FORECASTING ON TELECOMMUNICATION REVENUE BASED ON SERVICE TAX USING THE ARIMA MODEL

THERINA OLIVIA RAJATURAI

FAKULTI SAINS DAN TEKNOLOGI
UNIVERSITI MALAYSIA TERENGGANU

2009

FORECASTING ON TELECOMMUNICATION REVENUE BASED ON SERVICE TAX USING THE ARIMA MODEL

By
THERINA OLIVIA RAJATURAI

A Final Year Project submitted in partial fulfillment of the requirements for the award of the degree of Bachelor of Science (Financial Mathematics)

DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE AND TECHNOLOGY
UNIVERSITY MALAYSIA TERENGGANU
2009

JABATAN MATEMATIK
FAKULTI SAINS DAN TEKNOLOGI
UNIVERSITI MALAYSIA TERENGGANU

PENGAKUAN DAN PENGESAHAN LAPORAN MAT 4499B

Adalah ini diakui dan disahkan bahawa laporan penyelidikan bertajuk **FORECASTING ON TELECOMMUNICATION REVENUE BASED ON SERVICE TAX USING THE ARIMA MODEL**. No. Matriks: **UK 13107** telah diperiksa dan semua pembetulan yang disarankan telah dilakukan. Laporan ini dikemukakan kepada Jabatan Matematik sebagai memenuhi sebahagian daripada keperluan memperolehi Ijazah Sarjana Muda Sains Matematik Kewangan, Fakulti Sains dan Teknologi, UMT.

Disahkan oleh:

....

Penyelia Utama

Nama: MUHAMAD SAFIIH BIN LOLA
Pensyarah
Jabatan Matematik
Fakulti Sains dan Teknologi
Universiti Malaysia Terengganu
21030 Kuala Terengganu

Cop Rasmi:

Tarikh: 05/05/09

....

Ketua Jabatan Matematik

Nama:

Cop Rasmi:

Tarikh: 5/5/09

DR. HJ. MUSTAFA BIN MAMAT
Ketua
Jabatan Matematik
Fakulti Sains dan Teknologi
Universiti Malaysia Terengganu
21030 Kuala Terengganu

DECLARATION

I hereby declare that this final year project entitled Forecasting on Telecommunication Revenue based on Service Tax Using ARIMA model is the result of my own research except cited in the reference.

Signature :

Name : THERINA OLIVIA RAJATURAI

Matrix No : UK13107

Date : 05.05.2009

ACKNOWLEDGEMENT

At the end of this thesis, I would first like to thank God for giving me the strength to finish up this thesis despite the entire tight schedule. Besides that I would also like to thank everybody who has been helping me through out my project. Firstly, I would like to state my deepest gratitude to Dr. Safiih who has been giving me the guidance from the beginning.

I would like to thank Mr.Zaily for helping in the usage of the software. He was a good instructor in Custom. I thank him for teaching the extra knowledge of forecasting. Not forgetting my wonderful parents whom have been giving moral support in order to finish up my project. Their encouragement has made me this far in finishing up my paper work.

Last but not least I would like to thank my friend, Rajatheeban and my other friends as well for giving encouragement and also helping me out to finish up this project. Finally, I thank God again for all the care and guidance that He has given me through all these people around me. Thank you.

FORECASTING ON TELECOMMUNICATION REVENUE BASED ON SERVICE TAX USING THE ARIMA MODEL

ABSTRACT

The service tax used in Custom is to generate the future income for the country. This research is regarding the revenue telecommunication industry in the governments and the development sector. The paper practically focuses in two characteristics which are the fixed lines and the cellular phones revenue that are generated in the custom in order to get the future income based on the service tax. Basically, Custom generated the tax whereby the paper deals with service tax. The model ARIMA in Box Jenkins is used to forecast the value for the year 2009 by using the software provided in Custom. This model basically is used in checking data, estimated parameters and checking diagnostic made to obtain which model is suitable to be used in the prediction. From the analysis the suitable model for ARIMA model used is ARIMA (0,1,3). Therefore, the results obtain shows that the forecasted results had a slight increase in Custom in generating the service tax based on the telecommunication.

PERAMALAN HASIL TELEKOMUNIKASI BERDASARKAN CUKAI PERKHIDMATAN DENGAN MENGGUNAKAN MODEL ARIMA

ABSTRAK

Hasil cukai yang digunakan di Kastam adalah untuk menjana hasil pendapatan Negara. Kajian ini dilakukan terhadap hasil industri telekomunikasi di sektor awam dan juga kawasan pembangunan. Kertas kerja ini juga difokuskan kepada dua ciri iaitu talian tetap dan penggunakan telefon bimbit yang digunakan untuk menjana telekomunikasi di Kastam dimana untuk mendapatkan hasil pendapatan pada masa hadapan berdasarkan hasil cukai. Di kastam, pihak kerajaan selalu menggunakan cukai perkhidmatan untuk menjana hasil pendapatan negara. Model ARIMA di Box Jenkins telah digunakan untuk menjana hasil bagi tahun 2009 dengan menggunakan perisian yang digunakan di Kastam. Model ini digunakan untuk pengecaman data, pemeriksaan data dan pemeriksaan diagnostic digunakan untuk mendapatkan hasil yang sewajarnya. Selain itu, daripada analisis yang dilakukan model yang sesuai digunakan adalah ARIMA (0,1,3). Dengan itu, keputusan didapati bahawa terdapat peningkatan dalam hasil peramalan di Kastam dalam menjana cukai perkhidmatan berdasarkan kepada telekomunikasi.

CONTENTS

LIST OF TABLES

LISTINGS OF GRAPHS

LISTING OF ABBREVIATION

Abbreviation

ACF	Autocorrelation Function
ANOVA	Analysis of variance
AR	Autoregressive
ARIMA	Autoregressive Integrated moving average model
ARMA	Autoregressive moving average
C_C	Post payment
CLAG	Subscription tax
C_P2	Establishment proportion
CPRABYR	Tax pre payment
ICT	Information communication technology
IRB	Inland Revenue Board
KLL	Subscription rate
MA	Moving average
MAD	Median Absolute Deviation
MAPE	Mean absolute percentage error
MdAPE	Median absolute percentage error
PACF	Partial autocorrelation function
PB	Percentage Better
PBILLAG	Subscription number
PDF	Probability Density Function
PMF	Probability Mass Function
PP	Rate Tax Establishment
RMSD	Root Mean Square Deviation
RMSE	Root Mean Square Error
RV	Random vector
SPSS	Statistical Package for Social Sciences

LIST OF APPENDIX

Appendix

CHAPTER 1

TELECOMMUNICATION

1.1 Introduction

One of the greatest inventions of mankind is the Telecommunication. Telecommunication is the assisted transmission of signals over a distance for the purpose of communication. In the earlier times, this may have involved the use of smoke signals, drums, semaphore, flags or heliograph. Telecommunication is acknowledged by many observers to be the hottest industry segment in the early twenty-first century.

In this modern era, telecommunication typically involves the use of the electronic transmitters such as the telephone, television, radio or computer. The early inventors in the field of telecommunication include Alexander Graham Bell, Guelielmo Marconi and John Logie Baird. Telecommunication is an important part of the world economy and the telecommunication industry's was estimated to be $1.2 trillion in 2006 (Wikipedia, 2008).

The form of telecommunication is by the network, telephone exchange, communication devices, satellite communication and data communication. Communication of the data involves the fax done by the computer, voice such as the telephone, mobile phones and also video the video and web conference. It is practically a large distance through the medium such as fiber optics, internet and radio waves is known as the telecommunication.

1.2 Tax Revenue

According the telecommunication Service tax proposes a 5% tax on all revenues generated from telecommunications services billed to most residents. The tax will apply to all wire and wireless telecommunication providers offering telecommunications services within the city. The telecommunication service tax will apply to all revenues of all providers of telecommunication services within the city. Currently some providers pay 7% of local exchange, while some providers pay 7% of all revenues earned within the city. Wireless providers pay no tax for operating within the city. The telecommunication service tax will require all providers to pay an equal tax of 5% of the revenue earned within the city (Telecommunication Service Tax, 2002).

There tax revenues are practically involving both direct and indirect taxes. The direct taxes are collected by the Inland Revenue Board (IRB) and include taxes such as income tax on individuals and corporations, petroleum income tax, stamp duty and real property gains tax. But in this research we practically deal with the indirect taxes. The indirect taxes are practically collected mainly by the Royal Customs and Excise Department. Indirect taxes are import duties, export duties, excise duties, sales tax, service tax, property taxes, entertainment tax and road tax.

There are various types of taxes which by brings to generating of the revenue. Taxes are compulsory, unrequited payments in cash or in kind made by institutional units to government units. These taxes are described as the unrequited because the government provides nothing in return to the unit.

The taxes are divided into various which are the taxes on the products, other taxes on production, wealth, capital taxes. Taxes on income, profits and capital gains, taxes on payroll and workforce, taxes on property, goods and services on international trade and transactions. The service tax involves the import tariffs and excise taxes. These often constitute the most important revenue sources in developing especially in Custom. They generate the revenue in Custom.

1.3 Problem Statement

In this modern era, telecommunication is one of the important sources for generating the country's income. The upcoming of the country in economics, it increases the usage of using the telecommunication.

