THE APPLICATION OF DISAGGREGATION METHODS TO THE UNEMPLOYMENT RATE OF TURKEY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

UTKU GÖKSEL TÜKER

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
STATISTICS

SEPTEMBER 2010

Approval of the thesis:

THE APPLICATION OF DISAGGREGATION METHODS
TO THE UNEMPLOYMENT RATE OF TURKEY

submitted by **UTKU GÖKSEL TÜKER** in partial fulfillment of the requirements for the degree of **Master of Science in Department of Statistics, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** _______________

Prof. Dr. H. Öztaş Ayhan
Head of Department, **Department of Statistics** _______________

Assist. Prof. Dr. Ceylan Talu Yozgatlıgil
Supervisor, **Department of Statistics, METU** _______________

Examining Committee Members:

Assist. Prof. Dr. Berna Burçak Başbuğ Erkan ____________________
Department of Statistics, METU

Assist. Prof. Dr. Ceylan Talu Yozgatlıgil ____________________
Department of Statistics, METU

Assist. Prof. Dr. Özlem İlk ____________________
Department of Statistics, METU

Assist. Prof. Dr. Zeynep Işıl Kalaylıoğlu ____________________
Department of Statistics, METU

Cem Baş (M.Sc.) ____________________
TUIK

Date: 15.09.2010

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Utku Göksel TÜKER

Signature:

ABSTRACT

THE APPLICATION OF DISAGGREGATION METHODS
TO THE UNEMPLOYMENT RATE OF TURKEY

Tüker, Utku Göksel
M. Sc., Department of Statistics
Supervisor: Assist. Prof. Dr. Ceylan Talu Yozgatlıgil

September 2010, 97 pages

Modeling and forecasting of the unemployment rate of a country is very important to be able to take precautions on the governmental policies. The available unemployment rate data of Turkey provided by the Turkish Statistical Institute (TURKSTAT) are not in suitable format to have a time series model. The unemployment rate data between 1988 and 2009 create a problem of building a reliable time series model due to the insufficient number and irregular form of observations. The application of disaggregation methods to some parts of the unemployment rate data enables us to fit an appropriate time series model and to have forecasts as a result of the suggested model.

Key words: Disaggregation methods, time series data, interpolation approaches, regression based methods, modeling procedures.

ÖZ

TÜRKİYE İŞSİZLİK ORANI ÜZERİNE
DAĞITIM YÖNTEMLERİ UYGULAMASI

Tüker, Utku Göksel

Yüksek Lisans, İstatistik Bölümü

Tez Yöneticisi: Yrd. Doç. Dr. Ceylan Talu Yozgatlıgil

Eylül 2010, 97 sayfa

Bir ülkenin işsizlik oranını modellemek ve ileriye yönelik öngörülerde bulunmak, hükümet politikaları üzerine önlem almak açısından çok önemlidir. Türkiye İstatistik Kurumu (TÜİK) tarafından sağlanan işsizlik oranı verileri bir zaman serisi modeli kurmak için düzgün formda değildir. 1988 ile 2009 yılları arası işsizlik oranı verisi, yetersiz miktarda ve düzensiz yapıda gözlemlere bağlı olarak güvenilir bir zaman serisi modeli kurmak açısından sorun yaratmaktadır. İşsizlik oranı verilerinin bir kısmına dağıtım yöntemleri uygulamak uygun zaman serisi modelini kurmamızı ve bu model için öngörüler elde etmemizi sağlayacaktır.

Anahtar Kelimeler: Dağıtım yöntemleri, zaman serisi verileri, enterpolasyon yaklaşımları, regresyona dayalı yöntemler, modelleme prosedürleri.

To my family

and

my love…

ACKNOWLEDGEMENTS

I would like to express my earnest gratefulness to my supervisor Assist. Prof. Dr. Ceylan Yozgatlıgil for accepting me to work with, guidance, assistance, welcoming behavior and unlimited patience in my stressed times during the thesis study.

I also would like to thank my examining committee members, Assist. Prof. Dr. Berna Burçak Başbuğ Erkan, Assist. Prof. Dr. Özlem İlk and Assist. Prof. Dr. Zeynep Işıl Kalaylıoğlu for their acceptance of reviewing my thesis study by spending their invaluable time. I owe most of my statistical background to them through the period of my undergraduate education. Therefore, it is again proud-making to present one of my studies to my teachers one more time.

I owe my special gratitude to another member of my examining committee Cem Baş for arranging a time to attend my thesis review despite all the busy state of affairs he is in.

In addition, I would like to thank The Scientific and Technological Research Council of Turkey (TÜBİTAK) for the financial support during my graduate education.

I also convey my thanks to my family and my love, Neşe, for their never-ending support, patience, love and standing by my side in every circumstance. Without them, it has always been hard to cope with the situations that I am in, even the most unimportant one. They were, are and will always be the meaning of my life. I owe everything that I possess to them.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

ABBREVIATIONS

WN	White Noise
MA	Moving Average
AR	Autoregressive
ARMA	Autoregressive Moving Average
ARCH	Autoregressive Conditional Heteroskedasticity
ARIMA	Autoregressive Integrated Moving Average
SARIMA	Seasonal Autoregressive Integrated Moving Average
AUR	Actual Unemployment Rate
UR	Unemployment Rate (Disaggregated)
gnp	GNP (Gross National Product) (in purchasers' value)
gs	Government Services
ind	Industry
agr	Agriculture
trad	Trade

CHAPTER 1

INTRODUCTION

In statistics, modeling is a vital issue in cases where a considerable and solid mechanism is desired to be realized. A model can tell how the population of the data sampled has behaved up to now or will behave in the future. The idea of modeling lies on the basis of data usage. By the use of observations in data representing the population, a mathematical model is constructed with some conditions assumed. These conditions include the reasons which affect the behavior of the data but can not be explained in a scientific manner. In statistics, these conditions are called assumptions. Actually, a statistical model is beyond a simple mathematical model because the assumptions being used somehow confirm that the data behavior fits the population behavior.

A statistical model gives the historical and future behavior of the data. Each observation in the data is very important for the construction of a model. Few observations would also give us a model. However, would it be really reliable for statistical inference? Statisticians usually prefer to have enough observations to construct a model. More observations in a sample give more information about the population, since a sample is a representation of that population. With one observation there is nothing to say about the population. On the other hand, if the ninety percent of the observations in a population is taken as a sample, it will be more likely to get a realistic idea about the population. In statistics, the appropriate model to be fitted contains coefficients of explanatory variables that are determining the values taken by the dependent variable. These coefficients, known as parameters, are estimated by means of observations in the sample. Each of the estimated coefficients corresponding to these parameters is named as an estimate. Estimates are the sample statistics and these sample statistics reveal meaningful estimates as long as they represent the population with enough information. There are different types of models in statistics such as linear regression models, logistic regression models and time series models. They give close fitted values obtained by sample statistics to the actual population observations once they are built with sufficient information. Sufficient information is provided by the sample size in hand selected. Linear and logistic regression models do not have to consider the time dependency unlike the time series models. Therefore, the existence of time dependency inevitably makes the sufficient number of sample observations much more crucial for time series models, because of the sequential structure they hold.

The data to be used in building a time series model can be collected monthly or quarterly and the number of observations in the collected data may be enough. In the same way, the data may be collected annually with a large sample size and they could still not constitute a problem in building a model. However, it is not always easy to have sufficient number of observations. For example, a five-year-collected annual data creates a problem in model identification and specification. In such cases, it is better to look for some solutions in order to conduct a modeling analysis instead of refusing to build a model due to the existence of few observations.

Today, the studies conducted on such cases in order to have adequate number of observations are known as “Time Series Disaggregation Methods”. Although these methods are generally used for deriving quarterly figures from annual ones, they can also be useful in deriving monthly observations from quarterly figures. While a direct derivation of monthly data from annual data is available, someone can obtain quarterly observations at first and then get the monthly ones as the latter step.

In this study, we have considered the data of “Unemployment Rate of Turkey” (prepared by the Turkish Statistical Institute (TURKSTAT)) between the years of 1988 and 2009. However, from 1988 to 1999, the data are available only for the months of April and October. From 2000 to 2004, the quarterly figures are observed. And from 2005 to 2009, the data are collected monthly up to September, 2009. The irregular structure of data in hand has led us to use some of the disaggregation methods and derive the quarterly figures via the use of annual ones so that we can build a reliable model through sufficient number of observations.

One of the reasons of choosing the unemployment rate data of Turkey in our analysis comes from the importance of it for a country. A model estimation and having forecasts for the unemployment rate of a country may be helpful in taking precautions on governmental policies for the future. In literature it is likely to reach some studies focusing on the unemployment rate and its value. For instance, the study of Montgomery et al. (1998) is directly focused on forecasting the unemployment rate of U.S. Holden and Peel (1979) investigated the factors of unemployment rate in some countries. As another example, an analysis of how unemployment rate in European countries is affected by demographic and educational structure was conducted by Biagi and Lucifora (2008). Another reason for the idea of working on the unemployment rate data stems from the fact that it is an indicator of both economic and social structure of a country. Numerous papers support this ascertainment. On the economic side, it was stated by Barro (1977) that unemployment rate is considered to be a real economic variable of a country. Jackman et al. (1990) pointed out that increase in unemployment rate is a governmental problem and governments should build up new policies to tackle this issue. Another point that researchers believe is that unemployment plays role in migration. According to a study by Da Vanzo (1978), unemployed people together with people not satisfied with their jobs are much more likely to migrate compared to ones having active job. Similar to Da Vanzo (1978), McCormick (1997) studied the relation between regional unemployment and migration in the UK and finds out that unemployed people migrate to places with relatively high employment rates. One other scope that we face as an impact of unemployment is crime. In a study by Cantor and Land (1985), they made an effort to understand the link between crime and unemployment. In this study, seven indexes of crime rates were investigated to see whether unemployment leads to crime. Results revealed that being unemployed motivates crime. Many articles declare that the most common drawbacks resulting from unemployment are mental and psychological problems. Linn et al. (1985) and Kessler et al. (1988) reported that unemployed people are more likely to have the symptoms of depression and anxiety. Study by Clark and Oswald (1994) worked on the hypothesis that people may choose to be unemployed owing to the attractive aids from governments in Britain. Study ends in rejecting the hypothesis as well as stating that unemployment has more severe effects than divorce or marital separation. As it is understood from research examples above, unemployment is a very leading factor in economic and social life. Due to this fact discovering the structure of unemployment and being able to forecast it is very crucial. Therefore my intention to analyze unemployment rate is clarified.

The studies on time series disaggregation methods are very new in the literature and have few or no case conducted in Turkey. Although the study made by Sürmeli (1979) is comprised of mostly mathematical disaggregation methods, the method including a statistical model based approach that is examined has a drawback. The newer and developed related series approaches were not available when the Sürmeli conducted his study. In addition, the analysis he conducted is on solely the disaggregation methods and has no use of model identification, estimation and selection process being applied on real data. That is why the analyses that have been made in this thesis study are thought to have substantial contributions to the literature.

The organization of this thesis study has been made by mentioning the history of disaggregation methods in chapter two. Then we have given some idea about the basic time series concepts and the theoretical background of the disaggregation methods applied in the third chapter by dividing it into two sections. After that, we have applied some disaggregation methods in order to obtain the disaggregated data in the first section of chapter four. In the second section of the same chapter, we have focused on modeling procedures through the use of generated data and had forecasts of the models built. Finally, we have finished our study by presenting the conclusions that we have reached through what we have done in chapter five.

CHAPTER 2

HISTORICAL BACKGROUND

2.1. History of Disaggregation Methods

The history of time series disaggregation methods dates back to 1962. Either mathematical or statistical procedures have been developed since then. While the mathematical procedures are generally focusing on smoothing procedures, the statistical ones use model based approaches in the process of obtaining disaggregated data.

The first step to disaggregation methods was taken by Friedman (1962). Friedman claimed that the unobserved observations to be disaggregated would be estimated by the use of related variables and the observations in hand. However, the estimated observations given by the method he suggested did not meet some restrictions. These were called accounting restrictions. To illustrate, the consumption of food in a country is a flow variable, i.e. the sum of monthly disaggregated data should give the annual consumption of that year. If this restriction is not met, the disaggregated data is no longer valid for a model to be built. Hence, the method suggested by Friedman was incomplete.

Afterwards, these restrictions were given importance and taken into account by ongoing studies. Lisman and Sandee (1964) proposed a method of linear interpolation which was considering the trend and the seasonal components of a time series and providing the disaggregated data with the help of matrix calculations. This method was satisfying the accounting restrictions stated as well, but was not giving the disaggregated observations for the first and last year due to its application. Moreover, there were methods applying least squares approaches for getting disaggregated data. Boot et al. (1967) and Denton (1971) suggested methods based on minimization criterion of the differences between higher frequency (disaggregated) observations. Nevertheless, in the simulation study conducted by Kladroba (2005), the least squares methods had very high expenditure and uncertainty in addition to the incompatible results they gave.

In addition to these disaggregation procedures, there have been many studies generating sub-annual values via the help of only aggregated data. The methods of Cohen et al. (1971), Harvey and Pierse (1984), Al-Osh (1989) and Gomez and Maravall (1994) can be given as illustrations of such studies. These methods were criticized due to not using related variables, because they were using just some mathematical calculations and failing to involve more information than the aggregated data. Thus, there was a tendency to use more “statistical” criteria in disaggregation methods. Balmer (1975) interpreted the procedures using both mathematical and statistical criteria through the use of related series suggested by Vangrevelinghe (1966) and Ginsburgh (1973). These methods were using adjustment of the series interpolated with related series. However, Balmer stated that these methods should be improved in their estimation technique unless the uncorrelatedness assumption presented in the residuals is satisfied.

The most popular univariate regression-based related series approach applied today was proposed by Chow and Lin (1971) and its extensions (Chen, 2007). The method of Chow and Lin (1971) uses indicators in disaggregation process without ignoring the accounting restrictions. This fact and the effective results given by Chow/Lin method have a considerable effect on its popularity. Then, Fernandez (1981) and Litterman (1983) had considerable contributions on the development of this method. According to the study of Miralles et al. (2003) on the performance of the Chow/Lin method, the estimates revealed by the procedure were robust. Together with this, the studies on the accuracy of estimates had also been suggested through the methods of Fernandez (1981) and Litterman (1983), since these procedures are distinct from the one of Chow/Lin in the assumption that they are making for the error terms while generating the disaggregated series. Nevertheless, Guerrero (2003) pointed out that these methods are not away from subjectivity in using autocorrelation structure of the time series data. Moreover, he added that Guerrero (1990) and Wei and Stram (1990) do not have such a drawback and found solutions with paying attention to the autocorrelation structure of the data. These methods were using an ARIMA based approach to obtain disaggregated series with the help of related series. Nonetheless, in Tour Europe, La Defense, Paris (OECD/Eurostat Workshop); these time series procedures were said to have problem in recovering the seasonal pattern generally by Di Fonzo (2003). That is why regression based methods are still widely used today.

Generally, there are three types of time series disaggregation methods; namely, linear interpolation approaches, model based approaches and least squares methods. Model based ones are divided into two: regression based and ARIMA based procedures. The linear interpolation and least squares approaches are mathematical procedures which are known as smoothing procedures, whereas the model based approaches pay attention to statistical criterion together with the mathematical procedures. In a general picture, we can classify the time series disaggregation methods as follows;

A) Linear Interpolation Approaches

- Lisman/Sandee (1964)

B) Model Based Approaches

B1) Regression Based Approaches

- Chow/Lin (1971)
- Fernandez (1981)
- Litterman (1983)

B2) ARIMA Based Time Series Procedures

- Al-Osh (1989)
- Guerrero (1990)
- Wei/Stram (1990)

C) Least Squares Methods

- Boot/Feibes/Lisman (1967)
- Denton (1971)

Because there are so many disaggregation methods, either mathematical or statistical; the question of which methods should be used inevitably arises. The procedures which are not getting use of related series approach have a drawback of generating high-frequency observations being far away to involve more information compared to the low-frequency (aggregated) observations (Miralles et al., 2003). As stated, the least squares methods were also found to be ineffective in providing disaggregated data close to actual values in the study of Kladroba (2005) although they had very high expenditure and uncertainty. Moreover, a least square procedure would most probably fail to generate a realistic disaggregated data for a series having seasonal component, because it gives quarterly figures close to each other due to using a minimization procedure through annual values.

Although ARIMA based procedures use statistical criterion, they were said to have problem in recovering the seasonal pattern. Therefore, it will not be wise to apply these procedures for seasonal series in the process of disaggregation.

The choice of methods used in this analysis was made by considering the appropriateness of the data structure as stated before. After the inspection of unemployment rate data of Turkey, the seasonal pattern in the series was suspected. Therefore, we did not prefer to use a method of least squares and time series procedures. Even though, the mathematical procedures fail to give disaggregated data involving more information than the aggregated data, we chose to apply the one being most popular among all and having low expenditure, namely; Lisman/Sandee (1964). Another reason why the preference has been made on the method of Lisman/Sandee among all mathematical procedures results from the fact that the seasonal pattern is taken into account during the process of generating disaggregated series. Moreover, for a model based approach, the method suggested by Chow/Lin (1971), Fernandez (1981) and Litterman (1983) had been preferred to be applied. Actually, the methods of Fernandez (1981) and Litterman (1983) are slightly different than the one proposed by Chow/Lin (1971). They are all the same except the distribution assumption of the error terms. Therefore, the one which gave the most reasonable results had been chosen.

CHAPTER 3

METHODOLOGY

In this chapter, first section gives general information about the basics of a stationary time series in addition to the procedures used in a time series analysis. Then, in the second section, the details of disaggregation methods used in the study are presented.

3.1. Basic Time Series Concepts

A time series is a set of observations generated sequentially in time which are statistically dependent observations. It not only is a function of time but also may be a function of its past observations. Each observation in a time series is crucial in realizing the major objectives of the time series analysis such as identification, estimation, diagnostic checks and forecasting and past observations would have effect on the series itself. A time series analysis aims modeling the stochastic mechanism that generates the observed series at first and then having forecasted future values based on history via the constructed model.

In modeling a time series, four cases can be observed;

- The series is dependent on an error term,
- The series is dependent on both an error term and the past observations of it,
- Or the series depends on both its past observations and an error term,
- Or the series depends on both its past observations and an error term together with the past observations of it.

For the first one, the series is known to follow a white noise process, WN. These series are accepted to be stationary and independently and identically distributed. The details of this process will be presented later on.

In the second case the time series follows a moving average model, MA. Such series have constant mean and variances. The observations are not affected by their past.

In the third case, the behavior of the series changes due to the values observed for the past observations of the series. The model of such series is known as autoregressive model, AR.

The last one corresponds to autoregressive moving average process, ARMA. This process is actually a combination of moving average and autoregressive processes. The details of this process can be seen later on as well.

For an observed time series $\{Y_t\}$, there are four main functions; namely,

Mean Function:

$$E(Y_t) = \mu_t \tag{3.1.1}$$

Variance Function:

$$E[(Y_t - \mu_t)^2] = Var(Y_t) = \sigma_t^2 \geq 0 \tag{3.1.2}$$

Autocovariance Function:

$$E[(Y_t - \mu_t)(Y_s - \mu_s)] = Cov(Y_t, Y_s) = \gamma_{t,s} \tag{3.1.3}$$

Autocorrelation Function:

$$Corr(Y_t, Y_s) = \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)}\sqrt{Var(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\sigma_t^2}\sqrt{\sigma_s^2}} = \rho_{t,s} \tag{3.1.4}$$

for $t, s = 0, \pm 1, \pm 2, \ldots$

Mean and variance of a series are given by (3.1.1) and (3.1.2), respectively. They are both expected to be constant at the first step of a time series analysis. The covariance and its derivative correlation within the series are obtained by (3.1.3) and (3.1.4). These functions are also expected to give results being independent of time. However, someone can ask why these expectations are needed. The answer lies in the common assumption of a time series analysis known as stationarity.

The most vital and common assumption in time series is stationarity. When the probabilistic properties of any time series process do not change with time, then the series is said to be stationary. There are two types of stationarity:

- Strict Stationarity
- (Covariance or Weak) Stationarity

The stationarity concept to be mentioned in this study corresponds to the latter. A time series is said to be (covariance) stationary if its first and second order moments are unaffected by a change of time origin. That is, for a stationary time series the following conditions are true;

$$E\left(Y_t\right) = \mu_t = \mu_{t-h} = E(Y_{t-h}) = \mu \,,$$

$$Var(Y_t) = \sigma_t^{\ 2} = \sigma_{t-h}^{\ \ 2} = Var(Y_{t-h}) = \sigma^2 \,,$$

$$Cov(Y_t, Y_s) = \gamma_{t,s} = \gamma_{t-h,s-h} = Cov(Y_{t-h}, Y_{s-h}) = \gamma_{|t-s|} \,,$$

$$Cov(Y_t, Y_t) = \gamma_{t,t} = \gamma_{|t-t|} = \gamma_0 = \sigma^2$$

for $t, s, h = 0, \pm 1, \pm 2 ...$

On the other hand, a time series is said to be strictly stationary if its statistical properties are unaffected by a change of time origin including the joint distribution of that time series.

