BUILDING COST INDEX FORECASTING WITH TIME SERIES ANALYSIS

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

BY

MUSTAFA ALPTEKİN KİBAR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
CIVIL ENGINEERING

AUGUST 2007

Approval of the thesis:

BUILDING COST INDEX FORECASTING
WITH TIME SERIES ANALYSIS

submitted by MUSTAFA ALPTEKİN KİBAR in partial fulfillment of the requirements for the degree of Master of Science by,

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

Prof. Dr. Güney Özcebe ______________________
Head of Department, Civil Engineering

Asst. Prof. Dr. Rıfat Sönmez ______________________
Supervisor, Civil Engineering Dept., METU

Examining Committee Members:

Prof. Dr. Talat Birgönül ______________________
Civil Engineering Dept., METU

Asst. Prof. Dr. Rıfat Sönmez ______________________
Civil Engineering Dept., METU

Assoc. Prof. Dr. İrem Dikmen Toker ______________________
Civil Engineering Dept., METU

Assoc. Prof. Dr. Murat Gündüz ______________________
Civil Engineering Dept., METU

Caner Anaç, M.S. ______________________
Cost Control Engineer, İÇTAŞ

Date: 09.08.2007

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 : MUSTAFA ALPTEKİN KİBAR

Signature :

ABSTRACT

BUILDING COST INDEX FORECASTING WITH TIME SERIES ANALYSIS

Kibar, Mustafa Alptekin
M.Sc.,Department of Civil Engineering
Supervisor: Asst. Prof. Dr. Rıfat Sönmez

August 2007, 98 pages

Building cost indices are widely used in construction industry to measure the rate of change of building costs as a combination of labor and material costs. Cost index forecast is crucial for the two main parties of construction industry, contactor, and the client. Forecast information is used to increase the accuracy of estimate for the project cost to evaluate the bid price.

The aim of this study is to develop time series models to forecast building cost indices in Turkey and United States. The models developed are compared with regression analysis and simple averaging models in terms of predictive accuracy. As a result of this study, time series models are selected as the most accurate models in predicting cost indices for both Turkey and United States. Future values of building cost indices can be predicted in adequate precision using time series models. This useful information can be used in tender process in estimation of project costs, which is one of the critical factors affecting the overall success of a construction project. Better cost estimates shall enable contractors to produce cash flow forecasts more acurately. Furthermore accurate prediction of future prices is very useful for owners in budget allocations; moreover can help investors to evaluate project alternatives adequately.

Keywords: Cost Index, Forecasting, Time Series, Regression Analysis

ÖZ

ZAMAN SERİLERİ ANALİZİ KULLANARAK BİNA MALİYET ENDEKSLERİNİN TAHMİN EDİLMESİ

Kibar, Mustafa Alptekin
Yüksek Lisans.,İnşaat Mühendisliği
Tez Yöneticisi: Asst. Prof. Dr. Rıfat Sönmez

Ağustos 2007, 98 Sayfa

Bina maliyet endeksleri, bina maliyetlerinin zaman içindeki değişiminin malzeme ve işçilik maliyetlerine bağlı olarak ölçmek için kullanılmaktadır. Bina maliyet endeksinin tahmini inşaat sektöründe işveren için de, müteahhit için de büyük önem arz etmektedir. Bu endeksin doğru tahmin edilmesi ihale aşamasında belirlenen proje maliyetinin gerçeğe daha yakın olmasını sağlar.

Bu çalışmanın amacı Türkiye'de ve A.B.D'de yaygın olarak kullanılan bina maliyet endekslerinin tahminlerini yapmak için zaman serisi modelleri geliştirmektir. Geliştirilen bu modeller regresyon analizi ve basit ortalama modelleri ile tahmin yakınlığı açısından karşılaştırılmıştır. Çalışmanın sonunda zaman serisi modelleri; hem Türkiye hem de A.B.D. bina maliyet endekslerinin tahmini için en uygun yöntem olarak seçilmiştir. bina maliyet endekslerinin gelecek zaman değerleri zaman serileri modelleri yardımıyla yeterli hassaslıkta tahmin edilebilir. Bu tahmin ihale aşamasında proje maliyetinin belirlenmesi için kullanılabilir. Daha iyi tahminler nakit akışlarının daha kesin bir şekilde yapılmasına olanak sağlar. Ayrıca maliyet fiyatlarının doğru tahmini bütçe planlanması ve alternatif projelerin değerlendirilmesi için önem arz etmektedir.

Anahtar Sözcükler: Maliyet Endeksi, Tahmin, Zaman Serileri, Regresyon Analizi

To my wife,

to my family

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Asst. Prof. Dr. Rıfat Sönmez without whom I would not be able to complete my thesis, for his patience, guidance, support and tolerance at every step of this study. It has always been a privilege to work with him. I would like to thank to the examining committee members for their valuable comments.

For the provision of good times throughout my life, my father Mehmet Kibar, my mother Ruhat Kibar, and my brothers Alperen and Aykut Kibar, who have never left me alone, deserve special emphasis. I would like to express my appreciation to my family members for their endless love and efforts that encouraged me to realize my goals.

I wish to thank to my father in law Sırrı Çakır, my mother in law Hava Çakır and sister in law Bağdagül Çakır for their spiritual support.

I would like to express my special thanks to my friend Caner Anaç, for his unlimited physical and spiritual aids that made this study done. And to my friend Beliz Özorhon, for her remarkable support throughout my master life.

I should thank to all my friends, for their patience during this study and for their sincere and continuous love.

I wish to thank to my bosses Faruk İnsel and Rauf Akbaba, for their patience, guidance, support and tolerance on my academic careeer.

Finally, I would like to give my very special thanks to my wife, Sibel Kibar, as being the One that makes me who I am.

TABLE OF CONTENTS

ABSTRACT v
ÖZ vii
ACKNOWLEDGEMENTS x
LIST OF FIGURES xiv
LIST OF ABBREVIATION xv
CHAPTER 1
INTRODUCTION 1
CHAPTER 2
BACKGROUND AND LITERATURE REVIEW 4
2.1. Background 4
2.1.1. Construction Price Indices 4
2.1.1.1. General Description of Cost Indices 5
2.1.1.2. Building Cost Index of Turkey (BMI) 6
2.1.1.3. Building Cost Index of USA, ENR’s BCI 7
2.1.2. Forecasting Methods 8
2.1.2.1. Quantitative Methods 8
2.1.2.1.1. Regression Models 10
2.1.2.1.2. Neural Network Models 11
2.1.2.1.3. Time Series Models 11
2.1.2.1.3.1. Simple Average and Exponential Smoothing Models 14
2.1.2.1.3.2. ARIMA Models 17
2.1.2.2. Qualitative Methods 19
2.1.3. Measures of Accuracy 20
2.2. Previous Studies on Construction Cost index Forecast 22

CHAPTER 3
METHODOLOGY AND DATA ANALYSIS 27
3.1 Introduction 27
3.2 Study on Turkey, BMI 27
3.2.1 Regression Analysis 28
3.2.2 Time Series Analysis 36
3.2.3 Simple Averaging 43
3.3 Study on United States, BCI 45
3.3.1 Regression Analysis 46
3.3.2 Time Series Analysis 50
3.3.3 Simple Averaging 57
3.4 Discussion on Results 60
CHAPTER 4
SUMMARY AND CONCLUSIONS 63
REFERENCES 66
APPENDIX …………………………………………………………………...70

LIST OF TABLES

Table 3.1 Initial Regression Model Characteristics for BMI................................ 31
Table 3.2 Final Regression Model Characteristics for BMI................................. 31
Table 3.3 Regression Model Inputs and Test Period Predictions for BMI........... 33
Table 3.4 MAPE Values for Final Regression Model for BMI............................ 33
Table 3.5 MAPE values for Prediction Performance and Closeness of Fit for BMI 37
Table 3.6 ARIMA Model Results for Predict Period for BMI 38
Table 3.7 Winters Multiplicative Model Results for Predict Period for BMI 39
Table 3.8 Forecasts by *Dave* and *Pave* Models for BMI...................................... 43
Table 3.9 MAPE Values for *Dave* and *Pave* Models for BMI 44
Table 3.10 Differentiated and Percent Change Values for Actual BMI in Forecast Horizon 45
Table 3.11 Initial Regression Model Characteristics for BCI 47
Table 3.12 Final Regression Model Characteristics for BCI................................ 47
Table 3.13 Regression Model Inputs and Test Period Predictions for BCI.......... 48
Table 3.14 MAPE Values for Final Regression Model for BCI........................... 48
Table 3.15 MAPE values for Prediction Performance and Closeness of Fit........ 51
Table 3.16 ARIMA Model for Predict Period for BCI... 52
Table 3.17 Winters Additive Model for Predict Period for BCI 53
Table 3.18 Dave and Pave Model Forecasts... 58
Table 3.19 MAPE Values for Simple Averaging Models 58
Table 3.20 Comparison of Models According to Their MAPE Values for BMI . 60
Table 3.21 Comparison of Models According to Their MAPE Values for BCI .. 61

LIST OF FIGURES

Figure 3.1 Regression Model for BMI.. 35
Figure 3.2 ARIMA (0,2,2)(0,1,1)s No Intercept Model for BMI......................... 40
Figure 3.3 Winters Multiplicative Model for BMI.. 41
Figure 3.4 Comparison of Forecasts Between ARIMA and Winters Multiplicative Models for BMI 42
Figure 3.5 Regression Model for BCI 49
Figure 3.6 ARIMA (2,1,0) (0,1,1)s No Intercept Model for BCI......................... 54
Figure 3.7 Winters Additive Exponential Smooting Model for BCI.................... 55
Figure 3.8 Comparison of Forecasts Between ARIMA and Winters Additive Models for BCI 56
Figure 3.9 Simple Average Models for BCI 59

LIST OF ABBREVIATION

a	Intercept in Regression Analysis
α	Level Smoothing Weight
A_t	Actual Value
ANN	Artificial Neural Network
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
ARIMA(p,d,q) (p’,d’,q’)	p describes the AR part, d describes the integrated part and q describes the MA part, while p’, d’, q’ describes AR, I, and MA parts in seasonal terms respectively.
b	Slope of Line in Regression Analysis
BCI	Building Cost Index Published by ENR for United States
BMI	Building Cost Index for Turkey
c	Constant in ARIMA
CCI	Construction Cost Index Published by ENR for United States
CP	Construction Permits in Turkey
Dave	Simple Average Model with Differentiation
δ	Season Smoothing Weight
$\Delta_i =$	Percent of Change of Period
e_t	Random Forecast Error
ECU	European Currency Unit
EURO	Euro
ENR	Engineering News Record
EUROSTAT	Statistical Office of the European Community
EXR	Exchange Rate

f_1	Regression Model Result
f_2	Time Series Model Result
f	Constant in ARIMA
φ	Damped Trend Smoothing Weight
γ	Trend Smoothing Weight
HS	Housing Starts in US
$I_i =$	Index of Period I
l	Integration Coefficient for Regression and Time Series Models
$L(t)$	Level Parameter
m	Mean
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ME	Mean Error
MR	Mortgage Rates in US
MSE	Mean Square Error
n	Total number of observations
OECD	Organization for Economic Co-operation and Development
P_t	Predicted Value
Pave	Simple Average Model with Percent Change
P-Value	Significance Factor
R^2	Coefficient of Determination
RMSE	Root Mean Square Error
s	The Seasonal Length
s	Standard Deviation
$S(t)$	Season Parameter
SAS	Software Developed by SAS Institute

SPSS	Software Developed by SPSS Inc.
t	Time
$T(t)$	Trend Parameter
TCMB	Central Bank of the Republic of Turkey
TL	Turkish Lira
TURKSTAT	Turkish Statistical Institute
TÜFE	Consumer Price Index for Turkey
U	Theil U Inequality Coefficient
US	United States
USD	United States Dollar
ÜFE	Producer Price Index for Turkey
X	Independent Variable in Regression Analysis
XEU	European Currency Unit Symbol
Y	Dependent Variable in Regression Analysis
Y_t	Univariate Time Series Under Investigation
$Y'_t(k)$	Model-Estimated k-step Ahead Forecast at Time t for Series Y
YTL	New Turkish Lira

CHAPTER 1

INTRODUCTION

Building cost indices are widely used in construction industry to measure the rate of change of building costs as a combination of labor and material costs. The indices are originated to a base year in which the predetermined quantities of certain price elements add up to a constant value (generally to 100).

The indices are used for cost escalation procedure of certain purposes. Most of the tenders in Turkey require a qualification stage. The contractors have to fulfill the qualification criteria stated on tender documents. One of the most frequent qualification criteria is the past performance. The bidder should have been completed an equivalent or similar job with certain amount of the tendered project. The cost indices are used to convert the past project values to the current values for comparison of tender prices. Cost indices are also useful in reflecting the trend of price variations in construction projects. They are being used for the calculation of the price escalations in certain contract types in Turkish Construction Industry. In these contract types an additional clause exists to formulate the escalation on tender price. Escalation formula is dependent on certain price indicators such as inflation rates, TÜFE (Consumer Price Index for Turkey), ÜFE (Producer Price Index for Turkey); or construction cost indices such as BMI (Building Cost Index for Turkey). By this clause, impact of inflation on the project cost is compensated by the owner to some extend. In some contracts this escalation clause does not exist or exists without any escalation for the project. In these kinds of contracts the impact of inflation is totally reflected to the contractor. Foreign currencies like USD or EURO may be used in payments instead of YTL which compensate contractor’s loss to a limited extend under the assumption that economic conditions remain stable.

The forecasted values of cost indices are useful indicators of the actual project costs. Therefore the cost index forecasts are useful tools for estimating the target project cost at tender stage, which can improve the accuracy of contingency used in the bid price. The forecast is also necessary for the client to have an accurate estimate fot the budget.

Forecasting can be done rather by quantitative or by qualitative methods. Quantitative methods require quantitative data to produce the forecast. Statistical methods are generally introduced depending on the quality of quantitative data. If the data includes information on one or more of independent variables which are causally related to the forecasted variable then the causal methods are employed. Most common causal methods for predicting future indices are regression and neural network models. These methods first find the relation between dependent and independent data, then use this information to predict the future value of dependent variable with the given future values of independent variables. This procedure necessitates the prediction of future values of independent variables. If the quantitative data includes only the past values of forecasted variable, then statistical methods like time series and simple averaging tools are employed. Qualitative methods, so called judgmental methods are used when there is not enough quantitative data. This intuitive and subjective tool is generally used in conjunction with the quantitative methods for long term forecasts to catch the trends in future data.

The aim of this study is to develop time series models to forecast building cost indices in Turkey and United States. The models developed will be compared with regression analysis and simple averaging models in terms of predictive accuracy. The selected cost indices for the study are BMI for Turkey (building cost index published by TURKSTAT in quarterly basis), and BCI for USA (building cost index published by Engineering News Record, ENR, in monthly basis). Forecasting methods and the selected cost indices will be explained in the next sections in detail.

The general outline of the thesis is listed below.

Chapter 2: General definitions of cost indices, chosen indices for the study will be followed by the forecasting methods with the detailed explanations of regression, neural network, time series and simple averaging methods. The measures of accuracy descriptions will be followed by a literature survey which will demonstrate the past performed studies on the subject matter.

Chapter 3: The conducted analysis on forecasting of BMI and BCI cost indices will be explained in detail in this chapter. The chapter will contain 3 analyses for each index. Regression analysis, time series analysis and simple averaging method analysis will be done separately for BMI and BCI representing Turkish and United States construction industries.

Chapter 4: Summary of the conducted study will be given in this chapter. Moreover the discussions of the analyses and results will be done and final models will be selected according to their ability to forecast BMI and BCI indices. Conclusion will be done with the recommendations on future studies.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

2.1. Background

In this chapter a background of cost indices, regression, neural network and time serried models, and measures of accuracy will be given to emphasize the importance of cost indices and for better understanding of forecasting tools. Previous cost index studies related to construction industry will also be reviewed in this chapter.

2.1.1. Construction Price Indices

Construction price indices are used in measuring the rate of change of construction costs. As declared by Williams (1994), in construction, the ability to predict trends in prices, both short- and long term, can result in more accurate bids. Models of this type can potentially allow contractors to incorporate expected price fluctuations into their bidding strategy. The ability to predict variations in prices can result in reduced construction costs by allowing material purchases to be better timed, to take advantage of short-term fluctuations in material prices. Owner organizations may also benefit the prediction of changes in construction cost indexes. Improved budgeting decisions could be made if the trend in construction prices could be forecast accurately. This could allow improved analysis of the effects of delaying or accelerating large construction expenditures. (Williams, 1994)

2.1.1.1. General Description of Cost Indices

The first cost indices were developed by Carli in 1750 to determine the effects of the discovery of America on the purchasing power of money in Europe (Ostwald, 1992). It was an inflation index which is a common type of cost index widely used in whole world. Construction cost indices are other types of cost indices, which are valid for construction industry. According to Grogan (1994), Engineering News Record (ENR)'s CCI (Construction Cost Index) is the oldest inflation index currently used by engineers. The publish date of CCI was 1921 although it started in 1909. This index was designed as a general purpose tool to chart basic cost trends (Grogan, 1994).

A study conducted by the Statistics Directorate of the OECD (1994 (a) , 1996, 1994 (b)) and EUROSTAT (1995, 1996) states that the demand for adequate construction price indices arises from the need to assess real changes in the output from these activities which cannot be derived solely through reference to regular building and construction statistics. Construction price indices are used in guaranteed value clauses in rental, leasing, and other contracts; adjustment of sales contracts for buildings under construction; and as a basis for indexation for insurance purposes. They are also used to deflate national accounts estimates of output of construction activities, and gross fixed capital formation in residential construction. Construction price indices are calculated by the statistical directorates of countries (The Statistics Directorate of the OECD, 1994 (a), 1996, 1994 (b) & EUROSTAT, 1995, 1996).

