0.000000
2	3.760532	19.16749	77.74910	1.830938	1.252466
3	4.596819	22.72051	74.18536	2.155916	0.938221
4	5.286614	23.97724	72.96231	2.202118	0.858332
5	5.901149	24.82090	72.07827	2.297359	0.803468
6	6.459979	25.49278	71.40943	2.339008	0.758787
7	6.974166	25.96023	70.93880	2.372835	0.728142
8	7.453521	26.32794	70.57048	2.397376	0.704213
9	7.903985	26.61614	70.28191	2.416520	0.685434
10	8.330218	26.84900	70.04894	2.431748	0.670309

Figure A1: Impulse Responses to One Standard Deviation Shocks

Figure A1 (a): Response of LF to Cholesky One S.D. Innovations

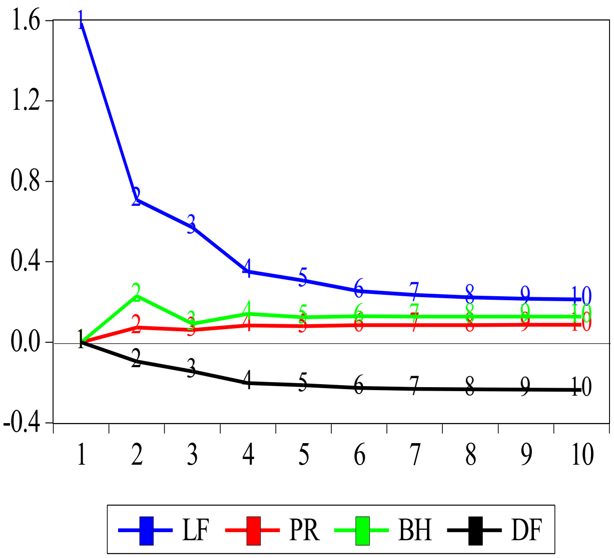


Figure A1 (b): Response of PR to Cholesky One S.D. Innovations

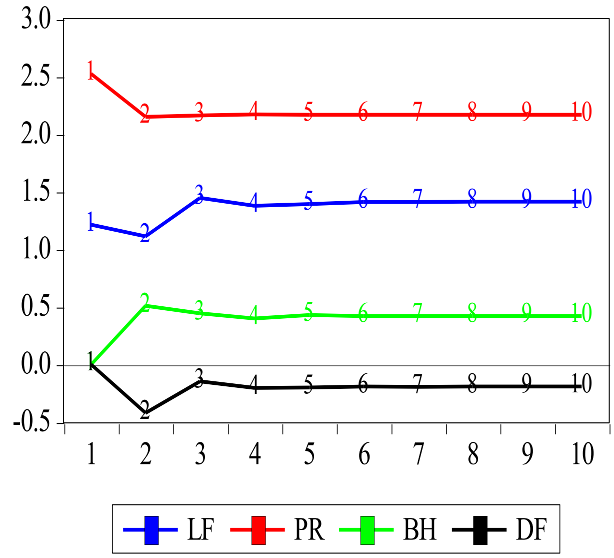