The observations of a stationary time series generally follow a stable movement and are spread among a horizontal mean line. The unexpected shifts or outliers are not seen for such series. These properties make the model identification and model estimation procedures easy and hence the forecasts to be made for the future observations become much more close to real values and reliable as opposed to those of nonstationary series.

In addition to these, partial autocorrelation function is also used in a time series analysis procedure and denoted as;

$$Corr(Y_t, Y_{t-h} \mid Y_{t-1}, Y_{t-2}, ..., Y_{t-(h-1)}) = \phi_{hh} . \qquad (3.1.5)$$

By partial autocorrelation function, the effect of intervening values presented in autocorrelation function is eliminated and direct correlation between Y_t and Y_{t-h} is given.

Partial autocorrelation function given in (3.1.5) and the autocorrelation function given in (3.1.4) are used in model identification procedure of a time series analysis.

3.1.1. White Noise Process

As pointed out before, a series following a WN process only depends on an error term and it is denoted as;

$$Y_t = \varepsilon_t$$

where $\varepsilon_t \overset{i.i.d.}{\sim} N(\mu_\varepsilon, \sigma_\varepsilon^{\ 2})$. The white noise process is the simplest time series model. A white noise process is composed of sequential, identical and independent random variables ε_t with a constant mean and variance. For a white noise process,

$$E(Y_t) = E(\varepsilon_t) = \mu_\varepsilon ,$$

$$Var(Y_t) = Var(\varepsilon_t) = \sigma_\varepsilon^{\ 2} ,$$

$$Cov(Y_t, Y_{t-h}) = 0$$

and

$$Corr(Y_t, Y_{t-h}) = \begin{cases} 1, h = 0 \\ 0, h \neq 0 \end{cases}$$

are true.

3.1.2. ARMA (Box – Jenkins, Autoregressive and Moving Average) Models

Autoregressive moving average models were first provided by George E. P. Box and Gwilym Jenkins. That is why autoregressive moving average models are also known as Box-Jenkins Models in the literature. A Box-Jenkins Model involves the past observations of both the series and the error terms. It is a model reflecting the properties of both autoregressive and moving average models. In other words, an ARMA model is a combination of AR and MA models.

General representation of an ARMA(p, q) model is;

$$(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p)Y_t = c+(1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q)\varepsilon_t$$

where p is the order corresponding to the AR term, whereas q is the one that corresponds to MA term. Moreover, ε_t is a 0-mean white noise process and c is the constant term of the model. Here B is the backshift operator and defined as;

$$B^h Y_t = Y_{t-h}\,.$$

- When $p=0$; the model is a MA(q) model.
- When $q=0$; the model is an AR(p) model.

MA models are always stationary, whereas this is not always the case for AR and ARMA models. Hence, there are stationarity conditions to be checked for these models.

On the other hand, AR models are always invertible. Invertibility is a concept in time series that provides the uniqueness of the autocorrelation functions of processes. That is, the autocorrelation functions of distinct invertible models are different from each other. Moreover, invertibility is also important in forecasting procedures. In general, forecast errors become smaller for invertible models.

MA models show cut-off property in their autocorrelation functions and exponential/oscillating decay in their partial autocorrelation functions as opposed to AR models. In AR models, the case is vice versa.

When neither *p* nor *q* is equal to 0, the ARMA model shows both cut-off and exponential/oscillating decay properties in its both autocorrelation and partial autocorrelation functions, since an ARMA model is actually a combination of AR and MA models.

3.1.3. Model Identification

The first step in model identification is taken by the use of time series plots. A time series plot can give us an idea about the stationarity of the series. If the observations lie among a horizontal mean line and do not spread widely among this line, then the series most probably has a constant mean and a homoskedastic variance. Because the time series plots can not be a proof of stationarity, the latter step in the progress is to look at the correlogram of the series at zero level.

A correlogram shows the significance and the movement of the lags of the series through its autocorrelation and partial autocorrelation functions. Hence, a correlogram is very helpful in model identification in giving an idea about the order of an ARMA model. Moreover, one can understand if the time series in hand is non-stationary by looking at the correlogram of that series. If the lags of autocorrelation function show slow linear decay and there is a significant spike at the first or second lag of the partial autocorrelation function while the other lags are within the white noise bands, then the nonstationarity is the case for the series. In addition to this, one can also test the nonstationarity of a series through some statistical procedures known in literature as “unit root tests”. The most common unit root tests used today are Augmented Dickey-Fuller Test proposed by Dickey and Fuller (1981), PP Test by Phillips-Perron (1988) and KPSS Stationarity Test by Kwiatkowski-

Phillips-Schmidt-Shin (1992). The basic difference between these tests comes from their estimated test equations.

When nonstationarity is the case, differencing or detrending-procedures are applied to the series to provide stationarity. If the stationarity is satisfied, the model estimation can be conducted. If the correlogram shows the basic characteristics of a MA model, the order of MA term is decided based on the autocorrelation function. If the AR model is more appropriate via looking at the correlogram, the order of AR term is identified by the partial autocorrelation function. In addition, the order of MA and AR terms of an ARMA model is also determined this way through the use of correlogram.

In addition to the specification of ARMA model orders through the correlogram, there are also other order identification tools used in time series analyses such as the minimum information criterion (MINIC) and the extended sample autocorrelation function (ESACF).

3.1.4. Model Estimation

In model estimation, assumptions play a vital role so that the model to be estimated is statistically reliable. In a zero mean ARMA model as;

$$(1-\phi_1 B-\phi_2 B^2-...-\phi_p B^p)Y_t=(1-\theta_1 B-\theta_2 B^2-...-\theta_q B^q)\varepsilon_t$$

where ε_t is a zero mean white noise process, there are three main assumptions, such as;

$$Cov(\varepsilon_i,\varepsilon_j)=0 \text{ for } i\neq j \tag{3.1.6}$$

$$Cov(\varepsilon_i,\varepsilon_i)=Var(\varepsilon_i)=\sigma^2 \tag{3.1.7}$$

$$\varepsilon_i \sim N(0,\sigma^2) \tag{3.1.8}$$

These assumptions actually come from the properties of a white noise process that is considered for error terms. A white noise process is comprised of independent and identically distributed observations. Therefore, these observations should be uncorrelated and homoskedastic as well. This fact corresponds to (3.1.6) and (3.1.7). The normality assumption in (3.1.8) plays an important role in the step of model parameter estimation, since the significance of the parameter estimates is decided based on the t-statistics calculated.

After constructing each reasonable model, these assumptions should be checked cautiously. These assumption check procedures are also known as diagnostic checks and comprised of several tests.

3.1.5. Diagnostic Checks

Jarque-Bera (1981) Normality Test

The normality of the error terms can be checked with Jarque-Bera Test. The null hypothesis for this test is;

H_0: Error terms are normally distributed.

The test statistic for Jarque-Bera Test is given by;

$$JB = n[\frac{s^2}{6} + \frac{(k-3)^2}{24}] \dot{\sim} \chi^2_{(2)},$$

where s is the skewness coefficient and k is the kurtosis coefficient;

$$s = \frac{E(\varepsilon_i^3)}{[E(\varepsilon_i^2)]^{3/2}},$$

$$k = \frac{E(\varepsilon_i^4)}{[E(\varepsilon_i^2)]^2}.$$

The rejection of null hypothesis brings the result of nonnormal distribution for the error terms and thus the model built becomes invalid. Either a transformation on the series or another model trial for the series can be the solution for such cases.

Serial L-M Correlation Test (Breusch (1978) – Godfrey (1978) Test)

For an AR(p) model as;

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t \text{ for } t = 0,1,2,...,n , \qquad (3.1.9)$$

the uncorrelatedness assumption made for the error terms can be checked by the Serial L-M Correlation Test. For this test, an artificial regression model is defined and it is given by;

$$\varepsilon_t = \gamma_0 + \gamma_1 y_{t-1} + ... + \gamma_p y_{t-p} + \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + ... + \rho_r \varepsilon_{t-r} + u_t ,$$

where u_t is the error term of this artificial regression model with zero mean and constant variance. Here, *r* is the order of serial correlation that we want to test. The null hypothesis to be tested is defined as;

$$H_0 : \rho_1 = ... = \rho_r = 0 . \qquad (3.1.10)$$

The test statistic of this null hypothesis is calculated by;

$$F_{stat} = \frac{R_a^{\,2} / r}{(1 - R_a^{\,2}) / ((n-p) - r)} \sim F(r, n-p-r) , \qquad (3.1.11)$$

where $R_a^{\,2}$ is the coefficient of determination obtained from the artificial regression model.

When the probability of the F-statistic calculated in (3.1.11) is greater than the specified significance level, the null hypothesis in (3.1.10) is failed to be rejected. This result brings the decision that no lag of the error term ε_t has a correlation with the error term itself up to the order *r*. Hence, the uncorrelatedness of the error terms is satisfied. On the other hand, if any correlation exists among the error terms, transformation can handle it. Otherwise, the model should be changed.

White Heteroskedasticity Test (White (1980) Test, with no cross terms)

For the AR(p) model stated in (3.1.9), heteroskedastic error terms can be checked by White Heteroskedasticity Test.

This test also requires an artificial regression model;

$$\varepsilon_t^{\,2} = \gamma_0 + \gamma_1 y_{t-1} + ... + \gamma_p y_{t-p} + \gamma_{p+1} y_{t-1}^2 + ... + \gamma_{2p} y_{t-p}^2 + u_t \,,$$

with the same error term u_t assumed in Serial L-M Correlation test. The null hypothesis is defined as;

$$H_0 : Var(\varepsilon_t) = E(\varepsilon_t^{\,2} \mid \underset{\sim}{y}) = \sigma^2 \,.$$

So the null hypothesis can actually be converted to;

$$\mathrm{H}_0 : \gamma_1 = \gamma_2 = ... = \gamma_{2p} = 0 \,.$$

The test statistic for this null hypothesis is calculated as;

$$F_{stat} = \frac{SS\,\mathrm{Re}\,g \big/ 2p}{SS\,\mathrm{Re}\,s \big/ n-2p-1} = \frac{MS\,\mathrm{Re}\,g}{MS\,\mathrm{Re}\,s} \sim F(2p, n-2p-1)$$

where $SS\,\mathrm{Re}\,g$ is the sum of squares of regression and $SS\,\mathrm{Re}\,s$ is the sum of squares of residuals. In addition, $MS\,\mathrm{Re}\,g$ and $MS\,\mathrm{Re}\,s$ are the mean of these functions, respectively.

When the null hypothesis is not rejected at the significance level specified, the decision that the variance of the error terms is constant is made. This leads homoskedasticity assumption to come true.

- If any two assumptions are not satisfied, the models constructed are not reliable for statistical inference.
- If the first two assumptions are satisfied but the third one is not, then autoregressive conditional heteroskedasticity (ARCH) models or other heteroskedastic models can be constructed.

3.1.6. Autoregressive Conditional Heteroskedasticity (ARCH) Models

The bases of ARCH models were first taken by Engle in 1982. These models are widespread in modeling, when the error terms are found to be uncorrelated but have non-constant variance. The general idea of ARCH models is to model the error variance. That is, any built model having heteroskedastic errors is not altered at all, however a solution for explaining the variance of error terms with a model is tried to be obtained. For an assumption of AR(p) model with heteroskedastic errors;

$$Y_t = \phi_1 Y_{t-1} + ... + \phi_p Y_{t-p} + \varepsilon_t , \qquad (3.1.12)$$

the conditional mean of Y_t given that all the lags of it are known is given by;

$$E(Y_t \mid Y_{t-1},...,Y_{t-p}) = \phi_1 Y_{t-1} + ... + \phi_p Y_{t-p} .$$

That is, the fitted model of AR(p) process reveals a mean equation of;

$$\hat{\mu} = \hat{\phi}_1 Y_{t-1} + ... + \hat{\phi}_p Y_{t-p} .$$

Nevertheless, the existence of non-constant variance of error terms in (3.1.12) brings the false inferences about the predicted and forecasted values of Y_t. Therefore, Engle put forward the assumption that the error terms have the following form.

$$\varepsilon_t = \sigma_t a_t .$$

Here, a_t is assumed to be the zero-mean white noise errors with variance 1 and independent of the lags of ε_t. For time t, the conditional variance of ε_t assuming that all the lags up to order *k* is known is provided by;

$$\begin{aligned} Var(\varepsilon_t) &= E(\varepsilon_t^2 \mid \varepsilon_{t-1},...,\varepsilon_{t-k}) = \sigma_t^{\ 2}, \\ \sigma_t^{\ 2} &= \theta_0 + \theta_1 \varepsilon_{t-1}^2 + ... + \theta_k \varepsilon_{t-k}^2 , \end{aligned} \qquad (3.1.13)$$

where $\theta_i > 0$ for $i=0,1,\ldots$ to have a nonnegative variance. The error variance is dependent on the lags of ε_t and takes values according to time change, therefore in (3.1.13) is named as the variance equation or the autoregressive conditional heteroskedasticity model of order *k*. Since $Var(\varepsilon_t)$ gives the optimal forecast of ε_t^2, we can define the relationship below as well.

$$\hat{\varepsilon}_t^2 = \hat{\theta}_0 + \hat{\theta}_1 \hat{\varepsilon}_{t-1}^2 + ... + \hat{\theta}_k \hat{\varepsilon}_{t-k}^2 .$$

By this relationship, it is not wrong to say that the residuals of AR(p) model follow an AR(k) model. So, as to understand if there exists an ARCH effect in the residuals up to order *k*, the null hypothesis

$$H_0 : \theta_1 = \theta_2 = ... = \theta_k = 0$$

can be tested by ARCH-LM test. The rejection of H_0 leads us to infer that ARCH effect is observed up to the k^{th} lag of residuals. Thus, the next step becomes to fit an ARCH(k) model in order to explain the variance of error terms ε_t.

3.1.7. ARIMA/SARIMA ((Seasonal) Autoregressive Integrated Moving Average) Models

ARIMA/SARIMA models are non-stationary models. A regular difference of order *d* is needed to make the series stationary. That is why, they are also said to be integrated of order *d*.

An ARIMA(p,d,q) model is in general form as;

$$(1-\phi_1 B-\phi_2 B^2-...-\phi_p B^p)(1-B)^d Y_t=(1-\theta_1 B-\theta_2 B^2-...-\theta_q B^q)\varepsilon_t$$

and a regular difference of order d make the series a stationary ARMA(p,q) process.

A SARIMA(p,d,q) x (P,D,Q) model with seasonal period *s* is;

$$\frac{(1-\phi_1 B-...-\phi_p B^p)(1-B)^d(1-B^s)^D(1-\Phi_1 B^s-...-\Phi_P B^{Ps})}{(1-\theta_1 B-...-\theta_q B^q)(1-\Theta_1 B^s-...-\Theta_Q B^{Qs})}Y_t=\varepsilon_t$$

and a regular difference of order *d* makes the series a stationary seasonal ARMA(p,q) process and a seasonal difference of order D makes this process a stationary ARMA(p,q) process. To be able to use this model, the series must show a periodic behavior.

As seen these two models differ from each other in one way, namely, the seasonal component.

As a SARIMA/ARIMA process is converted into a stationary ARMA process, the model identification and estimation processes are the same for an ordinary ARMA process.

3.1.8. Model Selection

If these assumptions are satisfied, the best model of all reasonable models estimated is chosen through information criteria. The most commonly used criteria are;

- Akaike's Information Criterion (AIC)
- Schwarz Bayesian Information Criterion (SBIC)

In general, the model with the minimum information criterion is preferred; therefore it is better to decide on a model by looking at both of the results given by these information criteria.

3.2. Time Series Disaggregation Methods

The inexistence of sub-annual values creates the problem of information loss about the population when they are needed to be used in the process of statistical modeling, since annual figures may not reflect the intra-year behavior of a series. To illustrate, unemployment rate may show periodical rise and fall in different times of a year in Turkey, however the sole annual observations may fail to reflect this seasonal behavior. In addition to this, the existence of sub-annual figures rather than only annual values means more observations in a sample and more observations give a better representative model for a population. Besides the statistical perspective, monthly or quarterly observations could be helpful in making decisions for vital cases. For example, sub-annual unemployment rate observations in a country may be needed in order to take sudden precautions in a year as governmental policy. Instead of building a model only with the annual figures, the interpolation methods and the procedures using indicator variables can be applied in the disaggregation of sub-annual values. The values to be disaggregated can be either quarterly or monthly. This study focuses on the implementation of methods in obtaining the quarterly figures.

The methods to be applied strictly rely on the structure of the data in hand. This structure is usually named as accounting restrictions or consistency conditions. For illustration, the annual value of a variable being interested can be the sum of monthly values of that year or the mean of quarterly figures may reveal the annual observation. This structural behavior of the data in hand should be reflected to the disaggregated series so that the accounting restrictions are met. The general picture of the data being utilized in the literature separates the structures into three as; flow, stock and index.

The flow variables are the ones corresponding to the first illustration and the index variables fall into the latter group. On the other hand, the case of stock variables are seen when the annual value is directly equal to the observation of a specific month or quarter of that year. It is easy to decide on that the structure of the unemployment rate data is index while the averages of sub-annual values expose the aggregated annual values.

For a given sample size of *n* and the variable of annual observations y_t, the relationship between y_t and the quarterly observations x_{it} for $i = 1,2,3,4$ can be defined as;

$$y_t = Cx_{it} ,$$

where the dimensions of y_t, x_{it} and C are *n*x1, 4*n*x1 and *n*x4*n*, respectively. The selection of matrix C needs caution so as to satisfy the accounting restrictions. The structure of the aggregated data specifies the components of the matrix C as follows:

For the flow variable;

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

For a stock variable having only the observation of the second quarter;

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

And for the index variable;

$$C = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}. \quad (3.2.1)$$

This matrix is known as the aggregation matrix, since the high-frequency observations are aggregated to the low-frequency values with the help of C. The aggregation matrix can be written in simple form as;

$$C = I \otimes q$$

where I is a *n*x*n* identity matrix, q is a 1x4 vector determined by the structure of variable being interested and $\otimes$ is the notation of Kronecker product.

For the flow variable;

$$q = [1 \quad 1 \quad 1 \quad 1].$$

For the stock variable having only the observation of the second quarter;

$$q = [0 \quad 1 \quad 0 \quad 0].$$

For the index variable;

$$q = [1/4 \quad 1/4 \quad 1/4 \quad 1/4].$$

Just like the aggregation matrix, the inverse relation between the annual and sub-annual values can be constructed with a matrix of disaggregation, H, such that;

$$x_{it} = Hy_t \,. \tag{3.2.2}$$

It is the intention of the interpolation methods to seek this proper disaggregation matrix. The following part gets use of a commonly used interpolation method; Lisman/Sandee (1964) method.

3.2.1. The Method of Lisman/Sandee (1964)

For an index variable, the aggregation matrix was stated in (3.2.1). Not surprisingly, the inverse relationship between the annual and sub-annual figures can be sustained by the disaggregation matrix;

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 1 & 1 & 1 \end{bmatrix}' = 4C' \tag{3.2.3}$$

since the equation (3.2.2) reveals,

$$x_{it} = \begin{bmatrix} y_1 & y_1 & y_1 & y_1 & \cdots & y_n & y_n & y_n & y_n \end{bmatrix}' . \tag{3.2.4}$$

Thus, the accounting restriction for index variables is met. Nonetheless, it is straightforward to utilize this disaggregation matrix, because the trend component and the seasonal behavior in the series are not taken into consideration with it. At the end of each year, there is a jump of values from the last quarter of that year to the first quarter of the succeeding year.

Lisman/Sandee (1964) proposed to implement a weighted structure placed in the quarterly figures instead of using the matrix in (3.2.3). That is, the jumps between the last quarter of the preceding year and the first quarter of the succeeding year are being coped with the help of weighted means of same quarters for the successive three years. The method does not directly discard the components in (3.2.4) but employ them in generating the disaggregated data. When these components are assumed to be the elements of a 4*n*x1 vector τ as in (3.2.4) such that;

$$\tau_{it} = \begin{bmatrix} \tau_{11} \\ \tau_{21} \\ \tau_{31} \\ \tau_{41} \\ \vdots \\ \tau_{1n} \\ \tau_{2n} \\ \tau_{3n} \\ \tau_{4n} \end{bmatrix}$$

and

$$\tau_{it} = y_t \text{ for } i = 1,2,3,4 \text{ \& } t = 1,2,...,n$$

a quarter i of a year t is suggested to be derived from the weighted mean of $\tau_{i,t-1}$, τ_{it} and $\tau_{i,t+1}$. In matrix notation; a generalization for a given year, say *t*=2, can be presented as follows;

$$x_{i2} = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \end{bmatrix} \begin{bmatrix} \tau_{i1} \\ \tau_{i2} \\ \tau_{i3} \end{bmatrix}. \tag{3.2.5}$$

In order to maintain a reasonable symmetry and smoothness for the disaggregated data, the weight coefficients can be defined as;

$$w_{11} = w_{43} = a, w_{21} = w_{33} = b, w_{31} = w_{23} = c, w_{41} = w_{13} = d,$$
$$w_{12} = w_{42} = e, w_{22} = w_{32} = f$$

so that (3.2.5) becomes

$$x_{i2} = \begin{bmatrix} a & e & d \\ b & f & c \\ c & f & b \\ d & e & a \end{bmatrix} \begin{bmatrix} \tau_{i1} \\ \tau_{i2} \\ \tau_{i3} \end{bmatrix}. \tag{3.2.6}$$

The existence of six unknowns defined in the disaggregation matrix brings the need of six independent equations so that the system can be solved.