Williams (1994) states that cost indices permit an estimator to forecast construction costs from the present to future periods without going through detailed costing. Because construction costs vary with time due to changes in demand, economic conditions, and prices, indexes convert costs applicable at a past date to equivalent costs now or in the future.

2.1.1.2. Building Cost Index of Turkey (BMI)

BMI is the building cost index for Turkey, published by Turkish Statistical Institute (TURKSTAT) on quarterly basis. The index covers all type of building structures namely, houses, apartments, shops, commercial buildings, medical buildings, schools, cultural buildings and administrative buildings. It covers more than 90 percent of construction activity of Turkish construction industry.

As noted in The Statistics Directorate of the OECD (1994 (a), 1996, 1994 (b)) and EUROSTAT (1995, 1996), the selection of items for inclusion in the index was made after extensive consultation with interested bodies, including the Finance and Industry Statistics Divisions within the TURKSTAT, the Chamber of Civil Engineers and of Architects, trade unions and a number of other institutions and associations. With the help of the Turkish Scientific and Technical Resource Institution and their publication Construction Unit Price Analysis, the items were selected and weights determined through detailed examination of bills of quantities for a sample of current projects representative in terms of regional distribution and project type of construction activity within the scope of the index. The index is calculated quarterly according to the Laspeyres formula and has base period 1991=100 (The Statistics Directorate of the OECD, 1994 (a), 1996, 1994 (b) & EUROSTAT, 1995, 1996).

Included in the index are costs of materials, labor and machinery. No taxes are included in the prices used in the calculation of the index, but the prices are net of discounts. Most of the cost data used is obtained through surveys of construction and other enterprises as well as from price lists. The data are collected from 24 provinces which have been chosen to represent all the regions of Turkey. In total, 295 items are priced from around 1.300 suppliers to construction firms (TURKSTAT, 2002).

2.1.1.3. Building Cost Index of USA, ENR's BCI

Among the wide variety of indices used in America ENR's CCI and BCI are the most common indices. Grogan (1994) stated that the engineering news record (ENR) index that started in 1909, is the oldest inflation index currently used by engineers. The publish date of Construction Cost Index CCI was 1921 although it started in 1909. The index was designed as a general purpose tool to chart basic cost trends. It remains today as a weighted aggregate index of the prices of constant quantities of structural steel, portland cement, lumber and common labor. This package of construction goods was valued at $100 using 1913 prices (Grogan, 2007).

The original use of common labor as a component of the CCI was intended to reflect wage rate activity for all construction workers. In the 1930s, however, wage and fringe benefit rates climbed much faster in percentage terms for common laborers than for skilled tradesmen. In response to this trend, ENR introduced its Building Cost Index, BCI in 1938 to weigh the impact of skilled labor trades on construction costs. (Grogan, 2007)

The drawbacks in using the BCI index is that it is not affected by crew productivity and does not explicitly consider the cost of equipment or management. Furthermore, if the project cost cannot be reasonably represented with the cost of cement, steel, and lumber, then the BCI may not be the best indicator of price variations. Despite these shortcomings, this index remains one of the best known and the most used indices in the industry today (Touran and Lopez, 2006).

According to Capano and Karshenas (2003) the BCI is more suitable to model cost of structures while CCI can be used to model projects where the labor cost is a high proportion of the total cost of the project.

The BCI is computed by combining 66.38 hour of skilled labor of bricklayers, carpenters, and structural ironworkers rates, 2500 pounds of standard structural steel shapes at the mill price prior to 1996 and the fabricated since 1996, 1.128 t of portland cement, and 1,088 board-ft of 2 x 4 lumber. The price of this combination was $100 in 1913.

2.1.2. Forecasting Methods

According to Touran and Lopez (2006) forecasting methods are used to produce numerical estimates of escalation, escalation factor, or cost escalation range from the relatively simple to complex and sophisticated techniques. Touran and Lopez (2006) also note that forecasting techniques are used to forecast one of three periods: (1) short term (next 3 months); (2) medium term (4 months–2 years); and (3) long term (more than 2 years). Estimating the increase in price over the long term is almost impossible because of the many uncertainties beyond the control of all parties (Westney, 1997). Touran and Lopez claims that the same is true of long-term construction projects with multiyear schedules and start dates in the future. Despite this difficulty, the owners of large long-term projects need to come up with the estimated cost of these projects. The more prudent way to approach these problems is to calculate a range of possible costs rather than a single figure (Touran and Lopez, 2006). Forecasting methods for escalation factors can be grouped into two major categories: (1) quantitative methods and (2) qualitative methods (Makridakis et al., 1998).

2.1.2.1. Quantitative Methods

Quantitative methods are used when sufficient quantitative information is available. Most of the forecasting techniques for escalation, escalation factor, and cost escalation are quantitative methods (Touran and Lopez, 1996). Taylor and Bowen (1987) suggests two quantitative forecasting categories, (1) the causal method and (2) the time-series method (statistical method). The causal method assumes that the predicted variable is controlled by one or more independent variables and the causal relationship is applied to predict the dependent variable (Wang and Mei, 1998). If sufficient accurate information is available on the future of the other variables (i.e. the independent variables), it can be used to predict the future value of the variable to be forecast (i.e. the dependent variable) (Kress, 1985). Runeson (1988) used an ordinary least-squares multiple regression method to establish a model for predicting building price. Koehn and Navvabi (1989) derived a multivariate linear cost formula for the construction cost function based on the interaction between economic and construction industry variables. Akintoye and Skitmore (1993) constructed a reduced-form simultaneous equation to predict construction tender price indices. Williams (1994) and Hanna and Chao (1994) employed neural network models which is also a causal method for predicting future cost indices.

Statistical methods utilize time-series analysis and curve fitting methods to forecast the variable in the future (Hanna and Blair, 1993). The time-series method was developed by Yule (1927) with his autoregression technique and Slutsky (1937) with his moving average technique. Brown (1970) improved the moving average technique and he further developed the exponential smoothing technique and various exponential smoothing models. Durbin (1970) improved the autoregression technique and derived the partial autocorrelation function from the estimation of the autocovariance and autocorrelation functions. Moreover, Box and Jenkins (1976) integrated the autoregression and moving average techniques and then developed a mixed model. Snyder (1982) and Kress (1985) regarded this mixed model as the best quantitative method for short term predictions.

2.1.2.1.1. Regression Models

Regression analysis is a forecasting tool in which the dependent variable is expressed in terms of the independent variables. The regression method's accuracy depends upon a consistent relationship with the independent variables (Touran and Lopez, 2006). Moreover according to Ng et al. (2000), regression models provide accurate predictions when price levels are steady, such that moving downward or upward. Linear regression models are the most common regression models which attempt to model the relationship between two variables by fitting a linear equation to observed data. A linear regression line has an equation of the form $Y=\alpha + \beta X$, where X is the explanatory variable (independent variable) and Y is the dependent variable. The slope of the line is β, and α is the intercept (the value of y when x = 0). Statistically speaking α can be divided into a constant and an error term in which the equation is expressed as follows:

$$Y = a' + b X + e$$

The error term e is assumed to be normally distributed with an expected value of 0. The two important indicators of regression models are the R^2 and the P-Value. The R^2 is the coefficient of determination which is the ratio of the regression sum of squares to the total sum of squares. It is an indicator of the fit of the explanatory variables to the dependent variable. The value of R^2 close to 1 indicates a good model with R^2 ranging from 0 to 1. Significance level, P-value is a test statistic designating the significance of the independent variables. Usually a P-value less than 0.1 designate a significant independent variable. Although the regression analysis are called in casual methods, it is also an statistical method, which contains statistical agents such as P-value and R^2.

2.1.2.1.2. Neural Network Models

Neural networks are part of the causal or explanatory methods (Touran and Lopez, 2006). Neural networks are fundamentally based on simple mathematical models of the way the human brain is believed to work (Makridakis et al., 1998).

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurones) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurones. This is true of ANNs as well. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyse. This expert can then be used to provide projections given new situations of interest and answer "what if" questions (Stergiou and Siganos, 2007).

2.1.2.1.3. Time Series Models

Statistics Glossary (Statistics Glossary, 2007) defines the time series as a sequence of observations which are ordered in time. Time series analysis accounts for the fact that data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for. (NIST/SEMATECH e-Handbook of Statistical Methods, 2007). According to

Touran and Lopez (2006) time series system uses the pattern in the historical data to extrapolate that pattern into the future, but it makes no attempt to discover the factors affecting the behavior. Makridakis et al. (1998) suggest that there are two main reasons to utilize a system as a black box. First, the system may not be understood, and even if it were understood it might be extremely complex to assess the relationships that govern its behavior. Second, the main objective of the system is not to know how it occurs but to forecast what will occur. Wang and Mei (1998) states that, since construction costs are tightly related to the labor and material costs, it is very difficult to cover the independent variables fully. From a practical viewpoint, the time-series method is easier to apply when appraising variations in future construction costs.

Statistics Glossary (Statistics Glossary, 2007) sets out the main features of time series as follows.

Trend Component

Trend is a long term movement in a time series. It is the underlying direction (an upward or downward tendency) and rate of change in a time series, when allowance has been made for the other components. A simple way of detecting trend in seasonal data is to take averages over a certain period. If these averages change with time we can say that there is evidence of a trend in the series. There are also more formal tests to enable detection of trend in time series.

Seasonal Component

The seasonal component, often referred to as seasonality, is the component of variation in a time series which is dependent on the time of year. It describes any regular fluctuations with a period of less than one year.

Cyclical Component

It is a non-seasonal component which varies in a recognizable cycle.

Irregular Component

The irregular component is that left over when the other components of the series (trend, seasonal and cyclical) have been accounted for.

Smoothing

Smoothing techniques are used to reduce irregularities (random fluctuations) in time series data. They provide a clearer view of the true underlying behaviour of the series.

In some time series, seasonal variation is so strong it obscures any trends or cycles which are very important for the understanding of the process being observed. Smoothing can remove seasonality and makes long term fluctuations in the series stand out more clearly. The most common type of smoothing technique is moving average smoothing although others do exist. Since the type of seasonality will vary from series to series, so must the type of smoothing.

Exponential Smoothing

Exponential smoothing is a smoothing technique used to reduce irregularities (random fluctuations) in time series data, thus providing a clearer view of the true underlying behavior of the series. It also provides an effective means of predicting future values of the time series (forecasting). Exponential Smoothing assigns exponentially decreasing weights as the observation get older.

Moving Average Smoothing

A moving average is a form of average which has been adjusted to allow for seasonal or cyclical components of a time series. Moving average smoothing is a smoothing technique used to make the long term trends of a time series clearer. When a variable, like the number of unemployed, or the cost of strawberries, is graphed against time, there are likely to be considerable seasonal or cyclical components in the variation. These may make it difficult to see the underlying trend. These components can be eliminated by taking a suitable moving average.

By reducing random fluctuations, moving average smoothing makes long term trends clearer.

Differencing

Differencing is a popular and effective method of removing trend from a time series. This provides a clearer view of the true underlying behavior of the series.

Autocorrelation

Autocorrelation is the correlation (relationship) between members of a time series of observations, such as weekly share prices or interest rates, and the same values at a fixed time interval later. More technically, autocorrelation occurs when residual error terms from observations of the same variable at different times are correlated (related). (Statistics Glossary, 2007)

2.1.2.1.3.1. Simple Average and Exponential Smoothing Models

Simple average method as the name implies is basically taking the average of data. The simple average is suitable for data that fluctuate around a constant or have a slowly changing level and do not have a trend or seasonal effects (Touran and Lopez, 2006). The fundamental principle of the exponential smoothing is that the values of the variable in the latest periods have more impact on the forecast and therefore should be given more weight (Kress, 1985). This method implies that as historical data get older, their weight will decrease exponentially (Touran and Lopez, 2006).

Formulations on some exponential smoothing models are given below as taken from the SPSS program help file.

Notation for Time Series models:

Y_t (t=1, 2, ..., n)	Univariate time series under investigation.
n	Total number of observations.
$Y'_t(k)$	Model-estimated k-step ahead forecast at time t for series Y.
s	The seasonal length.

Notation Specific to Exponential Smoothing Models

α	Level smoothing weight
γ	Trend smoothing weight
φ	Damped trend smoothing weight
δ	Season smoothing weight

Simple Exponential Smoothing

Simple exponential smoothing has a single level parameter and can be described by the following equations:

$$L(t) = aY(t) + (1-a)L(t-1)$$

$$Y'_t(k) = L(t)$$

It is functionally equivalent to an ARIMA(0,1,1) process.

Brown's Exponential Smoothing

Brown's exponential smoothing has level and trend parameters and can be described by the following equations:

$$L(t) = aY(t) + (1-a)L(t-1)$$

$$T(t) = a(L(t) - L(t-1)) + (1-a)T(t-1)$$

$$Y'_t(k) = L(t) + ((k-1) + a^{-1})T(t)$$

It is functionally equivalent to an ARIMA(0,2,2) with restriction among MA parameters.

Holt's Exponential Smoothing

Holt's exponential smoothing has level and trend parameters and can be described by the following equations:

$$L(t) = aY(t) + (1-a)(L(t-1) + T(t-1))$$

$$T(t) = g(L(t) - L(t-1)) + (1-g)T(t-1)$$

$$Y'_t(k) = L(t) + kT(t)$$

It is functionally equivalent to an ARIMA(0,2,2).

Damped-Trend Exponential Smoothing

Damped-Trend exponential smoothing has level and damped trend parameters and can be described by the following equations:

$$L(t) = aY(t) + (1-a)(L(t-1) + jT(t-1))$$

$$T(t) = g(L(t) - L(t-1)) + (1-g)jT(t-1)$$

$$Y'_t(k) = L(t) + \sum_{i=1}^{k} j^i T(t)$$

It is functionally equivalent to an ARIMA(1,1,2).

Simple Seasonal Exponential Smoothing

Simple seasonal exponential smoothing has level and season parameters and can be described by the following equations:

$$L(t) = a(Y(t) - S(t-s)) + (1-a)L(t-1)$$

$$S(t) = d(Y(t) - L(t)) + (1-d)S(t-s)$$

$$Y'_t(k) = L(t) + S(t+k-s)$$

It is functionally equivalent to an ARIMA(0,1,(1,s,s+1))(0,1,0) with restrictions among MA parameters.

Winters' Additive Exponential Smoothing

Winters' additive exponential smoothing has level, trend and season parameters and can be described by the following equations:

$$L(t) = a(Y(t) - S(t-s)) + (1-a)(L(t-1) + T(t-1))$$

$$T(t) = g(L(t) - L(t-1)) + (1-g)T(t-1)$$

$$S(t) = d(Y(t) - L(t)) + (1-d)S(t-s)$$

$$Y'_t(k) = L(t) + kT(t) + S(t+k-s)$$

It is functionally equivalent to an ARIMA(0,1,s+1)(0,1,0) with restrictions among MA parameters.

Winters' Multiplicative Exponential Smoothing

Winters' multiplicative exponential smoothing has level, trend and season parameters and can be described by the following equations:

$$L(t) = a(Y(t)/S(t-s)) + (1-a)(L(t-1) + T(t-1))$$

$$T(t) = g(L(t) - L(t-1)) + (1-g)T(t-1)$$

$$S(t) = d(Y(t)/L(t)) + (1-d)S(t-s)$$

$$Y'_t(k) = (L(t) + kT(t))S(t+k-s)$$

There is no equivalent ARIMA model.

2.1.2.1.3.2. ARIMA Models

ARIMA processes are mathematical models used for forecasting. ARIMA is an acronym for Auto Regressive, Integrated, Moving Average. Each of these phrases describes a different part of the mathematical model.

ARIMA processes have been studied extensively and are a major part of time series analysis. They were popularized by George Box and Gwilym Jenkins in the early 1970s; as a result, ARIMA processes are sometimes known as Box-Jenkins models. Box and Jenkins (1970) effectively put together in a comprehensive manner the relevant information required to understand and use ARIMA processes.

The ARIMA approach to forecasting is based on the following ideas:

1) The forecasts are based on linear functions of the sample observations;
2) The aim is to find the simplest models that provide an adequate description of the observed data. This is sometimes known as the principle of parsimony.

Each ARIMA process has three parts: the autoregressive (or AR) part; the integrated (or I) part; and the moving average (or MA) part. The models are often written in shorthand as ARIMA(p,d,q) where p describes the AR part, d describes the integrated part and q describes the MA part.

AR: This part of the model describes how each observation is a function of the previous p observations. For example, if p = 1, then each observation is a function of only one previous observation. That is,

$$Y_t = c + f Y_{t-1} + e_t$$

where Y_t represents the observed value at time t, Y_{t-1} represents the previous observed value at time t − 1, e_t represents some random error and c and f are both constants. Other observed values of the series can be included in the right-hand side of the equation if p > 1:

$$Y_t = c + f_1 Y_{t-1} + f_1 Y_{t-2} + ... + f_p Y_{t-p} + e_t$$

I: This part of the model determines whether the observed values are modeled directly, or whether the differences between consecutive observations are modeled instead. If d = 0, the observations are modeled directly. If d = 1, the differences between consecutive observations are modeled. If d = 2, the differences of the differences are modeled. In practice, d is rarely more than 2.

MA: This part of the model describes how each observation is a function of the previous q errors. For example, if q = 1, then each observation is a function of only one previous error. That is,

$$Y_t = c + f_1 e_{t-1} + e_t$$

Here e_t represents the random error at time t and e_{t-1} represents the previous random error at time t − 1. Other errors can be included in the right-hand side of the equation if q > 1.

Combining these three parts gives the diverse range of ARIMA models.