Figure A2: Variance Decompositions

Figure A2 (a): Variance Decomposition of LF

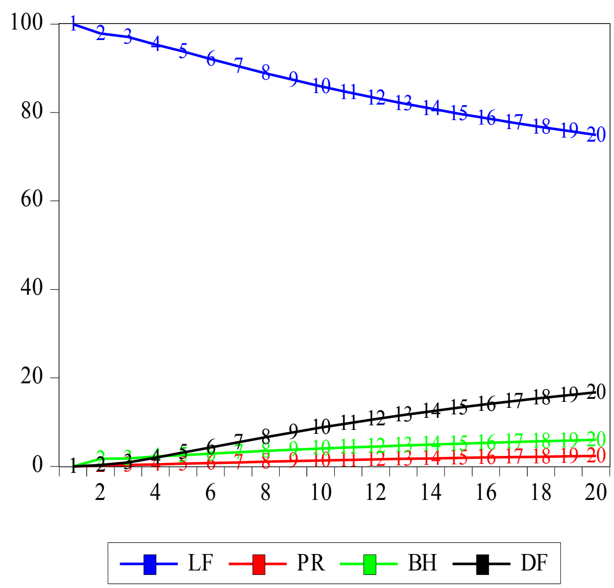


Figure A2 (b): Variance Decomposition of PR

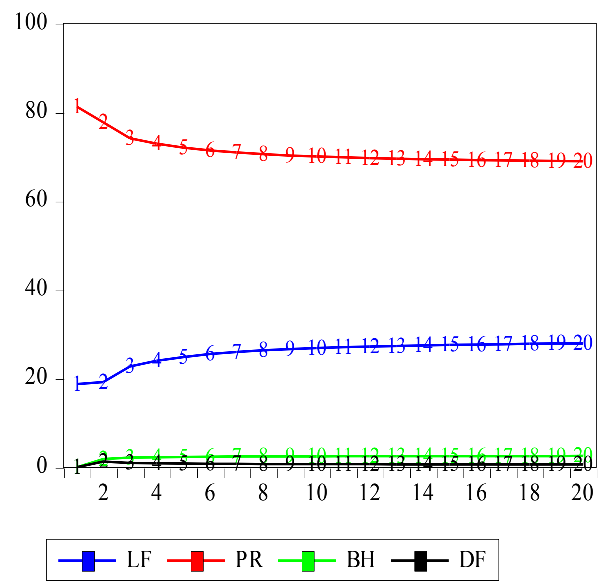


Figure A3: Graph of Actual, Fitted and Residual Plot of Load Factor

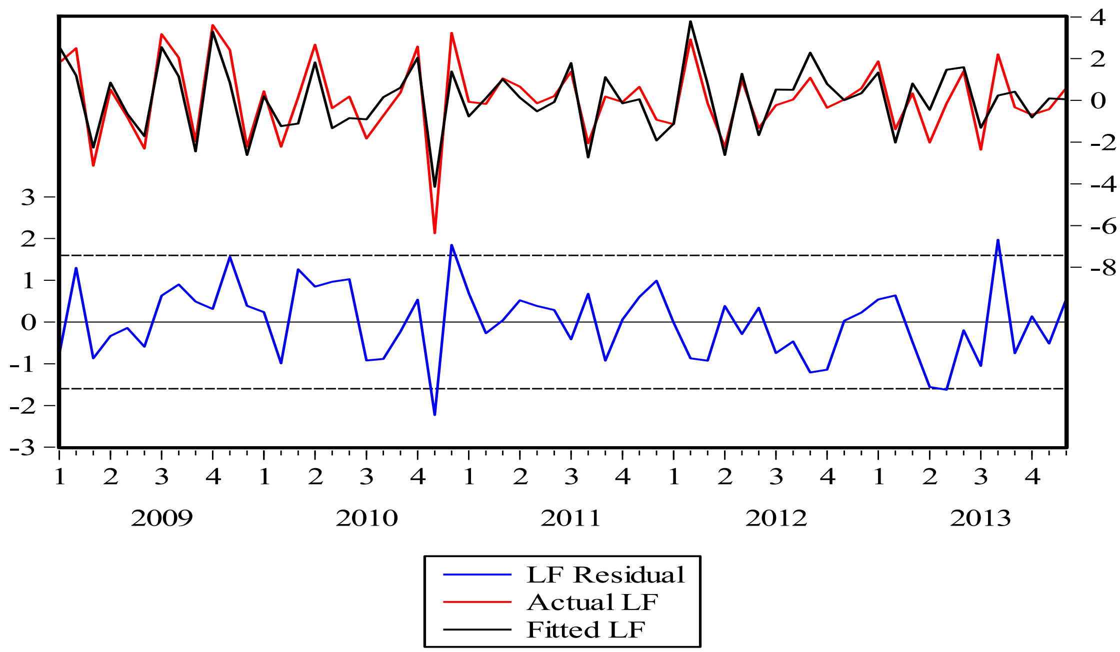


Figure A4: Graph of Actual, Fitted and Residual Plot of Passenger Revenue

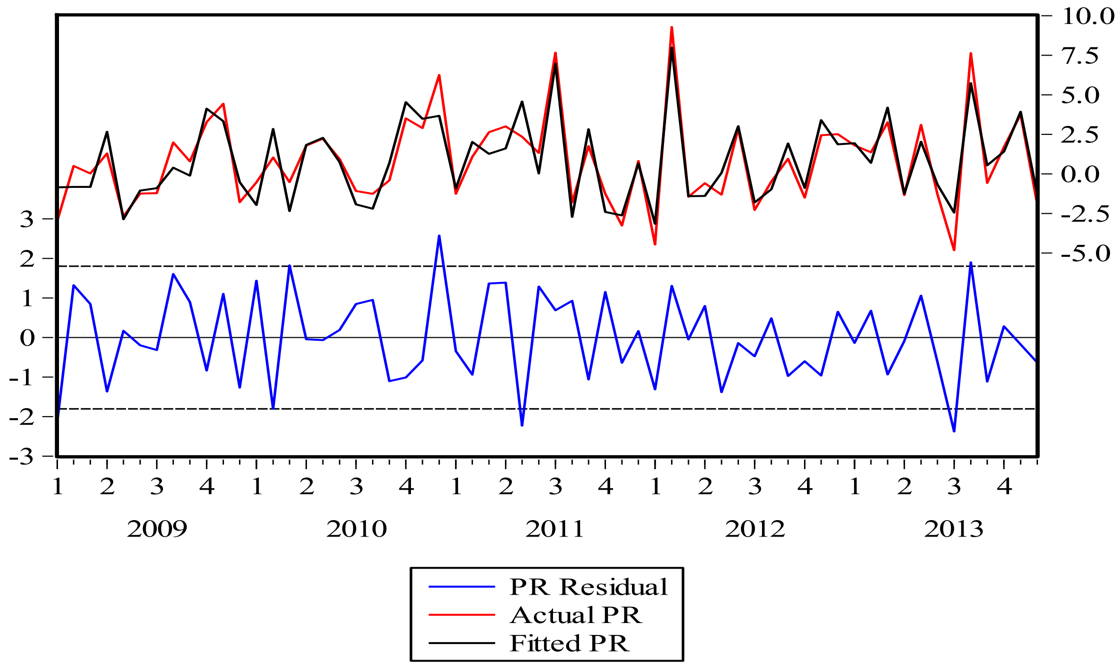


Figure A5: Histogram Normality Tests for Residuals of VEC Models

Figure A5 (a): Normality Histogram for LF Residuals

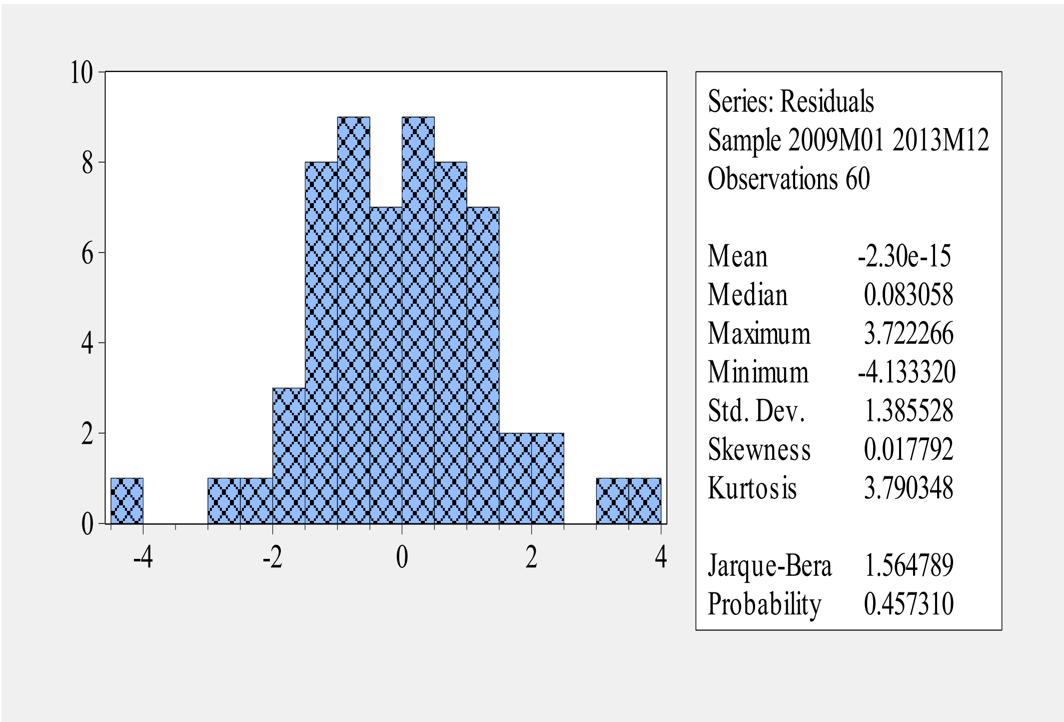
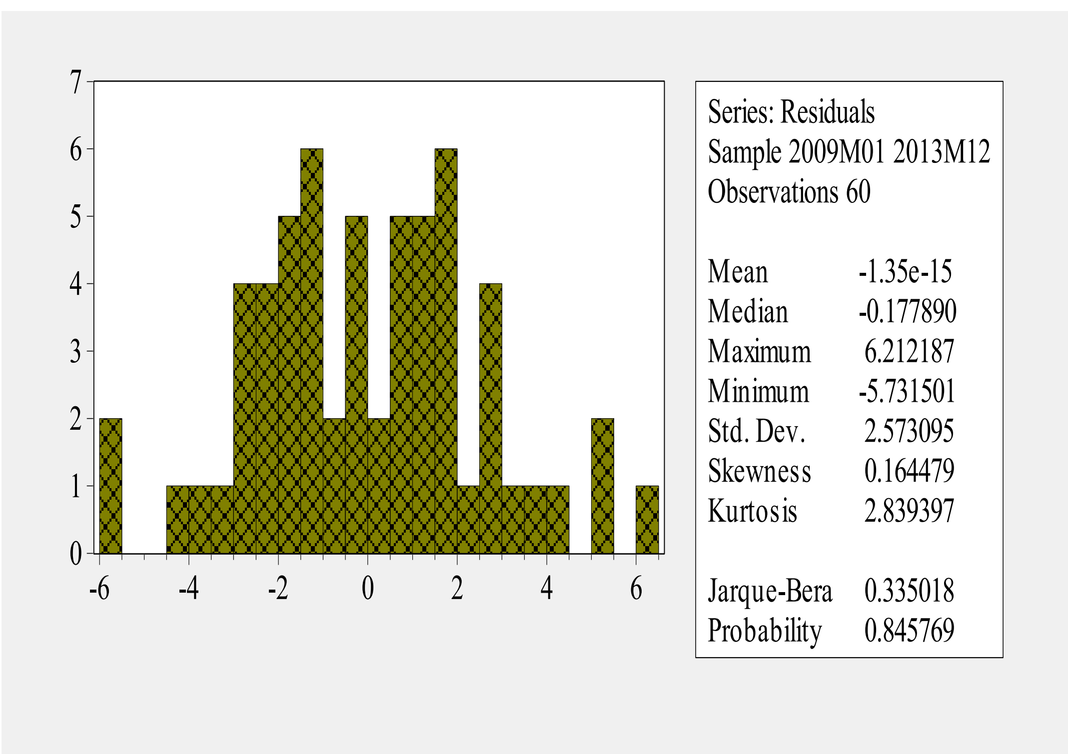


Figure A5 (b): Normality Histogram for PR Residuals



ANNEX

First of all, for those readers who shall raise a question ―*Why it would be necessary to include an Annex if Appendix is already existed in the study*?" please consider the following piece of note.

While they both serve as supplements to a main document, an appendix usually contains data that is referenced in the body of a document but cannot be completely contained in the main text, while an annex is a separate document containing data additional to the original text. An appendix can't exist without the main document, while an annex can be read independently.

Annex-1: Further on Seasonal Adjustment Philosophies

Most seasonal adjustments are based on either of following two broad procedures.