Basically, three issues determine these needed equations, namely; accounting restriction, trend component and seasonal behavior.

The mean of sub-annual values to be generated should exactly give the annual figure for a year according to the accounting restriction. For quarters of second year, the following four equations can be written;

$$x_{12} = a\tau_{11} + e\tau_{12} + d\tau_{13}, \tag{3.2.7}$$

$$x_{22} = b\tau_{21} + f\tau_{22} + c\tau_{23}, \tag{3.2.8}$$

$$x_{32} = c\tau_{31} + f\tau_{32} + b\tau_{33}, \tag{3.2.9}$$

$$x_{42} = d\tau_{41} + e\tau_{42} + a\tau_{43}. \tag{3.2.10}$$

In order not to violate the accounting restriction,

$$y_2 = \frac{\sum_{i=1}^{4} x_{i2}}{4} \Rightarrow \sum_{i=1}^{4} x_{i2} = 4y_2 \tag{3.2.11}$$

is needed to be satisfied.

The sum of (3.2.7), (3.2.8), (3.2.9) and (3.2.10) gives:

$$\sum_{i=1}^{4} x_{i2} = 4y_2 = (a+b+c+d)\tau_1 + 2(e+f)\tau_2 + (a+b+c+d)\tau_3,$$

and for $i = 1,2,3,4$ and $t = 1,2,3$,

$$\tau_t = \tau_{it} = y_t. \tag{3.2.12}$$

The only way to meet the accounting restriction in (3.2.11) is to assume

$$e + f = 2 \tag{3.2.13}$$

and

$$a + b + c + d = 0. \tag{3.2.14}$$

It is not surprising to obtain (3.2.13), since it comes from (3.2.12). Moreover, the quarters in the first and third years have no effect all alone in determining the figures of the second year. That is why (3.2.14) is being reached.

When the trend component in the annual series is considered, three more independent equations can be obtained. The subtraction of (3.2.7) from (3.2.8) gives;

$$x_{22} - x_{12} = (b-a)\tau_1 + (f-e)\tau_2 + (c-d)\tau_3 \, . \qquad (3.2.15)$$

In the same way, the difference of (3.2.8) and (3.2.9) reveals;

$$x_{32} - x_{22} = (c-b)\tau_1 + (b-c)\tau_3 \, . \qquad (3.2.16)$$

If a difference of δ is assumed to exist between successive years, the following equation becomes true;

$$y_t - y_{t-1} = \delta \Rightarrow \tau_t - \tau_{t-1} = \delta \, .$$

For a difference of δ between successive years, the successive quarterly values to be generated are assumed to have a difference of $\delta/4$ due to the smoothness satisfied. This can be proved easily as;

$$y_t - y_{t-1} = \frac{\sum_{i=1}^{4}(x_{it} - x_{i,t-1})}{4}$$

$$\Rightarrow (x_{4t} + x_{3t} + x_{2t} + x_{1t}) - (x_{4,t-1} - x_{3,t-1} - x_{2,t-1} - x_{1,t-1}) = 4(y_t - y_{t-1})$$

$$\Rightarrow (x_{4t} - x_{4,t-1}) + (x_{3t} - x_{3,t-1}) + (x_{2t} - x_{2,t-1}) + (x_{1t} - x_{1,t-1}) = 4(y_t - y_{t-1})$$

$$\Rightarrow k + k + k + k = 4k = 4(y_t - y_{t-1})$$

$$\Rightarrow k = (x_{it} - x_{i,t-1}) = \delta$$

$$\Rightarrow (x_{it} - x_{i,t-1}) = (x_{it} - x_{i-1,t}) + (x_{i-1,t} - x_{i-2,t}) + (x_{i-2,t} - x_{i-3,t}) + (x_{i-3,t} - x_{i,t-1})$$

$$\Rightarrow (x_{it} - x_{i-1,t}) = \frac{\delta}{4}$$

Then, (3.2.15) can be restated as;

$$\frac{\delta}{4} = x_{22} - x_{12} = \tau_2(b - a + f - e + c - d) + \delta(a - b + c - d)$$

and (3.2.16) turns into;

$$\frac{\delta}{4} = x_{32} - x_{22} = 2\delta(b-c) \, .$$

Therefore, the following three independent equations are maintained;

$$b - a + f - e + c - d = 0 \, , \qquad (3.2.17)$$

$$a - b + c - d = 0.25 \, , \qquad (3.2.18)$$

$$b - c = 0.125 \, . \qquad (3.2.19)$$

Now that there is only one equation left to be claimed, the seasonal behavior of the annual series plays a role in acquiring it. For an assumption of regular seasonal movement existing in the annual figures, the following equality can hold;

$$y_t - y_{t+1} = y_t - y_{t-1}. \qquad (3.2.20)$$

That is, $\{y_t\}$ are assumed to follow a cyclical movement. If this movement is considered as a sine curve and the quarters of any year, say y_2, are assumed to be lying on this curve symmetrically;

$$\begin{aligned} x_{12} &= r\sin(\pi/8) \\ x_{22} &= r\sin(3\pi/8) \\ x_{32} &= r\sin(5\pi/8) \\ x_{42} &= r\sin(7\pi/8) \end{aligned}$$

becomes true where r is the amplitude of the sine curve. Therefore, $x_{12} = x_{32}$ and $x_{22} = x_{42}$. By (3.2.12) and (3.2.20), the following equality holds;

$$\tau_t - \tau_{t+1} = \tau_t - \tau_{t-1}.$$

Considering the seasonal behavior, the subtraction of (3.2.7) from (3.2.8) turns into when the difference of $\tau_t - \tau_{t+1}$ is assumed to be ξ;

$$\begin{aligned} & x_{22} - x_{12} = (b-a)\tau_1 + (f-e)\tau_2 + (c-d)\tau_3 \\ & \Rightarrow x_{22} - x_{12} = (b-a)(\tau_2 - \xi) + (f-e)\tau_2 + (c-d)(\tau_2 - \xi) \\ & \Rightarrow x_{22} - x_{12} = (b-a+f-e+c-d)\tau_2 + (a-b-c+d)\xi \\ & \Rightarrow r\sin(3\pi/8) - r\sin(\pi/8) = (b-a+f-e+c-d)\tau_2 + (a-b-c+d)\xi \\ & \Rightarrow 0.541r = (a-b-c+d)\xi \end{aligned} \qquad (3.2.21)$$

Therefore, a relationship between r and ξ can give us a solution for the sixth equation needed. In order to satisfy the accounting restriction, the average of the quarterly figures should give the annual value. Under the assumption of a sine curve followed by the quarter values of years, the quarters of year y_t for $t = 1$ take values as follows.

$$x_{11} = r\sin(-\pi/8)$$
$$x_{21} = r\sin(-3\pi/8)$$
$$x_{31} = r\sin(-5\pi/8)$$
$$x_{41} = r\sin(-7\pi/8)$$

Therefore, the relationship below can be built:

$$y_2 - y_1 = \tau_2 - \tau_1 = \frac{\sum_{i=1}^{4} x_{i2} - \sum_{i=1}^{4} x_{i1}}{4} = \xi$$
$$\Rightarrow \frac{2[r\sin(\pi/8) + r\sin(3\pi/8)]}{4} - \frac{2[r\sin(-\pi/8) + r\sin(-3\pi/8)]}{4} = \xi$$
$$\Rightarrow 1.307r = \xi$$
$$\Rightarrow r = 0.765\xi$$

By using this relationship and (3.2.21), our sixth equation can be derived.

$$0.541r = (a-b-c+d)\xi$$
$$\Rightarrow 0.541(0.765)\xi = (a-b-c+d)\xi \qquad (3.2.22)$$
$$\Rightarrow a-b-c+d = 0.414$$

With these six independent equations, six unknowns can be found and the matrix in (3.2.6) can be constructed. Thus, the generation of quarterly figures becomes nothing but a matrix calculation.

3.2.2. The Method of Chow/Lin (1971)

After the interpolation methods and the approaches acquiring the disaggregated data by only mathematical calculations were criticized due to the lack of statistical criteria utilization, the ongoing studies started to be focusing on the use of indicator variables in the process of disaggregation. The most common method using the related series approach was put forward by Chow and Lin in 1971. The method they proposed was on behalf of constructing a relationship between the annual observations and the set of observed high frequency related series and after that using this relationship in generating the sub-annual figures. That is, the indicator variables are being regressed on the disaggregated data through the help of annual observations.

Once the related series are determined, the sub-annual figures can be explained by the observed *m* indicator variables such that;

$$X = Z\beta + a$$

where X is the 4*n*x1 vector of quarterly data to be disaggregated, Z is the 4*n*x*m* vector of indicators and β is their *m*x1 vector of coefficients. a is assumed to be the error terms distributed with zero mean and constant variance. By the use of aggregation matrix, the following statement becomes true;

$$Y = CX = CZ\beta + Ca\ .$$

The annual figures are known and the related series are observed. The estimation of β is the necessary part. Chow/Lin estimates the coefficient vector through the generalized least squares and finds $\hat{\beta}$ as the linear unbiased estimator with minimum variance such that;

$$\hat{\beta} = (Z^{'}C^{'}(CVC^{'})^{-1}CZ)^{-1}Z^{'}C^{'}(CVC^{'})^{-1}Y$$

where V is the variance-covariance matrix of a. The estimation of the β coefficient vector leads the estimated vector of disaggregated series to;

$$\hat{X} = Z\hat{\beta} + VC^{'}(CVC^{'})^{-1}(Y - CZ\hat{\beta})\ .$$

The error terms were said to have constant mean and variance. In addition to this, the method defines two distinct assumptions for the distribution of the error terms a. The first assumption is that the error terms are uncorrelated within themselves. In this case, the variance matrix of the error terms becomes;

$$V = \sigma_a^2 \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Nevertheless, this theoretical assumption is rarely met by the real data. Therefore, a second case is being put forward for the error terms. It argues that an AR(1) model can reflect the behavior of the series;

$$a_t = \phi_1 a_{t-1} + \varepsilon_t .$$

Then, the variance-covariance matrix of the error terms turns into;

$$V = \sigma_a^2 \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{4n-1} \\ \rho & 1 & \rho & \cdots & \rho^{4n-2} \\ \rho^2 & \rho & \ddots & \cdots & \rho^{4n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{4n-1} & \rho^{4n-2} & \cdots & \rho & 1 \end{bmatrix} . \qquad (3.2.23)$$

When the error terms neither are uncorrelated nor follow an AR(1) process, two variants of the Chow/Lin method can be utilized; Fernandez (1981) and Litterman (1983).

3.2.3. The Method of Fernandez (1981)

The method of Fernandez is based on the same steps conducted in the approach of Chow/Lin. It sets the relationship between the observed related series and the annual figures and finds the disaggregated series through this relationship. The method estimates the vector of coefficients β by generalized least squares, nevertheless the difference comes from the assumption that the error terms follow a random walk model instead of a white noise or an AR(1) model.

Fernandez criticized the method of Chow/Lin in two aspects. The variance-covariance matrix in (3.2.23) is dependent on an unknown parameter ρ and the estimation of ρ can be achievable through applying the method of ordinary least squares to the residuals only when adequate number of observations is available. In addition to this drawback, the AR(1) model proposed by Chow/Lin may fail to cope with the serial correlation existing in the residuals. Because of these two probable conditions, Fernandez claimed that the residuals may follow a model as;

$$a_t = a_{t-1} + \varepsilon_t \qquad (3.2.24)$$

where ε_t are the zero mean white noise error terms. By using this equation;

$$X_t - X_{t-1} = (Z_t - Z_{t-1})\beta + (a_t - a_{t-1}) \qquad (3.2.25)$$

becomes true. In order to sustain the equality (3.2.24), a matrix *D* is defined such that;

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 & 0 \\ 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix} \qquad (3.2.26)$$

so that $Da = \varepsilon$ is satisfied. Hence, the equation (3.2.25) can be rewritten as follows;

$$DX = DZ\beta + Da\,.$$

Since we have only the annual figures in hand, we need to utilize the relationship between the indicator variables and the annual figures and this relationship can be set up with the equation below;

$$Y = CX = CZ\beta + (CD^{-1})Da\,.$$

The only thing that is needed to be found is the estimate of coefficient matrix. The generalized least squares method gives an estimator for β as;

$$\hat{\beta} = (Z^{'}C^{'}(C(D^{'}D)^{-1}C^{'})^{-1}CZ)^{-1}Z^{'}C^{'}(C(D^{'}D)^{-1}C^{'})^{-1}Y$$

with the solution for the disaggregated data;

$$\hat{X} = Z\hat{\beta} + (D^{'}D)^{-1}C^{'}(C(D^{'}D)^{-1}C^{'})^{-1}(Y - CZ\hat{\beta})\,.$$

3.2.4. The Method of Litterman (1983)

When a white noise, an AR(1) or a random walk processes are not suitable for the sub-annual residuals and do not remove the serial correlation among the annual residuals, the method proposed by Litterman may be valuable in the process of disaggregation. Litterman found that the serial correlation existing in the annual residuals may be the result of that they follow an

ARIMA(1,1,0) model. The relationship between the sub-annual figures and the indicator variables is just the same as in the other two regression-based methods and can be written as;

$$X = Z\beta + a \, .$$

However, the error term a in this equation are thought to be fitted through an ARIMA(1,1,0) model. That is;

$$a_t = (\phi + 1)a_{t-1} - \phi a_{t-2} + \varepsilon_t \qquad (3.2.27)$$

can be true for this case. The equation (3.2.27) can be expressed in a following open form;

$$a_t = a_{t-1} + K_t \, ,$$

$$K_t = \phi K_{t-1} + \varepsilon_t \, .$$

As seen, the method of Litterman (1983) is actually an improved version of the method of Fernandez (1981) and the equation (3.2.27) brings the sole difference between them. The relationship between a_t and K_t can be built by the same matrix D in equation (3.2.26). In addition to this, the relationship between K_t and ε_t can be defined with the matrix below.

$$\Omega = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -\phi & 1 & 0 & 0 & \cdots & 0 \\ 0 & -\phi & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\phi & 1 & 0 \\ 0 & 0 & \cdots & 0 & -\phi & 1 \end{bmatrix} .$$

Hence, the relationship between the observed related series and the annual figures is;

$$Y = CX = CZ\beta + (CD^{-1}\Omega^{-1})\Omega D a \, .$$

Litterman estimates the β coefficient through the generalized least squares method as well.

$$\hat{\beta} = (Z^{'}C^{'}(C(D^{'}\Omega^{'}\Omega D)^{-1}C^{'})^{-1}CZ)^{-1}Z^{'}C^{'}(C(D^{'}\Omega^{'}\Omega D)^{-1}C^{'})^{-1}Y \, .$$

Therefore, the solution for the disaggregated series is obtained by;

$$\hat{X} = Z\hat{\beta} + (D^{'}\Omega^{'}\Omega D)^{-1}C^{'}(C(D^{'}\Omega^{'}\Omega D)^{-1}C^{'})^{-1}(Y - CZ\hat{\beta}) \, .$$

So as to generalize the regression based methods, Chow/Lin (1971), Fernandez (1981) and Litterman (1983) are all utilizing the same structure of relationship between the annual figures and the observed related series and maintain the disaggregated series through the help of this relationship. The only difference that they are being distinct is the assumptions made on the distribution of the annual residuals.

The appropriate specification of the error distribution gives the best disaggregated data. In addition, the choice of indicator variables becomes also vital in these methods, since it would not be surprising to be witness of disaggregated series being distant from actual values when any unrelated series is expected to be the explanatory variable of the annual figures.

CHAPTER 4

ANALYSES

This chapter is separated into two sections. The generation of quarterly figures is conducted in the first one. After obtaining the disaggregated series, modeling procedure and the forecasts of these models are presented in the second section.

4.1. Disaggregation of Quarterly Figures

This section focuses on one mathematical procedure; Lisman/Sandee and three regression-based methods; Chow/Lin, Fernandez and Litterman. While the application of the method of Lisman/Sandee is given in the first part of this section, disaggregation procedure with the regression-based methods is done in the second part.

4.1.1. Disaggregation of Unemployment Rate of Turkey with the Method of Lisman/Sandee (1964)

The system of equations being used in the method of Lisman/Sandee was stated before together with why and how these equations are derived. In this section of the study, the method is applied to the unemployment rate data of Turkey.

The derivation of equations actually includes the main part of the method. The only thing left for generating the disaggregated series is to solve these equations and to get into matrix calculations. (3.2.13), (3.2.14), (3.2.17), (3.2.18), (3.2.19) and (3.2.22) can be represented in matrix form;

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 0.25 \\ 0.125 \\ 0.414 \end{bmatrix}.$$

The coefficient matrix on the left-most side is full rank, since it comes from six independent equations for six unknowns. Therefore, the inverse of this matrix becomes helpful in generating these six unknowns and reveals the following equation system;

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} 0.25 & 0 & 0 & 0.5 & 0.5 & 0.25 \\ 0.25 & 0 & 0 & 0 & 0.5 & -0.25 \\ 0.25 & 0 & 0 & 0 & -0.5 & -0.25 \\ 0.25 & 0 & 0 & -0.5 & -0.5 & 0.25 \\ 0 & 0.5 & 0.5 & 0 & 0 & -0.5 \\ 0 & 0.5 & -0.5 & 0 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 0 \\ 0.25 \\ 0.125 \\ 0.414 \end{bmatrix}.$$

So, we get $a = 0.291$, $b = -0.041$, $c = -0.166$, $d = -0.084$, $e = 0.793$, $f = 1.207$ and the matrix in (3.2.6) becomes;

$$\begin{bmatrix} 0.291 & 0.793 & -0.084 \\ -0.041 & 1.207 & -0.166 \\ -0.166 & 1.207 & -0.041 \\ -0.084 & 0.793 & 0.291 \end{bmatrix}.$$

The utilization of this matrix gives the disaggregated series x_{it}. For illustration, the quarters of year 1989 can be obtained by;

$$\begin{bmatrix} x_{1,1989} \\ x_{2,1989} \\ x_{3,1989} \\ x_{4,1989} \end{bmatrix} = \begin{bmatrix} 0.291 & 0.793 & -0.084 \\ -0.041 & 1.207 & -0.166 \\ -0.166 & 1.207 & -0.041 \\ -0.084 & 0.793 & 0.291 \end{bmatrix} \begin{bmatrix} y_{1988} \\ y_{1989} \\ y_{1990} \end{bmatrix}.$$

In the same way the quarters of 1990 is got through;

$$\begin{bmatrix} x_{1,1990} \\ x_{2,1990} \\ x_{3,1990} \\ x_{4,1990} \end{bmatrix} = \begin{bmatrix} 0.291 & 0.793 & -0.084 \\ -0.041 & 1.207 & -0.166 \\ -0.166 & 1.207 & -0.041 \\ -0.084 & 0.793 & 0.291 \end{bmatrix} \begin{bmatrix} y_{1989} \\ y_{1990} \\ y_{1991} \end{bmatrix}.$$

All the annual figures and the disaggregated series for the unemployment rate of Turkey are presented in Appendix A and Appendix B.

As stated before, not all the unemployment rate data of Turkey are observed for quarters except from 2000 to 2008. However, the line graph of these years clearly exposes the seasonal structure existing in the series as below.

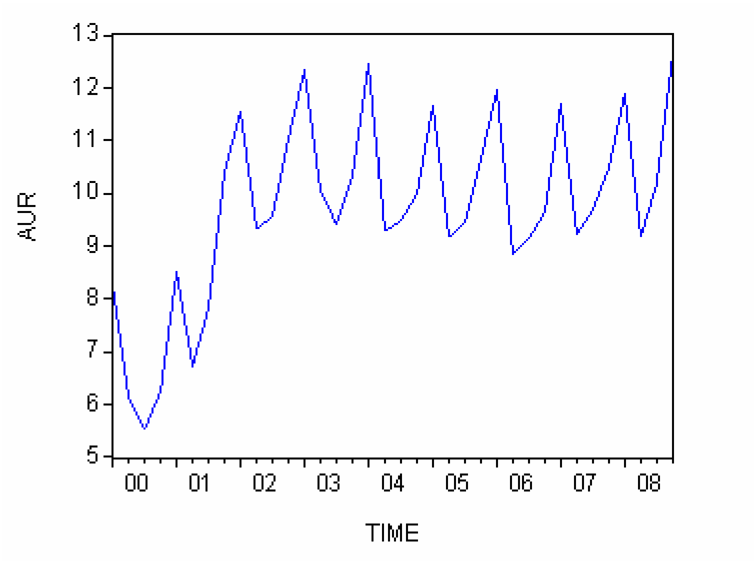


Figure 1. Actual Quarterly Values of the Unemployment Rate of Turkey between 2000 and 2008

The method of Lisman/Sandee considers the seasonal pattern in the derivation of equations; hence it is expected to see the seasonal structure in the generated disaggregated series as well. Nonetheless, the seasonality is not caught by the series and this is most probably due to the smooth structure achieved by the application of the method. The time series plot of the generated series from year 1989 to 2007 and the comparison of actual observations and the disaggregated data are the following.