There are also ARIMA processes designed to handle seasonal time series, and vector ARIMA processes designed to model multivariate time series. Other variations allow the inclusion of explanatory variables.

ARIMA processes have been a popular method of forecasting because they have a well-developed mathematical structure from which it is possible to calculate various model features such as prediction intervals. These are a very important feature of forecasting as they enable forecast uncertainty to be quantified. (Box and Jenkins (1970), Makridakis et al. (1998), Pankratz (1983), cited Hyndman (2001)

2.1.2.2. Qualitative Methods

Qualitative forecasting methods, in contrast with quantitative methods, do not require data in the same way (Touran and Lopez, 2006). The inputs required depend on the specific method and are in essence the product of judgment and accumulated knowledge (Blair et al., 1993; Hanna and Blair, 1993; and Makridakis et al., 1998). They can be used separately but are more often used in conjunction with quantitative methods. Qualitative methods are also called subjective methods (Blair et al., 1993) and judgmental methods (Kress, 1985). Blair et al. (1993) recommend the use of qualitative methods in long term forecast (forecast of duration over 2 years) because statistical methods, in general, are not suitable for it; statistical methods cannot predict a shift in the trend. Although forecaster's intuition may frequently prove to be more reliable than any mathematical method (Chatfield, 1975), it would be difficult to calculate a confidence level for the forecast. Subjective and intuitive estimates are widely used in construction estimating, especially when there is insufficient historical data. (Touran and Lopez, 2006)

2.1.3. Measures of Accuracy

Fitzgerald and Akintoye (1995) suggest that the quality of the forecasts produced by organizations can be assessed with the use of quantitative methods which measure statistical error; in essence, determine the magnitude of the forecast error e_t. The measures of forecast accuracy compare the predicted values with those that were observed as shown below:

$$e_t = P_t - A_t$$

where e_t is the forecast error, A_t is the actual value and P_t is the predicted value. (Fitzgerald and Akintoye, 1995)

Makridakis and Hibon (1984) have identified the most common measures of accuracy as the mean square error (MSE), Theil's U-coefficient and the mean absolute percentage error (MAPE). Other measures of accuracy include root mean square error (RMSE), mean error (ME), mean absolute error (MAE) and graphical presentation (Holden and Peel, 1988; Treham, 1989). All of these except graphical presentation are regarded as non-parametric measures of forecast accuracy (Fitzgerald and Akintoye, 1995).

Fitzgerald and Akintoye (1995) define the error types as follows:

Mean Error (ME)

This measures the presence of bias in forecasts rather than the precision of estimates. It is an arithmetic mean of forecast errors which permits negative and positive error to offset one another. This is represented as follows:

$$ME = \sum_{t=1}^{n} (e_t) / n$$

Mean Absolute Error (MAE)

This is a better measure of forecast precision; it ignores the signs of the forecast error and considers only the absolute magnitude. This is presented as follows:

$$MAE = \sum_{t=1}^{n} |e_t| / n$$

Mean Square Error (MSE)

The most frequently employed measures of forecast accuracy are based on the mean squared error of forecast. Like MAE, the mean squared error measures the magnitude of forecast errors. It can be used in two different ways: to aid in the process of selecting a forecasting model and to monitor a forecasting system in order to detect when something has gone wrong with the system. However, MSE penalizes a forecasting technique much more for large errors than for small errors.

$$MSE = \sum_{t=1}^{n} e_t^2 / n$$

Root Mean Square Error (RMSE)

This is calculated by taking the square root of MSE. This is, by mathematical necessity, always greater than the MAE when the forecast errors are not all of the same size. This can be expressed as a percentage of the mean of the actual values of the variable and interpreted as percentage error (RMSE%).

Theil U Inequality Coefficient

This is

$$U = \sqrt{\frac{(1/n)\sum_{t=1}^{n}(e_t)^2}{(1/n)\sum_{t=1}^{n}(A_t)^2}}$$

U attains its smallest value when forecasts are perfect and is in most cases confined to the closed interval between zero and unity (Theil, 1978). The advantage of *U* over RMSE and MSE is that its denominator acts as a scaling factor to take account of the size of the variables to be predicted. This method, which weighs error relative to the actual movements of the predicted variable, produces the most appropriate way to standardize for differences between either time intervals or variables with different base years (McNees and Ries, 1983). This advantage makes *U* more useful for comparing forecast accuracy across different forecast spans or horizons.

Mean Absolute Percentage Error (MAPE)

This is calculated as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}|e_t| / A_t (100)$$

Fitzgerald and Akintoye (1995)

2.2. Previous Studies on Construction Cost index Forecast

Forecast of cost indices have been done ever since the cost indices introduced. Many studies have been conducted for this purpose using different kinds of forecasting tools such as regression analysis, neural networks, simulation and time series analysis. The purpose of this chapter is to present information about previous studies regarding the forecasting tools.

Touran and Lopez (2006) studied on cost escalation modeling in large infrastructure projects. According to Touran and Lopez (2006) budgeting for cost escalation is a major issue in the planning phase of large infrastructure projects. A system was introduced by the authors for modeling the escalation uncertainty in large multiyear construction projects. The system uses a Monte Carlo simulation approach and considers variability of project component durations and the uncertainty of escalation factor during the project lifetime and calculates the distribution for the cost. It was claimed that the system could be used by planners and cost estimators for the budgeting effect of cost escalation in large projects with multiyear schedules.

Touran and Lopez (2006) introduced an escalation factor as the rate of change of the BCI from year to year using the equation below.

$$\Delta_i = \left[(I_i / I_{i-1}) - 1 \right] \times 100\%$$

Where D_i = percent of change of period i, I_i = index of period I, and I_{i-1} = index of the previous period (i-1). A positive value of D_i is an indication of increase in cost. In contrast, if the value of D_i is negative, that is because period i has experienced a deflation (Touran and Lopez, 2006). Therefore D_i was defined as the escalation factor that was tried to model. In the paper an analysis on the historical trend in BCI values was given, designating rate of change of index low and high inflation periods in United States. Touran and Lopez (2006) proposed to use a normal distribution representing the escalation factor in order to model the uncertainty in the value of index. Within this uncertainty it was argued that there should be a somewhat relation between that years escalation factor with the preceding years. To examine the hypothesis the correlation coefficient between successive indices was found as 0.9828.

In the light of these findings Touran and Lopez (2006) constructed their model as follows. First the mean (μ) and the standard deviation (σ) of the normal distribution was defined for the escalation rate. For every iteration of the simulation, a random value for inflation was generated for the first period. In the subsequent periods, the generated values for the previous period was to serve as the mean of the normal distribution used to model the inflation rate assuming the same standard deviation. As is claimed this was done to give a higher weight to the value of escalation in the period immediately before the period of interest. Touran and Lopez (2006) have also argued that the proposed approach was more or less similar to the method of simple average but incorporating the random variability of the escalation factor. It was also stated that the Estimation of (μ) and (σ) for the normal distribution can be related to the project duration. Such that a long horizon of time series should be chosen if the project duration is long, and vice versa is true. In the conclusion part it was concluded that the proposed model for cost escalation can provide a powerful tool to assess the impact of the escalation factor.

Williams (1994) investigated the usage of back-propagation neural-networks in cost index prediction. Williams (1994) has constructed two neural-network models one for predicting one-month change other for predicting six-month change for ENR construction index. Selected input variables for the study were as follows:

Percentage change in the construction cost index for one month

- The six-month percentage change in the construction cost index
- The prime lending rate
- The six-month percentage change in the prime lending rate
- The six-month change in the prime rate
- Number of housing starts for the month
- Percentage change in housing starts for one month
- Percentage change in housing starts for the preceding six-month period
- The month of the year

The reasoning for the selected variables was also made by Williams (1994), including the prime lending rate and housing starts. Williams (1994) points out the complexity of the relationship between construction prices and the prime lending rate. It was stated that an increase in interest rates raises the cost of capital projects whereas it reduces the same for industrial and commercial building. To reflect the level of activity in the construction industry housing starts were included in the model. It is suggested by Williams (1994) that in periods of high housing starts, with high demand for construction materials, it would be expected that increases in the cost index would be higher. In development of neural-network model, Williams (1994) have made necessary data transformations dictated by the neural network program such as changing the cost index data into a percentage change form. Williams (1994) have also suggested normalizing the input data namely the percentage change in the index, percentage change in housing starts and housing starts, to obtain better results. The neural-network model was constructed using the Neuroshell program (from Ward Systems Group, Inc., Frederick, Md.) which implements a three-layer back-propagation model. As a result of the constructed model Williams (1994) have stated that the neural networks are being produced poor predictions of the changes in the construction cost index. In the one-month model, the difference between the actual percentage change and the predicted percentage change was less than or equal to 0.25 in only 20 of 63 cases. For the six-month model, the difference between the actual percentage change and the predicted change was less than or equal to 0.5 for 20 of 66 cases (Williams, 1994). The author has also constructed an exponential smoothing and a linear regression model for comparison. According to the modeling studies the sum of the squares of errors (SSE) was found for exponential smoothing, linear regression and neural-network as 2.45, 2.65 and 5.31 respectively. Williams (1994) claims the overall accuracy of exponential smoothing and regression techniques of being higher than the neural-network model although they are unable to react to large variations in the index.

Ng et al. (2004) produced a study for integrating regression analysis and time series. in the study TPI (Tender Price index for Honk Kong) was chosen as the basis of interest. The independent variables of regression analysis were chosen from a previous study of Ng et al. (2000) and correlation analysis was conducted before starting the regression analysis. In regression analysis part an automated stepwise procedure has been followed with multivariate regression. Time series model has been decided to be an ARIMA model considering the simple exponential smoothing as an inadequate model for TPI prediction, and MA(2) model has been found to be the most suitable for the data. The regression and time series models has then been integrated by linear combinations by considering the forecast made by RA and TS as f_1 and f_2 respectively. From this, a new forecast of these quantities has been produced by:

$$f_3 = l \cdot f_1 + (1 - l) \cdot f_2$$

Where l is the weighting which was restricted to the range (0 – 1). Goodness-of-fit statistics has been used as assistance in assessing the fit of a model (Ng et al. 2004). Ng et al. (2004) have followed an iterative procedure to find λ. They have performed Back-cast testing to examine the forecast accuracy. According to Ng et al. (2004) the results of back-cast testing was confirmed that the integrated RA-TS model outperforms both the individual RA or TS forecasts, therefore the integrated model should have a high potential of improving the forecasting accuracy of TPI movement even under a rapidly changing environment.

CHAPTER 3

METHODOLOGY AND DATA ANALYSIS

3.1 Introduction

The aim of this study is to use time series analysis to improve the accuracy of construction cost index predictions. Time series models will be developed for cost indices in Turkey and United States. The selected cost indices for Turkey and United States are BMI (building cost index, published by TURKSTAT in quarterly basis), and BCI (building cost index, published by Engineering News Record, ENR, in monthly basis) respectively. A literature survey was conducted to investigate the forecasting methods and comparison techniques for these methods. As a result of this survey three main methods are determined. One is the regression analysis which is a causal method that uses independent variables for prediction of the dependent variable. The others are time series analysis and simple averaging methods which are both statistical methods that are using the historical data to achieve the future estimates. The results of time series models developed will be compared with the results of regression and simple averaging methods.

3.2 Study on Turkey, BMI

Kahraman (2005) conducted a study to compare the accuracy of construction cost indices used in Turkey. The aim of the study was to compare the cost indices (existing cost indices as well as alternative cost indices introduced by Kahraman) in terms of their adequacy for the representation of variations in the building costs in Turkey. The adequacy of the indices in representing the building costs were examined using regression analysis in which the subject price indices were used as independent variables and unit cost of the projects as dependent variables. Separate single variable regression models were formed for each index for comparison. The prediction performances of the constructed regression analyses were compared and two indices were selected as the most influential. One was the $PBPI_4$, a new index produced by Kahraman (2005), and the other was BMI, the building price index published by TURKSTAT; with the MAPE values 34.333 and 35.637 respectively.

As a conclusion to Kahraman's study, BMI can be used adequately for representing building cost variations in Turkey. Therefore in this study BMI is used for representing Turkish construction cost indices.

The BMI data was obtained from TURKSTAT covers an interval from the first quarter of 1991 to the fourth quarter of 2005. More current data was not be able to obtained from the institute. 60 sequential records are divided into two groups. First group, which consists of 52 records, represents the analysis data that have been used to form the prediction model. The second group of 8 records is used to test the prediction accuracy of the constructed model.

In the following subsections, the analysis results of three selected methods are given.

3.2.1 Regression Analysis

Parsimonious models are considered to be used for regression analysis. These models fit the data adequately without using any inconsistent variable that creates noise. P-values as explained earlier are used to eliminate the unnecessary variables, while the R^2 is used for the determination of best fit. Moreover the variables with illogical coefficients are also eliminated (i.e. minus signed ones).

Two important factors are considered as independent variable in regression analysis. First one is the number of construction permits (CP) given in that month in whole Turkey. The monthly data obtained from TURKSTAT is converted to the quarterly data by taking the average of three subsequent monthly data. Foreign currency exchange rates is considered as the second important factor that has influence on building costs in Turkey. US Dollar and Euro can be considered as the most significant currencies among all. A basket of US Dollar and Euro (i.e. EXR = 1 USD + 1 Euro) is used for this purpose. US Dollar exchange rates with Turkish Lira (TL) or New Turkish Lira (YTL) (after January 2005) are obtained from the official web site of Central Bank of the Republic of Turkey,TCMB, (TCMB, 2007) Also the exchange rates of Euro after January 1999 (Adaptation date of Euro) are obtained from this web site. The values from 1991 to 1999 are calculated from the cross rates between USD and ECU (European Currency Unit) and obtained from answers.com web site (Answers.com, 2007).

As defined in Sauder School of Business web site (Sauder School of Business, 2007) the European currency unit, ECU (XEU as the symbol), was an artificial "basket" currency that was used by the member states of the European Union as their internal accounting unit. The ECU was conceived on 13th March 1979 by the European Economic Community, the predecessor of the European Union, as a unit of account for the currency area called the European Monetary System. The ECU was also the precursor of the new single European currency, the Euro, which was introduced in 1999. Euro was replaced with ECU at par (that is, at a 1:1 ratio) on January 1, 1999 (Sauder School of Business, 2007).

The exchange rates obtained are the selling rates taken at the mid (i.e. 15^{th} day) of each month. After obtaining all the monthly data for Euro (ECU equivalent or actual) and USD values are added up to obtain the basket of exchange rates (EXR). As in the case of monthly construction permits data, the monthly exchange rate basket values are converted into quarterly format using averaging of three subsequent months.

Table A1 in Appendix shows the quarterly data for BMI, CP, and EXR from 1991-1st to 2005-4th quarters.

The time dependencies of the data are considered to add to regression models like time series models. This is achieved by using the previous time series of independent variables CP and EXR. This dependency is limited to four quarters representing the effects of the selected variables to BMI in one year period. Inclusion of time dependencies converts the regression model into a somewhat multivariate time series model, because the independency between the variables is one of the main assumptions of regular regression models. A stepwise procedure is followed for determining the parsimonious regression model. First model is constructed including four time series data of both variables namely, $CP_{(t-1)}$, $CP_{(t-2)}$, $CP_{(t-3)}$, $CP_{(t-4)}$, $EXR_{(t-1)}$, $EXR_{(t-2)}$, $EXR_{(t-3)}$ and $EXR_{(t-4)}$. The insignificant and irrational variables are then eliminated examining P-values and regression coefficients. The variable with P-value greater than 0.15 and regression coefficient with negative value is eliminated. A new regression model is constructed using the remaining variables of the previous elimination. The procedure is repeated until a model having all significant variables with positive coefficients is obtained.

As a result of the stepwise iteration the remaining variables of adequate regression model are found as $EXR_{(t-1)}$ and $EXR_{(t-4)}$. Construction permits turn out to be insignificant in BMI forecasting. Table 3.1 and Table 3.2 represent the characteristics of the initial and the final regression models respectively.

Table 3.1 Initial Regression Model Characteristics for BMI

Dependent Variable	Independent Variables	Coefficients	R^2	P-Value
BMI	$EXR_{(t-1)}$	7,095.3	0.9672	2.871E-16
	$EXR_{(t-2)}$	-312.6		0.8932
	$EXR_{(t-3)}$	1,565.6		0.5031
	$EXR_{(t-4)}$	2,412.9		0.1270
	$CP_{(t-1)}$	-0.0061		0.9236
	$CP_{(t-2)}$	-0,0211		0.7438
	$CP_{(t-3)}$	0.0701		0.2786
	$CP_{(t-4)}$	0.0129		0.8432

Table 3.2 Final Regression Model Characteristics for BMI

Dependent Variable	Independent Variables	Coefficients	Lower 95%	Upper 95%	R^2	P-Value
BMI	$EXR_{(t-1)}$	7,571.6	6,315.7	8,827.4	0.9920	6.109E-16
	$EXR_{(t-4)}$	3,355.2	1,841.5	4,868.9		5.216E-05

Prediction using regression models necessitates the forecast of the independent variables. Therefore the regression process becomes a two phase prediction. At the first phase, independent variables are predicted. In the second phase, dependent variable is predicted with the aid of prior defined independent variables.

For the constructed parsimonious regression model, the forecast of BMI is dependent on only one variable EXR. Considering the time dependency of EXR data the prediction of EXR values is decided to be performed using time series analysis. The quarterly EXR records are given as an input to time series modeling software SAS. The best fit result is taken from ARIMA (0,1,1)(1,0,0) model. The closeness of fit and prediction performances for the model is measured by MAPE as 9.98% and 4.98% respectively.