Though we know the importance of telecommunication but we still take things lightly. Most citizens prefer to waste their time on their cell phones, internet and any other ways involving telecommunication. They never think of the effect what will happen in the future. Despite that, there might be delaying of phone calls. If the telecommunication is not been used wisely the revenue generated will be slowed down than before.

Not only can that, the usage of internet, if downloaded in the wrong webpage cause the various types of virus attack. In this case it damages the files in the system. It definitely causes low generated in revenue. Therefore, one set of forecasting should be done in order to set the revenue needed to generate the service tax in the coming years. By doing so we will be fully prepared on the future income that we are about to receive. This study is done by the time series analysis.

1.4 Research Scope

Through out this research, the data used is from the Custom, Putrajaya. This proposal concerned about the data for the past 8 years in Custom, Putrajaya from the month of January 2000 until the month of December 2008. The data considers the service tax that will generate the future revenue of custom by analyzing the past data.

1.5 Objectives

The main objective of this research is that:

a) To forecast an appropriate ARIMA model to forecast on the future income of the revenue based on the telecommunication used in Custom.
b) To forecast an appropriate ARIMA model based on the monthly values of telecommunication.
c) To identify the appropriate parameter estimated.
d) To forecast the revenue based on the yearly telecommunication.

1.6 Significance of the Research

The revenue tax relating to telecommunication is one of the main important roles in the upcoming sector in the industry. With the technology raising the usage of the telecommunication is in high demand causes the revenue to increase.

The accurate forecasting of the revenue gives the future revenue income which by will generate the future income of our country. This is to ensure that there will not be any damages in the internal system such as the internet usages. Not only that, the research is done in order to upgrade the telecommunication in order to generate a proper future income.

Furthermore, we have to consider the constraints that effect to generate the revenue. The climatic weather, operational constraints in case if by any chance the instruments that they are using are not giving a good respond such as the cable, environmental and equipment limitation. These are the constraints that we have to consider in order to have the appropriate model that we need to forecast the revenue.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Over the years, telecommunication has been a rapid spiral over the years. However, the tax services and the system use by the government have not kept pace with the government change. The government practically deals with the indirect taxes (Gomperts, 2000). According to Eisenbach, (1999), the system of telecommunication has not kept up with the evolution of the technology and the industry. Nowadays there is a burden on the telecommunication providers they have to undergo through many taxes which are the telecommunication tax structure, using the principles of tax efficiency, competitive neutrality, tax equity and administrative. Therefore by here it is quite difficult to generate the future income with the leveling of the tax burdens.

Certain research has been done on how to improve the forecasting system in telecommunication. This study has been done in order in order to find out which forecast gives a better and accurate value in the future based on service tax. This is done by considering the tax burden that is been laid upon the telecommunication.

2.2 Information Flow

In this section we focus more on the tax attributes of securities. The objective of doing so is that to explain the tax exempts and taxable investors as well as the structure decisions of corporations (Thomas, 1997).

The proposal made presume of the weak environmental and argues the fact that the interference in the revenue forecasting processes is another instrument to conceal the extraction of public resources (Hellman, 2001).

The telecommunication system means any transmission or emission overseas from Montserrat are the inhabitant of the Caribbean island of signs, signals, writing, images and sound or intelligence of any nature cable wire, radio visual or any other magnetic system (Malaysian Tax Law, 2006).

Tax to be collected by persons providing telecommunication services and paid to the Accountant General. The tax collected by the person who authorized or licensed to provide and operate the telecommunication services from Headquarters (Malaysian Tax Law, 2006).

2.3 Research about Telecommunication

Robert (2002) discuss about the telecommunication demand forecasting states that telecommunication is increasingly deregulated and competetitive market place. This paper practically studies about the development of a competitive service. Various models have been used to describe the dynamic market involving telecommunication which is the differentials, use of standard econometric methods and diffusion model. Therefore, this paper emphasis more on which techniques are suitable for forecasting the demand for new products and services. Therefore there is a further study done based on this paper done by Detlef (2000) on how to improve the forecasting technique.

Detlef (2000) states that the success of telecommunication services in the presence of network. This journal talks about how to improve the telecommunication than the previous as stated above. To improve the forecasting techniques using the method such as the diffusion method. This journal forecast more on the services and technology. The resulting from diffusion explains of diffusion including critical mass, path dependency, lock in and inefficiency. It involves the non linear and stochastic diffusion process modeling regarding the network. The result, the development is forecasted. According to the result the prediction has highly success of any telecommunication service should be expressed in the form of diffusion path.

Whereas, Constantinos *et al.* (2006) states, seasonal decomposition and forecasting of telecommunication data. Studies also have been done using telecommunication which is forecasting on the monthly outgoing calls in a University. The total number of calls has been analyzed. Practically there are 3 different methods that are used which are the Decomposition, Exponential smoothing Method and SARIMA Method. The outcomes were to predict the future demands of telecommunication network of the University. The study made use of the historical data; practically the managers will predict the telecommunication by creating a reasonable accurate forecast. Different types of calls need to be separated due to different restrictions to use different average prices. The planning states the infrastructure investment and the call volume strategy are the topics where by the managerial decisions will depend on the forecast. Here we know that the managers will use the forecast value in order to make another step whether to invest or not into their strategy.

This case study is about classification, relationship and forecasting models for telecommunication services[Duk Bin Jun,(1997)]. This study suggest that there will forecast in the future market size for telecommunication services. The new services exist to satisfy the customers' needs. The forecast considers the various relationships among telecommunication services, such as the competiveness and complementariness. This section basically applied of the forecasting models proposed by Peterson and Mahajan to services in Korea. This paper do not have complete example of the application due to the lack of data. By this the total demand is forecasted and the proportion of a particular network or service market by predicting the diffusion rate.

The journal is about forecasting the telecommunication service subscribers in substitute and competitive environments by Duk B. Jun, (1999). This journal basically talks about how the telecommunication is effected by the rising of the competitive environmental and the service used to subscribe. The model used in this paper was proposed by Jun and Park. The method are basically are multigenerational diffusion model and the mix of the variables such as the price and advertising in the regression analysis. This diffusion model is very useful in analyzing complicated phenomena such as the competitive diffusion process. Furthermore it also provides the flexibility to include marketing mix variables as in the regression model.

Ananda *et al.* (1998) states that this paper investigates the long run demand for money and short run dynamics of the long run money demand function. This paper was mainly used in Sri Lanka during the post 1977 period. This paper suggests that formulating a sound of monetary aggregate should be controlled by the monetary authority and the factors involve in it. There are a few formulas which have been used in this paper. The technique used is to interpolate. Practically, changing the yearly data into the required months. The following formula used to calculate the required interpolation is shown in the methodology.

Mohan,S *et al.* (1995) discussed about the forecasting of the monthly inflows into a reservoir system in the monsoon climatic conditions using a multiplicative seasonal ARIMA model based on 25 years of data. This paper deals with the uncertainty analysis which is using the stochastic model. This paper take accounts on the optimal policies for the use of available water and the operation of water resources systems. The inflows are important therefore it is computed by mathematical programming techniques.

Lukasz *et.al.* (2009), discuss about the telecommunication policy. That is the consumer loss of the minimum duration for mobile telephone calls. This paper involves the consumer who use the telephones are from the category of fixed lines. The telephone policies used are the local, long distance call, international and also fixed line used. Not only that the internet connection is also used. For this paper the econometric model has been used. They take account on a few categories which are there are the age differences, gender usage, the residents using the cell phone, the call attribution used of the day, workday or weekend. By using the model concerning they have narrowed down to 800 individuals. There are few estimates that have been done which are concerning the price, adding to the model there a few models. Basically, there 3 models concerning the Econometric.

Adnan Al-mutawkkil (2009), discuss about the developments and broadcasting in telecommunication. It involves one of the major things in this country which is the ICT. ICT has been became the main growth factor for most countries nowadays. There are few categories involve which are the fixed lines, the internet and the mobile network. This study shows the global position of each place or country and also shows out the failures in developing the ICT infrastructure. The method used for further research is the factor analysis approach. The method is used to maintain the objective to this reduction technique. Not only that to determine the percentage used by each of the categories stated before. This paper has used the eigenvalue in order for the reduction approach. The results that are obtained show that it is significant as overall, though the broadcasting in this case study shows that there is the no correlation. Based on this, further study should be done based on telecommunication which will relate to the ICT development. There are many other factors that prompt to the changes in the telecommunication using various techniques and also models.

Gary *et al.* (2007), discuss about the usage of the linear models to forecast the telecommunication. This paper considers all kinds of environmental issues that effect the telecommunication. The data used are 29 monthly series. The data used had been further research and the results that have been gained positive trend. The data are basically related to monthly. The monthly series have stronger correlation with time. The forecasting techniques used are the curving fitting models, autoregressive models, autoregressive integrated moving average model (ARIMA), space model concerning with Kalman filtering and regression models. The researchers did about 18 observations. They also have used a few error terms which are the mean absolute percentage error (MAPE), percent better (PB) and median absolute percentage error (MdAPE). PB shows that the Holt provides the most reliable forecast. Whereas the other error term such as the MAPE in this case shows the forecast accuracy.

CHAPTER 3

METHODOLOGY

3.1 Introduction

The data used is yearly data based on service tax. The data was gathered from the Headquarters in Putrajaya, Selangor. It is practically covers all the kinds of indirect taxes. The indirect taxes do not involve the income tax. The data is based on the past 8 years starting from the year 2000 to 2008. So far the data that have been collected are from January 2000 to December 2008. Altogether there are 96 data. This set of data will later on will be converted to the accurate model in order to get the accurate forecast in the coming year. The purpose of this forecasting is to get the accurate forecast for the future revenue of the telecommunication based on the service tax.