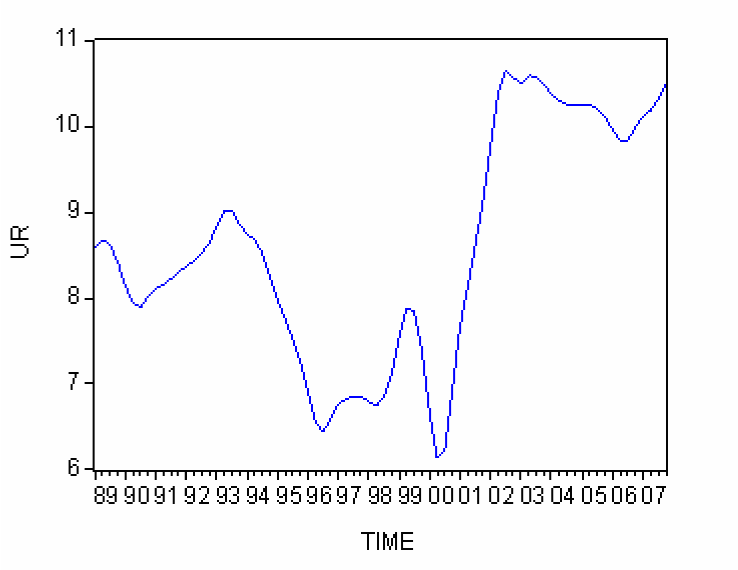


Figure 2. Disaggregated Series (Lisman/Sandee) between 1989 and 2007

Table 1. A Comparison of Disaggregated Series (Lisman/Sandee) and Actual Quarterly Values of Unemployment Rate of Turkey between 2000 and 2007

Year	Quarter	Disaggregated Series	Actual Values	Difference
2000	1	6.67	8.25	-1.58
2000	2	6.13	6.12	0.01
2000	3	6.22	5.53	0.69
2000	4	6.94	6.25	0.7
2001	1	7.66	8.49	-0.83
2001	2	8.12	6.73	1.39
2001	3	8.61	7.82	0.79
2001	4	9.11	10.4	-1.3
2002	1	9.76	11.55	-1.79
2002	2	10.39	9.32	1.07
2002	3	10.66	9.56	1.11
2002	4	10.57	11.05	-0.48
2003	1	10.5	12.32	-1.81
2003	2	10.59	10.03	0.56
2003	3	10.58	9.41	1.18
2003	4	10.49	10.33	0.16

Table 1 (continued)

Year	Quarter	Disaggregated Series	Actual Values	Difference
2004	1	10.38	12.45	-2.07
2004	2	10.3	9.28	1.03
2004	3	10.26	9.47	0.79
2004	4	10.25	9.99	0.26
2005	1	10.26	11.66	-1.4
2005	2	10.25	9.17	1.08
2005	3	10.2	9.44	0.76
2005	4	10.11	10.64	-0.54
2006	1	9.96	11.95	-1.99
2006	2	9.83	8.84	0.99
2006	3	9.84	9.14	0.7
2006	4	9.99	9.64	0.35
2007	1	10.12	11.68	-1.56
2007	2	10.19	9.23	0.96
2007	3	10.32	9.68	0.64
2007	4	10.51	10.48	0.04

4.1.2. Disaggregation of Unemployment Rate of Turkey with the Regression Based Methods

In this part, three regression based methods stated before are applied to the annual series. Since the application of the methods was not easily realizable without software, a time series disaggregation program, ECOTRIM, has been utilized in the process of generation and the higher frequency observations have been obtained.

In the specification of related series, the kinds of economic activities of gross national product in constant prices (1987 base) by production were examined. These activities were agriculture, industry, trade, government services, GDP (in purchasers' value) and the GNP (in purchasers' value). They were all obtained from the database of the Turkish Statistical Institute; however, unlike the unemployment rate data, the quarterly observations of these variables were available from years 1988 to 2006. Therefore, the regression based methods were applied to the annual data of the unemployment rate for these years. Many combinations of these variables were tried as related series through the program and all three regression based methods were applied to each of these combination groups. While arranging the groups of indicator variables, the following cases were considered:

- The variables GDP (in purchasers' value) and GNP (in purchasers' value) were slightly different from each other, therefore only one of them was chosen to be the candidate of any combination.
- While deciding on the group of variables as related series, the disaggregated data obtained for each combination were compared to the actual quarter values of unemployment rate data of Turkey, i.e. from the first quarter of 2000 to last quarter of 2006, since the quarters of the years until 2000 are not observed. The comparison of the disaggregated data and the actual values was based on the sum of squared distances. The combination group with the minimum sum of squared distances was chosen to be the related series.
- First, all the variables were assumed to be the indicators. Then, the variables having the same kind of movement were inspected and got into combination groups separately. For instance, GNP (in purchasers' value) and trade follow the same kind of movement and thus, they were generally not used together in the combination groups.
- In this manner, different numbered groups of variables were constructed and a general table was obtained.

In Table 2, the first combination includes all the variables and in this group gnp and trad seem to have same movement. Therefore, for the second and third combination group, they were tried separately and gnp was found to be more effective than trad. After inferring this result, the combination groups with three variables were built. While the variable ind does not have the same seasonal behavior as in gnp and trad, it was also tried in combination four like it is distinct from them. Nonetheless, the trad and gnp became much effective than ind and among all, the sixth group, the one with gnp, gave the best result. For the next step, the groups of two were constructed. Since trad and gnp were more likely to reveal better results compared to ind, agr was utilized with them, respectively. As expected, gnp exposed better results. Then, gs was combined with gnp and the minimum sum of squared distance for this case was less than that of eighth group. In order not to fail in choosing the best combination group, the effectiveness of gs-trad and gs-ind were also examined as well. Surprisingly, the combination of gs and trad brought about better estimates than gs and gnp. Although each of gnp, trad and ind gave small sum of squared distances on their own in the last three groups, none of them were decided to be the better than group ten. All in all, the desired results and the best disaggregated series were obtained as the variables trade and government services were used as related series.

Table 2. A Comparison of Regression-Based Methods for the Combination Groups of Economic Activities

	Sum of Squared Distances of Methods			**Decision**	
Combination	**Fernandez**	**AR(1) Min SSR**	**Litterman Min SSR**	**Minimum Sum of Squared Distance**	**Method**
1. gnp-gs-ind-agr-trad	200.68	229.89	202.91	200.68	Fernandez
2. gs-ind-agr-trad	354.13	760.28	331.73	331.73	Litterman
3. gnp-gs-ind-agr	200.95	160.32	213.84	160.32	AR(1)
4. gs-ind-agr	437.26	520.41	537.31	437.26	Fernandez
5. gs-agr-trad	383.96	620.05	467.06	383.96	Fernandez
6. gnp-gs-agr	353.80	420.58	462.83	353.80	Fernandez
7. trad-agr	276.45	621.72	489.22	276.45	Fernandez
8. gnp-agr	257.36	413.69	470.35	257.36	Fernandez
9. gnp-gs	19.09	20.31	20.36	19.09	Fernandez
10. gs-trad	15.17	26.83	13.16	13.16	Litterman
11. gs-ind	21.67	29.62	16.22	16.22	Litterman
12. trad	15.23	46.41	13.80	13.80	Litterman
13. gnp	17.67	44.76	22.55	17.67	Fernandez
14. ind	22.05	39.29	16.98	16.98	Litterman

Table 3 and Figure 3 show how the actual observations differ from the disaggregated data.

Table 3. A Comparison of Disaggregated Series (Litterman) and Actual Quarterly Values of Unemployment Rate of Turkey between 2000 and 2006

Year	Quarter	Disaggregated Series	Actual Values	Difference	Squared Distance
2000	1	8.04	8.25	-0.21	0.05
2000	2	6.49	6.12	0.37	0.14
2000	3	5.1	5.53	-0.43	0.18
2000	4	6.34	6.25	0.09	0.01
2001	1	8.35	8.49	-0.14	0.02
2001	2	8.14	6.73	1.41	1.98
2001	3	7.49	7.82	-0.33	0.11
2001	4	9.52	10.4	-0.89	0.79
2002	1	10.83	11.55	-0.71	0.51
2002	2	10.48	9.32	1.16	1.34
2002	3	9.39	9.56	-0.16	0.03
2002	4	10.67	11.05	-0.38	0.14
2003	1	11.74	12.32	-0.58	0.34
2003	2	10.85	10.03	0.82	0.68
2003	3	9.17	9.41	-0.24	0.06
2003	4	10.4	10.33	0.07	0.01
2004	1	11.26	12.45	-1.19	1.42
2004	2	10.33	9.28	1.06	1.12
2004	3	9.17	9.47	-0.3	0.09
2004	4	10.43	9.99	0.43	0.19
2005	1	11.55	11.66	-0.11	0.01
2005	2	10.34	9.17	1.17	1.36
2005	3	8.89	9.44	-0.55	0.3
2005	4	10.04	10.64	-0.61	0.37
2006	1	11.26	11.95	-0.68	0.47
2006	2	9.87	8.84	1.03	1.06
2006	3	8.56	9.14	-0.58	0.34
2006	4	9.93	9.64	0.29	0.09

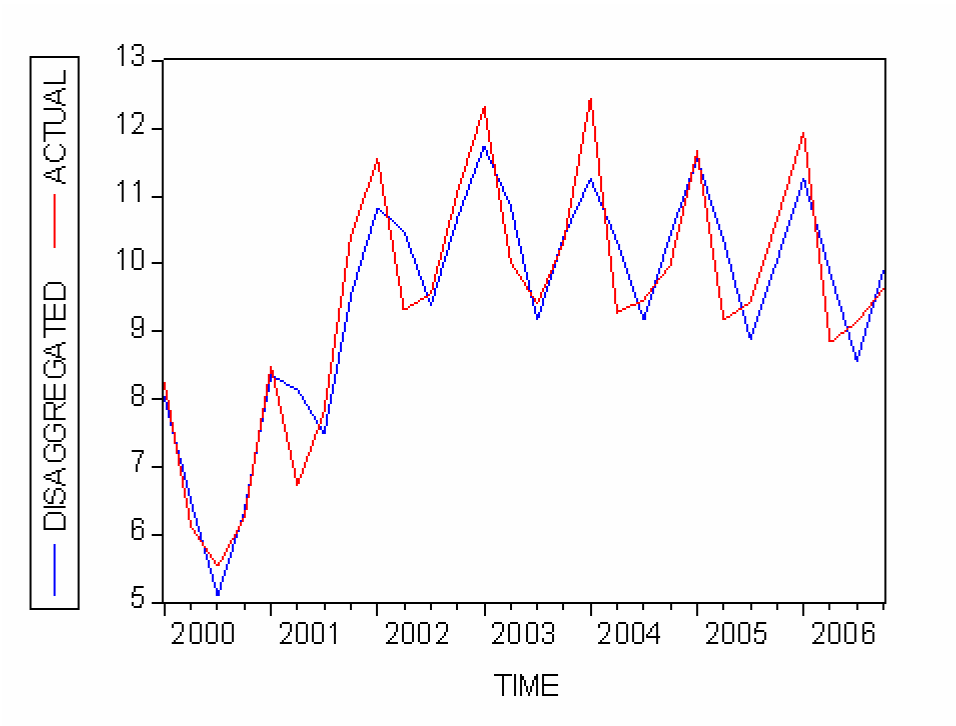


Figure 3. Disaggregated Series (Litterman) and Actual Quarter Values of Unemployment Rate of Turkey between 2000 and 2006

The details about the values disaggregated for each combination groups of variables are presented in Appendix C.

4.2. Modeling Processes

Now that, there are two different series generated by the methods of Lisman/Sandee and Litterman, these series can be modeled separately as in parts 4.2.1. and 4.2.2.

4.2.1. Modeling the Disaggregated Data Generated by Method of Lisman/Sandee

For the data generated by the method of Lisman/Sandee, we first examined if any transformation is needed to be applied to the series. According to the results of Box-Cox in Table 4, AIC and SBC give the minimum values when lambda is 1. Therefore, no transformation is needed for the series. Our next step is to get a general idea about the structure of the series through the time series plot.

Table 4. Box-Cox Results of Disaggregated Series (Lisman/Sandee)

LAMBDA	LOGLIK	RMSE	AIC	SBC
1.0	67.0291	0.009930	-122.058	-108.074
0.5	63.4034	0.010354	-114.807	-100.822
0.0	58.1302	0.011005	-104.260	-90.276
-0.5	51.4023	0.011900	-90.805	-76.82
-1.0	43.5045	0.013088	-75.009	-61.025

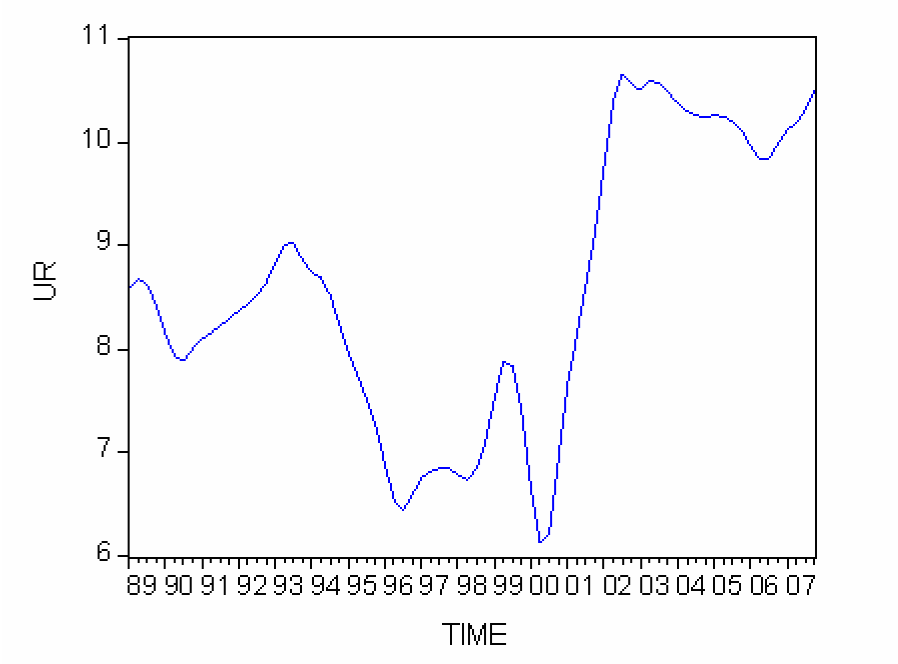
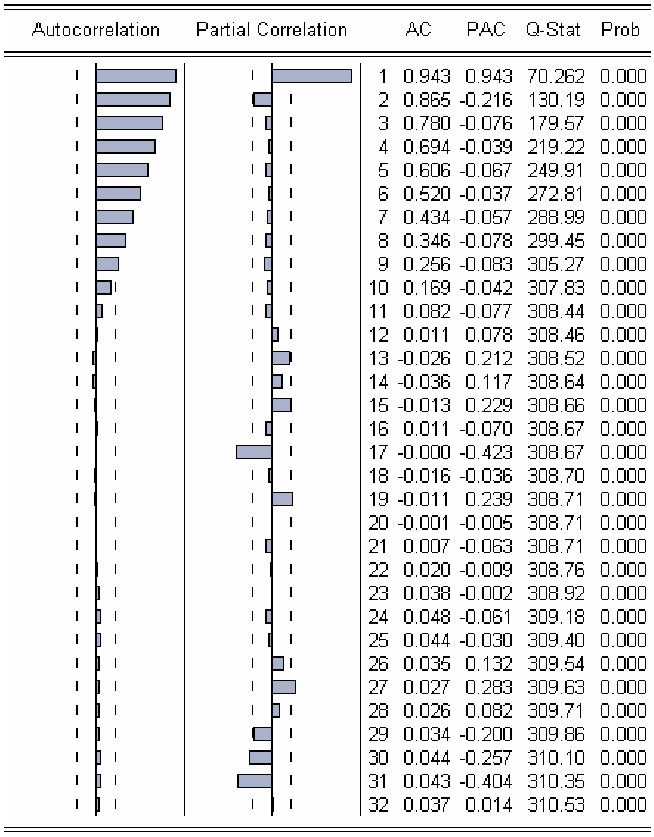


Figure 4. Time Series Plot of Disaggregated Series (Lisman/Sandee)

By the time series plot of unemployment rate data, the series has neither an exact trend nor a seasonal component. However, the observations do not spread among a horizontal mean line. Therefore, this visual inspection can give a clue about nonstationarity for the series.



Lag	AC	PAC	Q-Stat	Prob
1	0.943	0.943	70.262	0.000
2	0.865	-0.216	130.19	0.000
3	0.780	-0.076	179.57	0.000
4	0.694	-0.039	219.22	0.000
5	0.606	-0.067	249.91	0.000
6	0.520	-0.037	272.81	0.000
7	0.434	-0.057	288.99	0.000
8	0.346	-0.078	299.45	0.000
9	0.256	-0.083	305.27	0.000
10	0.169	-0.042	307.83	0.000
11	0.082	-0.077	308.44	0.000
12	0.011	0.078	308.46	0.000
13	-0.026	0.212	308.52	0.000
14	-0.036	0.117	308.64	0.000
15	-0.013	0.229	308.66	0.000
16	0.011	-0.070	308.67	0.000
17	-0.000	-0.423	308.67	0.000
18	-0.016	-0.036	308.70	0.000
19	-0.011	0.239	308.71	0.000
20	-0.001	-0.005	308.71	0.000
21	0.007	-0.063	308.71	0.000
22	0.020	-0.009	308.76	0.000
23	0.038	-0.002	308.92	0.000
24	0.048	-0.061	309.18	0.000
25	0.044	-0.030	309.40	0.000
26	0.035	0.132	309.54	0.000
27	0.027	0.283	309.63	0.000
28	0.026	0.082	309.71	0.000
29	0.034	-0.200	309.86	0.000
30	0.044	-0.257	310.10	0.000
31	0.043	-0.404	310.35	0.000
32	0.037	0.014	310.53	0.000

Figure 5. Correlogram of Disaggregated Series (Lisman/Sandee)

As seen in Figure 5, the autocorrelation function of the series shows a linear decay and only the first lag of partial autocorrelation function is out of significance boundaries with a significant spike. This kind of movements in the correlogram is widespread for nonstationary processes.

Table 5. Augmented Dickey-Fuller Test Results on Disaggregated Series (Lisman/Sandee)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.075438	0.7212
Test critical values:	1% level	-3.525618	
	5% level	-2.902953	
	10% level	-2.588902	

*MacKinnon (1996) one-sided *p*-values.

According to the unit root test results in Table 5, it is also for sure that the series is nonstationary. A probability of 0.7212 leads us fail to reject the existence of a unit root for the series.

A regular difference will most probably be a solution of making the series stationary. It will be better to look at the correlogram and the unit root test of the differenced series for the next step in order to understand whether a regular difference to be taken will cope with the nonstationarity.

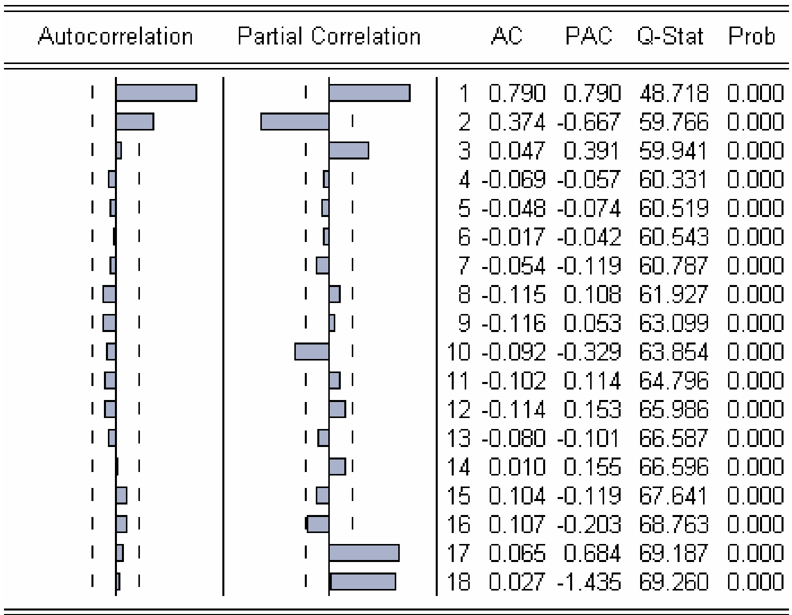


Figure 6. Correlogram of Differenced Disaggregated Series (Lisman/Sandee)

Table 6. Augmented Dickey-Fuller Test Results on Differenced Disaggregated Series (Lisman/Sandee)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.516956	0.0102
Test critical values:	1% level	-3.525618	
	5% level	-2.902953	
	10% level	-2.588902	

*MacKinnon (1996) one-sided *p*-values.