Annex-1.1 Model Based Methods

The aim behind model-based adjustment is to provide models for the components. Models for each of the components are usually estimated simultaneously using the Kalman filter or related techniques (see Harvey (1990) for further discussion of these methods). Each component is assigned its own model which determines the broad spectral characteristics (i.e. the cycles which are present in the original series) of that component. Each component is also assigned its own innovation variance which determines the relative strength of each component. The contributions of the components to the overall model are usually estimated using maximum likelihood methods.

Major computational differences between the various methods in the model based family are usually due to model specification. In some cases, the components are modeled directly. In other cases, the original time series is modeled and the component models are decomposed from that model. The decomposition of the component series is not generally unique. STAMP ((Koopman et al.(1995)) and TRAMO/SEATS (Gómez and Maravall (1997)) are two software packages that use model based methods. Peter 2005 can also be referred for more discussion of these seasonal adjustment methods.

Annex-1.2 Filter Based Methods(X-11 Style Seasonal Adjustment)

Generally the philosophy of X-11 is to remove a band of spectral power (range of cycles) around the seasonal harmonics (from the original data): these harmonics occur at 12, 6, 4, 3, 2.4 and 2 cycles per year in a monthly time series. The width of this band varies according to the characteristics of the data, but the band is always limited to “near” the seasonal harmonics regardless of the spectrum of the original data. In addition, the trend is always required to be “smooth”. To produce a seasonally adjusted series from a monthly time series, events that occur every 12, 6, 4, 3, 2.4 and 2 months need to be removed. These correspond to seasonal frequencies of 1, 2 (half-yearly), 3, 4 (quarterly), 5 and 6 cycles per year. To seasonally adjust a quarterly series, events which occur every 4 and 2 quarters are removed. The trend is produced by smoothing the seasonally adjusted estimates. This can be done by removing the shorter term cycles in the seasonally adjusted series. In filter based methods, no explicit model of a time series components is required. Features of the adjustment process derive directly from the properties of the filters used. Filter based methods can be thought of as X11 style methods. Methods include X11, X11ARIMA88, X12-ARIMA, STL(Cleveland et al. 1990), SABL(Cleveland et al. 1982), and SEASABS(Australian Bureau of Statistics 2005). X11 style methods are generally iterative. The trend, seasonal and irregular components are estimated separately in a computation loop. For example, in X11 a preliminary estimate of the trend is computed. The original data is then detrended to give an estimate of the combined seasonal and irregular component. This is then smoothed month by month to yield a separate estimate of the seasonal component. The computations continue through three iterations.

Major computational differences between the various methods in the X11 family are usually due to different techniques used at the ends of the time series. For example, some methods use asymmetric filters at the ends while others extend the series using ARIMA models and apply symmetric filters to the extended series. For more information one can refer Australian Bureau of Statistics (2005) and Hungarian Central Statistical Office (2007).

Annex-1.3 Comparison of the Two Approaches

Among the differences between the two competing approaches are the following.

- First, the two approaches use a different order when specifying the time series components. Under the model-based approach, the irregular component is typically specified first. This is particularly true of packages like TRAMO/SEATS where the irregular is assumed to have the largest variance possible given the spectrum for the original. In the X-11 style approach, the irregular is defined last as the residual after the seasonal and trend components are defined.

- Second, the X-11 approach aims to remove all spectral power at the seasonal frequencies whereas the model-based approach theoretically leaves some power at the seasonal frequencies in the irregular component and also in the trend (consistent with the model specifications for these components).
- Third, under the model-based approach the mix of trend, seasonal and irregular for a given model is defined by the variance of the innovations for each component model, often called the ―*hyper parameters*”. In X-11, the size of the bands around the seasonal harmonics are determined by the seasonal filter length, and the frequency cut off for the trend is determined by the length of the trend filter.

- Fourth, Model-based methods of seasonal adjustment generally use an additive decomposition, Z_t(original series) = S_t(seasonal part) + N_t(Non-Seasonal Part) whereas X-11 style methods use mostly multiplicative methods.

- Fifth, in model based approaches user can specify a model for the trend component. For example, $T_t = T_{t-2} + T_{t-1} + \varepsilon_t$. In filter based approaches trend component defined as cycles longer than a certain length. For this type of data (with strong trend and seasonal components), there are many plausible adjustments matching perfectly defensible definitions of seasonality. Generally, when there is no smoothness criteria associated with the trend, provided that the X-11 style adjustment exploits the ―known” model (as is done by the popular X11-ARIMA and X12-ARIMA variants) to obtain the end-series estimates, the X-11 solution to the adjustment method is just as valid as the model-based solution.

Moreover, the X-11 solution may be preferable for an official statistical agency due to the simplicity of removing a band around the seasonal harmonics for all data (and the simplicity of explaining to users what has been done to the data).For more detailed comparison one can see Andrew (1999).

Annex-2 *X12-ARIMA* Seasonal Adjustment Program

X12-ARIMA was developed by US Census Bureau as an extended and improved version of the X11-ARIMA method of Statistics Canada (Dagum, 1980). It belongs to the methodological lineage of the Census Bureau's X-11 program (Shiskin, 1967) and Statistics Canada's X11-ARIMA and X-11-ARIMA/88 (Dagum, 1988) programs. This software package estimates seasonality mainly by applying moving average filters to a possibly modified version of the input series. The modifications might include adjustments for extreme values, trading day effects, or holiday effects also estimated by the program. The filters are chosen from a fixed set of filters, partially or—in X-11-ARIMA/88 and X-12-ARIMA, possibly completely—automatically, on the basis of certain signal-to-noise ratios.

The major improvements in X-12-ARIMA fall into four general categories as follow.