The second phase of regression analysis is performed with the predicted EXR values obtained from phase one. The regression model given in Table 3.2 is constructed as follows to forecast the future values of BMI.

$$BMI_t = 7571.6 \times EXR_{t-1} + 3355.2 \times EXR_{t-4}$$

The predicted BMI values are calculated from the above formula for both fit period and test period. For the test period two predictions are made, one is the predictions made using the actual EXR values in forecast horizon; other is the predictions made using the predicted EXR values from ARIMA model. Regression model inputs and predictions for fit period are given in Table A2 in Appendix. Regression model inputs and predictions for test period are given in Table 3.3. Closeness of fit and prediction performance mean square error (MAPE) terms are given in Table 3.4.

Table 3.3 Regression Model Inputs and Test Period Predictions for BMI

PREDICTION WITH ACTUAL EXR VALUES						
Quarter	Actual BMI	$EXR_{(t-1)}$	$EXR_{(t-4)}$	Predict	Upper 95%	Lower 95%
2004-1	36,775.4	31.697	34.285	35,502.8	44,673.3	26,332.4
2004-2	38,279.0	30.071	32.429	33,649.0	42,334.3	24,963.6
2004-3	39,695.2	32.090	29.670	34,252.0	42,773.2	25,730.7
2004-4	40,439.9	32.873	31.697	35,524.9	44,451.3	26,598.5
2005-1	41,482.1	33.225	30.071	35,245.9	43,970.4	26,521.4
2005-2	42,162.8	30.733	32.090	34,036.5	42,753.6	25,319.3
2005-3	43,332.3	30.878	32.873	34,409.0	43,262.8	25,555.1
2005-4	43,770.3	29.792	33.225	33,704.8	42,475.6	24,934.0
PREDICTION WITH PREDICTED EXR VALUES						
Quarter	Actual BMI	$EXR_{(t-1)}$	$EXR_{(t-4)}$	Predict	Upper 95%	Lower 95%
2004-1	36,775.4	3.1697	3.4285	35,502.8	44,673.3	26,332.4
2004-2	38,279.0	3.2528	3.2429	35,509.3	44,503.2	26,515.4
2004-3	39,695.2	3.2859	2.9670	34,834.2	43,452.1	26,216.4
2004-4	40,439.9	3.3350	3.1697	35,886.1	44,872.4	26,899.8
2005-1	41,482.1	3.2989	3.2528	35,891.6	44,958.3	26,824.8
2005-2	42,162.8	3.2841	3.2859	35,890.6	44,988.9	26,792.3
2005-3	43,332.3	3.2782	3.3350	36,010.6	45,175.8	26,845.4
2005-4	43,770.3	3.2695	3.2989	35,823.6	44,923.3	26,724.0

Table 3.4 MAPE Values for Final Regression Model for BMI

Closeness of Fit	Prediction Performance Using Actual EXR Records	Prediction Performance Using Predicted EXR Values
14.4%	14.9%	12.2%

The predictions of the regression model is started at 4 periods beyond 1992 1st period, therefore the closeness of fit value is calculated for 48 periods. For the 48 fit periods the actual BMI values falls beyond the 95% confidence limits for 13 periods. In other words, the level of accuracy of the fit model is only 73%. Prediction performance with actual EXR data also show a depleted accuracy of value 80% (6 of 8 within limits), while the prediction performance with predicted EXR data shows 100% accuracy within the 95% confidence limits.

Examining Table 3.4 it is seen that the prediction performance of the model used predicted EXR values is better than that of used actual values. The error on prediction of EXR values had appositive effect on regression model accuracy. This should not be considered as a success of two phase prediction model. The error of the first phase prediction does not always deduct the error of the second phase.

Figure 3.1 represents the regression model predictions. This figure includes actual BMI (Actual), predicted BMI (Predicted), upper and lower confidence limits (Upper 95% and Lower 95%). It also includes predicted BMI using predicted EXR (Pred-Predict), upper and lower confidence limits using predicted EXR (Pred-Upper, Pred-Lower). Moreover $EXR_{(t-1)}$ and $EXR_{(t-4)}$ are also included in to figure for sake of visual comparison. EXR1* and EXR4* data as seen in Figure 3.1 is obtained by multiplication of actual $EXR_{(t-1)}$ and $EXR_{(t-4)}$ values with coefficients of lower 95% (6315 and 1841 respectively). This adjustment is performed only for visualization of the effects of EXR values on predicted data. Examining the figure it is seen that the predicted data does not provide a good fit on the actual data, especially in late periods. With the start of low inflation periods in Turkey, the exchange rates fluctuate on a stationary mean. But the building prices did not followed this trend; they continue their linear trend also on low inflation periods. Therefore the adequacy of predicting BMI values with regression analysis on exchange rates drop considerably with the loss of correlation between the two variables.

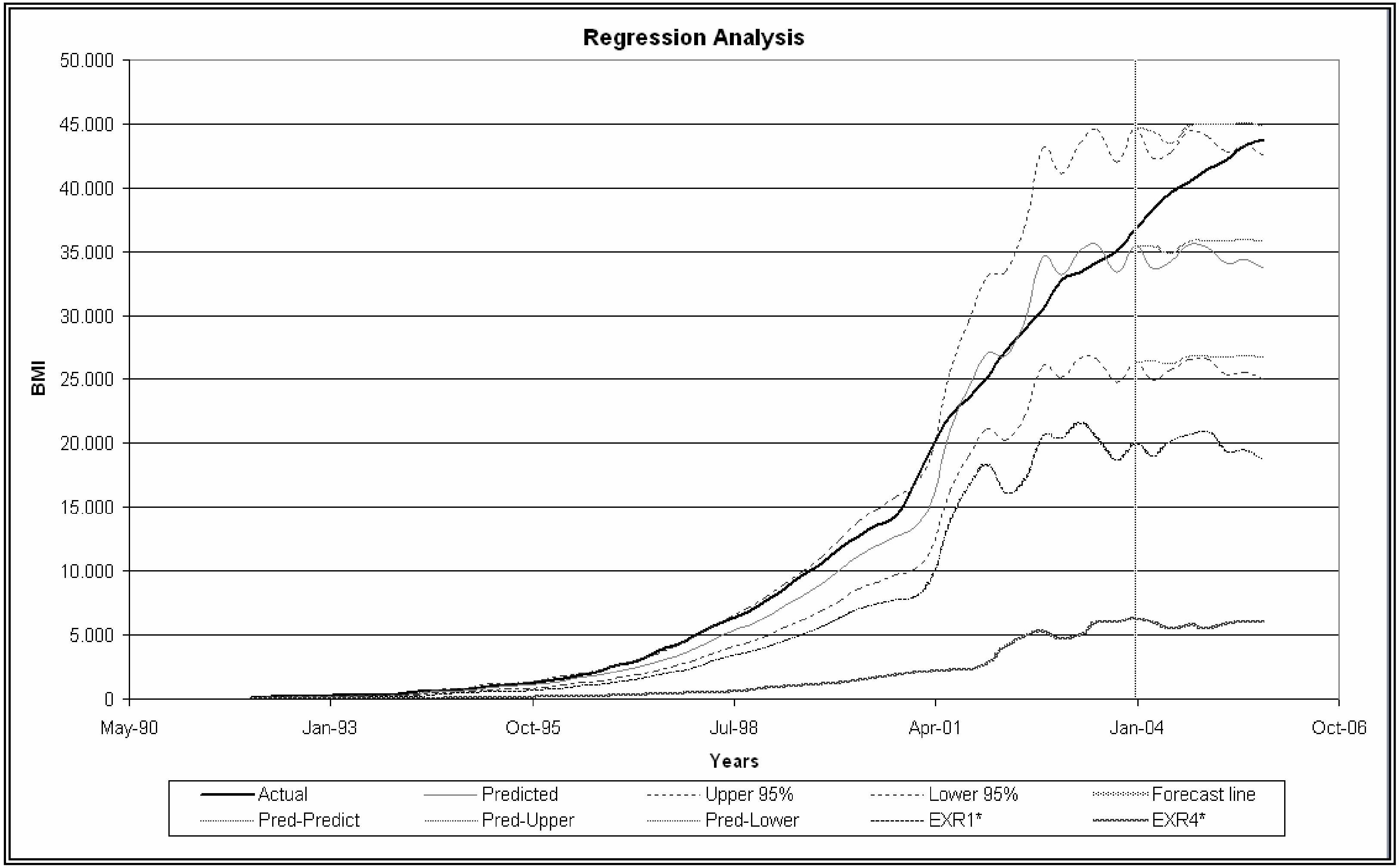

Figure 3.1 Regression Model for BMI

3.2.2 Time Series Analysis

Time series analysis is performed using the SAS software, developed by SAS Institute Inc. SAS is a useful tool for time series analysis and other statistical analyses. The main advantage of the software is that it applies an automatic fit procedure in which it finds the best fit model among the selected given models according to the given error criterion. Moreover the program uses powerful optimization tools for finding model parameters. By default the program performs 42 analyses for 42 time series models and lists the models in sorted order according to their closeness of fit error values (Default model list can be found in Table A3 in Appendix). The software also gives lower and upper limit forecasts for an adjustable confidence limit.

Entering the quarterly BMI records from 1991 1st to 2003 4th as input variable while setting MAPE as the error criterion and 95% as the confidence limit, SAS performed the time series analysis using the default models and finds the according MAPE values. As a result of processing; 6 models could not be fitted to the data. Closeness of fit MAPE values for remaining 36 models was found with the given distribution below:

20 models ranging from 2.9 to 4.6 percent

8 models ranging from 10 to 50 percent

4 models ranging from 100 to 900 percent

4 models ranging from 1500 to 3500 percent

Prediction performances are calculated manually for the first 20 models which have relatively low error values. Prediction performance MAPE values for these models are distributed as follows:

6 Models ranging from 1.2 to 1.7 percent

6 Models ranging from 3.3 to 5.2 percent

4 Models ranging from 6.7 to 15 percent

4 Models ranging from 22 to 52 percent

Table 3.5 shows the 6 time series models having the prediction performance MAPE values between 1.2 to 1.7 while having the closeness of fit MAPE values between 3.1 to 4.0.

Table 3.5 MAPE values for Prediction Performance and Closeness of Fit for BMI

No	Time Series Model	Prediction Performance	Closeness of Fit
1	ARIMA (0,2,2) (0,1,1)s No Intercept	1.20%	3.06%
2	Winters Multiplicative Exponential Smoothing	1.24%	3.51%
3	ARIMA (2,1,2) (0,1,1)s No Intercept	1.36%	3.29%
4	ARIMA (2,0,1) (0,1,1)s No Intercept	1.37%	3.35%
5	Log ARIMA (0,1,1) (0,1,1)s No Intercept	1.56%	4.04%
6	Airline Model	1.67%	4.03%

One exponential smoothing and one ARIMA model are selected for further investigation. ARIMA (0,2,2) (0,1,1)s No Intercept model is a Box Jenkins autoregressive integrated moving average model with zero auto regressive and two moving average terms after doubly differentiating, plus one seasonal moving average term after seasonally differentiating, with no constant term. The model parameters as optimized by SAS are as follows:

Moving Average, Lag 1 = 0.24607

Moving Average, Lag 2 = 0.28895

Seasonal Moving Average, Lag 4 = 0.73579

The predictions of the ARIMA model is started 6 periods beyond at 1992 3rd period, therefore the closeness of fit value is calculated for 46 periods. For the 46 periods of fit period the actual BMI values falls beyond the 95% confidence limits for 2 periods (one below the lower limit, one above the upper limit). For test data, 8 of 8 actual values lay between the 95% confidence limits.

The actual BMI data together with predicted data and the upper/lower confidence intervals for ARIMA Model for fit period is presented in Table A4 in Appendix, for predict period is presented in Table 3.6

Table 3.6 ARIMA Model Results for Predict Period for BMI

Forecast Period				
2004-1	36,775.4	36,557.8	37,409.3	35,706.3
2004-2	38,279.0	37,733.9	39,453.1	36,014.8
2004-3	39,695.2	38,889.3	41,443.7	36,334.8
2004-4	40,439.9	39,815.1	43,242.6	36,387.5
2005-1	41,482.1	41,366.7	45,860.6	36,872.9
2005-2	42,162.8	42,542.8	48,221.5	36,864.1
2005-3	43,332.3	43,698.0	50,628.1	36,767.8
2005-4	43,770.3	44,623.7	52,871.9	36,375.4

Figure 3.2 illustrates the ARIMA (0,2,2) (0,1,1)s No Intercept Model forecasts.

Winters Multiplicative Exponential Smoothing is the other time series method chosen for investigation. This method has level, trend and season parameters. The optimized values for level, trend and season parameters are as follows:

Level Smoothing Weight = 0.78725

Trend Smoothing Weight = 0.48565

Seasonal Smoothing Weight = 0.99900

Examining the Winters Multiplicative model predictions, out of 52 fit periods, 4 periods has actual BMI value outside the 95% confidence interval. For test data, 8 of 8 actual values lay between the 95% confidence limits.

The actual BMI data together with predicted data and the upper/lower confidence intervals for Winters Multiplicative Model for fit period is presented in Table A5 in Appendix, for predict period is presented in Table 3.7

As given in Table 3.5 the MAPE values in prediction performance are 1.20% and 1.24% and in closeness of fit are 3.06% and 3.51% for ARIMA and Winters Multiplicative Models respectively. As examining the comparison graph drawn with the actual data, and the forecasted data of two models for forecast horizon, the situation seems to differ on behalf of the Winters Multiplicative Method (Figure 3.3). By visual inspection Winters Multiplicative Model forecasts seem to have a better fit to the actual data. It follows a similar curvature with the actual data; both actual and Winters Multiplicative data show a increase at the end of forecast horizon while ARIMA data keeps its linearly increasing trend.

The Winters Multiplicative Exponential Smoothing Model forecasts are illustrated in Figure 3.3.

Table 3.7 Winters Multiplicative Model Results for Predict Period for BMI

Forecast Period				
2004-1	36,775.4	36,566.8	37,611.9	35,521.8
2004-2	38,279.0	38,005.2	39,618.5	36,391.9
2004-3	39,695.2	39,504.0	41,806.5	37,201.5
2004-4	40,439.9	39,919.6	42,936.8	36,902.4
2005-1	41,482.1	41,478.4	45,503.7	37,453.0
2005-2	42,162.8	42,944.1	47,880.4	38,007.8
2005-3	43,332.3	44,476.1	50,411.7	38,540.5
2005-4	43,770.3	44,790.8	51,624.2	37,957.3