The autocorrelation function and partial autocorrelation function of the differenced series do not show any behavior of a standard nonstationary series. Moreover, the Augmented Dickey-Fuller Test supports the stationarity of the differenced series as well (Table 6, *p*-value=0.0102). Therefore, the correlogram of the new series constructed by a regular difference of the original series can be used in model identification. There is significance for the first two lag of the autocorrelation function of the new series, DIF, since they are both out of significance boundaries. Whereas the partial autocorrelation function shows cut off after lag 3. Both functions have a movement of oscillating decay. Therefore, an ARMA(3,2) model can be fitted to this new series.

Table 7. First Model for Differenced Disaggregated Series (Lisman/Sandee)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DIF(-1)	1.926344	0.113061	17.03802	0.0000
DIF(-2)	-1.347228	0.200037	-6.734885	0.0000
DIF(-3)	0.376727	0.112017	3.363113	0.0013
MA(2)	-0.813464	0.090360	-9.002511	0.0000

R-squared	0.873808	Mean dependent var	0.029028
Adjusted R-squared	0.868240	S.D. dependent var	0.269466
S.E. of regression	0.097813	Akaike info criterion	-1.757576
Sum squared resid	0.650575	Schwarz criterion	-1.631095
Log likelihood	67.27274	Durbin-Watson stat	2.013678

The table above shows the parameter estimates and their significance of the fitted ARMA(3,2) model for the series, DIF. The probability column at the end corresponds to the *p*-values of the *t*-statistics calculated for each parameter in the model. At a significance level of 0.05, all of the parameter estimates are statistically significant. However, the intercept and the MA(1) term have been dropped of the model, since their effects to the model were insignificant.

Once the model is constructed, the next thing that should be considered is to check the assumptions made at the beginning of the analysis. For error terms, it was accepted that they are normally distributed. In addition, a homoskedastic (constant) variance and uncorrelatedness within the error terms were also considered as true. These three assumptions are very vital in validation of the suggested model.

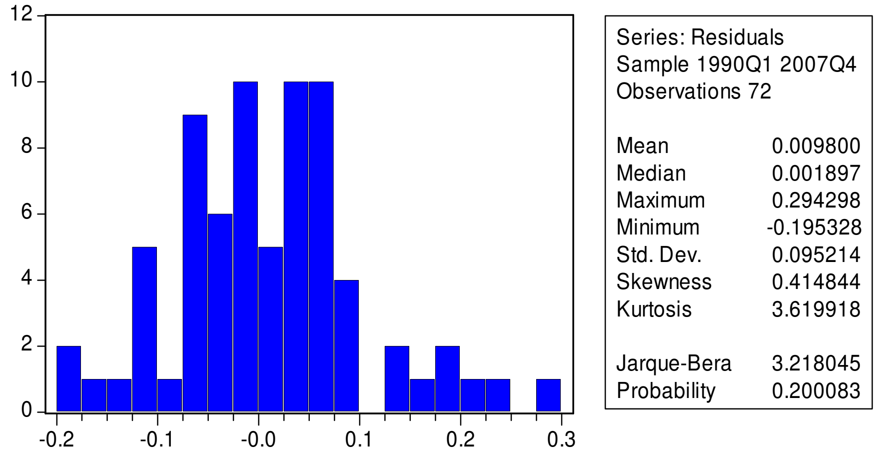


Figure 7. Jarque-Bera Normality Test Results for Residuals of First Model

The Jarque-Bera Test Statistic has a *p*-value of 0.200083, which is the proof of a normal distribution of error terms at a significance level 0.05. Also, we have conducted the Shapiro-Wilk Test and the normality of error terms is not violated (*p*-value=0.126). Hence, it can be concluded that there is no doubt about the normality assumption made on the error terms.

Table 8. Bresuch-Godfrey Serial Correlation LM Test Results for Residuals of First Model

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.973216	Probability	0.428463
Obs*R-squared	3.399201	Probability	0.493370

The *p*-value corresponding to *F*-statistic above is 0.428463. It is greater than 0.05. The existence of a serial correlation between the error terms is rejected. The uncorrelatedness assumption is satisfied as well.

Table 9. White Heteroskedasticity Test Results for Residuals of First Model

White Heteroskedasticity Test:

F-statistic	6.129263	Probability	0.000039
Obs*R-squared	26.01647	Probability	0.000221

However, the result of the White Heteroskedasticity Test is not as expected. The *p*-value of the *F*-statistic is strictly less than the alpha level 0.05 and the assumption that the error terms have a constant variance is rejected.

Table 10. ARCH Test Results for Residuals of First Model

ARCH Test:

F-statistic	4.075121	Probability	0.000512
Obs*R-squared	25.76599	Probability	0.002231

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CONSTANT	0.004565	0.002403	1.899874	0.0629
RESID^2(-1)	-0.063426	0.134922	-0.470095	0.6402
RESID^2(-2)	0.110939	0.131460	0.843896	0.4025
RESID^2(-3)	0.288637	0.130012	2.220080	0.0307
RESID^2(-4)	0.560409	0.135853	4.125103	0.0001
RESID^2(-5)	0.229370	0.152738	1.501726	0.1391
RESID^2(-6)	0.015741	0.135906	0.115820	0.9082
RESID^2(-7)	-0.178948	0.130038	-1.376114	0.1746
RESID^2(-8)	-0.228237	0.131249	-1.738963	0.0878
RESID^2(-9)	-0.182381	0.134542	-1.355573	0.1810

The variance of the series may not be constant, but can be modeled. That is, an ARCH model can explain the behavior of the variance for the series. In order to decide if there is an ARCH effect in the residuals, a test is conducted as in Table 10. The F-statistic gives a *p*-value of 0.000512. Thus, an ARCH effect is seemed to exist in the residuals. When the *p*-values of the squared residuals are considered, the effect is not observed after the fourth lag. Therefore an ARCH(4) model can be considered as the variance equation.

Table 11. Second Model for Differenced Disaggregated Series (Lisman/Sandee)

	Coefficient	Std. Error	z-Statistic	Prob.
DIF(-1)	1.635243	0.072188	22.65260	0.0000
DIF(-2)	-1.116942	0.072823	-15.33773	0.0000
DIF(-3)	0.269038	0.051189	5.255772	0.0000
MA(2)	0.013286	0.172565	0.076994	0.9386
Variance Equation				
CONSTANT	0.000507	0.000499	1.016882	0.3092
RESID(-1)^2	0.006232	0.035798	0.174080	0.8618
RESID(-2)^2	1.035630	0.518839	1.996053	0.0459
RESID(-3)^2	-0.012337	0.023709	-0.520337	0.6028
RESID(-4)^2	0.159248	0.183411	0.868257	0.3853

R-squared	0.843598	Mean dependent var	0.029028
Adjusted R-squared	0.823738	S.D. dependent var	0.269466
S.E. of regression	0.113131	Akaike info criterion	-2.436044
Sum squared resid	0.806319	Schwarz criterion	-2.151461
Log likelihood	96.69759	Durbin-Watson stat	1.683987

Although, the MA(2) term in the mean equation has a *p*-value of 0.9386 which is greater than 0.05 significance level, the model fitted representing mean and variance equation seems appropriate.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.277	0.277	5.7700	
		2	-0.066	-0.155	6.0999	0.014
		3	-0.110	-0.051	7.0374	0.030
		4	-0.097	-0.063	7.7677	0.051
		5	-0.095	-0.074	8.4905	0.075
		6	0.004	0.036	8.4920	0.131
		7	0.206	0.188	11.982	0.062
		8	-0.040	-0.195	12.112	0.097
		9	-0.001	0.117	12.112	0.146
		10	0.052	0.031	12.346	0.194
		11	-0.068	-0.109	12.748	0.238
		12	-0.237	-0.180	17.735	0.088
		13	-0.092	0.038	18.498	0.101
		14	0.048	-0.029	18.708	0.132
		15	-0.033	-0.039	18.807	0.172
		16	-0.019	-0.060	18.840	0.221
		17	0.035	0.011	18.959	0.271
		18	0.023	0.033	19.009	0.328
		19	-0.069	-0.027	19.488	0.362
		20	-0.035	-0.063	19.617	0.418
		21	-0.054	-0.049	19.923	0.463
		22	-0.046	0.033	20.146	0.512
		23	0.008	-0.026	20.154	0.573
		24	0.079	-0.002	20.845	0.591
		25	0.029	-0.027	20.938	0.642
		26	-0.041	-0.008	21.135	0.685
		27	-0.059	-0.079	21.547	0.713
		28	-0.086	-0.069	22.449	0.714
		29	-0.033	0.002	22.580	0.754
		30	-0.040	-0.058	22.782	0.786
		31	0.010	-0.025	22.794	0.824
		32	-0.001	-0.043	22.794	0.856

Figure 8. Correlogram of Residuals for Second Model

In Figure 8, the lags of the correlogram for the standardized residuals are all within the significance boundaries. Therefore, the error terms represented by these residuals follow a white noise process. Most probably the residuals have no more ARCH effect, nevertheless, checking this through ARCH test will make the things more clear.

According to the ARCH test in Table 12, there is no remaining ARCH effect in the residuals (*p*-value=0.914878). The *p*-values corresponding to the lags of squared residuals also support this result.

Table 12. ARCH Test Results for Residuals of Second Model

ARCH Test:

F-statistic	0.426682	Probability	0.914878
Obs*R-squared	4.256296	Probability	0.893748

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CONSTANT	1.218361	0.429888	2.834137	0.0065
STD_RESID^2(-1)	0.049999	0.135377	0.369334	0.7134
STD_RESID^2(-2)	-0.048306	0.136410	-0.354123	0.7247
STD_RESID^2(-3)	-0.010491	0.136124	-0.077070	0.9389
STD_RESID^2(-4)	0.129538	0.135621	0.955148	0.3438
STD_RESID^2(-5)	-0.178764	0.129045	-1.385283	0.1718
STD_RESID^2(-6)	-0.016554	0.131406	-0.125979	0.9002
STD_RESID^2(-7)	-0.107418	0.131680	-0.815749	0.4183
STD_RESID^2(-8)	0.050883	0.139123	0.365745	0.7160
STD_RESID^2(-9)	-0.027459	0.138000	-0.198976	0.8430

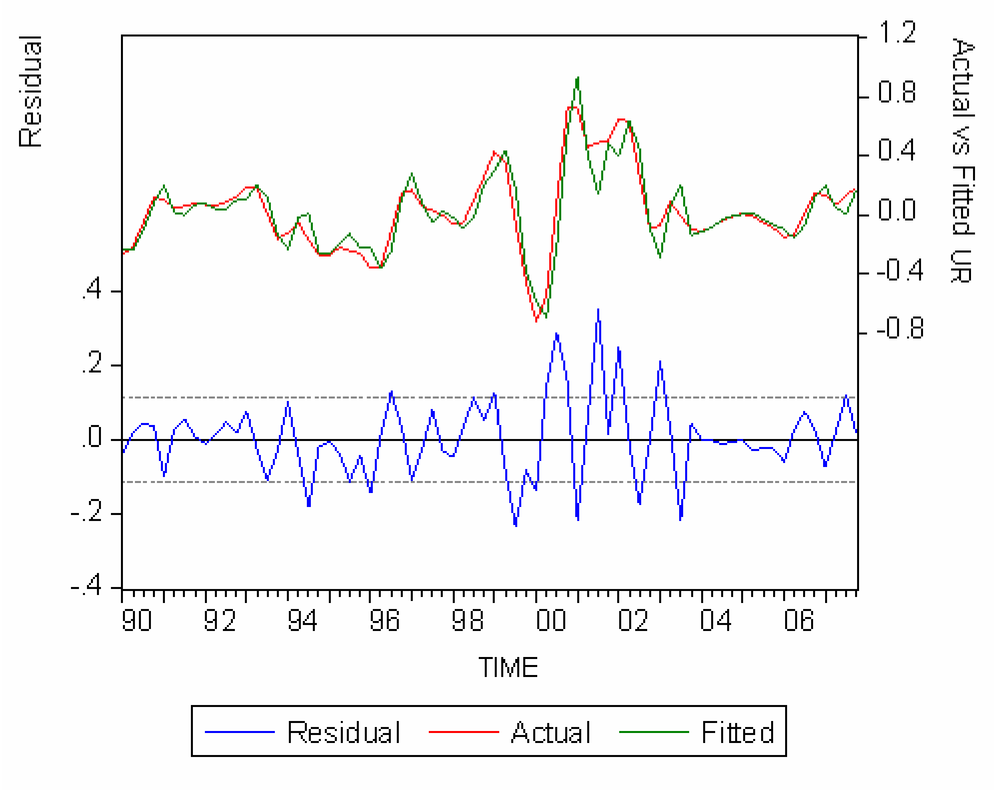


Figure 9. Actual-Fitted-Residual Graph for the Second Model

According to Figure 9, the fitted values are close to the actual values and the second model seems to have been fitted well. The mean equation for the series is as follows;

$$DIF_t = 1.635DIF_{t-1} - 1.117DIF_{t-2} + 0.269DIF_{t-3} + \varepsilon_t + 0.013\varepsilon_{t-2}.$$

While the variance equation is;

$$\sigma_t^2 = 0.001 + 0.006\varepsilon_{t-1}^2 + 1.036\varepsilon_{t-2}^2 - 0.012\varepsilon_{t-3}^2 + 0.159\varepsilon_{t-4}^2.$$

Model 3 for the Differenced Disaggregated Series (Lisman/Sandee)

As stated, the MA(2) term in the second model built before seemed insignificant. Omitting that term would give a better model as well. Therefore, it is advantageous to attempt in fitting an ARIMA(3,1,0) model, too.

Table 13. Third Model for Differenced Disaggregated Series (Lisman/Sandee)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DIF(-1)	1.700811	0.104227	16.31833	0.0000
DIF(-2)	-1.368697	0.159507	-8.580768	0.0000
DIF(-3)	0.499471	0.103869	4.808656	0.0000
R-squared	0.858604	Mean dependent var		0.029028
Adjusted R-squared	0.854505	S.D. dependent var		0.269466
S.E. of regression	0.102784	Akaike info criterion		-1.671593
Sum squared resid	0.728960	Schwarz criterion		-1.576731
Log likelihood	63.17733	Durbin-Watson stat		1.749639

All the coefficients of the ARIMA(3,1,0) model is statistically significant, so the important thing is to examine the diagnostic checks for this model.

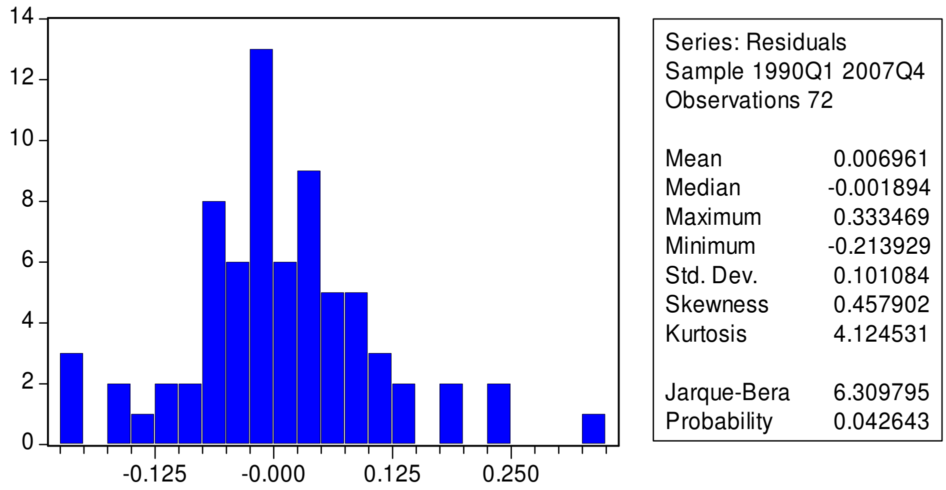


Figure 10. Jarque-Bera Normality Test Results for Residuals of Third Model

The *p*-value of the Jarque-Bera Test Statistic is 0.042643. Even if it is slightly less than the significance level of 0.05, the normality assumption of the error terms is not satisfied. Even if the homoskedasticity and the uncorrelatedness assumptions made will be satisfied by White Heteroskedasticity and Serial Correlation LM tests, respectively, the existence of nonnormal error terms makes the results of the *t*-statistics calculated and the model constructed invalid. Therefore, there is no need to look at the other diagnostic checks. This model is totally invalid and statistically unreliable.

Model 4 for the Disaggregated Series (Lisman/Sandee)

In the second model constructed, the nonstationarity seemed to be emerging from heteroskedastic variance of the series. If the same steps made in that model are also applied to the original series, a new valid model can be obtained too. Without taking a regular difference for the series of unemployment rate of Turkey, an ARMA(3,2) model can be fitted.

However when an ARMA(3,2) model with an intercept is fitted for the series, it is concluded that the intercept and MA(2) term are statistically insignificant in explaining the unemployment rate. Therefore, the final model constructed has become an ARMA(3,1) process without an intercept as below.

Table 14. Fourth Model for Disaggregated Series (Lisman/Sandee)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
UR(-1)	2.049548	0.112246	18.25947	0.0000
UR(-2)	-1.500723	0.209753	-7.154726	0.0000
UR(-3)	0.452259	0.112484	4.020666	0.0001
MA(1)	0.876660	0.060143	14.57623	0.0000
R-squared	0.995264	Mean dependent var		8.565068
Adjusted R-squared	0.995058	S.D. dependent var		1.387469
S.E. of regression	0.097539	Akaike info criterion		-1.763886
Sum squared resid	0.656461	Schwarz criterion		-1.638381
Log likelihood	68.38183	Durbin-Watson stat		2.008468

All the terms in the model seem significant. Now that the model is appropriate, the diagnostic checks become important for the next step.

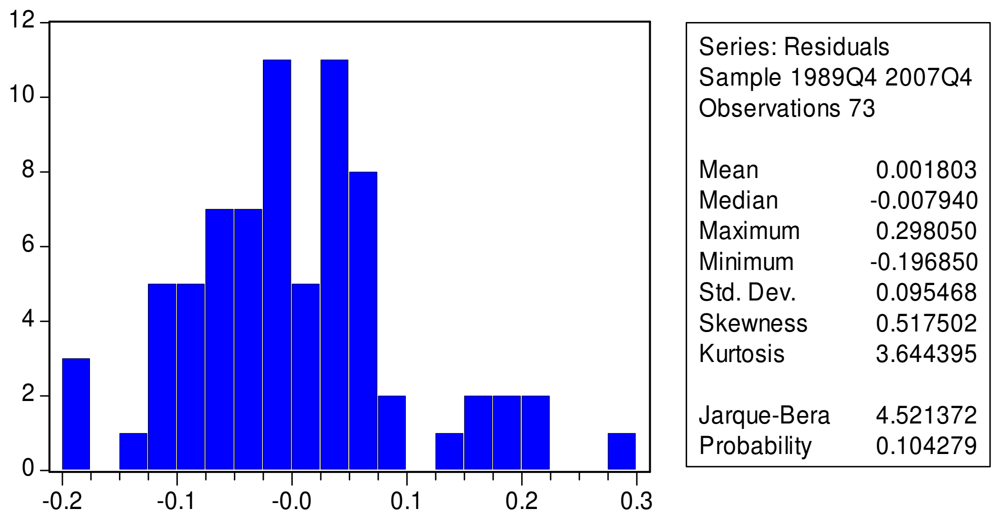


Figure 11. Jarque-Bera Normality Test Results for Residuals of Fourth Model

By looking at the Jarque-Bera Test Statistic and its *p*-value, it can be said that the distribution of the error terms is normal. A Shapiro-Wilk Test also reveals a normal distribution (*p*-value=0.082). Thus, the normality assumption made on the error terms in the model is not violated.

Table 15. Bresuch-Godfrey Serial Correlation LM Test Results for Residuals of Fourth Model

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.897959	Probability	0.470430
Obs*R-squared	3.797642	Probability	0.434084

Breusch-Godfrey Serial Correlation LM Test gives a *p*-value of 0.470430. There is not any serial correlation between the residuals up to the order four. Hence, the uncorrelatedness assumption of the error terms is also satisfied.

Table 16. White Heteroskedasticity Test Results for Residuals of Fourth Model

White Heteroskedasticity Test:

F-statistic	1.844998	Probability	0.103785
Obs*R-squared	10.48539	Probability	0.105644

In a surprising manner, the table above gives a homoskedastic variance for the error terms. However, it was expected that the homoskedasticity of the error terms was not going to be satisfied and a new ARCH model was going to be constructed in explaining the variance of the series. In order to be sure about if the series have homoskedastic variance, an ARCH test should be conducted for the residuals.

Table 17. ARCH Test Results for Residuals of Fourth Model

ARCH Test:

F-statistic	4.525266	Probability	0.000184
Obs*R-squared	27.51636	Probability	0.001148

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CONSTANT	0.004375	0.002297	1.904567	0.0622
RESID^2(-1)	-0.091993	0.132736	-0.693051	0.4912
RESID^2(-2)	0.105289	0.129656	0.812067	0.4203
RESID^2(-3)	0.311328	0.127269	2.446226	0.0177
RESID^2(-4)	0.572769	0.133633	4.286129	0.0001
RESID^2(-5)	0.243736	0.150975	1.614414	0.1123
RESID^2(-6)	0.072734	0.133705	0.543989	0.5887
RESID^2(-7)	-0.210398	0.127141	-1.654838	0.1038
RESID^2(-8)	-0.222696	0.129389	-1.721144	0.0910
RESID^2(-9)	-0.222164	0.132297	-1.679280	0.0989

The ARCH test above gives that there is an ARCH effect in residuals up to order 4. Even though, White Heteroskedasticity Test did not say so, an ARCH(4) model is needed to be built as the variance equation of the series.