1) new modeling capabilities using regARIMA models—regression models with ARIMA errors—for estimating other calendar or disturbance effects with built-in or user-defined regressors;
2) new diagnostics for modeling, model selection, adjustment stability, and for the quality of seasonal adjustment;
3) additional capabilities to make it easier to adjust large numbers of series and determine which have problematic adjustments; and
4) a new user interface.

The article by Findley, Monsell, Bell, Otto, and Chen (1998) gives a detailed overview.

Annex-2.1 Statistical Tests for Seasonality

When performing seasonal adjustment by using X-12 ARIMA, it is important to test for three basic conditions: (1) whether the observed series is seasonal, (2) if the seasonal effects can be estimated reliably, and (3) no residual seasonality is left in the adjusted series. These conditions are tested with a set of standard tests built-in the program to assess the quality of the original series. A brief description of the tests follows.

Annex-2.1.1 An F-test for the Presence of Seasonality

This test is based on a one-way analysis of the variance of the seasonal-irregular (SI) ratios in case of multiplicative models or SI differences for additive models. It evaluates the presence of stable seasonality using F_s as a quotient of two variances: the "between months (quarters)" variance, which is mainly due to the seasonal component and the "residual" variance, which is mainly due to the irregular component. So the test checks for the equality of the monthly means; that is, it tests the hypotheses

H_0: $m_1=m_2=\ldots=m_{12}$

H_1: $m_p \neq m_q$ for at least one pair (p,q) .

Where, m_1, m_2…, m_{12} are the monthly means of the SI component (the de-trended series) found in the seasonal adjustment output. It assumes that the SI values are independently distributed as normal with means m_i and common standard deviation σ. However, while this could be true conceptually for the underlying true SI ratios, the *estimates* of the SI ratios are actually dependent and heteroscedastic, which affects the behavior of the resulting F-statistic. The traditionally attempted solution to this problem is to not use a critical value from the F-distribution, but to instead use a cut-off value of 4.0 as a general. But since data are correlated in the time series, a cut-off value of 7.0 is recommended by the authors of X-11-ARIMA (Lothian and Morry, 1978) and (McDonald-Johnson et al. 2006), with values greater than 7 indicating that the series is seasonal. Since several of the basic assumptions in the F-test are probably violated, the value of the F-ratio to be used for rejecting the null hypothesis, i.e. no significant seasonality present, is tested at the one per thousand probability level.

Annex-2.1.2 Kruskal-Wallis Chi-squared Test

This is another nonparametric test for the presence of stable seasonality based on Kruskal-Wallis statistic, *K*, which evaluates the equality of median values across different months (a value of this statistic falling into the rejection region means that median values are not constant across months).

Annex-2.1.3 An F-Test for the Presence of Moving Seasonality

The moving seasonality test is based on a two-way analysis of variance performed on the SI ratios (differences). It tests for the presence of moving seasonality characterized by gradual changes in the seasonal amplitude but not in the phase. The F-test is printed in Table D8 of the X12- ARIMA printout.

The total variance of the SI ratios (differences) is considered to be the sum of:

1. σ_m^2, the ―*between months (quarters)*" variance, which primarily measures the magnitude of the seasonality. It is equal to the sum of squares of the differences between the average for each month of the SI and the total average, corrected by the corresponding degrees of freedom.
2. σ_y^2 , the ―*between years*" variance, which primarily measures the year to year movement of seasonality. It is equal to the sum of squares of the differences between the annual average of the SI for each year and the total average of the SI for the whole series, corrected by the corresponding degrees of freedom.
3. σ_r^2, the ―*residual*" variance, which is equal to the total variance minus the ―between months (quarters)" variance and the ―between years" variance.

The F ratio for the presence of moving seasonality is the quotient between the ―*between years*" variance and the ―*residual*" variance. A test value falling in the rejection region means that the seasonal-irregular component of the series is not stable across years.

To calculate the variance in an additive model the absolute values of S + I are used, otherwise the annual average is always equal to zero. For a multiplicative model, the SI ratios are replaced by absolute deviations from 100, i.e. by | SI - 100 |. Contrary to the previous test, for which a high value of F is a good indication of the presence of measurable seasonality, a high value of F corresponding to moving seasonality reduces the probability of a reliable estimate of the seasonal factors.

Annex-2.1.4 A Combined Test for the Presence of Identifiable Seasonality

This test combines the previous test for the presence of moving seasonality with the F-test for the presence of stable seasonality and the Kruskal-Wallis Chi-squared test.

The main purpose of this test is to determine whether the seasonality of the series is ―*identifiable*" by X-12 or not. For example, if there is little stable seasonality and most of the process is dominated by a rapidly moving seasonal component, the chances are that the seasonal component will not be accurately estimated, because they will not be properly identified by X12-ARIMA.

The test basically consists of combining the F values obtained from the three prescribed tests as follows:

1. If the F_S-test for the presence of stable seasonality at the 0.1% level of significance fails, the null hypothesis, i.e. seasonality is not identifiable, is accepted.
2. If (1) passes but the F_M-test for the presence of moving seasonality at the 5% level of significance fails, then this F_M value is combined with the FS value from (1) to give

   $$T_1 = \frac{7}{F_M - F_S} \quad \text{and} \quad T_2 = \frac{3F_M}{F_S}$$

   A simple average of the two T's is calculated. If this average is greater than or equal to one, the null hypothesis, i.e. *identifiable seasonality is not present*, is accepted.
3. If (1) passes and the FM test passes but one of the two T statistics fails, or the Kruskal-Wallis test fails at the 1% level, then the program prints ―*identifiable seasonality probably present*".
4. If the F_S, F_M and Kruskal-Wallis chi-squared values all pass, then the null hypothesis, i.e. identifiable seasonality is not present, is rejected. The program prints ―identifiable *seasonality present*".