3.2 Time Series Pattern

Most time series patterns can be described in terms of two basic classes of components which are the trend and seasonality. The former represents a general systematic linear or most often is the nonlinear component that changes over time and does not repeat or at least does not repeat within the time range captured by the data. Basically most pattern in time series data are overall trend. Sometimes there is also pattern called the multiplicative seasonality indicates that the relative amplitude seasonal changes is constant over time, thus it also will be related to the trend. Therefore to choose the trend one should be able to analyze the data given.

3.3 Autoregressive Process

The Autoregressive process which by in most time series consist of elements that are serially dependent in the sense that one can estimate a coefficient or a set of coefficients that describe consecutive elements of the series from specific, time lagged (previous) elements. The equation is given below:

$x_t = \xi + a_1 * x_{(t-1)} + a_2 * x_{(t-2)} + a_3 * x_{(t-3)} + \ldots + \varepsilon$

where:

ξ	is a constant (intercept)
a_1, a_2, a_3	are the autoregressive model parameters
ε	random error component

3.4 Stationary Requirement

An autoregressive model only will be stable only if the parameters are within a certain range. For an instant if there is only one autoregressive parameter then is must fall within the interval of $-1 < a < 1$. If not the past events will accumulate and the values of successive x_t' will move towards to infinity meaning that the series would not be stationary.

3.5 Moving Average Process

This process is independent from the autoregressive process, each element in the series can also be affected by the past error (or random shock) that cannot be accounted for by the autoregressive components, that is:

$$x_t = \mu + \xi_t - b_1 * \varepsilon_{(t-1)} - b_2 * \varepsilon_{(t-2)} - b_3 * x_{(t-3)} - \ldots - \varepsilon$$

where:

μ	is a constant
b_1, b_2, b_3	are the moving average model parameters
ξ	random error component

3.6 Invertibility Requirement

There is a duality between the moving average process and the autoregressive process (Bowermann, 1990), that is, the moving average equation above can be rewritten into autoregressive form of infinite order. This only is done if the moving average parameters follow certain conditions, if the model is invertible. Otherwise, the series will not be stationary.

3.7 Statistical Stationary

A stationary time series is one whose statistical properties such as mean, variance, autocorrelation are all constant over time. Most statistical forecasting methods are based on assumption that the time series can be rendered approximately to stationary. Another reason for trying to stationary a time series is to be able to obtain meaningful sample statistics such as mean, variances and correlations with other variables. However, sometimes de-trending is not sufficient to make series stationary, in which case it may be necessary to transform it into a series of period to period or to season to season differences. If the mean, variance and autocorrelation of the original series are not constant in time, even after the de-trending, perhaps the statistics of the changes in the series between periods or between seasons will be constant. This series is called to be the difference-stationary.

3.8 First Difference

This difference of the time series will change from one period to the next. If *Y(t)* denotes the value of the time series *Y* at period *t*, then the first difference of *Y* at period t is equal to *Y(t)-Y(t-1)*.

3.9 Autocorrelation

It is a mathematical tool for finding repeating patterns, such as the presence of a periodic signal which has been buried under noise or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies (Wikipedia, 11/25/08). The autocorrelation function (ACF) of a random process describes the correlation between the processes at different points in time. Let X_t be the value of the process at time *t*. Where by *t* may be an integer for a discrete-time process or a real number for continuous-time process.

If X_t has mean μ and variance σ^2 then the definition of the ACF is:

$$\mathrm{R}(t,s) = \frac{E[(X_t - \mu)(X_s - \mu)]}{\sigma^2}$$

where:

E is the expected value

Since the variance may be zero (constant process) or infinite. If the function *R* is well defined its value must lie in the range [-1,1] with indicating perfect correlation and -1 indicating perfect anti-correlation.

Say X_t is second-order stationary then the ACF depends only on the difference between *t* and *s* and can be expressed as a function of a single variable. This gives a simpler form:

$$\mathrm{R}(k) = \frac{E[(Xi - \mu)(X_{i+k} - \mu)]}{\sigma^2}$$

where:

k is a lag. $|t\text{-}s|$

3.10 Box Jenkins Methodology

The box Jenkins methodology, named after the statisticians George Box and Gwilym Jenkins, applies autoregressive moving average (ARMA) or ARIMA models to find the best fit of a time series to past values of this time series, in order to make the forecast. The original model uses an iterative three-stage modeling approach. The first is the model identification followed by parameter estimation and lastly is the model checking.

3.11 Model Identification

This model and model selection is making sure that the variables are stationary, identifying seasonality in the dependent series use seasonally differencing it if necessary, and using plots of the autocorrelation and partial autocorrelation functions of the dependent time series to decide the autoregressive or moving average component should be used in the model. Certain emphasize should be done on the importance of selecting the models based on sound scientific principles modeling the underlying data throughout their book on model selection.

3.12 The Usage of the Interpolation

Below are 4 equations used in order to interpolate the data for the fixed lines and cellular phones used in Custom. The data provided are from the yearly computed from the year 2000 to 2008. From here we have to use the interpolate method to get the monthly figures. The method of interpolation used to derive quarterly figures from annual data, it is therefore the outline given from the Goldstein and Khan (1976), is follows: let the x_{t-1} , x_t and x_{t+1} are the 3 successive annual observations of variable x. if the quadratic function that passes through the three points the equation will be as below. These are the functions used to interpolate the required data.

$$\int_{1}^{1.25} (as^2 + bs + c)ds = 0.05468x_{t-1} + 0.23438x_t - 0.039067x_{t+1}$$

$$\int_{1.25}^{1.50} (as^2 + bs + c)ds = 0.00781x_{t-1} + 0.26563x_t - 0.02344x_{t+1}$$

$$\int_{1.50}^{1.75} (as^2 + bs + c)ds = -0.02344x_{t-1} + 0.26562x_t + 0.00781x_{t+1}$$

$$\int_{1.75}^{2} (as^2 + bs + c)ds = -0.0391x_{t-1} + 0.23437x_t + 0.05469x_{t+1}$$

Interpolation is a method of constructing new points within the range of a discrete set of a known data points. Interpolation is merely using a function as shown above to calculate new data points. The interpolation is a method used to gain in simplicity which offsets the error. The Gaussian process is a powerful non linear interpolation. These can also be used for the regression such as fit in the noisy data.

3.13 Usage of Software

The software used in this section is the Forecast pro and SPSS. Forecast pro is gained from Custom, Putrajaya. This software is a line of comprehensive forecasting solutions for professionals in manufacturing, planning and inventory control, marketing and sales. This forecast combines a built expert system with a few forecasting techniques which are the exponential smoothing, Box Jenkins and also the other forecasting techniques. It mainly focused on the Box Jenkins method. This software provides quick and accurate forecast. It also gives a full range of diagnostics for each model and also graphs.

Besides that, the other software used is the Statistical Package for the Social Sciences, SPSS. This software is used to estimate the multiply regression that has been used. There are few statistical that have taken into account in order to know the relationship gained using the other software. The other statistical term that is considered is the simple coefficient of determination and also the simple correlation coefficient. If the correlation is more than 0.75 or 75% the strength of the relationship between the y and x are good. The way to measure the simple coefficient of determination and also the simple correlation coefficient is as shown. These manual methods can be calculated from the ANOVA table.

$$r^2 = \frac{Explained\ variation}{total\ variation}$$

$$r = \sqrt{r^2}$$

3.14 Parameter Estimation

This estimation is using the econometric computation algorithms o arrive at coefficient which fits the selected ARIMA model. The most common methods use the maximum likelihood estimation or non linear least square estimation.

The first is a se of statistical samples taken from a random vector (RV) of size N. Put into a vector,

$$a = \begin{bmatrix} a[0] \\ a[1] \\ .. \\ .. \\ .. \\ a[N-1] \end{bmatrix}$$

Secondly, the corresponding M parameters,

$$b = \begin{bmatrix} b[1] \\ b[2] \\ .. \\ .. \\ .. \\ b[M] \end{bmatrix}$$

Which need to be established with their probability density function (*pdf*) or probability mass function *(pmf)*. The symbol represents as follows *p(a/b)*.

The probability density function (*pdf*) of the noise for one sample *w[n]* is shown below,

$$P(w[n]) = \frac{1}{\alpha\sqrt{2\pi}} exp\left(-\frac{1}{2\alpha^2} w[n]^2\right)$$

and the probability of x[n] becomes

$$P(x[n];A) = \frac{1}{\alpha\sqrt{2\pi}} \exp\left(-\frac{1}{2\alpha^2}(x[n]-A)^2\right)$$

3.15 Error Usage

The first error used is Mean Absolute Percentage Error (MAPE)

$$\text{MAPE} = \frac{1}{n}\sum_{t-1}^{n} \left|\frac{A_t - F_t}{A_t}\right|$$

The lower the error tern signifies that the model is accurate.

The second error used is the Median Absolute Deviation (MAD)

$$\text{MAD} = median_i(\,|X_i - median_j(X_j)|\,)$$

The third error used is the Root Mean Square Deviation (RMSD)

$$(\theta_1,\theta_2) = \sqrt{MSE\ (\theta_1,\theta_2)} = \sqrt{E((\theta_1 - \theta_2)^2)} = \sqrt{\frac{\sum_{i=1}^{n}(x_{1,i} - x_{2,i})^2}{n}}$$

This section discuss about the significant of the model. The best model that is chosen will be the significant level lesser than 0.05. Therefore the lesser the MAPE the accurate and the model will be appropriate.