Table 18. Fifth Model for Disaggregated Series (Lisman/Sandee)

	Coefficient	Std. Error	z-Statistic	Prob.
UR(-1)	2.355516	0.099959	23.56482	0.0000
UR(-2)	-2.054731	0.198538	-10.34929	0.0000
UR(-3)	0.700076	0.102568	6.825459	0.0000
MA(1)	0.324845	0.107363	3.025683	0.0025
Variance Equation				
CONSTANT	0.001849	0.001066	1.734654	0.0828
RESID(-1)^2	-0.016802	0.057448	-0.292481	0.7699
RESID(-2)^2	0.693774	0.509771	1.360952	0.1735
RESID(-3)^2	-0.019031	0.030630	-0.621324	0.5344
RESID(-4)^2	0.231886	0.313261	0.740233	0.4592

Table 18 (continued)

R-squared	0.994225	Mean dependent var	8.565068
Adjusted R-squared	0.993503	S.D. dependent var	1.387469
S.E. of regression	0.111833	Akaike info criterion	-2.305869
Sum squared resid	0.800430	Schwarz criterion	-2.023483
Log likelihood	93.16420	Durbin-Watson stat	1.669578

The ARCH(4) model built as the variance equation is appropriate. The residuals of this equation should not include an ARCH effect, if the appropriateness of this model is the case. Therefore it would be better to look at the results of an ARCH test in order to see if there is still an ARCH effect left in the residuals.

Table 19. ARCH Test Results for Residuals of Fifth Model

ARCH Test:

F-statistic	0.732389	Probability	0.677443
Obs*R-squared	6.962295	Probability	0.641045

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CONSTANT	0.679737	0.292809	2.321437	0.0241
STD_RESID^2(-1)	0.119645	0.133186	0.898332	0.3730
STD_RESID^2(-2)	0.095126	0.134310	0.708259	0.4818
STD_RESID^2(-3)	0.044260	0.135066	0.327690	0.7444
STD_RESID^2(-4)	0.160035	0.118147	1.354539	0.1812
STD_RESID^2(-5)	-0.004440	0.119547	-0.037142	0.9705
STD_RESID^2(-6)	0.018420	0.118630	0.155275	0.8772
STD_RESID^2(-7)	-0.053842	0.118691	-0.453634	0.6519
STD_RESID^2(-8)	0.068666	0.117927	0.582273	0.5628
STD_RESID^2(-9)	-0.174235	0.117719	-1.480086	0.1447

There is no ARCH effect in the remaining residuals just as expected.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.264	0.264	5.2886	
		2	-0.090	-0.172	5.9175	0.015
		3	-0.165	-0.101	8.0355	0.018
		4	-0.003	0.066	8.0361	0.045
		5	0.129	0.090	9.3849	0.052
		6	0.049	-0.031	9.5828	0.088
		7	0.108	0.149	10.545	0.103
		8	-0.126	-0.189	11.878	0.105
		9	0.056	0.195	12.150	0.145
		10	0.144	0.073	13.943	0.124
		11	-0.007	-0.126	13.947	0.175
		12	-0.183	-0.140	16.944	0.110
		13	-0.152	0.010	19.066	0.087
		14	0.039	-0.022	19.204	0.117
		15	0.007	-0.043	19.208	0.157
		16	0.004	-0.038	19.210	0.204
		17	0.020	0.079	19.251	0.256
		18	0.039	0.077	19.403	0.306
		19	-0.051	-0.100	19.669	0.352
		20	0.026	0.080	19.738	0.411
		21	-0.029	-0.058	19.824	0.469
		22	-0.057	0.022	20.176	0.510
		23	-0.027	-0.032	20.255	0.567
		24	0.041	-0.002	20.446	0.615
		25	0.050	-0.009	20.735	0.654
		26	-0.027	0.032	20.818	0.703
		27	-0.052	-0.138	21.145	0.734
		28	-0.075	-0.003	21.833	0.746
		29	0.022	0.072	21.893	0.786
		30	-0.036	-0.106	22.061	0.818
		31	-0.025	0.003	22.140	0.849
		32	-0.033	-0.003	22.282	0.874

Figure 12. Correlogram of Residuals for Fifth Model

In addition, all the lags of the standardized residuals of the variance equation are inside the significance boundaries as given in Figure 12 above. They are following a white noise process as assumed.

The actual values and fitted values in Figure 13 are much closer for this model rather than the results of second model fitted. However, discrete results like information criteria imply better idea about which model is more meaningful, hence they should be examined cautiously.

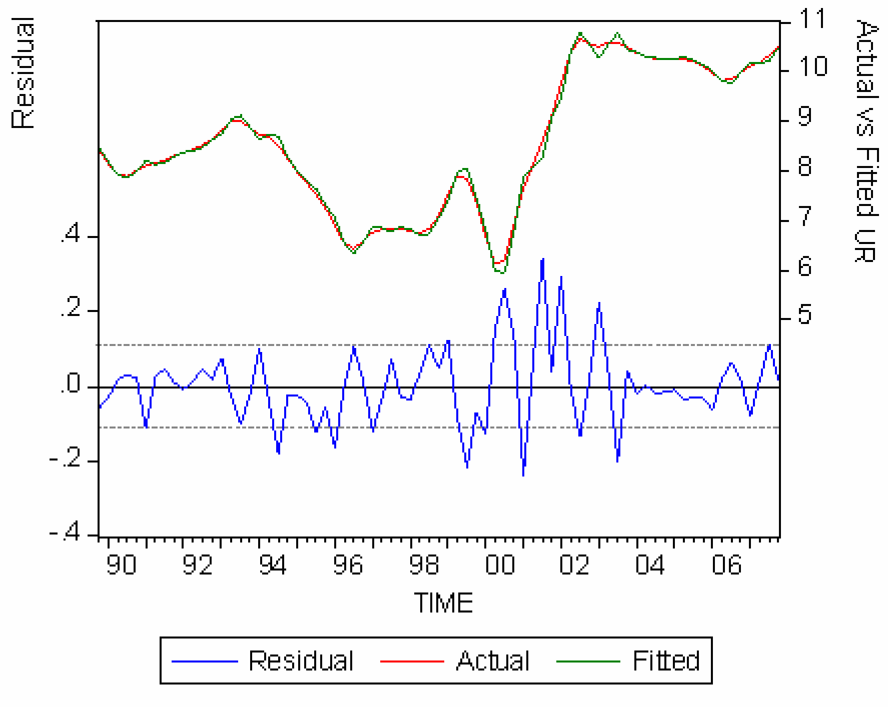


Figure 13. Actual-Fitted-Residual Graph for the Fifth Model

The mean equation for the series is as follows;

$$UR_t = 2.356UR_{t-1} - 2.055UR_{t-2} + 0.7UR_{t-3} + \varepsilon_t + 0.325\varepsilon_{t-1}\,.$$

While the variance equation is;

$$\sigma_t^2 = 0.002 - 0.017\varepsilon_{t-1}^2 + 0.694\varepsilon_{t-2}^2 - 0.019\varepsilon_{t-3}^2 + 0.232\varepsilon_{t-4}^2\,.$$

Model Selection and Forecasts

For the series generated by the method of Lisman/Sandee, three main models, second, third and fifth models, have been fitted. However, the diagnostics of the third model have led us not to use the fitted equation for statistical inferences. Therefore, it should be reasonable to use either the second or the fifth model. Both of the models include same number of parameters. Considering the parsimony, the model with smaller information criteria would give better forecasts. When the second model and the fifth model are compared, the former is better; since both Akaike and Schwarz Information Criteria corresponding to this model are smaller. Therefore, it is much reasonable to rely on the forecasts of the second model. The forecasts of quarters for the next three years and the available actual values corresponding to these quarters are given below.

Table 20. A Comparison of Forecasts of Second Model (Lisman/Sandee) and Actual Values of Unemployment Rate of Turkey for the Quarters of Years from 2008 to 2010

Year	Quarter	Forecasts	Actual Values	Distance
2008	1	10.7	11.88	1.18
2008	2	10.82	9.17	1.66
2008	3	10.87	10.18	0.69
2008	4	10.87	12.64	1.77
2009	1	10.83	16.12	5.29
2009	2	10.79	13.61	2.82
2009	3	10.77	13.43	2.66
2009	4	10.77	NA	NA
2010	1	10.78	NA	NA
2010	2	10.79	NA	NA
2010	3	10.8	NA	NA
2010	4	10.8	NA	NA

4.2.2. Modeling the Disaggregated Data Generated by Method of Litterman

As in part 4.2.1., the results of Box-Cox in Table 21 give the minimum AIC and SBC values when lambda is 1. Therefore, we have applied no transformation for our series generated by the method of Litterman.

Table 21. Box-Cox Results of Disaggregated Series (Litterman)

LAMBDA	LOGLIK	RMSE	AIC	SBC
1.0	-39.5682	0.16390	91.136	105.121
0.5	-45.0577	0.17960	102.115	116.1
0.0	-52.7680	0.20294	117.536	131.52
-0.5	-62.2999	0.23667	136.600	150.584
-1.0	-73.2018	0.28708	158.404	172.388

According to Figure 14, the unemployment rate of Turkey; UR, generated by the method of related series, has a regular seasonal movement. The trend component is not so obvious, thus the correlogram of this series would be helpful in deciding surely in the existence of it. The series seems stationary at first sight. The observations follow a stable movement among a horizontal mean line without any extreme candidate.

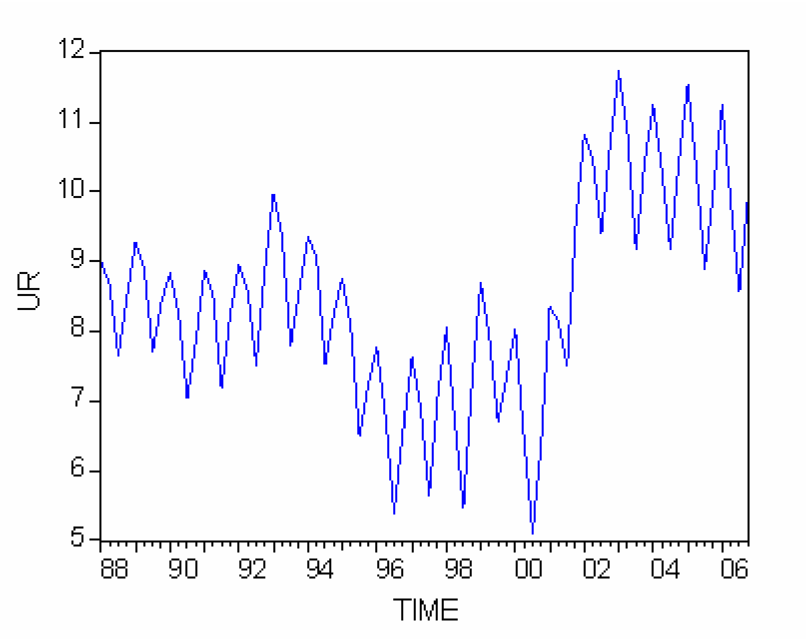


Figure 14. Time Series Plot of Disaggregated Series (Litterman)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.718	0.718	40.738	0.000
		2	0.435	-0.165	55.919	0.000
		3	0.594	0.739	84.566	0.000
		4	0.703	-0.260	125.30	0.000
		5	0.406	-0.305	139.07	0.000
		6	0.139	-0.232	140.71	0.000
		7	0.274	0.234	147.15	0.000
		8	0.352	-0.205	157.93	0.000
		9	0.090	0.175	158.66	0.000
		10	-0.142	-0.500	160.47	0.000
		11	-0.000	0.826	160.47	0.000
		12	0.148	-2.779	162.49	0.000

Figure 15. Correlogram of Disaggregated Series (Litterman)

Due to the existence of seasonal component in the series, the autocorrelation and partial autocorrelation functions of the correlogram have lags reflecting this periodical movement. It is generally easier in autocorrelation function to specify the period of the seasonality by looking at the lags compared to those of partial autocorrelation function. The n*4^{th} lags of autocorrelation function seem to be affected by seasonality. Therefore, a 4^{th} degree seasonal difference is expected to eliminate the seasonal effect in the series. It would be better for the next step to look at the time series plot and the correlogram of this new series to be generated by taking seasonal difference.

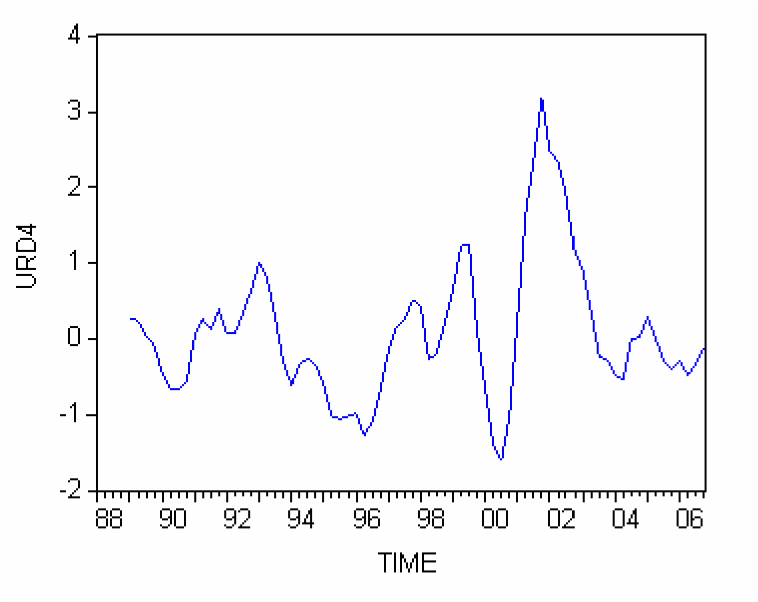


Figure 16. Time Series Plot of Seasonally Differenced Disaggregated Series (Litterman)

There is no sign of seasonality left in the new series, URD4, after the seasonal difference was taken. The inspection that the seasonal period is 4 has really come true. However, just a visual inspection with a time series plot is not enough. The correlogram should also be checked.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.878	0.878	57.850	0.000
		2	0.619	-0.661	87.053	0.000
		3	0.311	-0.094	94.513	0.000
		4	0.045	0.140	94.670	0.000
		5	-0.106	0.218	95.562	0.000
		6	-0.167	-0.270	97.820	0.000
		7	-0.167	-0.048	100.10	0.000
		8	-0.143	0.015	101.81	0.000
		9	-0.124	0.040	103.10	0.000
		10	-0.116	-0.116	104.27	0.000
		11	-0.077	0.367	104.79	0.000
		12	-0.046	-0.488	104.98	0.000

Figure 17. Correlogram of Seasonally Differenced Disaggregated Series (Litterman)

The autocorrelation function shows no lags with regular periodical movements. Therefore, the seasonal period that is assumed as 4 has been supported through this correlogram as well.

As the seasonal component and its period are determined, the model to be built can be specified. The model will be a SARIMA model, nevertheless, the order of autoregressive (and/or moving average) and seasonal autoregressive (and/or moving average) terms in the model are still unknown. Let our model be in the following form;

$$\frac{(1-\phi_1 B-...-\phi_p B^p)(1-B)^d(1-B^s)^D(1-\Phi_1 B^s-...-\Phi_P B^{Ps})}{(1-\theta_1 B-...-\theta_q B^q)(1-\Theta_1 B^s-...-\Theta_Q B^{Qs})} Y_t = \varepsilon_t . \quad (4.2.1)$$

First, the orders of regular and seasonal differences, d and D respectively, should be determined. In order to decide these orders, the Augmented Dickey-Fuller unit root test results in Table 22 can be helpful. In single mean results, regular difference is unnecessary for the series; whereas it seem that a seasonal difference would handle with a seasonal unit root. That is, the difference orders in the model can be thought as $d=0$ and $D=1$.

Table 22. Augmented Dickey-Fuller Test Results on Disaggregated Series (Litterman)

Augmented Dickey-Fuller Unit Root Tests					
Type	Lags	Tau	Pr<Tau	F	Pr>F
Single Mean	1	-3.6	0.008	6.48	0.0057
	4	-2.86	0.0552	4.1	0.0859

The correlogram of the original series, UR, can be used in determining the orders of autoregressive and moving average terms for the SARIMA model denoted in (4.2.1). However, n*4th lags should be neglected in deciding the order because the original series is not seasonally adjusted and its lags in the autocorrelation function still reflects the seasonal movement. Considering this, the autocorrelation function in Figure 15 shows oscillating decay and the partial autocorrelation function has significant spikes at many lags. However, when the seasonal effect in the partial autocorrelation is taken into consideration, the existence of many lags out of boundaries may be the result of the periodical movement. The first cut-off in the partial autocorrelation function seems after first lag, and an order of p=1 can be inferred for the autoregressive term in the SARIMA model.

The correlogram of the differenced series, URD4, can be used in the specification of the orders of seasonal autoregressive and seasonal moving average terms. The oscillating decay in the autocorrelation function and the cut-off property in the partial autocorrelation function for the series URD4 are the indicators for the existence of an order for the seasonal autoregressive term in the model. The first two lags of the partial autocorrelation function shows significance and then, a cut off is observed, therefore the order of seasonal autoregressive term can be considered as P=2.

Since no regular difference is taken and one seasonal difference has become sufficient, the model to be fitted can be defined as SARIMA(1,0,0)X(2,1,0) with seasonal period of 4 and presented in general form as;

$$(1-\phi_1 B)(1-B^4)(1-\Phi_1 B^4-\Phi_2 B^8)Y_t=\varepsilon_t . \quad (4.2.2)$$

If the equation (4.2.2) is written in open form as follows;

$$(1-\phi_1 B-(\Phi_1+1)B^4+\phi_1(\Phi_1+1)B^5-(\Phi_2-\Phi_1)B^8+\phi_1(\Phi_2-\Phi_1)B^9+\Phi_2 B^{12}-\phi_1\Phi_2 B^{13})Y_t=\varepsilon_t$$

the lags of 1, 4, 5, 8, 9, 12 and 13 of the series UR become the explanatory variables of the model. A model with an intercept can be built by these determining variables.

Table 23. Model Fitted for Disaggregated Series (Litterman)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
UR(-1)	1.140823	0.046915	24.31690	0.0000
UR(-5)	-0.376911	0.071839	-5.246642	0.0000
UR(-8)	0.877458	0.059871	14.65580	0.0000
UR(-9)	-0.792624	0.059947	-13.22205	0.0000
UR(-12)	0.153877	0.050457	3.049689	0.0034

R-squared	0.955169	Mean dependent var	8.495000
Adjusted R-squared	0.952130	S.D. dependent var	1.613250
S.E. of regression	0.352966	Akaike info criterion	0.830014
Sum squared resid	7.350512	Schwarz criterion	0.998677
Log likelihood	-21.56045	Durbin-Watson stat	1.603576

A constant term and 4^{th} and 13^{th} lags of the series did not have significant effect in explaining the model. Therefore, the inclusion of these terms in the model was unnecessary and we have fitted the model above. Although it is not a SARIMA model, it is an AR(12) model constructed through the help of SARIMA model structure. Not all the 12 lags of the AR model are significant, but these 5 lags are the ones having contribution in the explanation of the original series. It is essential that the diagnostic checks of this model are made so that the model built is statistically meaningful.

According to Figure 18, the Jarque-Bera Test Statistic reveals a normal distribution for the error terms in the model at 0.05 significance level. The normality assumption made for the error terms is not violated. (Also, Shapiro-Wilk; *p*-value=0.733).

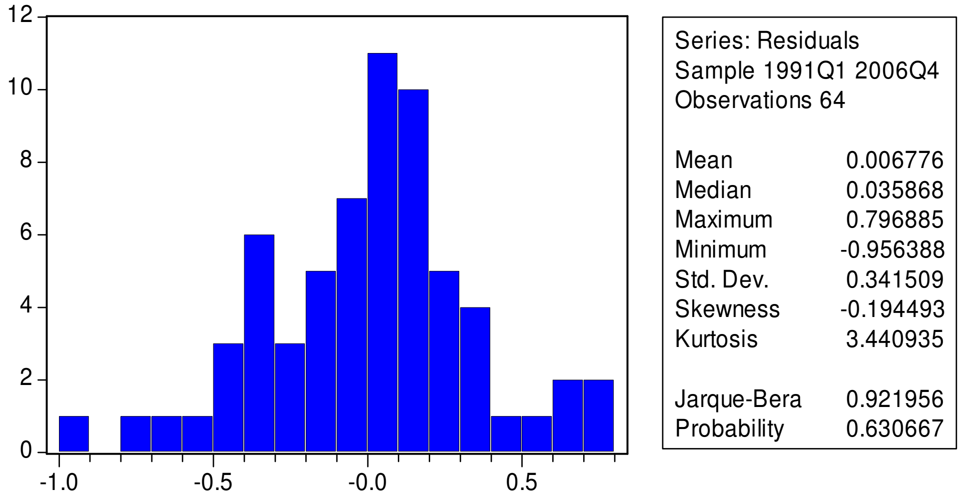


Figure 18. Jarque-Bera Normality Test Results for Residuals of The Model (Litterman)

Table 24. Bresuch-Godfrey Serial Correlation LM Test Results for Residuals of The Model (Litterman)

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.049288	Probability	0.404563
Obs*R-squared	6.772329	Probability	0.342416

The *F*-statistic calculated in Breusch Godfrey Serial Correlation LM Test has a probability of 0.404563 and this *p*-value is greater than the specified error rate 0.05. An existence of correlation between error terms is deniable. The uncorrelated error terms create no problem for the model.