Annex-2.1.5 An F-test for the Presence of Residual Seasonality

This is also an important F-test in X-12 ARIMA, performed for the whole length of the series as well as for the last three years. The effect of the trend is removed by a first-order difference of lag three for monthly series and lag one for quarterly series, that is $\hat{O}_t - \hat{O}_{t-s/4}$ where $\hat{O}_t$ are the values of table in X-12 ARIMA printout. Two F-ratios are printed at the end of the table, as well as a message indicating the presence or absence of residual seasonality for the last three years and the whole length of the series (Higginson, 1976). In general, the following overview can be made about the above tests.

If the F_S test supports the null hypothesis of no stable seasonality, time series are considered not to be seasonal; generally, a consistent conclusion is also provided by the *K* test, which shows that median values are constant across months. If, on the other hand, the F_S test rejects the null of no seasonality, assuming stability, seasonality is present. In the latter case, two outcomes can happen as far as the F_M test is concerned. If F_M accepts the null of no moving seasonality, stable seasonality is present, and the conclusion of ―*identifiable stable seasonality present*” is reached; the program can easily disentangle the seasonal component. On the contrary, a rejection of the absence by part of F_M means that the seasonal component is moving over years, and the process of disentangling seasonality is difficult because the presence of moving seasonality can cause distortion. Depending on the combination of different tests, the program leads to the conclusion of ―*identifiable stable seasonality not present*” or ―*identifiable stable seasonality probably not present*”; the appropriate conclusion (i.e., ―identifiable seasonality not present” or ―probably not present”) depends on the degree of moving seasonality relative to stable seasonality and has to be based on different combinations of tests. Such ―*negative*” conclusions are problematic if the ultimate goal is to disentangle seasonality.

Annex 2.2 Quality Control Statistics of Seasonal Adjustment

There should not be any seasonal effects in the published seasonally adjusted series or in the irregular component. If there is, in either of both of them, then the residual seasonality test will be significant to reject the null hypothesis of ―*no residual seasonality*”. Therefore, measures of goodness of the seasonal adjustment procedure should be computed. The Statistics Canada X-11 program had two statistics called Q_1 and Q_2 that provided an indication of the amount and nature of the irregular and seasonal components, respectively. In X-12 ARIMA they are reduced to a single Q-statistic, which results from the combination of several other measures. Most of them are obtained from the summary measures of Table F2 in the program‘s printout.

Here following are proposed key quality indicators (set of statistics), to measure the reliability of seasonal adjustment, that are combined to produce the Q-statistic (Peter L. 2005). All of these statistics together with their Q-statistic are incorporated in X-12 ARIMA program.

The main indicator for the quality of the seasonal adjustment is the so-called Q-test statistic. This test statistic is calculated as a weighted average of 11 different component test statistics called M1, … , M11. The weights are given below in Table An1 (for data length of more than 6 years) and Table An2 (for data length of less than 6 years). Mathematically,

$$\text{Q-statistic} = \frac{\sum_{i=1}^{11} W_i \times (Mi)}{W_i} \ .$$

Table An1: Weights Used in the Calculation of Q When the Series is Longer than 6 Years.

Statistics (Mi)	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
Weight (W_i)	13	13	10	5	11	10	16	7	7	4	4

TableAn2: Weights Used in the Calculation of Q When the Series is Shorter than 6 Years.

Statistics (Mi)	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
Weight (W_i)	17	17	10	5	11	10	30	0	0	0	0

It should be noticed that the weights in the second table is not a simple rescaling of the weights in the first table.

All test statistics are standardized so that they range between 0 and 3, and they have been normalized so that the acceptance area is the interval [0;1].That means only values less than one are considered acceptable. So that Q-statistic has to be lower than 1 to judge the de-seasonalized series as acceptable.

M1

The seasonal component and the irregular component cannot be separated suitably if the variation of the irregular component is too high when they are compared to the variation of the seasonal component.

M1 measures the relative contribution (to the variance of the percentage wise changes in the components of the original series) from the irregular component over a three-month long interval.

If the contribution from the irregular component is too large, it can be interpreted to be due to the circumstance that the variation of the irregular component dominates the variation of the seasonal component.

Define now $\bar{I}^2 = \frac{1}{n-1}\sum_{t=2}^{n}(I_t - I_{t-1})^2$, where I_t are the final estimated irregular factors, and n is the number of observations in the series. Using the same formula $\bar{C}$ and $\bar{S}$ is defined. Consider now the fraction

$$\frac{\bar{I}^2}{\bar{I}^2 + \bar{C}^2 + \bar{S}^2}$$

If this fraction exceeds 0.1 then the series is considered behaving unsatisfactorily with respect to the variation of the irregular factor seen in relation to the overall variation.

The fraction

$$M1 = \frac{\bar{I}^2}{\bar{I}^2+\bar{C}^2+\bar{S}^2} \times 10 \qquad \text{has to be smaller than 1.}$$

M2

M2 is equivalent to M1. The only difference is the procedure when removing the trend (in order to make the series stationary). A line is fitted to the trend-cycle estimates in order to obtain a trend estimate. This trend estimate is removed from the original series and hence a stationary original series is obtained, this series is denoted B1' in X-12. The finally estimated trend-cycle component is de-trended in the same manner. Denote this de-trended series D12'. Hereafter, a fraction corresponding to the contribution from the irregular component is calculated. If this test statistic exceeds 1, then the hypothesis of the non-dominating irregular component is rejected.