There is other error term that has been taken into consideration. Which are the MAD and RMSE. The lower the error term that will support the first statement done by the MAPE which is choosing the appropriate model.

3.15.1 Other Error Terms

Besides the MAPE, MAD and also the RMSE, there is also a few other that should be taken into consideration which are the R-Square, Durbin Watson and Ljung Box. When the R-Square is approaching 1 meaning that fitted model explains all the variability. This explains the relationship between the *x* and also predicted value *y*. In other words, it will be the response variable and also the regression. Lastly, is the

Durbin Watson, the best way to choose the model is taking the value lesser than 2 or 2. This is because there is evidence there is a positive correlation.

3.16 Model Checking

This model is checking whether the estimated model confirms to the specifications of a stationary Univariat process. In particular, the residuals should be independent from each other and constant in mean and variance of over time. There is also the Ljung-Box test which can be defined as follows

H_0 : The residuals are stationary
H_1 : Residuals are not stationary
The test statistic is:

$$Q = n(n+2) \sum_{j=1}^{h} \frac{\hat{p}_j^2}{n-j}$$

where:

N	sample size
$\hat{p}_j$	the sample autocorrelation at lag j
h	is the number of lag that have been tested

for the significance level α, the critical region for rejection of the hypothesis of randomness is rejected if $Q > \chi^2_{1-\alpha,h}$, where by $\chi^2_{1-\alpha,h}$ is the α-quantile of the chi-square distribution with h degrees of freedom. The residuals ill be the *ACF* and *PACF*, whereby if it is above the significant line in the bar chart it will not be stationary for the residuals.

3.16.1 Testing of Correlation Coefficient

In order to know that the relationship is strong, the correlation of coefficient should be taken into consideration. Therefore there is a test that needs to be done which is shown below:

$H_0: \rho = 0$ there is no correlation coefficient

$H_1: \rho \neq 0$ there is correlation relationship

Therefore when the ρ approaches 1 then the correlation of coefficient will be strong.

3.16.2 Testing Of Significance

From here there is hypothesis that should be followed that is concerning the significance.

The hypothesis is as follows as below:

H_0: $\beta = 0$ Equation is significant

H_1: $\beta \neq 0$ Equation is not significant

Hereby, the significance gained is lesser than 0 so the equation is significant. Besides that there is also the F model. If the F(model) is bigger than the value from the table therefore it is significant. It is stated as shown:

$$F\text{ (model)} > F_{[\alpha]}$$

Therefore the model will be significant.

CHAPTER 4

RESULTS & DISCUSSION

4.1 Introduction

This section discuss about the appropriate model chose from the results obtained. From here, the things that are considered are the MAPE, MAD and RMSE. Besides the error term the other thing that are also taken into account are R- Square followed by the Ljung Box and also the Durbin Watson method. This therefore state that the closer it approaches 1, the larger is the proportion of the total variation that is explained by the model and the greater is the utility of the model in predicting y.

The appropriate equation model for telecommunication is discussed. Continued by the parameter estimated which was discussed in the methodology. Therefore from here the chosen parameter will be chosen out from the best ARIMA model chosen. On the other hand, the graphs chosen are explained in the section. The graphs which explain more on the Partial Autocorrelation Function and Autocorrelation function to determine the stationary of the error term. Lastly will be the forecasted value and the results obtained as the final result.

4.2 Model Analysis

This section is practically processing and transforming the data or information with the results obtained. The data that is obtained are practically stationary therefore further forecasting could be done. On the other hand, if the data is not stationary than it would be difficult to forecast. Meaning that, data could not be forecasted. The most important step need to be taken in order to continue to analysis the model, the model must be stationary.

4.2.1 ARIMA models

There are various types of ARIMA models. The best one is chosen by considering all the error term, R-Square and also the Durbin Watson method.

Table 1: ARIMA models

Model ARIMA	R-Square	Ljung Box	Durbin Watson	MAPE <0.05	MAD	RMSE
(2,0,2)*(1,0,1)	0.7855	21.11	2.105	0.3954	1.01E+07	1.86E+07
(1,1,1)*(0,1,1)	0.8303	15.92	1.906	0.2483	9.65E+06	1.56E+07
(0,1,1)*(0,1,0)	0.9273	18.93	2.355	0.03113	2.79E+06	4.71E+06
(0,1,3)*(0,1,0)	0.9976	18.41	2	0.01138	9.61	27.36
(0,1,3)*(0,1,1)	0.9974	17.05	2	0.01046	0.03	0.1003
(0,1,3)*(0,1,0)	0.9966	33.22	2	0.01095	1.65	4.59
(0,1,3)*(0,1,0)	0.9968	21.3	1.957	0.01498	1.32	2.43
(0,1,3)*(1,1,0)	0.9979	19.23	2	0.01141	8.08	24.20
(0,1,3)*(0,1,0)	0.9958	12.72	1.963	0.01657	2.78	5.40
(0,1,3)*(0,1,0)	0.9947	11.89	1.875	0.01763	4.56	11.11

From table 1, the ARIMA table which is gained. The ARIMA model that is chosen is the ARIMA (0,1,3)*(0,1,1) because from here all the error term are considered which are the R-square, Durbin Watson, Ljung Box, MAPE, MAD and RMSE.

Any model that result the lowest MAPE, MAD and RMSE produce converged process for parameters estimate the best selected as the best model. The MAPE is 0.01046 produces lesser than the significant level which is lesser than 0.05. Whereas, the other errors MAD and RMSE should result the lowest value as well. Compare to the other ARIMA model 0.03 and 0.1003 are the lowest error model compared to others. Each whereby are the MAD and RMSE. When R^2 approaches 1 meaning that fitted model explains all the variability there is a response variable and regression.

MAPE is the mean absolute percentage error, MAD is the mean absolute deviation and RMSE is the root mean square deviation. From the model we can see that the lowest in this error will be more significant. Therefore we can conclude that the lowest is 0.01046 for MAPE, 0.03 for MAD and 0.1003 for RMSE.

The Ljung-Box from the formula is 12.0170. Therefore, from above value it can be said that the alternative hypothesis is accepted because the calculated hypothesis is 17.05 > 12.0170. The alternative hypothesis is accepted whereby it is stated that the residuals are not stationary.

As for the Durbin Watson, if it is equal to 2 or approximate to 2 it shows that there is a positive serial autocorrelation. In this analysis the Durbin Watson is equals to 2. The correlation is a way to measure how associated or related two variables are. The purpose of doing the correlation is to allow in making the prediction about one variable based on the known variable. Therefore from the results obtained from the table 1 the appropriate model chosen is ARIMA (0,1,3)*(0,1,1).

4.3 The model of Telecommunication Revenue

Telecommunication revenue generated = fixed lines + cellular phones used in Custom
= (subscription tax + establishment proportion Post Payment + subscription number) + (tax Pre payment + subscription rate + rate tax Establishment)

The original equation will be the:

$Y= \alpha_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7$

The equation that is gained from the software is as shown. The equation is taken from the ANOVA table which is the Table 4.

4.3.1 The model Equation

Table 2: Equation Telecommunication Model

Model	Unstandardized Coefficients	Significant
Constant	5.80E+08	0.104
Cpybayar	710794.8	0.156
cp2	-2.00E+08	0.165
bbillag	-138503	0.93
cc	473873	0.169
pp	-47716.7	0.931

The equation is as shown below:

$y = 5.8E{+}08 + 710794.8x_2 + (-2.0E{+}08)\ x_3 + (-138503)\ x_5 + 473873x_6 + (-47716.7)\ x_7$

There has been multicollinearity. Multicollinearity is a statistical phenomenon in which two or three variables are highly correlated. There is a large change in the estimated regression coefficients when a predictor variable is added or deleted. In this case the variables x_1 and x_4 have been deleted.

4.3.2 Testing of Correlation Coefficient

In order to know that the relationship is strong, the correlation of coefficient should be taken into consideration. From the analysis the alternative hypothesis is accepted. Whereby from the results, the R-square has correlation relationship.

Table 3: The correlation of the variable x and y

Model	R	R Square	Adjusted R Square	F Change
1	0.96	0.922	0.791	7.066

4.3.3 Testing of Significance

From here there is hypothesis that should be followed that is concerning the significance. Hereby, the significance gained is lesser than 0 so the equation is significant. Besides that there is also the F model. If the F(model) is bigger than the value from the table therefore it is significant. It is stated as shown:

F (model) > $F_{[\alpha]}$

The numerator will be 5 and the denominator will be 8. Therefore the value is 3.69. From Table 4 the F model is 7.066 > 3.69. The H_0 is rejected meaning that the model is significant.

Table 4: Significance of the ANOVA table.

	Model	Sum of Squares	df	Mean Square	F	Sig
1	Regression	2.63E+16	5	5.25E+15	7.066	0.069
	Residuals	2.23E+15	3	7.44E+14		
	Total	2.85E+16	8			

4.4 Parameter Estimated

From the best model that was chosen the estimated parameter will be the moving average is 0.4894 as for the integrated is -0.5240.

Table 5: Shows the estimated parameter for Autoregressive, Moving Average and Integrated.