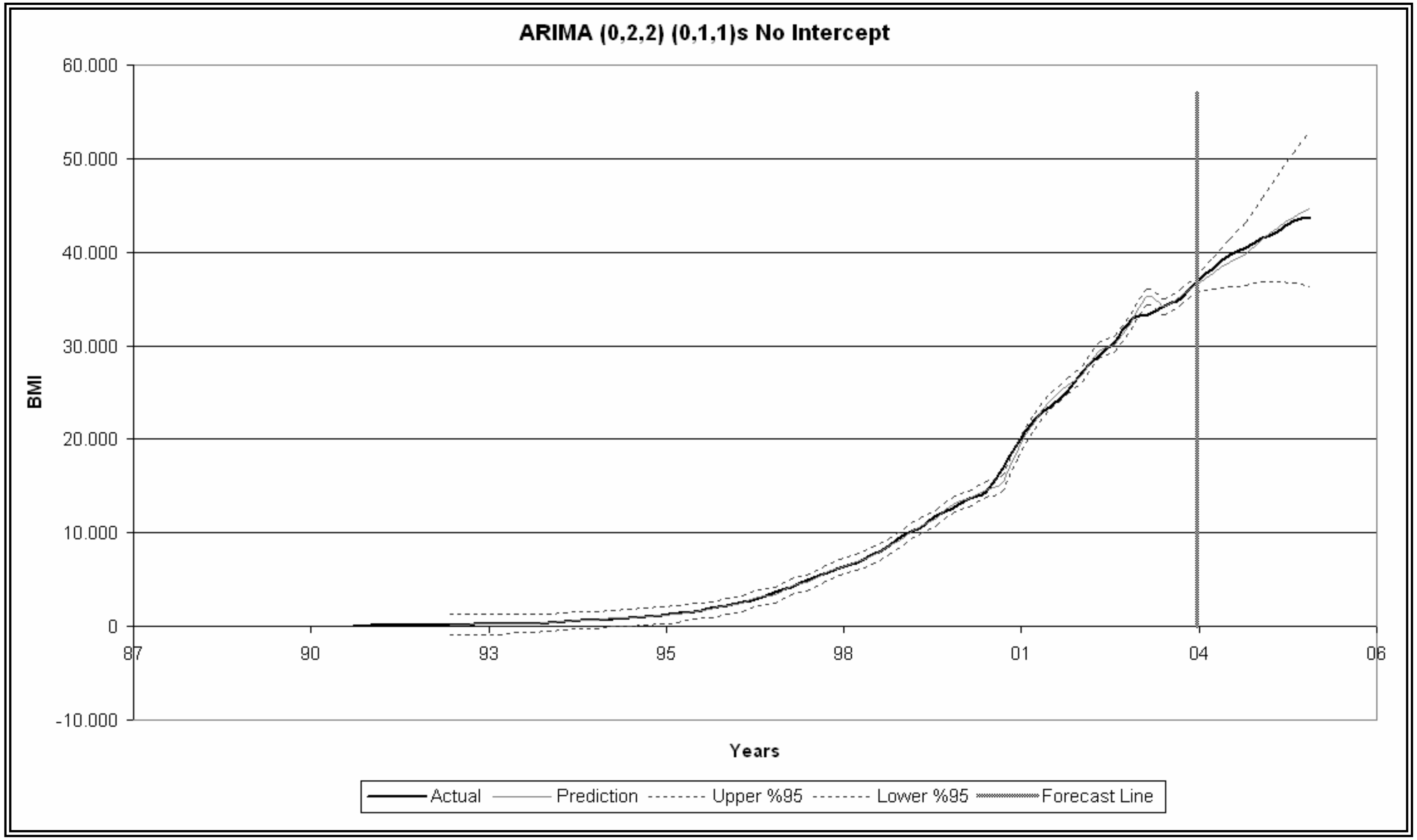

Figure 3.2 ARIMA Model for BMI

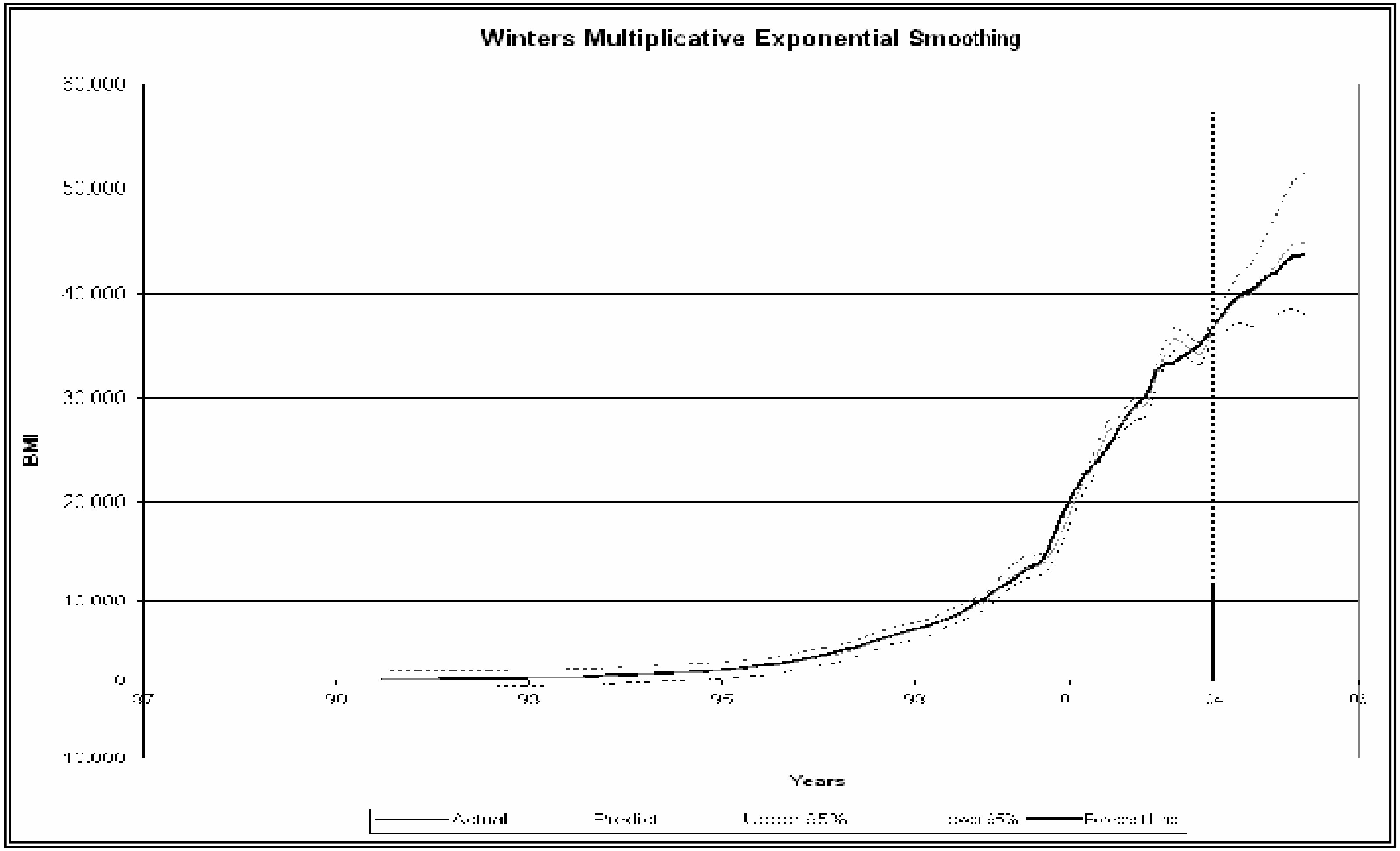

Figure 3.3 Winters Model for BMI

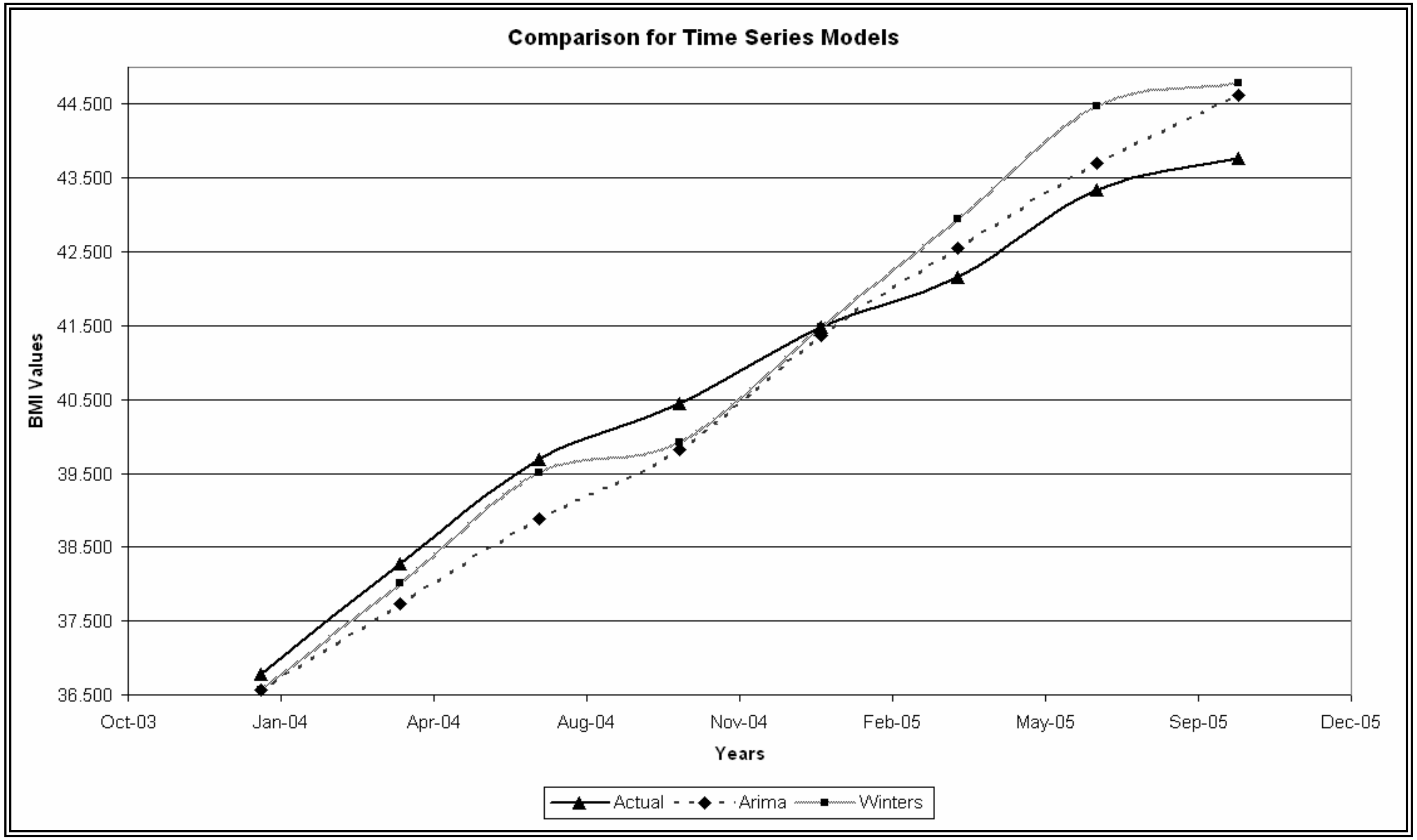

Figure 3.4 Comparison of Forecasts Between ARIMA and Winters Multiplicative Models for BMI

3.2.3 Simple Averaging

Two simple averaging models are constructed for the prediction of BMI values. Differentiated average and percent average models are named as *Dave* and *Pave*. *Dave* is the average of difference of the successive values, while *Pave* is the average of percent change of the successive values. The formulations for the models are given as:

$$Dave = \frac{1}{n-1}\sum_{i=1}^{n}\left(BMI_{i+1} - BMI_i\right)$$

$$Pave = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{BMI_{i+1}}{BMI_i} - 1\right)\times 100$$

After finding *Dave* and *Pave* values with according formulas the forecast of the future BMI values are done by the below formulas:

$$BMI_{t+1} = BMI_t + Dave$$

$$BMI_{t+1} = BMI_t \times \left(\frac{Pave}{100} + 1\right)$$

Differentiated and percent change BMI values as well as predictions by the *Dave* and *Pave* models for fit period are given in Table A6 in Appendix. Predictions in forecast horizon are tabulated in Table 3.8. Closeness of fit and prediction performance MAPE values is listed in Table 3.9.

Table 3.8 Forecasts by *Dave* and *Pave* Models for BMI

Quarter	Actual BMI	*Dave* Forecast	*Pave* Forecast
2004-1	36,775.4	35,734.2	39,524.3
2004-2	38,279.0	36,419.8	44,571.6
2004-3	39,695.2	37,105.4	50,263.4
2004-4	40,439.9	37,791.0	56,682.0
2005-1	41,482.1	38,476.6	63,920.3
2005-2	42,162.8	39,162.2	72,082.9
2005-3	43,332.3	39,847.8	81,287.9
2005-4	43,770.3	40,533.4	91,668.4

Table 3.9 MAPE Values for *Dave* and *Pave* Models for BMI

	Closeness of fit	Prediction Perfromance
Dave	103.50%	6,32%
Pave	3,92%	51.60%

BMI values change approximately from 82 to 35,000 in the test period (from 1991 to 2003), which causes high variations in *Dave* values (i.e. from 10 to 2,800), whereas *Pave* values range between 2 to 32 numerically. Therefore the high variations in *Dave* model was expected to bring huge errors in forecast. This expectation is also verified by the closeness of fit errors. But interestingly it was observed that *Dave* model gave better forecasts than *Pave* model with the prediction performance error values of 6.32% and 51.60%. Structure of data is investigated to comment on this unexpected result. Examining Table A6 in Appendix, it is observed that on low inflationary periods after 2001, the BMI values tend to increase more slowly therefore the percent change values tend to slow down. Therefore taking an average of the all test data including the high inflation periods bring huge error terms. On the other hand, inclusion of past periods with smaller differentiation values (because of the relative differences in size) lower the average difference which results in a good fit to slowly increasing future data. Table 3.10 shows differentiated and percent change values for actual BMI records in forecast horizon.

Table 3.10 Differentiated and Percent Change Values for Actual BMI in Forecast Horizon

Quarter	Actual BMI	Differentiated BMI	Percent Change BMI
2004-1	36,775.4	1,726.8	4.93%
2004-2	38,279.0	1,503.6	4.09%
2004-3	39,695.2	1,416.2	3.70%
2004-4	40,439.9	744.7	1.88%
2005-1	41,482.1	1,042.2	2.58%
2005-2	42,162.8	680.7	1.64%
2005-3	43,332.3	1,169.5	2.77%
2005-4	43,770.3	438.0	1.01%
	40.742.1	**1,090.2**	**2.82%**
	Average BMI	***Dave***	***Pave***

3.3 Study on United States, BCI

BCI is the building cost index used in United States published by Engineering News Record, ENR, in monthly basis. According to Touran and Lopez (2006), BCI is one of the most important, oldest, and commonly used cost index in construction industry in USA. BCI will be used in this study to represent the building cost forecasts in American construction industry.

The BCI data was obtained from ENR journals and the ENR home page (ENR, 2007) from January 1978 to June 2007. The monthly data is converted in a quarterly format to be consistent with the BMI study, moreover to bring out the seasonal characteristics of the data. Three subsequent monthly data was averaged to obtain the quarterly data. 106 of the 118 data (from 1978 1st quarter to 2004 2nd quarter) is used for the model constructions while the remaining 12 data (2004 3rd quarter to 2007 2nd quarter) is chosen as test data for prediction accuracy.

3.3.1 Regression Analysis

Two factors are determined to be used as independent variables in BCI regression models. First one, housing starts data (HS) is published by US Census Bureau on monthly basis. The data is obtained from the Census Bureau web page (Census, 2007). Second one mortgage rates (MR) is published by Federal Reserve on monthly basis and obtained from its web page (Federal Reserve, 2007). Both monthly data are converted to quarterly data by averaging the three subsequent monthly data.

Table A7 in Appendix shows the quarterly data for BCI, HS, and MR from 1978-1^{st} to 2007-2^{nd} quarters.

The final regression model is found from the first constructed model by the stepwise procedure as explained in part 4.2.1. The first model including 8 variables such that four time series of two variables ($HS_{(t-1)}$, $HS_{(t-2)}$, $HS_{(t-3)}$, $HS_{(t-4)}$, $MR_{(t-1)}$, $MR_{(t-2)}$, $MR_{(t-3)}$, $MR_{(t-4)}$) is reduced to 5 variables in the final model with the elimination criteria according to P-values and negative coefficients. The initial and final regression model characteristics are represented in Tables 3.11 and 3.12 respectively.

Table 3.11 Initial Regression Model Characteristics for BCI

Dependent Variable	Independent Variables	Coefficients	R^2	P-Value
BCI	$HS_{(t-1)}$	6.0404	0.9806	0.00389
	$HS_{(t-2)}$	3.9860		0.08468
	$HS_{(t-3)}$	0.4000		0.85787
	$HS_{(t-4)}$	3.4567		0.07322
	$MR_{(t-1)}$	118.8544		2.49E-07
	$MR_{(t-2)}$	10.4902		0.71410
	$MR_{(t-3)}$	20.7883		0.46707
	$MR_{(t-4)}$	47.8599		0.02984

Table 3.12 Final Regression Model Characteristics for BCI

Dependent Variable	Independent Variables	Coefficients	R^2	P-Value
BCI	$HS_{(t-1)}$	6.1965	0.9803	0.00145
	$HS_{(t-2)}$	4.0822		0.01562
	$HS_{(t-4)}$	3.6807		0.00976
	$MR_{(t-1)}$	131.8116		7.50E-15
	$MR_{(t-4)}$	64.4618		4.29E-05

MAPE for closeness to fit is obtained as 12.14% while MAPE for prediction performance is obtained as 19.35% using the actual values of independent variables (Table 3.14). Since the MAPE values are very high the regression model is not considered as adequate. Therefore further analysis on the prediction of independent variables is not performed.

Regression model inputs and predictions for fit period are given in Table A8 in Appendix. Regression model inputs and predictions for test period are given in Table 3.13. Closeness of fit and prediction performance MAPE vaues are given in Table 3.14.

Table 3.13 Regression Model Inputs and Test Period Predictions for BCI

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
2004-3	4047	3514	5481	1548
2004-4	4127	3569	5606	1532
2005-1	4118	3311	5192	1430
2005-2	4184	3367	5277	1457
2005-3	4216	3585	5626	1543
2005-4	4302	3648	5748	1547
2006-1	4334	3499	5481	1516
2006-2	4335	3505	5488	1521
2006-3	4364	3645	5689	1600
2006-4	4445	3505	5462	1548
2007-1	4424	3158	4901	1415
2007-2	4438	3036	4686	1386

Table 3.14 MAPE Values for Final Regression Model for BCI

Closeness of Fit	Prediction Performance Using Actual EXR Records
12.14%	19.35%

Below Figure 3.5 shows actual BCI and predicted BCI for both fit period and test period.

As in BMI regression figure the independent variables are shown on the graph by multiplication of the actual values with coefficients of lower 95% for better visualization. Therefore HS1* and MR1* are the values obtained by multiplication of $HS_{(t-1)}$ and $MR_{(t-1)}$ values with 2.44 and 103.34 respectively. As it is seen from figure the predicted data does not show an adequate fit on the actual data.