M3

In the test statistics M3, a comparison between the magnitude of the monthly/quarterly change in the irregular component is compared to the magnitude of the monthly/quarterly change in the trend-cycle. The aim of the seasonal adjustment is removing the seasonal component from the original series in order to give an estimate of a seasonally adjusted series. As X-12 is an iterative process, it is mainly important that not only is the seasonal factor estimated thoroughly throughout the steps that lead to the final seasonal adjustment, but also the trend-cycle and the irregular component have to be estimated thoroughly. If the movement of the irregular factor from one period to another is dominating in the CI-series, then it is difficult to separate these two components and the quality of the seasonal adjustment is therefore low.

The so-called $\tilde{I}/\tilde{C}$ rate is therefore calculated;

$$\frac{\tilde{I}}{\tilde{C}} = \frac{\sum_{t=2}^{n} |I_t - I_{t-1}| / I_{t-1}}{\sum_{t=2}^{n} |C_t - C_{t-1}| / C_{t-1}} \ .$$

If the $\tilde{I}/\tilde{C}$ rate is high, then the variation in the seasonally adjusted series is only due to the irregular component. The rate is considered high if it exceeds 3, and the corresponding test statistic for monthly series is therefore:

$$M3 = \frac{\left|\tilde{I}/\tilde{C} - 1\right|}{2}.$$

And for quarterly series it equals:

$$M3 = \frac{\left|\tilde{I}/\tilde{C} - 1/3\right|}{2/3} \ .$$

M4

The test statistic M4 examines the extent of 1st order autocorrelation in the irregular component. A basic assumption in the F-test statistics in X-12 is that the irregular component is a series of white noise. That is, a series of independently identically distributed error terms with mean 0, a constant variance and the covariance between these terms are 0. X-12 uses a so-called sign-test denoted ADR (Average Duration of Run) in order to test the randomness in the finally estimated residuals.

This non-parametric test is based on the number of turning points. The test is designed to test the randomness in the estimated residuals against a hypothesis saying that the error terms are a AR(1)-process. For a white noise process (with an infinite number of observations) the ADR equals 1.5. The M4 test statistic is based on the formula by Bradley on approximation of the normal distribution:

$$M4 = \frac{\left|\frac{n-1}{ADR} - \frac{2(n-1)}{3}\right|}{\sqrt{\frac{16n-29}{90}}} \times \frac{1}{2.58} \quad ,$$

where 2.58 is the 1% significance level in a two-sided normal distribution test. If M4 exceeds 1 then autocorrelation is present in the residuals.

M5

M5 is an indication of the number of periods that it takes the average absolute changes in the trend-cycle to dominate the corresponding change in the irregular component. The test statistic is equivalent to M3 in the sense that it examines the relative magnitude of the changes in the irregular factor and the trend-cycle components. For k = 1, ...,12 (or k=1,2,3,4 for quarterly series) the fraction

$$\frac{\frac{1}{n-1}\sum_{j=k+1}^{n}(I_t - I_{t-1})^2}{\frac{1}{n-1}\sum_{j=k+1}^{n}(C_t - C_{t-1})^2} \equiv \frac{\bar{I}(k)}{\bar{C}(k)}$$

is calculated. Thereafter, the Monthly Cyclical Dominance (MCD) is derived (for quarterly series QCD is derived) as

$$\frac{\bar{I}(k)}{\bar{C}(k)} \leq 1 \wedge \frac{\bar{I}(k-1)}{\bar{C}(k-1)} > 1 \Rightarrow MCD = k \ .$$

MCD only takes integer values. It is constructed to be rather ‗silly' since it does not notice how close to 1 the I/C ratio is. In order to make MCD 'wiser', a new test statistic MCD' has been constructed in order to solve this problem.

MCD' interpolates linearly the I/C-ratio in order to determine its equalization to 1. The exact formula for MCD' will not be given, but it has to be noticed that whenever MCD' takes a value exceeding 6 then it is not acceptable. Therefore, M5 has the form:

$$M5 = \frac{|MCD' - 0.5|}{5} \quad \text{for monthly series, and}$$

$$M5 = \frac{|QCD' - 0.17|}{1.67} \quad \text{for quarterly series.}$$

Therefore, it holds that M5>1 when MCD's or QCD's are unacceptably high.

M6

In the test statistic M6, the magnitude of the yearly changes in the irregular factor is compared to the magnitude of the yearly change in the seasonal factor. When the aim of the seasonal adjustment is taken into consideration, it is seen that it is very important to be able to identify the seasonal factors. In order to separate the irregular factor from the seasonal component X-12 uses a 3 x 5 moving average on the SI-ratio. Experiences have shown though, that if the yearly change in the irregular factor is too small (compared to the yearly change in the seasonal component), that is when the I/S-ratio is low, then at 3 x 5 is insufficient when it comes to following the seasonal movement. On the other hand, the 3 x 5 seasonal filter is too flexible when the I/S-ratio is too high, and the seasonal factors that have been determined contain some of the irregular movement. Empirical studies have shown that the 3 x 5 moving average is superior when the I/S-ratio lies in the interval [1.5 , 6.5]. Therefore,

$$M6 = \frac{\left|\frac{\bar{I}}{\bar{S}} - 4\right|}{2.5}.$$

Whenever M6 is larger than 1 the hypothesis is rejected, but the problem of a high value of M6 can be solved by using a 3x1 moving average whenever the I/S-ratio is smaller than 1.5, or by using the stable seasonal pattern (stable seasonality option) of the relation exceeds 6.5.