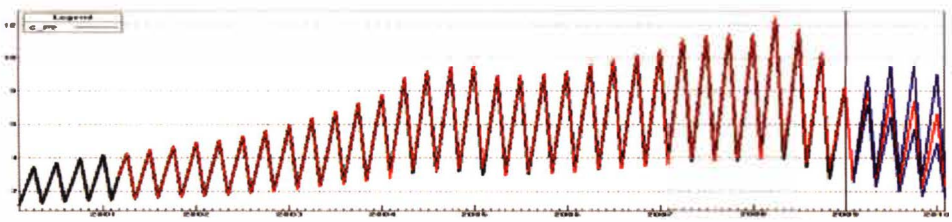
Model ARIMA	AR	Parameter Estimated MA	Integrated
(2,0,2)*(1,0,1)	a[1] = 0.0066	b[1] = 0.0670	-
	a[2] = 0.9842	b[2] = 0.5083	-
(1,1,1)*(0,1,1)	a[1] = -0.5745	b[1] = 0.8835	-
(0,1,1)*(0,1,0)	-	b[1] = 0.9117	-
(0,1,3)*(0,1,0)	-	-	b[3] = -0.5504
(0,1,3)*(0,1,1)	-	b[12] = 0.4894	b[3]= -0.5240
(0,1,3)*(0,1,0)	-	-	b[3] = -0.6551
(0,1,3)*(0,1,0)	-	b[1] = -0.1040	b[3] = -0.4924
		b[2] = -0.0570	
(0,1,3)*(1,1,0)	-	-	b[3] = -0.5229
(0,1,3)*(0,1,0)	-	b[1] = -0.0423	b[3] = -0.4167
		b[2] = 0.0306	
(0,1,3)*(0,1,0)	-	b[1] = -0.0370	b[3] = -0.4580
		b[2] = 0.0811	

From here the ARIMA (0,1,3) equation will be as follows as below, $X_t = 0.4894Z_t - (0.5240)Z^3 + \varepsilon_t$. The following is the moving average and also the integrated. The integrated is in the power of 3. From here the error term will be from MAPE, therefore it will be 0.01046. As overall, the equation will be written as:

$X_t = 0.4894Z_1 - (0.5240)Z^3 + 0.01046$

4.5 Forecasted values

The graph shows that forecasted value for the year 2009. It has a slight increase the gradually it decreases over the time period of time.

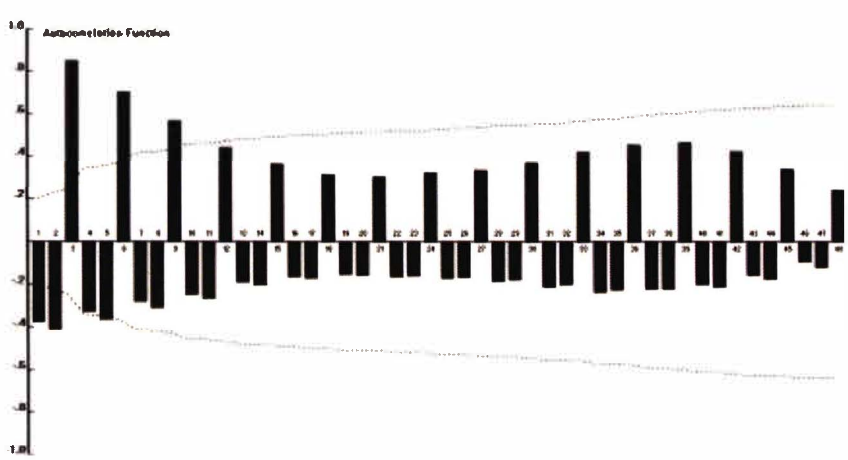


Graph 1: Upper and Lower Forecasted value

Bacically there are two parts, the first part which is the black lines indicates that it is a fitted value. Whereas the red lines indicated the actual value. During the period of 2009 onwards there are practically are 3 lines which are the upper value of 95%, the forecast results and also the lower value of 95%.

4.6 Autocorrelation Function

This shows the ACF, whereby it denotes the moving average movement. This is to analysis the bar chart it is not smooth enough. This is to verify that the residuals are not significant and are not random.

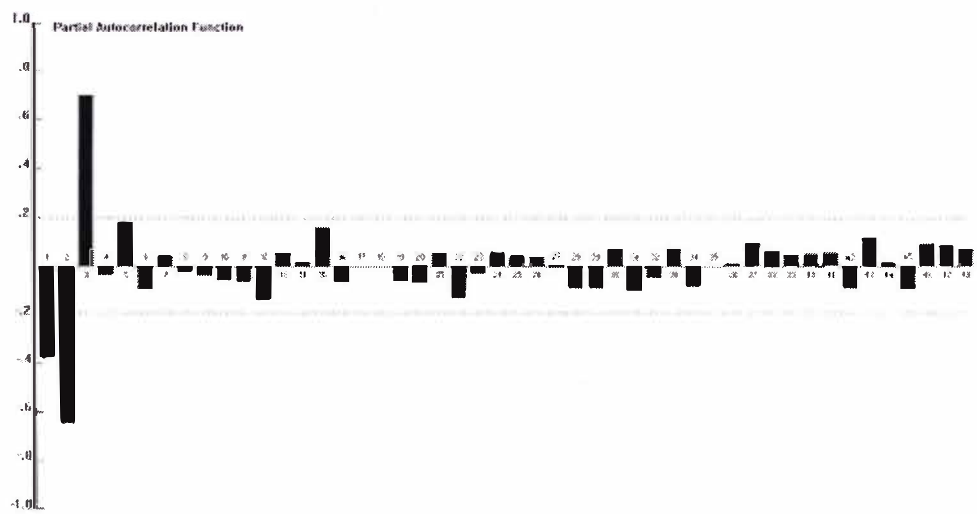


Graph 2: Autocorrelation Function

This is an autocorrelation functions (ACF) obtained, it is a Moving Average (MA). From here the bar chart is in not in the given range. There is a various gape between the upper line and also the lower. The upper and lower are practically brings about the same meaning but normally to see the data is random or not. The bar chart either the upper or lower side therefore it will bring the same results as whichever side that is being analyzed.

4.7 Partial Autocorrelation Function

This shows the PACF, whereby it denotes the autoregressive movement. This is to analysis the bar chart it is not smooth enough. This is to verify that the residuals are not significant and are not random.



Graph 3: Partial Autocorrelation Function

This is a partial autocorrelation functions (PACF) obtained from the software, it is a Autoregressive (AR). From here the bar chart is in not in the given range. There is a various gape between the upper line and also the lower. The upper and lower are practically brings about the same meaning but normally to see the data is random or not. The bar chart either the upper or lower side therefore it will bring the same results as whichever side that is being analyzed.

4.8 Anually Forecasted Value

Below shows that the annully forecasted value for the year 2009. The table shows that forecasted value for the month, followed by the lower and upper value and lastly is the annually forecasted value.

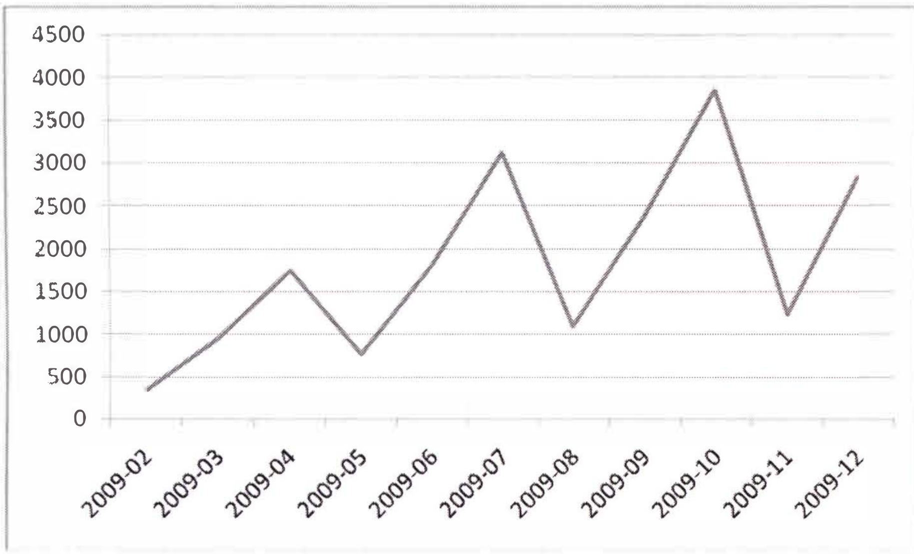
Table 6: Forecasted Value

		ARIMA MODEL (0,1,3) * (0,1,1)		
Date	**Forecast**	**Lower 95%**	**Upper 95%**	**Annually**
2009-02	2667	2505	2839	
2009-03	5333	4881	5827	
2009-04	8000	7177	8916	
2009-05	2608	2257	3013	
2009-06	5216	4386	6201	
2009-07	7823	6420	9533	
2009-08	2462	1977	3066	
2009-09	4924	3876	6256	
2009-10	7386	5708	9556	
2009-11	2211	1680	2911	
2009-12	4423	3307	6130	5916

The table above is a Box Jenkins Forecast Analyses for Telecommunication Revenue based on Service Tax. The year 2009-12 indicates that is the forecasted value for that year which is 5,916 million. The value has a slight increase through out the year after a gradual drop in the year of 2008. There has been only a slight change in the graph and also from the table which is obtained above.