In Table 25, a *p*-value of 0.594466 results in the rejection of heteroskedastic variance for the error terms and the constant variance assumption made has been satisfied.

Table 25. White Heteroskedasticity Test Results for Residuals of The Model (Litterman)

White Heteroskedasticity Test:

F-statistic	0.838286	Probability	0.594466
Obs*R-squared	8.740273	Probability	0.556917

Table 26. ARCH Test Results for Residuals of The Model (Litterman)

ARCH Test:

F-statistic	1.124024	Probability	0.366128
Obs*R-squared	10.09489	Probability	0.342858

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CONSTANT	0.094553	0.056387	1.676863	0.1005
RESID^2(-1)	-0.166439	0.149052	-1.116648	0.2701
RESID^2(-2)	-0.055849	0.151722	-0.368100	0.7145
RESID^2(-3)	0.001691	0.150309	0.011248	0.9911
RESID^2(-4)	0.081603	0.146040	0.558772	0.5791
RESID^2(-5)	0.320035	0.137359	2.329922	0.0244
RESID^2(-6)	0.215084	0.144737	1.486035	0.1442
RESID^2(-7)	-0.129389	0.148359	-0.872139	0.3878
RESID^2(-8)	0.059100	0.150033	0.393910	0.6955
RESID^2(-9)	-0.066575	0.149377	-0.445685	0.6580

Even if all the assumptions made for the error terms in the model have not been violated, the residuals still need to be controlled for the existence of an ARCH effect. The homoskedastic variance of the error terms is supported by the probability of the *F*-statistic calculated in ARCH test, since it is considerably greater than the specified significance level, 0.05.

The acceptance of all assumptions made before modeling inevitably emerges the expectation of a white noise process for the error terms. The correlogram below shows the significance boundaries for the lags of standardized residuals. None of the lags exceeds these boundaries which meets the expectation that the error terms are following a white noise process.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.192	0.192	2.4825	0.115
		2	-0.009	-0.047	2.4874	0.288
		3	-0.217	-0.214	5.7365	0.125
		4	0.013	0.104	5.7482	0.219
		5	0.034	0.009	5.8312	0.323
		6	0.062	0.003	6.1074	0.411
		7	0.006	0.019	6.1098	0.527
		8	-0.100	-0.106	6.8577	0.552
		9	0.117	0.190	7.9175	0.542
		10	-0.007	-0.081	7.9214	0.637
		11	-0.077	-0.129	8.3945	0.678
		12	-0.074	0.063	8.8414	0.716
		13	0.081	0.065	9.3854	0.743
		14	0.127	0.068	10.739	0.706
		15	0.032	-0.028	10.829	0.765
		16	-0.135	-0.143	12.435	0.714
		17	-0.078	0.086	12.984	0.737
		18	0.078	0.066	13.548	0.758
		19	0.125	0.001	15.005	0.722
		20	0.033	0.023	15.111	0.770
		21	-0.102	-0.083	16.142	0.762
		22	-0.145	-0.088	18.265	0.690
		23	-0.061	-0.026	18.649	0.721
		24	0.086	0.051	19.424	0.729
		25	0.079	0.083	20.103	0.741
		26	0.039	0.012	20.273	0.778
		27	0.047	0.026	20.522	0.808
		28	-0.042	-0.073	20.727	0.837

Figure 19. Correlogram of Residuals for The Model (Litterman)

According to Figure 20, the fitted values have caught the seasonal movement and seemed to have been fitted very well. The actual values are so close to those of the fitted model. This model is the best one among all the models fitted throughout the analysis.

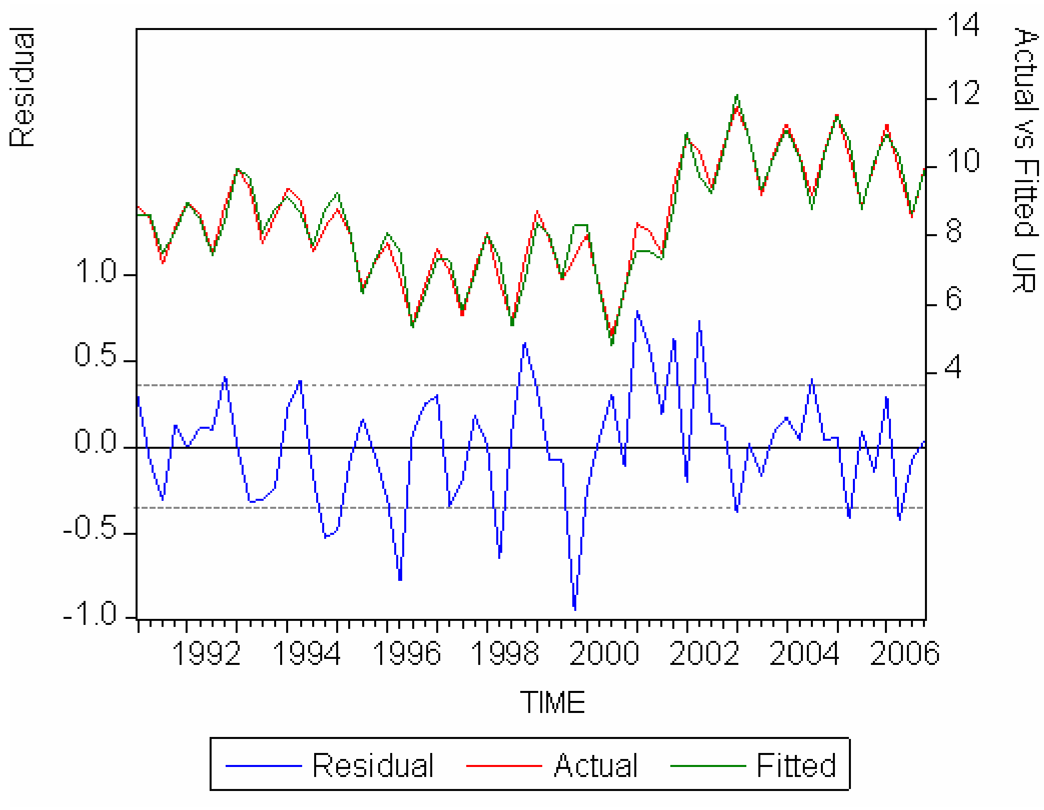


Figure 20. Actual-Fitted-Residual Graph for The Model (Litterman)

Table 27. A Comparison of Forecasts of The Model (Litterman) and Actual Values of Unemployment Rate of Turkey for the Quarters of Years from 2007 to 2010

Year	Quarter	Forecasts	Actual Values	Forecast Error
2007	1	11.14	11.68	0.53
2007	2	9.98	9.23	0.75
2007	3	8.68	9.68	1
2007	4	10.04	10.48	0.43
2008	1	11.41	11.88	0.46
2008	2	10.15	9.17	0.98
2008	3	8.87	10.18	1.31
2008	4	10.32	12.64	2.31
2009	1	11.63	16.12	4.49
2009	2	10.41	13.61	3.2
2009	3	9.07	13.43	4.36
2009	4	10.47	NA	NA
2010	1	11.82	NA	NA
2010	2	10.49	NA	NA
2010	3	9.13	NA	NA
2010	4	10.56	NA	NA

As seen, the forecasts of this model start from the first quarter of 2007. Since every quarterly observations of the related variables were available from 1988 to 2006, it was not possible to generate the disaggregated data after 2006. The quarterly disaggregated data are obtained only for this time interval. That is why, it is not surprising to think that the remarkable increasing behavior of the original series has not been caught by this model and the forecast errors grow larger year by year.

CHAPTER 5

CONCLUSION

Aim of this thesis study is firstly to apply some selected time series disaggregation methods to a real data in order to generate quarterly figures from annual ones. Since higher number of observations would be better representative of the population, the study focuses on this purpose as the first step. Then, some reasonable models are built and forecasts for these generated data are obtained as the latter step.

As the data in our thesis study, we have considered the unemployment rate of Turkey provided by the Turkish Statistical Institute from 1988 up to 2009. However, the data are not collected based on regular time intervals. All of the quarterly observations of the data were available from 2000 to 2008, whereas up to 2000, just two values corresponding to the months of April and October were provided. Moreover, the first three quarters in 2009 were not usable in the study owing to the lack of last quarter. So far, the data used in this irregular form and results obtained by this data are suspicious.

As disaggregation methods, one mathematical and three statistical procedures were examined and applied to the data. The mathematical method was the method of Lisman/Sandee and the statistical ones were regression-based methods of Chow/Lin, Fernandez and Litterman.

For selecting the indicator variables to be utilized in statistical methods, the economic activities of gross national product in constant prices (1987 base) by production were studied on and several combination groups of these activities were built. Then, the generated series provided through the regression-based methods for these groups of combinations were compared to the actual quarterly figures.

According to the comparison of results, the closest disaggregated series to the actual observations were maintained by the method of Litterman when the economic activities, trade and government service, were used as the indicator variables.

The series obtained by the procedure of Lisman/Sandee did not show any seasonal behavior, whereas the method of Litterman gave series exposing a seasonal structure as in the original series.

For these two separate series, different kinds of models have been fitted. Since the series are distinct from each other in their structure, the final models constructed for these were also different. While an ARCH model defined the behavior of the series generated by the method of Lisman/Sandee, an AR model inferred by the help of SARIMA model was fitted for the series provided by the procedure of Litterman.

The model fitted to the disaggregated series of Lisman/Sandee gave high forecast errors when compared to the actual observations. On the other hand, the forecasts of the model fitted to the disaggregated series of Litterman were not that bad, since the first six forecasted quarter values can be accepted as close. However, for the other succeeding forecasted values, the forecast errors are greater. Although the results say so, it should also be considered that the forecasting procedure reveals good results in short run predictions. Also, the effect of the economic crises that we have on the unemployment rate cannot be predicted before. The use of other explanatory variables and considering structural dynamic models may help us to predict this kind of huge crisis before. Although good forecasts with small errors have not been realized, our main purposes have been achieved by obtaining quarterly series from 1988 to 2006 using disaggregation methods and then, modeling the series using univariate time series models.

Therefore, for our study, it is not wrong to say that the results of Litterman were more preferable than that of Lisman/Sandee. The series of Litterman both caught the seasonal structure and gave better forecasts when compared to the series of Lisman/Sandee. If the actual observations had not followed a seasonal behavior, the considerable superiority of the method of Litterman over the method of Lisman/Sandee may not have been noticeable as in this study. That is, for a nonseasonal series, the method of Lisman/Sandee can be as efficient as the regression-based methods. Just like the method of Lisman/Sandee, least square methods and ARIMA procedures could give good results due to the structure of the series as well. Thus, time series disaggregation methods should be wisely pondered according to the observed series in hand and it is essential that the decision of which of these methods should be applied is given this way.

This study is thought to have contributions to the literature, especially for Turkey, because this study is comprised of both disaggregating and modeling procedures unlike the other studies conducted in Turkey. In addition, selected disaggregation methods were applied to the unemployment rate data of Turkey for the first time and the generated data have been maintained close to the observed actual values. For the future studies, anyone who wants to work with Turkish unemployment rate data can use the data that we generated. Moreover, other disaggregation methods can also be tried to see whether it is possible to generate closer quarterly, or even monthly, values. However, considering the forecasts in this study, it will be better to focus on the forecasting procedures in a more detailed manner such that an intervention analysis can be conducted to be able to see the effect of economical crisis on the series.

REFERENCES

Al – Osh, M. (1989). A Dynamic Linear Model Approach for Disaggregating Time Series Data. Journal of Forecasting, 8, 85 – 96.

Balmer, D.W. (1975). An Observation on the Disaggregation of Annual Time Series Data. Applied Statistics, 24, 352 – 354.

Barro, R.J. (1977). Unanticipated Money Growth and Unemployment in the United States. The American Economic Review, 67, 101 – 115.

Bera, A.K. and Jarque, C.M. (1981). Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals: Monte Carlo Evidence. Economic Letters C.7, 4, 313 – 318.

Biagi, F. and Lucifora, C. (2008). Demographic and Education Effects on Unemployment in Europe. Labour Economics, 15, 1076 – 1101.

Breusch, T.S. (1978). Testing for Autocorrelation in Dynamic Linear Models. Australian Economic Papers, 17, 334 – 355.

Boot, J.C.G., Feibes, W. and Lisman, J.H.C. (1967). Further Methods of Derivation of Quarterly Figures from Annual Data. Journal of the Royal Statistical Society, Series C, 16, 65 – 75.

Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis, Forecasting and Control. San Francisco: Holden – Day.

Cantor, D. and Land, K.C. (1985). Unemployment and Crime Rates in the Post – World War II United States: A Theoretical and Empirical Analysis. American Sociological Review, 50, 317 – 322.

Chen, B. (2007). An Empirical Comparison of Methods for Temporal Disaggregation at the National Accounts. Bureau of Economic Analysis, August 30.

Chow, G.C. and Lin, A. (1971). Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series. The Review of Economics and Statistics, 53, 372 – 375.

Clark, A.E. and Oswald, A.J. (1994). Unhappiness and Unemployment. The Economic Journal, 104, 648 – 659.

Cohen, K.J., Muller, W. and Padberg, M.W. (1971). Autoregressive Approaches to Disaggregation of Time Series Data. Journal of the Royal Statistical Society, Series C, 20, 119 – 120.

Cryer, J.D. and Chan, K.S. (2008). Time Series Analysis with Applications in R, 2nd ed. New York: Springer.

Da Vanzo, J. (1978). Does Unemployment Affect Migration? Evidence from Micro Data. The Review of Economics and Statistics, 60, 504 – 514.

Denton, F.T. (1971). Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization. Journal of the Americal Statistical Association, 66, 99 – 102.

Di Fonzo, T. (2003). Introduction to the Principles of Benchmarking and Temporal Disaggregation for Economic Series. Application of Advanced Temporal Disaggregation Techniques to Economic Statistics, OECD/Eurostat Workshop, 27 November 2003, Paris.

Dickey, D.A. and Fuller, W.A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. Econometrica, 49, 1057 – 1072.

Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation. Econometrica, 50, 987 – 1007.

Fernandez, R.B. (1981). A Methodological Note on the Estimation of Time Series. The Review of Economics and Statistics, 63, 471 – 476.

Friedman, M. (1962). The Interpolation of Time Series by Related Series. Journal of the American Statistical Association, 57, 729 – 757.

Ginsburgh, V.A. (1973). A Further Note on the Derivation of Quarterly Figures Consistent with Annual Data. Applied Statistics, 22, 368 – 374.

Godfrey, L.G. (1978). Testing Against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables. Econometrica, 46, 1293 – 1301.

Gomez, V. and Maravall, A. (1994). Estimation, Prediction and Interpolation for Nonstationary Series with the Kalman Filter. Journal of the American Statistical Association, 89, 611 – 624.

Guerrero, V.M. (1990). Temporal Disaggregation of Time Series. An ARIMA-based Approach. International Statistical Review, 58, 29 – 46.

Guerrero, V.M. (2003). Monthly Disaggregation of a Quarterly Time Series and Forecasts of Its Unobservable Monthly Values. Journal of Official Statistics, 19, 215 – 235.

Harvey, A.C. and Pierse, R.G. (1984). Estimating Missing Observations in Economic Time Series. Journal of the American Statistical Association, 79, 125 – 131.

Holden, K. and Peel, D.A. (1979). The Determinants of the Unemployment Rate: Some Empirical Evidence. Journal of the Royal Statistical Society, Series D, 28, 101 – 107.

Jackman, R., Pissarides, C. and Savouri, S. (1990). Labor Market Policies and Unemployment in the OECD. Economic Policy, 5.

Kessler, R.C., Turner, J.B. and House J.S. (1988). Effects of Unemployment on Health in a Community Survey: Main, Modifying and Mediating Effects. Journal of Social Issues, 44, 69 – 85.

Kladroba, A. (2005). The Temporal Disaggregation of Time Series, University of Duisburg – Essen, December 2005.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992). Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root. Journal of Econometrics, 54, 159 – 178.

Linn, M.W., Sandifer, R. and Stein, S. (1985). Effects of Unemployment on Mental and Physical Health. American Journal of Public Health, 75, 502 – 506.

Lisman, J.H.C. and Sandee, J. (1964). Derivation of Quarterly Figures from Annul Data. Journal of the Royal Statistical Society, Series C, 13, 87 - 90.

Litterman, R.B. (1983). A Random Walk. Markov Model for the Distribution of Time Series. Journal of Business and Statistics, 1, 169 – 173.

McCormick, B. (1997). Regional Unemployment and Labor Market Mobility in the UK. European Economic Review, 41, 581 – 589.

Miralles, J.M.P., Lladosa, L.E.V. and Vallés, R.E. (2003). On the Performance of the Chow-Lin Procedure for Quarterly Interpolation of Annual Data: Some Monte-Carlo Analysis. Spanish Economic Review, 5, 291 – 305.

Montgomery, A.L., Zarnowitz, W., Tsay, R.S. and Tiao, G.C. (1998). Forecasting the U.S. Unemployment Rate. Journal of the American Statistical Association, 93, 478 – 493.

Phillips, P.C.B. and Perron, P. (1988). Testing for a Unit Root in Time Series Regression. Biometrika, 75, 335 – 346.

Sürmeli, M.C. (1979). Disaggregation of Time Series Data. Master Thesis, Middle East Technical University, Ankara.

Vangrevelinghe, E. (1966). L'évolution a Court Terme de la Consommation des Ménages. Etudes et Conjonctures (I.N.S.E.E.), 9, 54 – 102.

White, H. (1980). A Heteroskedasticity – Consistent Covariance Matrix Estimator and A Direct Test for Heteroskedasticity. Econometrica, 48, 817 – 838.

Wei, W.W.S. and Stram, D.O. (1990). Disaggregation of Time Series Models. Journal of the Royal Statistical Society, Series B, 52, 453 – 467.