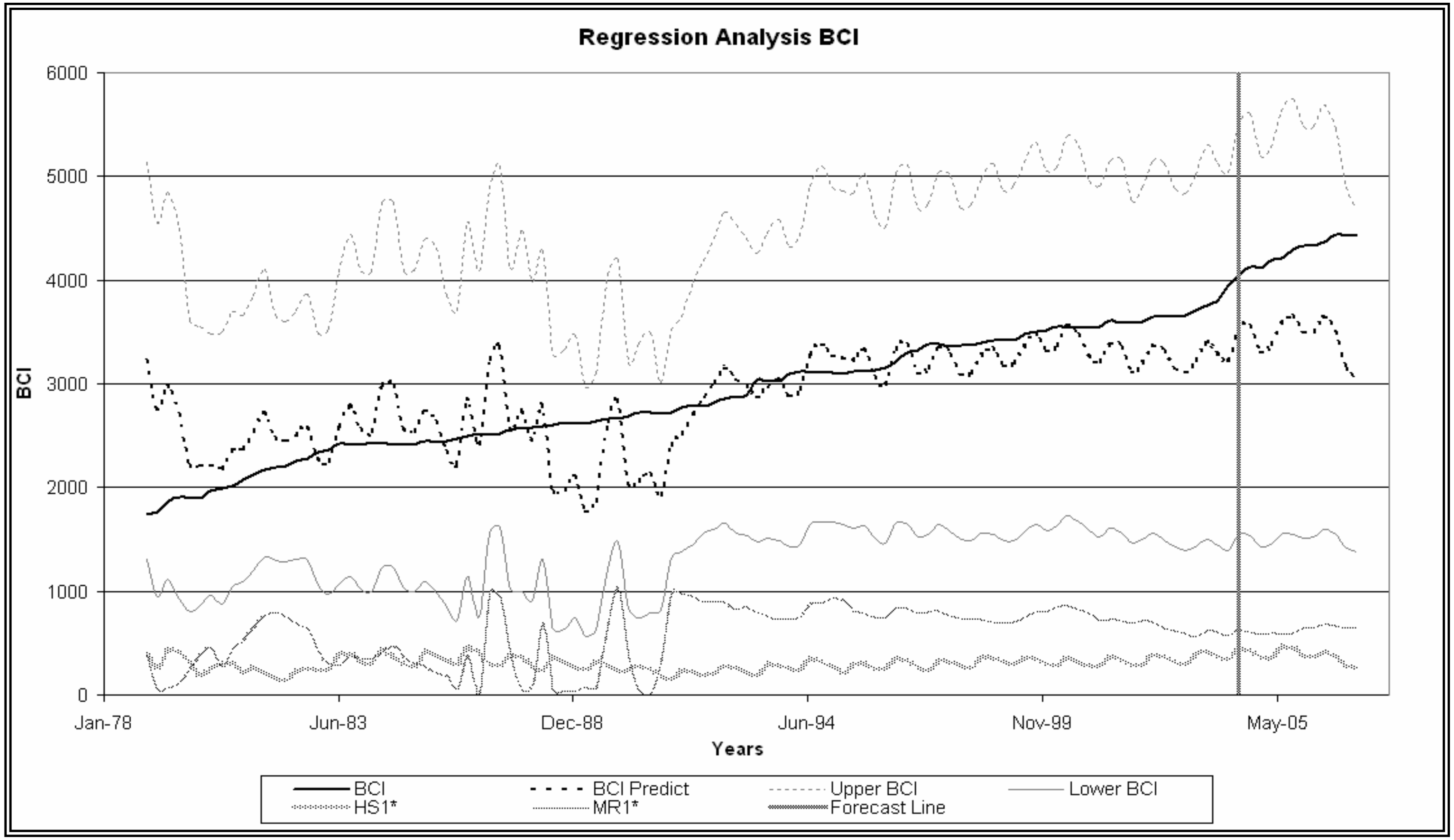

Figure 3.5 Regression Model for BCI

3.3.2 Time Series Analysis

Time series analysis for BCI data is performed by the SAS software. Using the quarterly BCI data from 1978 1st quarter to 2004 2nd quarter as input data, time series analysis is conducted by SAS software to find the best fit models. As a result of processing; 2 models could not be fitted to the data. Closeness of fit MAPE values for remaining 40 models was found with the given distribution below:

23 models ranging from 0.6 to 0.8 percent
7 models ranging from 0.80 to 1.0 percent
6 models ranging from 2.0 to 3.5 percent
4 models ranging from 19.0 to 19.6 percent

Prediction performances are calculated manually for the first 30 models which have relatively low error values. Prediction performance MAPE values for these models are distributed as follows:

2 Models ranging from 0.9 to 1.0 percent
11 Models ranging from 1.1 to 2.0 percent
10 Models ranging from 2.0 to 4.0 percent
7 Models ranging from 4.0 to 7.5 percent

Table 3.15 shows the 6 time series models having the prediction performance MAPE values between 0.9 to 1.5 while having the closeness of fit MAPE values between 0.6 to 0.8.

Table 3.15 MAPE values for Prediction Performance and Closeness of Fit

No	Time Series Model	Prediction Performance	Closeness of Fit
1	Winters Additive Exponential Smoothing	0.95%	0.67%
2	Log ARIMA (2,1,0)(0,1,1) No Intercept	0.97%	0.73%
3	Log Linear Holt Exponential Smoothing	1.11%	0.77%
4	Log ARIMA (2,1,2)(0,1,1) No Intercept	1.12%	0.71%
5	Log Linear Trend with Autoregressive Errors	1.24%	0.77%
6	ARIMA (0,2,2)(0,1,1) No Intercept	1.30%	0.68%

One exponential smoothing and one ARIMA model are selected for further investigation. ARIMA (2,1,0) (0,1,1)s No Intercept model is a Box Jenkins autoregressive integrated moving average model with two auto regressive and zero moving average terms after single differentiating plus one seasonal moving average term after seasonally differentiating with no constant term. The model parameters as optimized by SAS are as follows:

Autoregressive, Lag 1 = 0.23895

Autoregressive, Lag 2 = 0.05697

Seasonal Moving Average, Lag 4 = 0.73823

The predictions of the ARIMA model is started 5 periods beyond at 1979 2^{nd} period, therefore the closeness of fit value is calculated for 101 periods. For 101 fit periods the actual BCI values falls beyond the 95% confidence limits for 4 periods (one below the lower limit, three above the upper limit). For test data, 12 of 12 actual values lay between the 95% confidence limits.

The actual BCI data together with predicted data and the upper/lower confidence intervals for ARIMA Model are presented for fit period in Table A9 in Appendix, for predict period in Table 3.16.

Table 3.16 ARIMA Model for Predict Period for BCI

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
2004-3	4,047	4027	4106	3948
2004-4	4,127	4061	4189	3935
2005-1	4,118	4085	4256	3920
2005-2	4,184	4144	4353	3943
2005-3	4,216	4197	4451	3953
2005-4	4,302	4221	4517	3939
2006-1	4,334	4243	4578	3926
2006-2	4,335	4303	4677	3950
2006-3	4,364	4357	4779	3963
2006-4	4,445	4382	4848	3950
2007-1	4,424	4405	4913	3937
2007-2	4,438	4467	5021	3960

Figure 3.6 illustrates the ARIMA (2,1,0) (0,1,1)s No Intercept Model forecasts.

Winters Additive Exponential Smoothing is the other time series method chosen for investigation. This method has level, trend and season parameters. The optimized values for level, trend and season parameters are as follows:

Level Smoothing Weight = 0.99900

Trend Smoothing Weight = 0.16574

Seasonal Smoothing Weight = 0.99900

Examining Winters Multiplicative model predictions, out of 106 fit periods, 5 periods has actual BCI value outside the 95% confidence interval. For test data, 12 of 12 actual values lay between the 95% confidence limits.

The actual BCI data together with predicted data and the upper/lower confidence intervals for Winter Additive Model are presented for fit period in Table A10 in Appendix, for predict period in Table 3.17.

Table 3.17 Winters Additive Model for Predict Period for BCI

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
2004-3	4,047	4010	4064	3956
2004-4	4,127	4053	4135	3971
2005-1	4,118	4085	4194	3976
2005-2	4,184	4141	4276	4006
2005-3	4,216	4198	4360	4036
2005-4	4,302	4241	4430	4052
2006-1	4,334	4273	4491	4056
2006-2	4,335	4330	4576	4083
2006-3	4,364	4387	4662	4111
2006-4	4,445	4429	4736	4123
2007-1	4,424	4462	4799	4124
2007-2	4,438	4518	4888	4148

Figure 3.7 illustrates the Winters Additive Exponential Smoothing Model forecasts.

As given in Table 3.15 the MAPE values in prediction performance are 0.95% and 0.97% and in closeness of fit are 0.67% and 0.73% for Winters Additive Model and ARIMA Model respectively. Examining the comparison figure (Figure 3.8) both models seem to have a good fit on data and can be used adequately for BCI prediction.

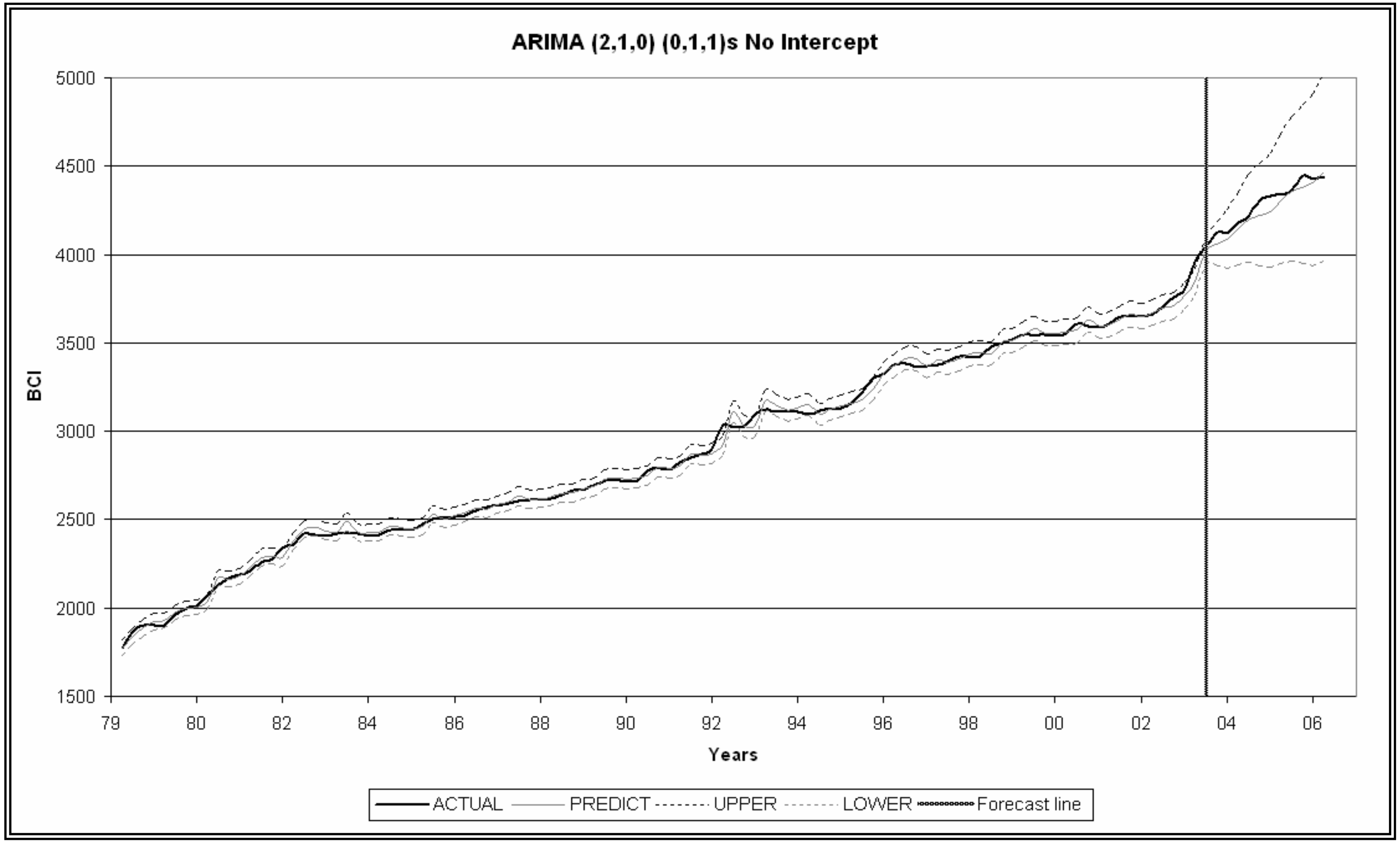

Figure 3.6 ARIMA Model for BCI

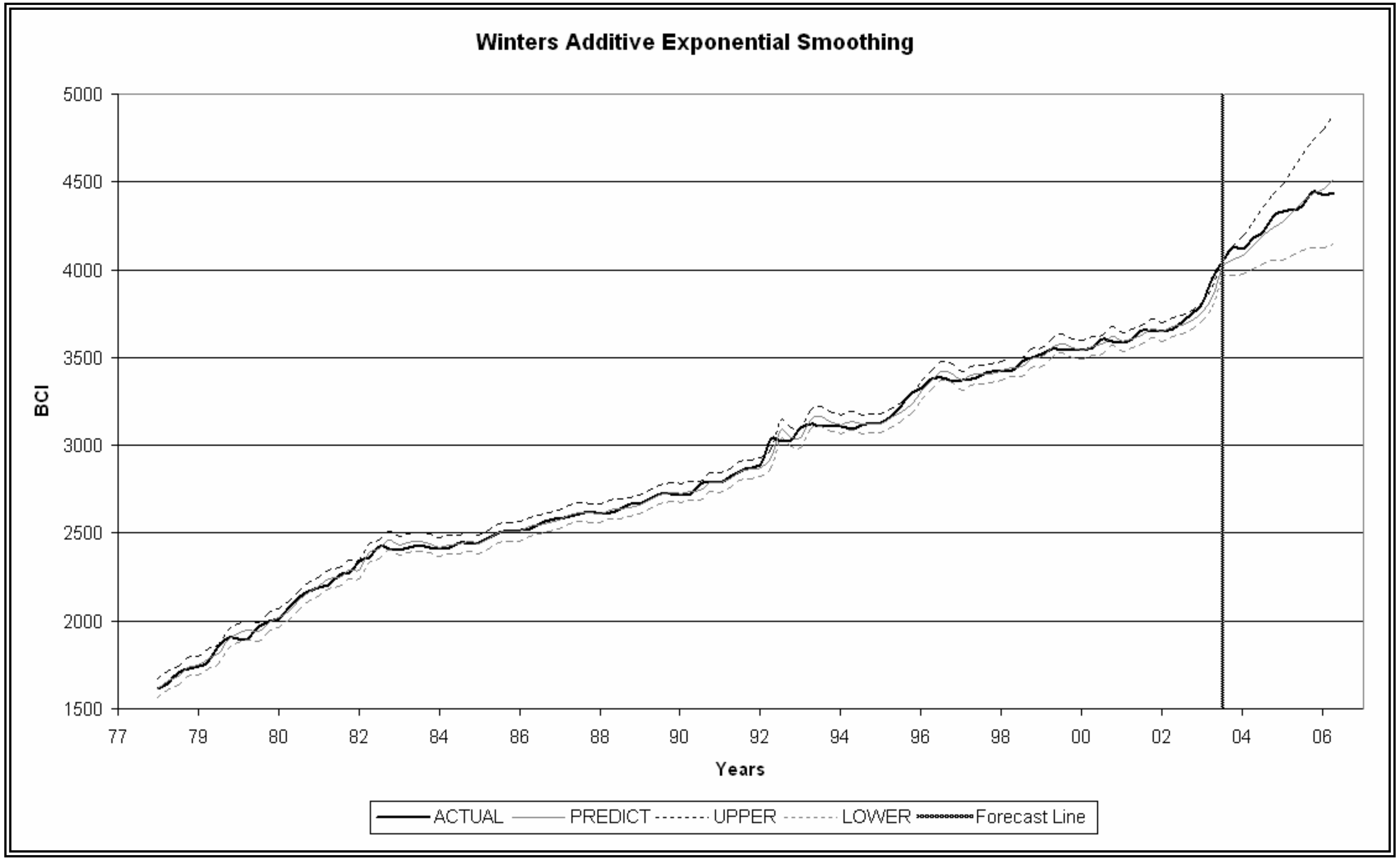

Figure 3.7 Winters Additive Model for BCI

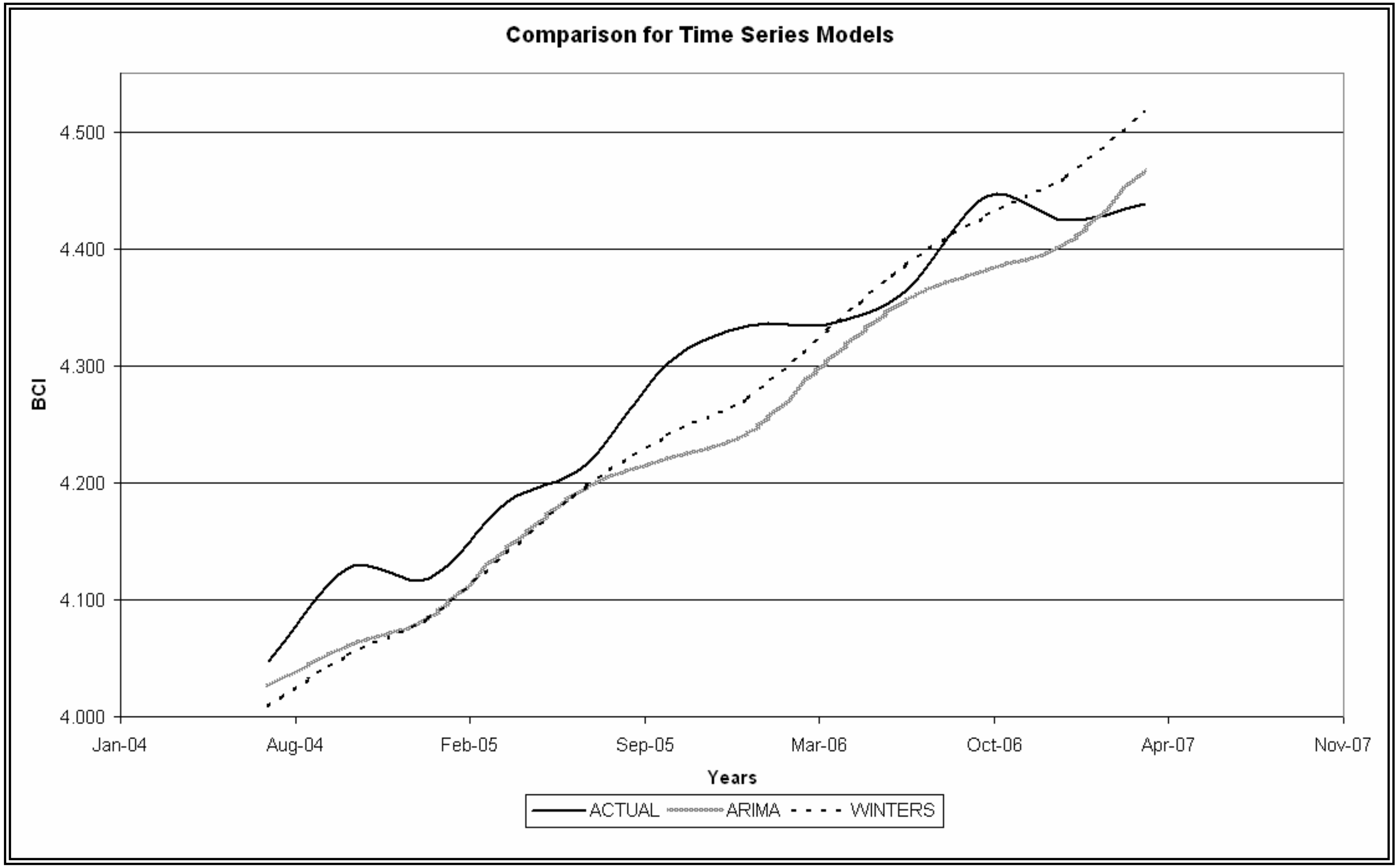

Figure 3.8 Comparison of Forecasts Between ARIMA and Winters Additive Models for BCI

3.3.3 Simple Averaging

Differentiated average; ***Dave*** and percent average; ***Pave*** models are constructed for BCI data. Formulations of the models are as follows:

$$Dave = \frac{1}{n-1}\sum_{i=1}^{n}\left(BCI_{i+1} - BCI_{i}\right)$$

$$Pave = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{BCI_{i+1}}{BCI_{i}} - 1\right)\times 100$$

After finding ***Dave*** and ***Pave*** values with according formulas the forecast of the future BCI values are done by the below formulas:

$$BCI_{t+1} = BCI_{t} + Dave$$

$$BCI_{t+1} = BCI_{t} \times \left(\frac{Pave}{100} + 1\right)$$

Differentiated and percent change BMI values as well as predictions by the *Dave* and *Pave* models for fit period are given in Table A11 in Appendix. Predictions in forecast horizon are tabulated in Table 3.18. Closeness of fit and prediction performance MAPE values is listed in Table 3.19.