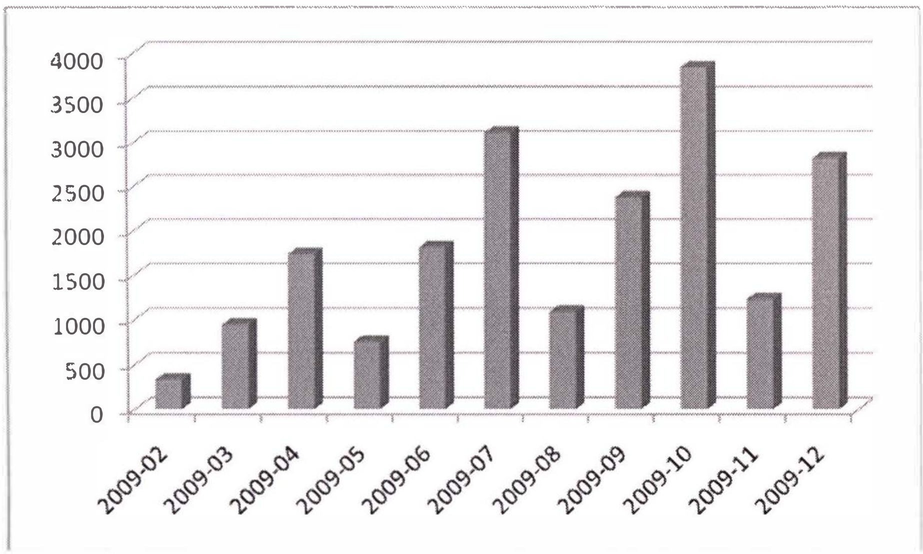
4.9 The Differences

The graph below shows that the difference between the upper value 95% and the lower value 95% for the forecasted value. The graph that is obtained here, there is a large drop and also a large fluctuation for the month February, March, August and November. This happens because there is a festival celebration for the month concerning and also there have been school holidays in the month of November. This explains the fluctuation for the concerning months. The most common festivals that are celebrated will be the Chinese New Year, Hari Raya Puasa and Deepavali. In the month of November, there is a school holiday for this month.



Graph 4: The Difference between upper 95% and lower 95%

The bar chart in graph 5 is the same explaination as the previous graph. This is another form of the graph. From here, it explains the same thing in the previous section. There is a great increase in the month of October because as the school holidays and the festivals are approaching most of the clients or customers whom are dealing with the telecommunication will have the concerning tax updated. Meaning the concern, will have the regarding service tax all paid before going on going on holiday.



Graph 5: The bar of the difference

CHAPTER 5

CONCLUSION

5.1 Introduction

In this paper we have forecasted for the year 2009. Practically I had altogether 96 data for my case study. Therefore I can simplify that the ARIMA model that I have obtained is ARIMA (0, 1, 3). This model have obtained the lowest 0.01046 though there are some mean absolute error percentage are lesser but by considering all the other factors I have obtained the given model.

The further details about the other errors concerning this project have been discussed chapter 4 which is the results. There has been multicollinearity in this model. The forecast which was made by the custom, they have to do further analysis and studying in order to see how is the future income that they have produced. Not only that by doing so they can improve their service as well. By doing this overall analysis they can improve the next future income.

Adding to it if the more data are used the more accurate the forecasted value will be. The data which have been used is to generate the future income in Custom.

Not only that, the ARIMA model can be used to be studies by using various types of other models. This is to test the accuracy of the results obtained. Due to it, there are 2 types of soft wares which is the Forecast Pro from the Customs and also SPSS to generate the equation.

5.2 Suggestion

In this paper work practically there are only 2 models that have been used which are the Forecast Pro and also the SPSS. Besides that there is only one model that is compared which is the ARIMA model. For the future research on telecommunication, besides ARIMA model, fuzzy model also can be used for further research. This research also can be done based on the econometric model it further will discuss about the econometric implementation. From here that can say how much does it benefit and the revenue gained from the economics.

Forecast pro has been good software where it is forecasted and compared with the error terms and choose the suitable and the smallest error value model. Besides this software there are various soft wares that should be taken into consideration which is the Eview and Autobox. The condition in order to identify the error terms, the data must be stationary or else the results will vary greatly. Few soft wares that are used should be compared with the error term in order to find the appropriate model that can be used for the topic. The lesser is the error term the appropriate will be the model gained.

REFERENCE

Absolute deviation (Wikipedia). http://en.wikipedia.org/wiki/Absoluteerror [18 February 2009].

Adnan Al-Mutawkkil, Almas Heshmati, Junseok Hwang. 2009. Development of telecommunication and broadcasting infrastructure indices at the global level. Telecommunication policy 33: 176-199.

Ananda Weliwita & E.M. Ekanayake.1998. Demand for money in Sri Lanka during the post 1977 period: a co integration and error correction analysis

Autocorrelation (Wikipedia). 2007. http://en.wikipedia.org/wiki/Autocorrelation [18 September 2008].

Bowermann, B.L., O'Connel, R.T. 1990. Forecasting and Time Series: *An Applied Approach, 3rd Ed. Duxubury Press, Belmont.*

Constantinos S. Hilas, Sotirios K. Goudos, John N. Saholos.2006. Seasonal Decomposition and Forecasting of Telecommunication Data: A comparative Study. Technology Forecasting & Social Change 73: 495-509.

Data Analysis. 2009. http://en.wikipedia.org/wiki/Data_analysis [2 April 2009].

Detlef Schoder, 2004.Forecasting the success of telecommunication services in the presence of network effects. Information economics and policy 12: 181-200.

Duk bin Jun, Seon Kyoung Kim, Myoung Hwan Park, Yoon Seo Park, Jae Ho Juhn, Chin Kyooh Lee, Young Jin Joo .1997. Classification, Relationship and Forecasting Models for Telecommunication Services.

Duk B. Jun, Seon K. Kim, Yoon S. Park, Myoung H. Park, Amy R. Wilson.2002. Forecasting Telecommunication Service Subscribers in Substitutive and Competitive Environments. International Journal of Forecasting 18: 561-581.

Eisenbach, J.A.1999. The High Cost of Taxing Telecom. The Progress and Freedom Foundation Report.

Farr R.E. 1988. Telecommunications Traffic, Tariffs and Costs :An Introduction for Managers, Peter Peregrinus.

Gary Madden & Joachim Tan. 2007.Forecasting telecommunication data with linear models.Telecommunication Policy 31:31-44

Goldstein, M. & Khan, M.S. 1976. Large versus small price changes in the demand for imports, IMF Staff Papers, **23,** 200± 25

Gomperts,V. 2000:Testimony before the Subcommittee on Oversight of the House Committee on Ways and Means, Hearing on Internet Tax Issues. May 16.

Hellman & Kaufmann.2001. and Fjeldstad and Tungodden .2003. The environmental interference in The Revenue Forecasting.

Hemant Kumar Sabat 2002. The evolving mobile wireless value chain and market structure.Telecommunication Policy 26: 505-535

Interpolation. 2009. http://en.wikipedia.org/wiki/interpolation [20 March 2009].

Jun, D.B., & Park, Y.S 1999. A choice-based diffusion model for multiple generations of products. Technological Forecasting and Social Change, 61, 45-58.

Lukasz Grzybowski & Pedro Pereira. 2009. The consumer loss of the minimum duration for mobile telephone calls. Telecommunication Policy 33: 200-206.

Malaysian Tax Law, Retrieved on Nov 2006, Access: 18/09/2008 http://www.citcindia.org/itr/Nov06/International_taxation_MalyasianTaxLaw.htm

Multicollinearity. 2009. http://en.wikipedia.org/wiki/multicollinearity [18 February 2009].

Peterson, R.A. & Vijay Mahajan. 1978. Multi-Product Growth Models, in Research Marketing, J. Sheth,ed. Greenwich, CT: JAI Press, Inc.,: 201-231.

Robert Fildes, V. Kumar.2002. Telecommunication demand forecasting. *International Journal of Forecasting 18: 489-522.*

S. Mohan & S. Vedula. 1995. Multiplicative Seasonal ARIMA Model for Longterm Forecasting of Inflows. Water resources Management 9:115-126.

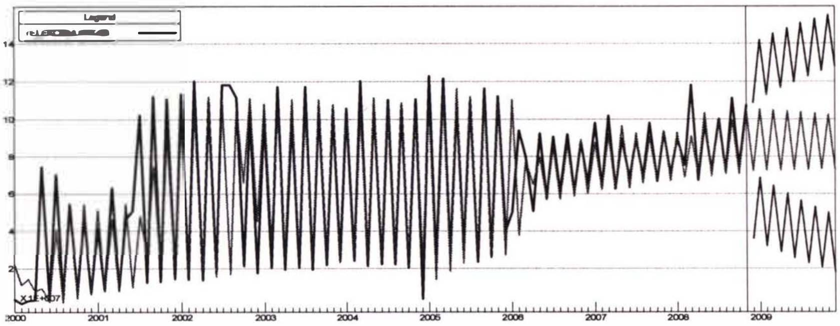
Statistical Package for the Social Sciences. 2009. http://en.wikipedia.org/wiki/SPSS http://en.wikipedia. org/wiki/SPSS [20 March 2009].

Telecommunication. 1 Jun 2008. http:// en.wikipedia.org/wiki/telecommunication [16 September 2008]

Telecommunication Service (tax) Act and Subsidiary Legislation. 2002. http:// www.service.org/acts_Subsidiary_legation [19 September 2008].