M7

M7 tests the degree of stable seasonality seen in relation to the degree of movable seasonality. If a time series that has been corrected for the trend-cycle (the SI-rate) only exhibits a little amount of stable seasonality seen in relation to the movable seasonality, then the identification of the stable seasonal variation is difficult. The test that is applied is a combination of the F-test applied on the final SI-ratios from X-11, and a test statistic that has been developed by Statistics Canada by J. Higginson (1975). This test indicates whether the seasonal pattern is identifiable by X-12 or not. The test from X-11 measures the magnitude of stable seasonal pattern in the time series. Denote this test statistic F_s, while Higgins' test examines whether movable seasonality is present in the series. Denote this test statistic F_m. The seasonal pattern is identifiable if the absolute error that is present in the final estimates for the seasonal factors is small. This error depends on both the F-test statistics mentioned above. If F_s is low then a high degree of disturbance is found and if furthermore F_m is high, more disturbances are to be found in the series, namely the disturbance due to the movable seasonality.

M7 is hence a combination of the two F-tests mentioned above.

$$M7 = \sqrt{\frac{1}{2}(\frac{7}{F_s} + 3\frac{F_m}{F_s})} \ .$$

About M8, . . . , M11

In the full length of a time series, only a constant seasonal component can be optimally estimated. This is due to the seasonal filters that are applied in X-12. As such the estimates of the seasonal factors contain a considerable error if the original series contains year-to-year movements.

Two types of movements are considered quite differently. That is, the ones exhibiting random fluctuations and the ones where changes prevail in the same direction throughout the years.

The magnitude of the first-mentioned movement can be measured, using the average year-to-year change in the seasonal factors. The magnitude of the second-mentioned movement can be measured by a simple arithmetic average of the changes. Such an average will namely give an indication of the magnitude of the systematic (linear) movement.

Random movements are measured by the test statistics M8 and M10. The test statistics M9 and M11 describe the magnitude of the linear movement. M8 and M10 use all data in the calculations, while M10 and M11 are calculated only on the basis of data from the latest periods.

The test statistics M10 and M11 were introduced since users of seasonally adjusted figures only are focused on the quality of the seasonal adjustment in the latest available years, and these test statistics concentrate only on the seasonal movement in the end of the series.

It is especially important to know whether there is an unambiguously determined linear movement in the seasonal factors for the most recent years. If this linearity is present then the estimates for the seasonal factors are disturbed significantly by the seasonal filters. It is the same disturbance that prevents the use of the seasonal factors from the most recent years when measuring the magnitude of the seasonal movement. This problem is solved by examining the three years before the most recent three year period. It is hoped that the seasonal movement remains the same in these new final years.

The test statistics M8, . . . , M11 are based upon the normalized seasonal factors:

$$S_t' = \frac{S_t - \bar{S}}{\sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(S_t - \bar{S})^2}} .$$

M8

The test statistic M8 measures the random fluctuation of the seasonal factors in the full range of the series. A high value indicates a high degree in the X-12 estimation of the seasonal factors. If the seasonal factors for each year are very different (and random) then the seasonal adjustment is not usable as a very unstable seasonal pattern has been determined.

The variation in the seasonal factors can be measured by the following;

$$|\Delta \bar{S}'| = \frac{1}{m(T-1)}\sum_{j=1}^{m}\sum_{i=1}^{T}\left|S'_{mi+j} - S'_{m(i-1)+j}\right| \quad ,$$

where, m is the number of observations within a calendar year (i.e. either 4 or 12), and T is the number of years.

As the average acceptable value for the variation in the seasonal factors is set equal to 10 %, the test statistic M8 is the following:

$$M8 = 10\left|\Delta \bar{S}'\right|$$

In the calculation of M8, data are only used from the years where the seasonal factors have been calculated without using extrapolation.

M9

The test statistic is used for testing the average linear movement in the seasonal factors in the full length period. When an average of the year to year changes is formed, the amount of systematic movement in the series is measured. If the only present fluctuations are random then this average will be close to zero. If most changes occur in the same direction then the mean absolute change is close to the average arithmetic change.

As it holds that (telescopic sum) $\sum_{i=1}^{n-1} \Delta S'_{mi+j} = S'_{m(n-1)} - S'_j$, and an acceptance limit of [0 , 1], it holds that:

$$M9 = \frac{\sum_{j=1}^{m}\left|S'_{m(n-1)+j} - S'_j\right|}{m(n-1)} \times 10\ .$$

M10 *and* *M11*

These test statistics are identical to M8 and M9, respectively, but are only calculated using the years n-2, n-3, n-4, and n-5, where n is the total number of total years. Users are typically often only interested in the latest available data, and this is the reason for these test statistics to provide information on the quality of the latest estimates of the seasonal factors. In the calculations, data are only used from the most recent years where the seasonal factors have been calculated without the use of extrapolations.

DECLARATION

I, the undersigned, declare that this thesis is my original work and has not been presented for a degree in any other university, and that all sources of materials used for the thesis have been duly acknowledged.

Declared by:

Name: ________________________________

Signature: _____________________________

Date: _________________________________

Confirmed by Advisor:

Name: ________________________________

Signature: _____________________________

Date: _________________________________