Table 3.18 Dave and Pave Model Forecasts

Quarter	Actual BMI	*Dave* Forecast	*Pave* Forecast
2004-3	4,047	3,975	3987
2004-4	4,127	3,998	4021
2005-1	4,118	4,020	4056
2005-2	4,184	4,042	4091
2005-3	4,216	4,064	4126
2005-4	4,302	4,087	4162
2006-1	4,334	4,109	4198
2006-2	4,335	4,131	4234
2006-3	4,364	4,153	4270
2006-4	4,445	4,176	4307
2007-1	4,424	4,198	4344
2007-2	4,438	4,220	4382

Table 3.19 MAPE Values for Simple Averaging Models

Simple Averaging Model	Prediction Performance	Closeness of Fit
Dave	4.17%	0.75%
Pave	2.24%	0.80%

Pave model has a better prediction performance than *Dave* model although *Dave* has a better performance on fitting the data. 0.05% of error on data fit can be neglected on 1.93% of error on prediction therefore *Pave* model can be said a better model among *Dave* model on prediction of BCI future values.

Figure 3.9 shows Dave and Pave model fits in single graph.

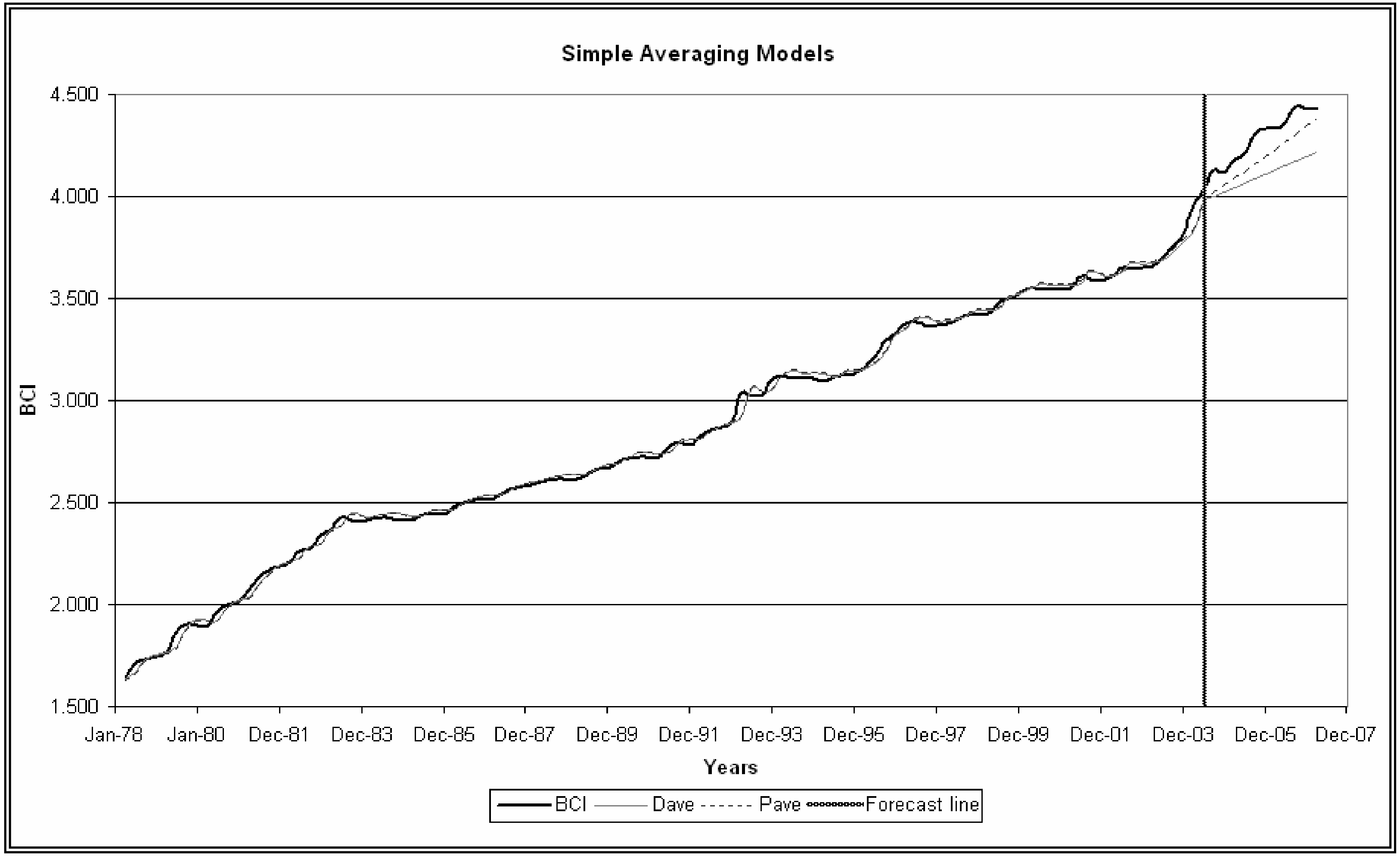

Figure 3.9 Simple Average Models for BCI

3.4 Discussion on Results

In BMI analysis 8 of the 60 data was selected for test data which counts to 13.3% of the total data whereas in BCI analysis this ratio drops to 10.2% with 12 periods of test data over a 118 total data. The both ratios are acceptable in statistical point of view.

Table 3.20 and 3.21 are prepared for the comparison of regression, time series and simple averaging models according to their accuracies to fit and predict the actual BMI and BCI data.

Table 3.20 Comparison of Models According to Their MAPE Values for BMI

MODEL	Closeness of Fit	Prediction Performance
Regression Model *	14.40%	14.90%
ARIMA (0,2,2) (0,1,1)s No Intercept	3.06%	1.20%
Winters Multiplicative Exponential Smoothing	3.51%	1.24%
ARIMA (2,1,2) (0,1,1)s No Intercept	3.29%	1.36%
ARIMA (2,0,1) (0,1,1)s No Intercept	3.35%	1.37%
Log ARIMA (0,1,1) (0,1,1)s No Intercept	4.04%	1.56%
Dave Model	103.50%	6.32%
Pave Model	3.92%	51.60%

* Prediction performance is calculated using the actual values of independent variables

Table 3.21 Comparison of Models According to Their MAPE Values for BCI

MODEL	Closeness of Fit	Prediction Performance
Regression Model *	12.14%	19.35%
Winters Additive Exponential Smoothing	0.67%	0.95%
Log ARIMA (2,1,0)(0,1,1)s No Intercept	0.73%	0.97%
Log Linear (Holt) Exponential Smoothing	0.77%	1.11%
Log ARIMA (2,1,2)(0,1,1)s No Intercept	0.71%	1.12%
Log Linear Trend with Autoregressive Errors	0.77%	1.24%
Dave Model	0.75%	4.17%
Pave Model	0.80%	2.24%

* Prediction performance is calculated using the actual values of independent variables

As given in the Tables regression models did not perform well in both of the index predictions. Moreover the two phase prediction theme would bring a second error term to be added. As a result of these findings regression models are not selected as adequate models.

Examining Table 3.20, the fit and predict error values in BMI forecast showed a cross match between *Dave* and *Pave* models. Dave model has a MAPE value of 103.50% for closeness of fit where as 6.32% for prediction performance, while the Pave model has 3.92% for closeness of fit where as 51.60% for prediction performance. This improper nature as discussed in part 3.2.3, is related to the change of inflation trends in Turkey in recent periods. As a result, the simple averaging models did not perform well for prediction of BMI cost index. *Dave* and *Pave* models give better results in BCI forecasting. The accuracy depends on the regular characteristics of the BCI data, as the data show a regular trend in time the averaging procedures gives better results than in BMI. Being simple and straight forward, and having MAPE value for prediction performance as 2.24%, *Pave* model can be used in rough estimates of BCI forecast.

In both BMI and BCI forecasts time series models performed better than regression and simple average models with relatively high percentages. Success of time series model among others is the ability to fit the seasonal and trend characteristics of BMI and BCI data. Moreover the exponential smoothing models gave very good results depending on their ability to fit the recent data. Exponential smoothing assigns exponentially decreasing weights as the observation get older; therefore recent trend changes in BMI and BCI were catch by exponential smoothing models. Seasonality is also an important pattern catch by the time series models. Examining the comparison tables (Tables 3.20, 3.21) it is seen that all of the top 5 time series models for BMI and BCI show seasonal characteristics. It is also noted that time series models for BCI give better results using Log transformation whereas the models for BMI give better results in regular form.

CHAPTER 4

SUMMARY AND CONCLUSIONS

Cost index forecast is crucial for the two main parties of construction industry, contactor, and the client. Forecast information is used to increase the accuracy of estimate for the project cost to evaluate the bid price. This research covered the studies performed on predicting the cost indices for Turkey and United States using three different models including regression, time series and simple averaging models.

General definitions of cost indices were given together with the history of the cost indices. Selected cost indices for Turkey: BMI, and United States: BCI, were further investigated for their origin, components and coverage. Forecasting methods were investigated in two main headers as quantitative and qualitative methods. The qualitative methods are used in the absence of quantitative data and when intuitive decision is needed. On the other hand quantitative methods require historical data to process and they are weak in catching the trends in future forecast. Common quantitative methods in use of cost index forecast are regression, neural network, and time series models. These models were explained in detail with the assumptions and the governing formulations. Measures of accuracies were also mentioned to set a basis for evaluation criteria for the comparison of models. Past studies on cost index forecasting were examined. The methodologies as well as the results of the past studies used as guidance for this research.

In this thesis, the steps and detailed explanations of the studies performed to meet the objective of this research are covered. The analysis performed for Turkey and United States were described in two separate parts. Also the sources of the obtained data, and necessary adjustments on the data were given in these parts. The methodologies of model construction were given for the regression, time series and simple averaging models. The model results were found and referred to the Appendix section. Mean absolute percentage error was selected as the measure of accuracy, and closeness of fit and prediction performance MAPE values were found for each model. In discussion of the results part MAPE values for all models were tabulated in a comparison table and necessary discussions were made for the adequacy of the models.

The regression models necessitate the prediction of the independent variables to perform the prediction for the cost index. This two phase forecast procedure reduces the accuracy of the regression model. Moreover despite the closeness of the R^2 values to 1.0 and P-values to 0 in regression models for both indices, closeness of fits and prediction performances were not good enough to select the regression model as the adequate model.

Simple averaging models which were introduced as differentiated average, *Dave* and percent average, *Pave* models did not perform well in prediction of BMI cost index. The change of the inflation trends in recent periods reduced the accuracy of simple averaging models in Turkish cost indices. Recommendation for future studies can be made that, assuming the low inflationary period persists; simple average models can be constructed including only the low inflationary periods when adequate time periods for modeling is reached. The simple averaging models have relatively good performance in prediction of BCI forecast. Especially *Pave* model is remarkable considering its simplicity and relatively good accuracy.

Time series models gave the best results in forecasting of both cost indices. The MAPE values for prediction performance in BMI forecast was found within the range of 1.2% to 1.5% for time series models while it was 14.90% for regression, 6.32% for *Dave* and 51.60% for *Pave* models. Similarly in BCI forecast prediction performance MAPE value ranges between 0.95% and 1.12% for time series models while it was 19.35% for regression, 4.17% for *Dave* and 2.24% for *Pave* models. Success of time series model among others is the ability to fit the seasonal and trend characteristics of the cost index data. Moreover exponential smoothing models assigns exponentially increasing weights to the recent data, which result in better fit.

In the beginning of the study monthly BCI data was converted into quarterly format as to comply with the quarterly BMI data. As a result, quarterly data fits well on seasonal time series models, therefore quarterly modeling is recommended for seasonal characteristics of cost indices. Alternately a new study can be constructed on the adequacy of monthly, quarterly and yearly data on prediction of future cost indices.

This study aimed to find the most suitable forecast tool to estimate the future construction cost indices in Turkey and United States. As a result of analysis performed and the discussions made, time series models are selected as the most reliable models in predicting cost indices for both Turkey and United States. Future values of construction cost indices can be predicted in adequate precision using time series models. This useful information can be used in tender process in estimation of project costs, which is one of the critical factors affecting the overall success of a construction project. Better cost estimates shall enable contractors to produce cash flow forecasts more acurately. Furthermore prediction of future construction prices is very useful for owners in budget allocations; moreover can help investors to evaluate project alternatives.

REFERENCES

Akintoye, S.A. and Skitmore, R.M. 1993. "A comparative analysis of the three macro price forecasting models", Construction Management and Economics, 12(3), 257-70.

Answers.com, Last updated June 25, 2007. http://www.answers.com/topic/euro, Last access June 25, 2007.

Blair, A. N., Lye, L. M., and Campbell, W. J., 1993. "Forecasting construction cost escalation", Canadian Journal of Civil Engineering, 20(4), 602–612.

Box, G.F.P. and Jenkins, G.M., 1970. "Time series analysis: forecasting and control", San Francisco, Holden-Day.

Box, G.F.P. and Jenkins, G.M., 1976 "Time series analysis: forecasting and control", San Francisco, Holden-Day.

Brown, R.G., 1970. "Smoothing, forecasting and prediction of discrete time series", Prentice-Hall, Englewood Cliffs, NJ.

Capano, G. D., and Karshenas, S., 2003. "Applying accepted economic indicators to predict cost escalation for construction." ASC Proc., 39th Annual Conf., 277–288.

Census, Last updated June 25, 2007. http://www.census.gov/const/startsua.pdf, , Last access June 25, 2007.

Central Bank of The Republic of Turkey, Last updated June 25, 2007. http://www.tcmb.gov.tr/, Last access June 25, 2007.

Chatfield, C., 1975. "The analysis of time series: Theory and practice" Chapman and Hall, London.

Durbin, J., 1970. Testing for serial correlation in least squares dependent variables" Econometrica, 38, 410-21.

ENR, 2007. Engineering News Record, Last updated June 25, 2007. http://enr.construction.com/Default.asp, Last access June 25, 2007.

EUROSTAT, Luxembourg, "Methodological aspects of construction price indices" 1996.

EUROSTAT, Luxembourg, "Industrial trends: National methods, Supp." December 1995.

Federal Reserve, Last updated June 25, 2007. http://www.federalreserve.gov/releases/h15/data.htm, Last access June 25, 2007.

Fitzgerald, E., and Akintoye, A., 1995, "The accuracy and optimal linear correction of UK construction tender price index forecasts", Construction Management and Economics, 13, 493-500.

Grogan, T., 1994. "Cost history: indexes hit by high lumber prices", Eng. News Record 28, 46.

Grogan, T., 2007. ENR FAQ, Last updated June 25, 2007. http://enr.construction.com/features/conEco/indexFAQ.asp, Last access June 25, 2007.

Hanna, A. S., and Blair, A. N., 1993. "Computerized approach for forecasting the rate of cost escalation." Proc., Comput. Civ. Build. Tech. Conf., 401–408.

Hanna, A. S., and Chao, L., 1994. "Quantification of cost uncertainties using neural network technique." Proc., 1st Congress on Computing in Civil Engineering, Washington, D.C., Vol. I, 41–46.

Holden, K. and Peel, D.A., 1988. "A comparison of some inflation, growth and unemployment forecasts", Journal of Economic Studies, 15, 45-52.

Hyndman, Rob J. 2001. ARIMA Processes, Last updated May 25, 2001. www-personal.buseco.monash.edu.au/~hyndman/papers/ARIMA.pdf, Last access June 25, 2007.

Kahraman, S. 2005. "Determination of a price index for escalation of building construction costs in Turkey", M.S. Thesis, Middle East Technical University, Graduate School of Natural and Applied Sciences, Ankara.

Kress, G., 1985. "Forecasting courses aimed at managers, not technicians", Journal of Business Forecasting, 14, 10-1.