Thomas A. Wilson & Micheal Smart. 1997. The tax attributes and Securities. http://en.wikipedia.org/wiki/taxatributes [20 September 2008].

Appendix A : Output Results for Forecast Model for Telecommunication



Forecast Model for Telekommunikasi (Range : 2000-2008)
ARIMA(2,0,2)*(1,0,1)

Term	Coefficient	Std. Error	t-Statistic	Significance
a[1]	0.0066	0.0178	0.3717	0.2891 <-
a[2]	0.9842	0.0179	55.0323	1.0000
b[1]	0.0670	0.0848	0.7900	0.5686 <-
b[2]	0.5083	0.0834	6.0978	1.0000
A[12]	0.1532	3.6662	0.0418	0.0333 <-
B[12]	0.1487	3.6605	0.0406	0.0323 <-
_CONST	508707.3094			

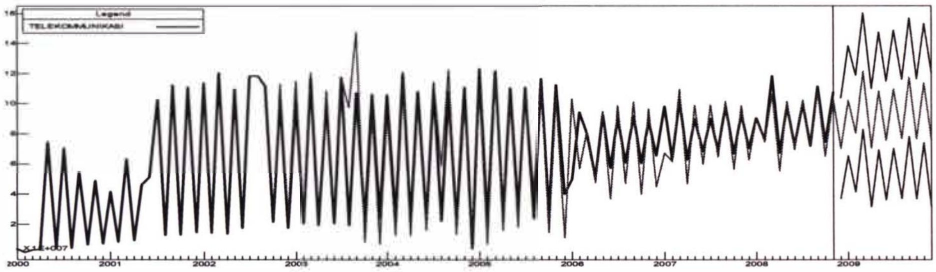
Try alternative model ARIMA(2,0,2)*(0,0,1)

Within-Sample Statistics

Sample size 107	Number of parameters 6
Mean 6.538e+007	Standard deviation 4.032e+007
R-square 0.7855	Adjusted R-square 0.7749
Durbin-Watson 2.105	Ljung-Box(18)=21.11 P=0.726
Forecast error 1.913e+007	BIC 2.119e+007
MAPE 0.3954	RMSE 1.859e+007
MAD 1.005e+007	

Forecasted Values

Date	2.5 Lower	Forecast	Quarterly	Annual	97.5 Upper
2008-12	35907900.000	72333896.000	254147392.000	1058473784.000	108759896.000
2009-01	68890320.000	105382632.000		141874944.000	
2009-02	32121868.000	72515552.000		112909232.000	
2009-03	64485252.000	104930920.000	282829104.000		145376592.000
2009-04	28736466.000	72622160.000		116507856.000	
2009-05	60335900.000	104262928.000		148189952.000	
2009-06	25775528.000	72788032.000	249673120.000		119800536.000
2009-07	56662524.000	103708128.000		150753728.000	
2009-08	23076688.000	72923512.000		122770336.000	
2009-09	53337208.000	103210608.000	279842248.000		153084000.000
2009-10	20623444.000	73063136.000		125502824.000	
2009-11	50186236.000	102647240.000		155108240.000	
2009-12	18305144.000	73181560.000	248891936.000	1061236408.000	128057976.000



Forecast Model for Telekommunikasi)2001, Sept – 2008)
ARIMA(1,1,1)*(0,1,1)

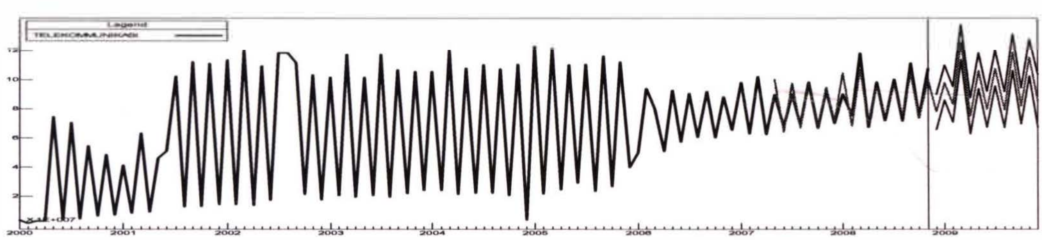
Term	Coefficient	Std. Error	t-Statistic	Significance
a[1]	-0.5745	0.0990	-5.8045	1.0000
b[1]	0.8835	0.0571	15.4657	1.0000
B[12]	0.4822	0.0829	5.8200	1.0000

Within-Sample Statistics

Sample size 87	Number of parameters 3
Mean 7.333e+007	Standard deviation 3.806e+007
R-square 0.8303	Adjusted R-square 0.8263
Durbin-Watson 1.906	Ljung-Box(18)=15.92 P=0.4015
Forecast error 1.586e+007	BIC 1.683e+007
MAPE 0.2483	RMSE 1.559e+007
MAD 9.653e+006	

Forecasted Values

Date	2.5 Lower	Forecast	Quarterly	Annual	97.5 Upper
2008-12	37026696.000	70148032.000	251961528.000	1056287920.000	103269368.000
2009-01	65283472.000	101713216.000		138142960.000	
2009-02	41487648.000	80026352.000		118565056.000	
2009-03	82656968.000	121342264.000	303081832.000		160027568.000
2009-04	31477796.000	70594312.000		109710832.000	
2009-05	69091864.000	108211992.000		147332128.000	
2009-06	36113124.000	75394344.000	254200648.000		114675560.000
2009-07	70033544.000	109356792.000		148680032.000	
2009-08	37166764.000	76590640.000		116014520.000	
2009-09	77826456.000	117314032.000	303261464.000		156801600.000
2009-10	36961548.000	76532744.000		116103944.000	
2009-11	73822240.000	113465064.000		153107888.000	
2009-12	31940764.000	76182464.000	266180272.000	1126724216.000	120424168.000



Forecast Model for Telekommunikasi 2006, april – 2008)
ARIMA(0,1,1)*(0,1,0)

Term	Coefficient	Std. Error	t-Statistic	Significance
b[1]	0.9117	0.0587	15.5234	1.0000

Within-Sample Statistics

Sample size 32	Number of parameters 1
Mean 8.168e+007	Standard deviation 1.777e+007
R-square 0.9273	Adjusted R-square 0.9273
Durbin-Watson 2.355	Ljung-Box(18)=18.93 P=0.6039
Forecast error 4.789e+006	BIC 4.976e+006
MAPE 0.03113	RMSE 4.714e+006
MAD 2.785e+006	

Forecasted Values

Date	2.5 Lower	Forecast	Quarterly	Annual	97.5 Upper
2008-12	65751172.000	77740584.000	259554080.000	1063880472.000	89730000.000
2009-01	85835232.000	97871280.000		109907328.000	
2009-02	70964472.000	83046976.000		95129480.000	
2009-03	113681864.000	125810640.000	306728896.000		137939408.000
2009-04	62821380.000	74996256.000		87171128.000	
2009-05	93939256.000	106160056.000		118380856.000	
2009-06	67246472.000	79513024.000	260669336.000		91779576.000
2009-07	95558608.000	107870744.000		120182880.000	
2009-08	66978716.000	79336272.000		91693824.000	
2009-09	106501984.000	118904792.000	306111808.000		131307600.000
2009-10	69449376.000	81897264.000		94345152.000	
2009-11	102797568.000	115290376.000		127783184.000	
2009-12	67363384.000	85427656.000	282615296.000	1156125336.000	103491928.000

Appendix B : Output Results for Establishment Proportion

Forecast Model for C_P2
ARIMA(0,1,3)*(0,1,1) with log transform

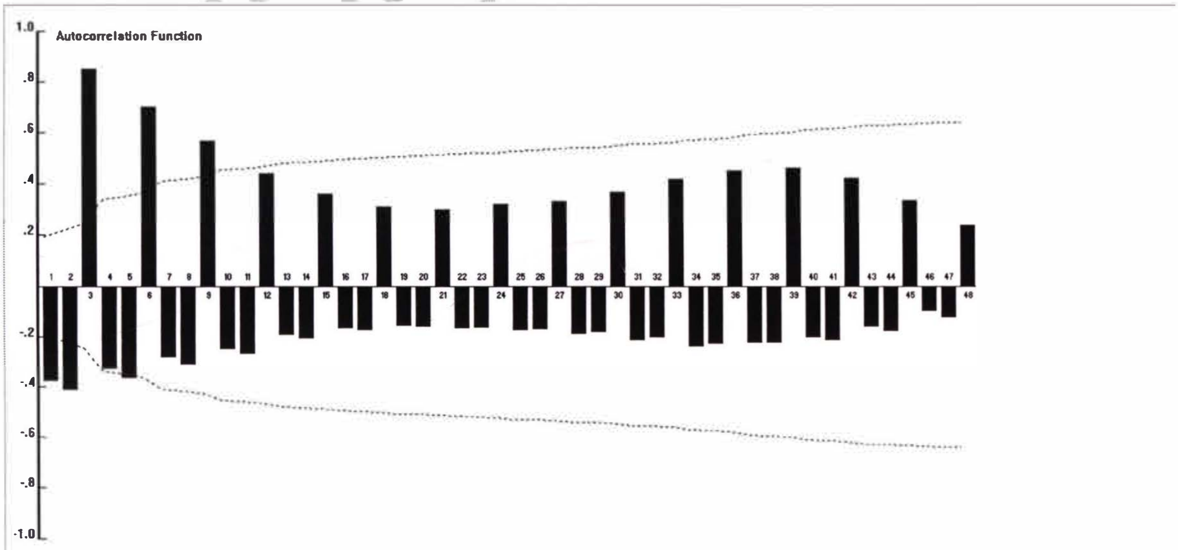
Term	Coefficient	Std. Error	t-Statistic	Significance
b[1]	0.0000	0.0945	0.0000	0.0000 <-
b[2]	0.0000	0.0945	0.0000	0.0000 <-
b[3]	-0.5240	0.1132	-4.6303	1.0000
B[12]	0.4894	0.1780	2.7491	0.9928