APPENDIX A

ANNUAL FIGURES OF THE UNEMPLOYMENT RATE OF TURKEY

Table 28. Unemployment Rate of Turkey between 1988 and 2008

Year	Unemployment Rate (in %)
1988	8.45
1989	8.57
1990	7.99
1991	8.20
1992	8.49
1993	8.93
1994	8.55
1995	7.63
1996	6.62
1997	6.82
1998	6.87
1999	7.66
2000	6.49
2001	8.37
2002	10.35
2003	10.54
2004	10.30
2005	10.20
2006	9.90
2007	10.28
2008	10.97

APPENDIX B

DISAGGREGATED SERIES ACCORDING TO METHODS AND ACTUAL QUARTER VALUES OF UNEMPLOYMENT RATE OF TURKEY

Table 29. Disaggregated Series and Actual Values between 1988 and 2009

Year	Quarter	Lisman/Sandee	Litterman	Actual Values
1988	1	NA	9	NA
1988	2	NA	8.66	NA
1988	3	NA	7.65	NA
1988	4	NA	8.49	NA
1989	1	8.59	9.28	NA
1989	2	8.68	8.9	NA
1989	3	8.62	7.7	NA
1989	4	8.42	8.41	NA
1990	1	8.15	8.83	NA
1990	2	7.94	8.25	NA
1990	3	7.89	7.04	NA
1990	4	8.01	7.85	NA
1991	1	8.11	8.88	NA
1991	2	8.16	8.51	NA
1991	3	8.22	7.17	NA
1991	4	8.3	8.24	NA
1992	1	8.37	8.95	NA
1992	2	8.43	8.59	NA
1992	3	8.52	7.52	NA
1992	4	8.64	8.89	NA
1993	1	8.83	9.97	NA
1993	2	9.01	9.38	NA
1993	3	9.02	7.79	NA
1993	4	8.86	8.58	NA
1994	1	8.74	9.36	NA
1994	2	8.69	9.06	NA
1994	3	8.52	7.53	NA
1994	4	8.25	8.24	NA
1995	1	7.98	8.76	NA
1995	2	7.76	8.04	NA
1995	3	7.52	6.48	NA

Table 29 (continued)

Year	Quarter	Lisman/Sandee	Litterman	Actual Values
1995	4	7.26	7.23	NA
1996	1	6.9	7.78	NA
1996	2	6.55	6.77	NA
1996	3	6.44	5.39	NA
1996	4	6.59	6.54	NA
1997	1	6.76	7.63	NA
1997	2	6.82	6.94	NA
1997	3	6.85	5.64	NA
1997	4	6.85	7.06	NA
1998	1	6.79	8.05	NA
1998	2	6.74	6.67	NA
1998	3	6.84	5.46	NA
1998	4	7.1	7.29	NA
1999	1	7.53	8.7	NA
1999	2	7.89	7.9	NA
1999	3	7.84	6.7	NA
1999	4	7.39	7.36	NA
2000	1	6.67	8.04	8.25
2000	2	6.13	6.49	6.12
2000	3	6.22	5.1	5.53
2000	4	6.94	6.34	6.25
2001	1	7.66	8.35	8.49
2001	2	8.12	8.14	6.73
2001	3	8.61	7.49	7.82
2001	4	9.11	9.52	10.4
2002	1	9.76	10.83	11.55
2002	2	10.39	10.48	9.32
2002	3	10.66	9.39	9.56
2002	4	10.57	10.67	11.05
2003	1	10.5	11.74	12.32
2003	2	10.59	10.85	10.03
2003	3	10.58	9.17	9.41
2003	4	10.49	10.4	10.33
2004	1	10.38	11.26	12.45
2004	2	10.3	10.33	9.28
2004	3	10.26	9.17	9.47
2004	4	10.25	10.43	9.99
2005	1	10.26	11.55	11.66
2005	2	10.25	10.34	9.17
2005	3	10.2	8.89	9.44
2005	4	10.11	10.04	10.64

Table 29 (continued)

Year	Quarter	Lisman/Sandee	Litterman	Actual Values
2006	1	9.96	11.26	11.95
2006	2	9.83	9.87	8.84
2006	3	9.84	8.56	9.14
2006	4	9.99	9.93	9.64
2007	1	10.12	NA	11.68
2007	2	10.19	NA	9.23
2007	3	10.32	NA	9.68
2007	4	10.51	NA	10.48
2008	1	NA	NA	11.88
2008	2	NA	NA	9.17
2008	3	NA	NA	10.18
2008	4	NA	NA	12.64
2009	1	NA	NA	16.12
2009	2	NA	NA	13.61
2009	3	NA	NA	13.43

APPENDIX C

DISAGGREGATED SERIES OF ECONOMIC ACTIVITY GROUPS ACCORDING TO REGRESSION BASED METHODS

Table 30. gnp-gs-ind-agr-trad

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	9.50	9.30	9.61	8.25
2000	2	8.34	8.93	8.33	6.12
2000	3	1.38	0.88	1.59	5.53
2000	4	6.74	6.85	6.44	6.25
2001	1	10.35	10.04	10.52	8.49
2001	2	9.95	10.55	9.82	6.73
2001	3	3.70	3.66	3.80	7.82
2001	4	9.49	9.25	9.36	10.40
2002	1	12.99	12.82	13.07	11.55
2002	2	12.37	12.65	12.53	9.32
2002	3	5.50	5.37	5.63	9.56
2002	4	10.53	10.54	10.15	11.05
2003	1	13.81	13.67	13.99	12.32
2003	2	12.53	12.88	12.60	10.03
2003	3	5.49	5.37	5.62	9.41
2003	4	10.34	10.24	9.95	10.33
2004	1	13.10	12.93	13.10	12.45
2004	2	12.33	12.75	12.55	9.28
2004	3	5.69	5.46	5.94	9.47
2004	4	10.07	10.04	9.60	9.99
2005	1	13.35	12.98	13.64	11.66
2005	2	12.68	13.25	12.78	9.17
2005	3	4.98	4.74	5.11	9.44
2005	4	9.81	9.84	9.29	10.64
2006	1	12.78	12.67	12.88	11.95
2006	2	12.52	12.89	12.86	8.84
2006	3	4.77	4.67	4.86	9.14
2006	4	9.54	9.38	9.02	9.64

Table 31. gs-ind-agr-trad

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	10.72	12.19	10.72	8.25
2000	2	8.55	9.76	8.51	6.12
2000	3	-0.30	-3.63	0.10	5.53
2000	4	7.00	7.65	6.64	6.25
2001	1	11.23	11.95	11.33	8.49
2001	2	10.01	11.06	9.86	6.73
2001	3	2.11	-0.35	2.39	7.82
2001	4	10.15	10.83	9.91	10.4
2002	1	13.90	14.99	13.89	11.55
2002	2	12.64	13.42	12.79	9.32
2002	3	3.81	1.05	4.14	9.56
2002	4	11.04	11.93	10.56	11.05
2003	1	14.64	15.62	14.75	12.32
2003	2	12.77	13.67	12.83	10.03
2003	3	3.65	0.70	4.00	9.41
2003	4	11.10	12.17	10.58	10.33
2004	1	14.18	15.48	14.07	12.45
2004	2	12.40	13.07	12.65	9.28
2004	3	3.86	0.72	4.33	9.47
2004	4	10.76	11.92	10.14	9.99
2005	1	14.39	15.26	14.62	11.66
2005	2	12.50	13.10	12.65	9.17
2005	3	3.14	0.02	3.48	9.44
2005	4	10.78	12.43	10.06	10.64
2006	1	14.07	15.77	14.01	11.95
2006	2	12.45	12.85	12.79	8.84
2006	3	2.87	-0.13	3.18	9.14
2006	4	10.22	11.12	9.63	9.64

Table 32. gnp-gs-ind-agr

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	9.50	8.73	9.63	8.25
2000	2	8.35	8.34	8.50	6.12
2000	3	1.37	2.36	1.30	5.53
2000	4	6.74	6.52	6.53	6.25
2001	1	10.35	10.05	10.40	8.49
2001	2	9.95	10.01	10.00	6.73
2001	3	3.70	4.50	3.69	7.82
2001	4	9.49	8.93	9.40	10.40
2002	1	12.98	12.56	13.06	11.55
2002	2	12.37	12.47	12.53	9.32
2002	3	5.49	6.39	5.48	9.56
2002	4	10.53	9.97	10.31	11.05
2003	1	13.81	13.43	13.99	12.32
2003	2	12.53	12.53	12.68	10.03
2003	3	5.48	6.53	5.45	9.41
2003	4	10.35	9.68	10.05	10.33
2004	1	13.10	12.53	13.09	12.45
2004	2	12.33	12.63	12.55	9.28
2004	3	5.69	6.76	5.74	9.47
2004	4	10.08	9.26	9.81	9.99
2005	1	13.34	12.88	13.51	11.66
2005	2	12.68	13.07	12.88	9.17
2005	3	4.97	5.98	4.92	9.44
2005	4	9.82	8.88	9.50	10.64
2006	1	12.78	11.99	12.92	11.95
2006	2	12.52	13.04	12.82	8.84
2006	3	4.77	5.80	4.72	9.14
2006	4	9.54	8.79	9.16	9.64

Table 33. gs-ind-agr

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	11.13	11.30	11.74	8.25
2000	2	8.81	9.14	9.12	6.12
2000	3	-1.16	-1.73	-1.92	5.53
2000	4	7.19	7.26	7.03	6.25
2001	1	11.35	11.66	11.71	8.49
2001	2	10.24	10.52	10.39	6.73
2001	3	1.51	0.96	0.90	7.82
2001	4	10.39	10.36	10.49	10.40
2002	1	14.15	14.45	14.56	11.55
2002	2	12.74	13.13	13.03	9.32
2002	3	3.12	2.53	2.47	9.56
2002	4	11.36	11.27	11.31	11.05
2003	1	14.87	15.13	15.40	12.32
2003	2	12.95	13.25	13.25	10.03
2003	3	2.89	2.35	2.15	9.41
2003	4	11.45	11.44	11.36	10.33
2004	1	14.50	14.77	14.89	12.45
2004	2	12.44	12.89	12.74	9.28
2004	3	3.04	2.50	2.40	9.47
2004	4	11.20	11.02	11.16	9.99
2005	1	14.57	14.84	15.14	11.66
2005	2	12.56	12.94	12.79	9.17
2005	3	2.36	1.75	1.60	9.44
2005	4	11.32	11.28	11.28	10.64
2006	1	14.53	14.75	15.09	11.95
2006	2	12.38	12.94	12.65	8.84
2006	3	2.12	1.52	1.37	9.14
2006	4	10.59	10.40	10.51	9.64

Table 34. gs-agr-trad

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	10.89	11.65	11.47	8.25
2000	2	8.64	9.42	8.91	6.12
2000	3	-0.68	-2.42	-1.37	5.53
2000	4	7.11	7.32	6.95	6.25
2001	1	11.22	11.98	11.53	8.49
2001	2	10.09	10.76	10.22	6.73
2001	3	1.89	0.34	1.36	7.82
2001	4	10.29	10.42	10.38	10.40
2002	1	13.97	14.78	14.33	11.55
2002	2	12.62	13.41	12.88	9.32
2002	3	3.54	1.89	2.98	9.56
2002	4	11.25	11.30	11.20	11.05
2003	1	14.69	15.46	15.17	12.32
2003	2	12.82	13.50	13.09	10.03
2003	3	3.33	1.69	2.68	9.41
2003	4	11.32	11.51	11.22	10.33
2004	1	14.29	15.12	14.64	12.45
2004	2	12.34	13.18	12.60	9.28
2004	3	3.50	1.85	2.94	9.47
2004	4	11.06	11.04	11.01	9.99
2005	1	14.41	15.19	14.94	11.66
2005	2	12.45	13.20	12.66	9.17
2005	3	2.82	1.05	2.15	9.44
2005	4	11.13	11.37	11.07	10.64
2006	1	14.28	15.12	14.80	11.95
2006	2	12.30	13.23	12.52	8.84
2006	3	2.57	0.83	1.91	9.14
2006	4	10.47	10.44	10.39	9.64

Table 35. gnp-gs-agr

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	10.73	10.90	11.42	8.25
2000	2	8.60	8.85	8.95	6.12
2000	3	-0.46	-1.00	-1.37	5.53
2000	4	7.10	7.21	6.96	6.25
2001	1	11.03	11.27	11.46	8.49
2001	2	10.08	10.28	10.27	6.73
2001	3	2.14	1.64	1.40	7.82
2001	4	10.24	10.29	10.36	10.40
2002	1	13.80	14.05	14.29	11.55
2002	2	12.53	12.82	12.87	9.32
2002	3	3.80	3.25	3.01	9.56
2002	4	11.25	11.26	11.21	11.05
2003	1	14.54	14.74	15.14	12.32
2003	2	12.75	12.97	13.10	10.03
2003	3	3.59	3.08	2.71	9.41
2003	4	11.28	11.38	11.21	10.33
2004	1	14.12	14.38	14.59	12.45
2004	2	12.25	12.57	12.60	9.28
2004	3	3.75	3.21	2.96	9.47
2004	4	11.07	11.03	11.04	9.99
2005	1	14.21	14.42	14.85	11.66
2005	2	12.41	12.65	12.70	9.17
2005	3	3.09	2.52	2.18	9.44
2005	4	11.10	11.22	11.08	10.64
2006	1	14.11	14.35	14.75	11.95
2006	2	12.21	12.60	12.53	8.84
2006	3	2.85	2.29	1.95	9.14
2006	4	10.45	10.38	10.38	9.64

Table 36. trad-agr

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	10.41	11.54	11.54	8.25
2000	2	8.32	9.57	8.97	6.12
2000	3	0.27	-2.47	-1.53	5.53
2000	4	6.97	7.32	6.97	6.25
2001	1	10.76	11.90	11.60	8.49
2001	2	9.80	10.86	10.27	6.73
2001	3	2.85	0.43	1.20	7.82
2001	4	10.09	10.30	10.42	10.40
2002	1	13.41	14.66	14.42	11.55
2002	2	12.32	13.59	12.93	9.32
2002	3	4.55	1.95	2.80	9.56
2002	4	11.10	11.19	11.22	11.05
2003	1	14.18	15.36	15.26	12.32
2003	2	12.53	13.62	13.14	10.03
2003	3	4.33	1.77	2.51	9.41
2003	4	11.12	11.43	11.26	10.33
2004	1	13.79	15.06	14.72	12.45
2004	2	12.10	13.45	12.65	9.28
2004	3	4.45	1.85	2.78	9.47
2004	4	10.86	10.82	11.05	9.99
2005	1	13.95	15.13	15.00	11.66
2005	2	12.14	13.36	12.71	9.17
2005	3	3.83	1.05	1.98	9.44
2005	4	10.89	11.27	11.12	10.64
2006	1	13.80	15.02	14.88	11.95
2006	2	11.95	13.43	12.59	8.84
2006	3	3.57	0.85	1.74	9.14
2006	4	10.30	10.33	10.41	9.64

Table 37. gnp-agr

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	10.28	10.75	11.44	8.25
2000	2	8.30	8.94	8.97	6.12
2000	3	0.43	-0.93	-1.42	5.53
2000	4	6.96	7.20	6.97	6.25
2001	1	10.61	11.21	11.48	8.49
2001	2	9.79	10.30	10.29	6.73
2001	3	3.04	1.81	1.35	7.82
2001	4	10.04	10.17	10.37	10.40
2002	1	13.28	13.89	14.32	11.55
2002	2	12.25	13.01	12.89	9.32
2002	3	4.75	3.39	2.95	9.56
2002	4	11.10	11.09	11.22	11.05
2003	1	14.06	14.59	15.17	12.32
2003	2	12.49	13.06	13.12	10.03
2003	3	4.53	3.23	2.66	9.41
2003	4	11.08	11.28	11.22	10.33
2004	1	13.65	14.31	14.61	12.45
2004	2	12.05	12.88	12.62	9.28
2004	3	4.63	3.28	2.91	9.47
2004	4	10.85	10.72	11.05	9.99
2005	1	13.80	14.38	14.88	11.66
2005	2	12.12	12.75	12.72	9.17
2005	3	4.04	2.60	2.12	9.44
2005	4	10.86	11.09	11.10	10.64
2006	1	13.67	14.23	14.78	11.95
2006	2	11.89	12.82	12.55	8.84
2006	3	3.78	2.35	1.90	9.14
2006	4	10.28	10.23	10.39	9.64

Table 38. gnp-gs

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	7.74	7.16	8.25	8.25
2000	2	6.68	6.51	6.79	6.12
2000	3	4.83	5.50	4.50	5.53
2000	4	6.71	6.79	6.42	6.25
2001	1	8.28	7.85	8.49	8.49
2001	2	8.46	8.31	8.44	6.73
2001	3	7.18	7.83	7.01	7.82
2001	4	9.58	9.51	9.55	10.40
2002	1	10.87	10.44	11.06	11.55
2002	2	10.56	10.42	10.67	9.32
2002	3	9.04	9.66	8.88	9.56
2002	4	10.91	10.86	10.77	11.05
2003	1	11.64	11.14	11.98	12.32
2003	2	10.95	10.76	11.08	10.03
2003	3	8.92	9.62	8.65	9.41
2003	4	10.65	10.65	10.45	10.33
2004	1	11.24	10.82	11.46	12.45
2004	2	10.40	10.32	10.52	9.28
2004	3	8.85	9.46	8.66	9.47
2004	4	10.69	10.58	10.55	9.99
2005	1	11.22	10.67	11.68	11.66
2005	2	10.58	10.43	10.63	9.17
2005	3	8.63	9.33	8.35	9.44
2005	4	10.38	10.38	10.16	10.64
2006	1	11.07	10.55	11.49	11.95
2006	2	10.13	10.10	10.09	8.84
2006	3	8.34	9.02	8.06	9.14
2006	4	10.08	9.95	9.98	9.64

Table 39. gs-ind

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	7.06	6.70	7.53	8.25
2000	2	6.19	6.26	6.20	6.12
2000	3	6.02	6.23	5.86	5.53
2000	4	6.70	6.77	6.37	6.25
2001	1	7.56	7.42	7.69	8.49
2001	2	8.00	8.07	7.89	6.73
2001	3	8.39	8.57	8.42	7.82
2001	4	9.54	9.44	9.50	10.40
2002	1	10.12	9.97	10.22	11.55
2002	2	10.05	10.19	10.07	9.32
2002	3	10.26	10.40	10.30	9.56
2002	4	10.96	10.83	10.80	11.05
2003	1	10.87	10.66	11.10	12.32
2003	2	10.47	10.52	10.53	10.03
2003	3	10.13	10.37	10.05	9.41
2003	4	10.69	10.61	10.49	10.33
2004	1	10.59	10.41	10.73	12.45
2004	2	9.90	10.12	9.94	9.28
2004	3	9.95	10.14	9.92	9.47
2004	4	10.75	10.52	10.59	9.99
2005	1	10.47	10.22	10.88	11.66
2005	2	9.98	10.14	9.91	9.17
2005	3	9.89	10.09	9.79	9.44
2005	4	10.47	10.35	10.23	10.64
2006	1	10.40	10.09	10.77	11.95
2006	2	9.49	9.82	9.29	8.84
2006	3	9.56	9.76	9.47	9.14
2006	4	10.17	9.95	10.10	9.64

Table 40. trad

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	7.51	6.30	7.95	8.25
2000	2	6.41	6.27	6.49	6.12
2000	3	5.43	6.67	5.03	5.53
2000	4	6.60	6.72	6.48	6.25
2001	1	8.12	7.14	8.29	8.49
2001	2	8.18	8.04	8.18	6.73
2001	3	7.71	9.09	7.36	7.82
2001	4	9.49	9.21	9.67	10.40
2002	1	10.59	9.59	10.90	11.55
2002	2	10.40	10.25	10.44	9.32
2002	3	9.61	10.93	9.22	9.56
2002	4	10.79	10.61	10.82	11.05
2003	1	11.38	10.32	11.72	12.32
2003	2	10.75	10.51	10.83	10.03
2003	3	9.48	10.93	9.05	9.41
2003	4	10.55	10.41	10.56	10.33
2004	1	11.02	10.10	11.25	12.45
2004	2	10.27	10.32	10.24	9.28
2004	3	9.39	10.62	9.06	9.47
2004	4	10.51	10.15	10.63	9.99
2005	1	11.07	9.93	11.41	11.66
2005	2	10.32	10.21	10.35	9.17
2005	3	9.21	10.59	8.81	9.44
2005	4	10.22	10.09	10.25	10.64
2006	1	10.85	9.70	11.18	11.95
2006	2	9.93	9.94	9.93	8.84
2006	3	8.87	10.26	8.47	9.14
2006	4	9.97	9.72	10.04	9.64

Table 41. gnp

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	7.69	6.35	8.23	8.25
2000	2	6.64	6.21	6.82	6.12
2000	3	5.00	6.70	4.38	5.53
2000	4	6.63	6.70	6.52	6.25
2001	1	8.23	7.19	8.48	8.49
2001	2	8.40	7.99	8.50	6.73
2001	3	7.38	9.08	6.85	7.82
2001	4	9.49	9.23	9.66	10.40
2002	1	10.75	9.63	11.14	11.55
2002	2	10.55	10.22	10.66	9.32
2002	3	9.26	10.93	8.69	9.56
2002	4	10.83	10.61	10.89	11.05
2003	1	11.56	10.35	12.01	12.32
2003	2	10.92	10.47	11.09	10.03
2003	3	9.12	10.92	8.49	9.41
2003	4	10.56	10.42	10.57	10.33
2004	1	11.17	10.13	11.48	12.45
2004	2	10.43	10.27	10.47	9.28
2004	3	9.03	10.62	8.52	9.47
2004	4	10.56	10.16	10.72	9.99
2005	1	11.20	9.98	11.61	11.66
2005	2	10.54	10.15	10.67	9.17
2005	3	8.82	10.59	8.22	9.44
2005	4	10.26	10.09	10.32	10.64
2006	1	11.03	9.75	11.47	11.95
2006	2	10.09	9.91	10.16	8.84
2006	3	8.51	10.25	7.93	9.14
2006	4	9.99	9.71	10.06	9.64

Table 42. ind

Year	Quarter	Fernandez	AR(1) Min SSR	Litterman Min SSR	Actual Values
2000	1	7.07	6.47	7.41	8.25
2000	2	6.20	6.35	6.18	6.12
2000	3	6.05	6.44	5.85	5.53
2000	4	6.64	6.70	6.52	6.25
2001	1	7.58	7.35	7.57	8.49
2001	2	8.00	8.12	7.91	6.73
2001	3	8.44	8.81	8.35	7.82
2001	4	9.47	9.22	9.67	10.40
2002	1	10.09	9.79	10.24	11.55
2002	2	10.10	10.38	9.99	9.32
2002	3	10.32	10.65	10.18	9.56
2002	4	10.87	10.57	10.97	11.05
2003	1	10.87	10.51	11.05	12.32
2003	2	10.50	10.61	10.47	10.03
2003	3	10.19	10.66	9.98	9.41
2003	4	10.61	10.38	10.67	10.33
2004	1	10.58	10.27	10.68	12.45
2004	2	9.99	10.44	9.80	9.28
2004	3	10.00	10.39	9.86	9.47
2004	4	10.62	10.09	10.85	9.99
2005	1	10.52	10.14	10.68	11.66
2005	2	10.01	10.34	9.88	9.17
2005	3	9.92	10.31	9.76	9.44
2005	4	10.36	10.02	10.49	10.64
2006	1	10.42	9.85	10.65	11.95
2006	2	9.53	10.12	9.29	8.84
2006	3	9.59	9.98	9.44	9.14
2006	4	10.07	9.67	10.24	9.64