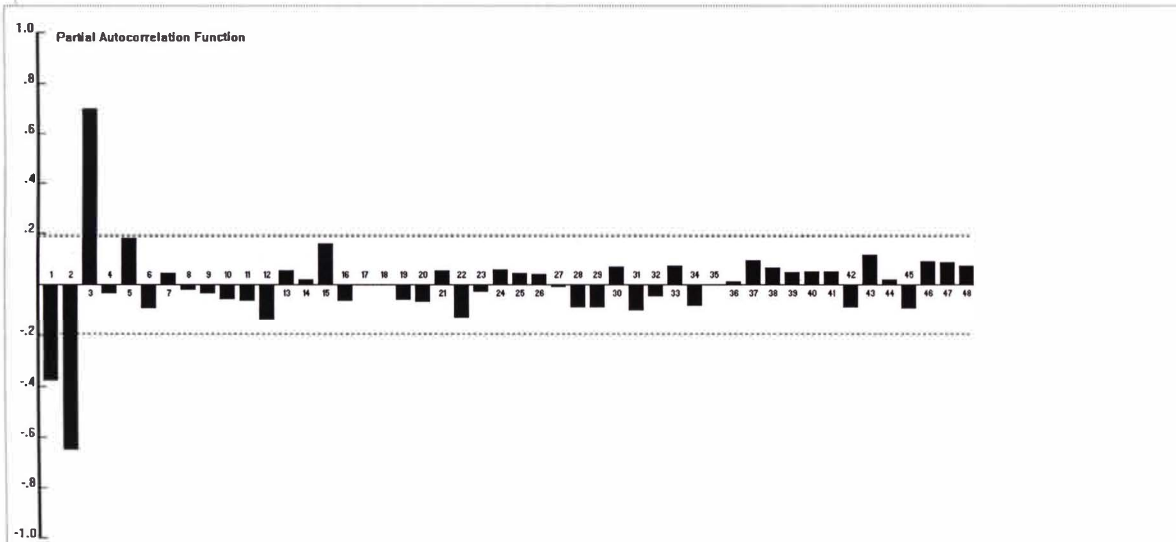
Within-Sample Statistics

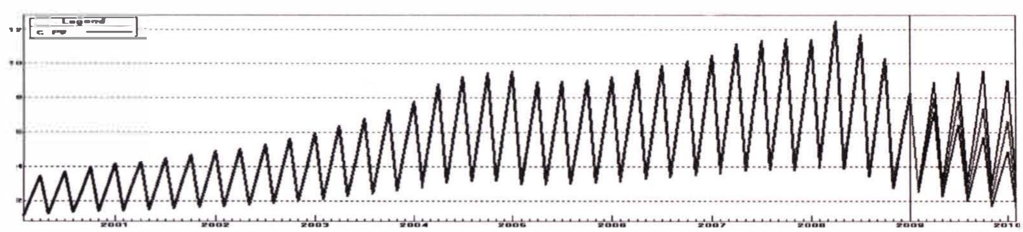
Sample size 108	Number of parameters 4
Mean 1.504	Standard deviation 0.593
R-square 0.9974	Adjusted R-square 0.9973
Durbin-Watson 2	Ljung-Box(18)=17.05 P=0.4804
Forecast error 0.03053	BIC 0.147
MAPE 0.01046	RMSE 0.1003
MAD 0.03301	

Forecasted Values

Date	2.5 Lower	Forecast	Quarterly	Annual	97.5 Upper
2009-02	2.505	2.667		2.839	
2009-03	4.881	5.333	16.250		5.827
2009-04	7.177	8.000		8.916	
2009-05	2.257	2.608		3.013	
2009-06	4.386	5.216	15.823		6.201
2009-07	6.420	7.823		9.533	
2009-08	1.977	2.462		3.066	
2009-09	3.876	4.924	15.209		6.256
2009-10	5.708	7.386		9.556	
2009-11	1.680	2.211		2.911	
2009-12	3.307	4.423	14.020	61.30	5.916
2010-01	4.885	6.634		9.010	
2010-02	1.616	2.251		3.137	


Autocorrelation Function


Partial Autocorrelation Function


Legend

Appendix C: Telecommunication data

Jan-00	3,515,914.96	Oct-02	21,337,257.84
Feb-00	1,560,302.06	Nov-02	103,310,556.40
Mar-00	2,987,701.46	Dec-02	17,505,380.24
Apr-00	3,159,391.72	Jan-03	101,718,552.70
May-00	74,492,298.13	Feb-03	20,497,704.76
Jun-00	3,963,294.85	Mar-03	117,222,307.10
Jul-00	70,469,335.36	Apr-03	19,389,064.90
Aug-00	4,280,367.16	May-03	101,718,552.70
Sep-00	54,599,648.40	Jun-03	20,497,704.76
Oct-00	6,444,648.64	Jul-03	117,222,307.10
Nov-00	48,777,773.05	Aug-03	19,389,064.90
Dec-00	6,982,705.37	Sep-03	106,864,245.40
Jan-01	41,510,225.67	Oct-03	21,847,716.17
Feb-01	8,384,020.02	Nov-03	105,791,355.60
Mar-01	63,284,158.70	Dec-03	24,168,215.72
Apr-01	9,356,973.82	Jan-04	105,595,524.00
May-01	46,218,038.12	Feb-04	24,147,440.20
Jun-01	50,986,061.90	Mar-04	120,216,900.00
Jul-01	102,324,493.30	Apr-04	21,468,334.72
Aug-01	12,696,491.42	May-04	107,868,762.40
Sep-01	112,184,069.60	Jun-04	22,161,174.10
Oct-01	12,918,015.12	Jul-04	110,387,131.80
Nov-01	110,947,840.70	Aug-04	22,170,654.83
Dec-01	14,495,427.16	Sep-04	107,404,501.70
Jan-02	113,406,338.40	Oct-04	20,583,927.15
Feb-02	14,262,862.60	Nov-04	110,541,826.30
Mar-02	120,350,659.40	Dec-04	3,963,045.98
Apr-02	13,791,444.95	Jan-05	122,820,036.70
May-02	109,476,771.20	Feb-05	21,707,625.52
Jun-02	17,290,012.34	Mar-05	121,628,948.10
Jul-02	118,241,440.00	Apr-05	24,682,373.92
Aug-02	118,241,440.00	May-05	110,270,197.00
Sep-02	111,583,035.50	Jun-05	29,144,512.89

Jul-05	110,457,014.80
Aug-05	23,694,958.58
Sep-05	116,194,154.00
Oct-05	26,873,203.43
Nov-05	112,095,505.50
Dec-05	40,017,687.40
Jan-06	50,138,990.67
Feb-06	94,079,948.04
Mar-06	79,986,802.28
Apr-06	50,783,741.73
May-06	92,587,386.01
Jun-06	57,283,515.91
Jul-06	90,624,162.45
Aug-06	60,622,025.44
Sep-06	91,761,399.66
Oct-06	60,319,360.95
Nov-06	88,254,907.74
Dec-06	65,306,754.51
Jan-07	98,193,261.23
Feb-07	63,089,087.24
Mar-07	102,041,550.18
Apr-07	62,501,295.87
May-07	89,625,041.30
Jun-07	67,858,978.34
Jul-07	90,017,775.92
Aug-07	68,848,463.11
Sep-07	98,307,834.28
Oct-07	66,672,339.83
Nov-07	93,005,611.58
Dec-07	70,053,515.40
Jan-08	90,184,207.08
Feb-08	75,359,901.91
Mar-08	118,123,564.10

Apr-08	67,309,181.70
May-08	98,472,986.32
Jun-08	71,825,954.76
Jul-08	100,183,669.93
Aug-08	71,649,196.11
Sep-08	111,217,717.38
Oct-08	74,210,189.98
Nov-08	107,693,303.50
Dec-08	248,990.47

Appendix D: Data Concerning Tax in Telecommunication

Year	Subsciption tax	Estab_prop	Post_pay	Sub_No	Tax_pay	Sub_Rate	Rate_Establi
2000	5122	4628	2599	1236	2523	3392	66.4
2001	7385	4710	3069	1310	4316	3400	65.1
2002	9053	4670	2961	1347	6092	3323	62.3
2003	11124	4572	2566	1378	8558	3194	58
2004	14611	4446	2555	1508	12057	2934	52.3
2005	19545	4366	2925	1527	16620	2839	49.5
2006	19464	4342	3368	1511	16096	2831	48.3
2007	21864	4352	3618	1489	18246	2863	48.4
2008	25167	4319	4506	1550	20662	2768	45.9

FORECASTING ON TELECOMMUNICATION REVENUE BASED ON SERVICE TAX USING THE ARIMA MODEL - THERINA OLIVIA RAJATURAI