Koehn, E. and Navvabi, M.H., 1989. "Economics and social factors in construction." Cost Engineering, 31, 15-8.

Makridakis, S., Wheelwright, S. C., and Hyndman, R. J., 1998. "Forecasting methods and applications", John Wiley & Sons, New York.

Makridakis, S. and Hibon, M., 1984. "Accuracy of forecasting: - an empirical investigation, In Makridakis, S. (ed.) The forecasting accuracy of major time series methods", New York, John Wiley and Sons, pp. 35-59.

McNees, S.K. and Ries, J., 1983. "The track records of macroeconomics forecasts", New England Economic Review, November/December, 5-18.

NIST/SEMATECH e-Handbook of Statistical Methods, Last Updated, July 18, 2006. http://www.itl.nist.gov/div898/handbook/, Last Access June 25, 2007.

Ng, S.T., Cheung S.O., Skitmore, R.M., Lam, K.C. and Wong, L.Y., 2000. "The prediction of tender price index directional changes", Construction Management and Economics, 18(7), 843-52.

Ng, S.T., Cheung S.O., Skitmore, R.M., and Wong, T.C.Y., 2004. "An integrated regression analysis and time series model for construction tender price index forecasting", Construction Management and Economics, June 2004, 22, 483-93.

OECD, Paris, Main Economic Indicators, Sources and Methods, 1996.

OECD, Paris, Producer Price Indices, Sources and Methods, 1994. (b)

OECD, Paris, Consumer Price Indices, Sources and Methods, 1994. (a)

Ostwald, P.F., 1992. "Engineering Cost Estimating", Englewood Cliffs, NJ, Prentice-Hall, Inc., 3rd ed., p.170.

Pankratz, A., 1983. "Forecasting with univariate Box–Jenkins models: concepts and cases", New York: John Wiley & Sons.

Runeson, K.G., 1988. "Methodology and method for price level forecasting in the building industry", Construction Management and Economics, 6(1), 49-55.

Sauder School of Business, Last updated January 30, 2006. http://fx.sauder.ubc.ca/ECU.html, Last access June 25, 2007.

Slutsky, E., 1937. "The summation of random causes as the source of cyclic processes", Econometrica, 5, 105-146.

Snyder, D., 1982. "Current approaches to time series forecasting", Business Forum, 7, 24-6.

Statistic Glossary, Last updated September 19, 1997. http://www.stats.gla.ac.uk/steps/glossary/time_series.html#timeseries, Last access June 25, 2007.

Stergiou and Siganos, Last updated May 12, 1997. http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html#Introduction%20to%20neural%20networks, Last access June 25, 2007.

Taylor, R.G. and Bowen, P.A., 1987. "Building price-level forecasting: an examination of techniques and application", Construction Management and Economics, 5(1), 21-44.

TCMB, Last updated June 25, 2007 http://www.tcmb.gov.tr/, Last access June 25, 2007.

Theil, H. 1978. "Introduction to Econometrics", Englewood Cliffs, NJ, Prentice-Hall, Inc.

Touran, A. And Lopez, R., 2006. "Modeling Cost Escalation in Large Infrastructure Projects." Journal of Construction Engineering and Management, ASCE, August 2006, 853-860.

Treham, B., 1989, "Forecasting growth in current quarter real GNP", Economic Review, Federal Reserve Bank of San Francisco, Winter, 39-51.

TURKSTAT, 2002. Fiyat Endeksleri ve Enflasyon, Sorularla istatistikler Dizisi 2.

Wang, C.H. and Mei, Y.H., 1998. "Model for forecasting construction cost indices in Taiwan", Construction Management and Economics, 16, 147-157.

Westney, R. E., 1997. "Engineer's cost handbook: Tools for managing project costs", Marcel Dekker Inc., New York.

Williams, T.P., 1994. "Predicting changes in construction cost indexes using neural networks", Journal of Construction Engineering and Management, June 1994, Vol. 120. No. 2.

Wilmot, C.G. and Cheng, G., 2003. "Estimating future highway construction costs", Journal of Construction Engineering and Management, ASCE, Vol. 129, No. 3, 272-279.

Yule, G.U., 1927. On a method of investigating periodicities in disturbed series with special reference to Wolfer's sunspot numbers", Philosophical Transactions of the Royal Society of London, A, 226, 267-98.

APPENDIX

REFERRED TABLES

Table A1: Quarterly Data for BMI, CP and EXR

Quarter	BMI	CP	EXR	Quarter	BMI	CP	EXR
1991-1	81.8	7,164	0.0076	1998-3	6,240.8	8,766	0.5802
1991-2	93.4	11,191	0.0088	1998-4	6,817.0	5,835	0.6445
1991-3	106.8	11,042	0.0099	1999-1	7,734.4	6,310	0.7264
1991-4	118.1	11,098	0.0112	1999-2	8,651.6	9,912	0.8166
1992-1	139.8	7,939	0.0128	1999-3	9,748.5	8,766	0.9008
1992-2	155.8	12,358	0.0152	1999-4	10,610.1	5,835	1.0154
1992-3	182.4	12,539	0.0170	2000-1	11,836.6	2,961	1.1228
1992-4	201.4	13,161	0.0183	2000-2	12,693.5	6,763	1.1839
1993-1	230.9	7,200	0.0198	2000-3	13,647.3	10,299	1.2355
1993-2	272.4	11,906	0.0223	2000-4	14,327.8	6,357	1.2749
1993-3	306.8	13,018	0.0249	2001-1	16,864.1	3,643	1.5283
1993-4	340.9	16,887	0.0286	2001-2	19,651.2	5,342	2.2293
1994-1	450.9	11,271	0.0381	2001-3	22,112.3	12,669	2.6503
1994-2	582.3	10,514	0.0705	2001-4	23,545.7	4,156	2.8989
1994-3	628.9	11,295	0.0721	2002-1	25,111.8	2,171	2.5578
1994-4	730.2	14,680	0.0812	2002-2	27,217.8	4,238	2.7154
1995-1	842.7	8,255	0.0932	2002-3	28,896.9	3,775	3.2765
1995-2	955.6	10,862	0.0995	2002-4	30,428.1	5,563	3.2415
1995-3	1,057.9	11,742	0.1068	2003-1	32,811.0	2,490	3.4285
1995-4	1,173.0	15,109	0.1222	2003-2	33,362.3	4,346	3.2429
1996-1	1,414.5	7,469	0.1461	2003-3	34,197.9	4,879	2.9670
1996-2	1,627.3	11,069	0.1716	2003-4	35,048.6	6,233	3.1697
1996-3	1,907.0	10,308	0.1937	2004-1	36,775.4	4,990	3.0071
1996-4	2,175.4	13,395	0.2244	2004-2	38,279.0	5,339	3.2090
1997-1	2,659.1	7,321	0.2587	2004-3	39,695.2	6,114	3.2873
1997-2	3,039.9	10,219	0.2946	2004-4	40,439.9	8,722	3.3225
1997-3	3,679.7	10,450	0.3414	2005-1	41,482.1	6,220	3.0733
1997-4	4,160.6	14,328	0.4020	2005-2	42,162.8	9,548	3.0878
1998-1	4,906.3	6,310	0.4684	2005-3	43,332.3	9,618	2.9792
1998-2	5,589.5	9,912	0.5330	2005-4	43,770.3	12,698	2.9717

Table A2: Regression Model Inputs and Model Fit Period Predictions

Quarter	Actual BMI	$EXR_{(t-1)}$	$EXR_{(t-4)}$	Predict	Upper 95%	Lower 95%
1991-1	81.8					
1991-2	93.4	0.0076				
1991-3	106.8	0.0088				
1991-4	118.1	0.0099				
1992-1	139.8	0.0112	0.0076	110.3	135.9	84.7
1992-2	155.8	0.0128	0.0088	126.4	155.8	97.0
1992-3	182.4	0.0152	0.0099	148.3	182.4	114.2
1992-4	201.4	0.0170	0.0112	166.3	204.6	128.0
1993-1	230.9	0.0183	0.0128	181.5	223.9	139.1
1993-2	272.4	0.0198	0.0152	200.9	248.8	153.0
1993-3	306.8	0.0223	0.0170	225.9	279.6	172.1
1993-4	340.9	0.0249	0.0183	249.9	308.9	191.0
1994-1	450.9	0.0286	0.0198	283.0	348.9	217.1
1994-2	582.3	0.0381	0.0223	363.3	444.9	281.7
1994-3	628.9	0.0705	0.0249	617.3	743.6	491.1
1994-4	730.2	0.0721	0.0286	641.9	775.7	508.0
1995-1	842.7	0.0812	0.0381	742.6	902.3	583.0
1995-2	955.6	0.0932	0.0705	942.2	1,166.0	718.4
1995-3	1,057.9	0.0995	0.0721	995.3	1,229.4	761.2
1995-4	1,173.0	0.1068	0.0812	1,081.1	1,338.1	824.0
1996-1	1,414.5	0.1222	0.0932	1,237.9	1,532.5	943.4
1996-2	1,627.3	0.1461	0.0995	1,440.0	1,774.1	1,106.0
1996-3	1,907.0	0.1716	0.1068	1,657.6	2,034.8	1,280.4
1996-4	2,175.4	0.1937	0.1222	1,876.6	2,304.9	1,448.4
1997-1	2,659.1	0.2244	0.1461	2,189.3	2,692.2	1,686.3
1997-2	3,039.9	0.2587	0.1716	2,534.5	3,119.2	1,949.9
1997-3	3,679.7	0.2946	0.1937	2,880.5	3,543.7	2,217.3
1997-4	4,160.6	0.3414	0.2244	3,337.8	4,106.3	2,569.4
1998-1	4,906.3	0.4020	0.2587	3,911.8	4,808.2	3,015.3
1998-2	5,589.5	0.4684	0.2946	4,535.0	5,569.1	3,500.8
1998-3	6,240.8	0.5330	0.3414	5,181.1	6,367.3	3,994.9
1998-4	6,817.0	0.5802	0.4020	5,741.8	7,079.0	4,404.6

Table A2: Continued

Quarter	Actual BMI	$EXR_{(t-1)}$	$EXR_{(t-4)}$	Predict	Upper 95%	Lower 95%
1999-1	7,734.4	0.6445	0.4684	6,451.4	7,969.9	4,933.0
1999-2	8,651.6	0.7264	0.5330	7,288.3	9,007.4	5,569.2
1999-3	9,748.5	0.8166	0.5802	8,129.6	10,033.4	6,225.8
1999-4	10,610.1	0.9008	0.6445	8,982.9	11,089.8	6,876.0
2000-1	11,836.6	10.154	0.7264	10,125.4	12,500.1	7,750.6
2000-2	12,693.5	11.228	0.8166	11,241.2	13,887.4	8,595.0
2000-3	13,647.3	11.839	0.9008	11,986.3	14,836.7	9,136.0
2000-4	14,327.8	12.355	10.154	12,761.5	15,850.2	9,672.9
2001-1	16,864.1	12.749	11.228	13,420.2	16,720.9	10,119.5
2001-2	19,651.2	15.283	11.839	15,543.8	19,255.3	11,832.4
2001-3	22,112.3	22.293	12.355	21,024.6	25,694.5	16,354.7
2001-4	23,545.7	26.503	12.749	24,344.5	29,602.7	19,086.2
2002-1	25,111.8	28.989	15.283	27,077.0	33,031.0	21,122.9
2002-2	27,217.8	25.578	22.293	26,846.3	33,433.0	20,259.5
2002-3	28,896.9	27.154	26.503	29,452.1	36,874.0	22,030.1
2002-4	30,428.1	32.765	28.989	34,534.6	43,037.5	26,031.6
2003-1	32,811.0	32.415	25.578	33,125.1	41,067.8	25,182.5
2003-2	33,362.3	34.285	27.154	35,069.8	43,485.9	26,653.7
2003-3	34,197.9	32.429	32.765	35,547.1	44,579.4	26,514.8
2003-4	35,048.6	29.670	32.415	33,340.7	41,973.5	24,707.8

Table A3: Default Time Series Models Used In SAS Software

No	Time Series Model
1	Mean
2	Linear Trend
3	Linear Trend with Autoregressive Errors
4	Linear Trend with Seasonal Terms
5	Seasonal Dummy
6	Simple Exponential Smoothing
7	Double (Brown) Exponential Smoothing
8	Linear (Holt) Exponential Smoothing
9	Damped Trend Exponential Smoothing
10	Seasonal Exponential Smoothing
11	Winters Additive Exponential Smoothing
12	Winters Multiplicative Exponential Smoothing
13	Random Walk with Drift
14	Airline Model
15	ARIMA (0,1,1)s No Intercept
16	ARIMA (0,1,1) (1,0,0)s No Intercept
17	ARIMA (2,0,0) (1,0,0)s
18	ARIMA (0,1,2) (0,1,1)s No Intercept
19	ARIMA (2,1,0) (0,1,1)s No Intercept
20	ARIMA (0,2,2) (0,1,1)s No Intercept
21	ARIMA (2,1,2) (0,1,1)s No Intercept
22	Log Mean
23	Log Linear Trend
24	Log Linear Trend with Autoregressive Errors
25	Log Linear Trend with Seasonal Terms
26	Log Seasonal Dummy
27	Log Simple Exponential Smoothing
28	Log Double (Brown) Exponential Smoothing

Table A3 Continued

No	Time Series Model
29	Log Linear (Holt) Exponential Smoothing
30	Log Damped Trend Exponential Smoothing
31	Log Seasonal Exponential Smoothing
32	Log Winters Additive Exponential Smoothing
33	Log Winters Multiplicative Exponential Smoothing
34	Log Random Walk with Drift
35	Log Airline Model
36	Log ARIMA (0,1,1)s No Intercept
37	Log ARIMA (0,1,1) (1,0,0)s No Intercept
38	Log ARIMA (2,0,0) (1,0,0)s
39	Log ARIMA (0,1,2) (0,1,1)s No Intercept
40	Log ARIMA (2,1,0) (0,1,1)s No Intercept
41	Log ARIMA (0,2,2) (0,1,1)s No Intercept
42	Log ARIMA (2,1,2) (0,1,1)s No Intercept

Table A4: ARIMA (0,2,2) (0,1,1)s No Intercept Model Results for Fit Period

Quarter	Actual	Predict	Upper 95%	Lower 95%
1991-1	81.8			
1991-2	93.4			
1991-3	106.8			
1991-4	118.1			
1992-1	139.8			
1992-2	155.8			
1992-3	182.4	173.6	1,304.3	-957.1
1992-4	201.4	205.6	1,323.0	-911.9
1993-1	230.9	230.4	1,333.6	-872.8
1993-2	272.4	255.8	1,358.6	-847.0
1993-3	306.8	317.5	1,268.9	-633.9
1993-4	340.9	335.1	1,281.6	-611.4
1994-1	450.9	387.1	1,330.6	-556.3
1994-2	582.3	549.1	1,492.4	-394.1
1994-3	628.9	691.7	1,588.1	-204.8
1994-4	730.2	678.3	1,572.8	-216.1
1995-1	842.7	871.7	1,765.2	-21.8
1995-2	955.6	958.6	1,851.9	65.2
1995-3	1,057.9	1,050.6	1,924.3	176.9
1995-4	1,173.0	1,174.8	2,047.5	302.0
1996-1	1,414.5	1,313.6	2,186.0	441.2
1996-2	1,627.3	1,640.0	2,512.3	767.7
1996-3	1,907.0	1,792.9	2,655.8	930.0
1996-4	2,175.4	2,177.3	3,039.8	1,314.8
1997-1	2,659.1	2,468.4	3,330.8	1,606.1
1997-2	3,039.9	3,094.5	3,956.8	2,232.3
1997-3	3,679.7	3,382.6	4,240.1	2,525.1
1997-4	4,160.6	4,269.9	5,127.2	3,412.6
1998-1	4,906.3	4,684.7	5,541.9	3,827.5
1998-2	5,589.5	5,599.1	6,456.3	4,741.9
1998-3	6,240.8	6,284.3	7,139.1	5,429.6
1998-4	6,817.0	6,867.6	7,722.2	6,013.0

Table A4 Continued

Quarter	Actual	Predict	Upper 95%	Lower 95%
1999-1	7,734.4	7,564.6	8,419.2	6,710.1
1999-2	8,651.6	8,585.7	9,440.3	7,731.2
1999-3	9,748.5	9,548.5	10,401.7	8,695.3
1999-4	10,610.1	10,729.4	11,582.6	9,876.2
2000-1	11,836.6	11,642.2	12,495.4	10,789.1
2000-2	12,693.5	13,021.3	13,874.4	12,168.1
2000-3	13,647.3	13,655.5	14,507.9	12,803.1
2000-4	14,327.8	14,599.7	15,452.1	13,747.3
2001-1	16,864.1	15,320.7	16,173.1	14,468.4
2001-2	19,651.2	18,980.0	19,832.3	18,127.6
2001-3	22,112.3	21,912.9	22,764.9	21,060.9
2001-4	23,545.7	24,186.1	25,038.1	23,334.1
2002-1	25,111.8	25,750.5	26,602.5	24,898.5
2002-2	27,217.8	26,998.8	27,850.8	26,146.8
2002-3	28,896.9	29,430.5	30,282.3	28,578.7
2002-4	30,428.1	30,265.5	31,117.3	29,413.8
2003-1	32,811.0	32,602.2	33,453.9	31,750.4
2003-2	33,362.3	35,223.0	36,074.7	34,371.2
2003-3	34,197.9	34,180.4	35,032.1	33,328.8
2003-4	35,048.6	35,249.2	36,100.8	34,397.6

Table A5: Winters Multiplicative Model Results for Fit Period

Quarter	Actual	Predict	Upper 95%	Lower 95%
1991-1	81.8	81.8	1,126.8	-963.2
1991-2	93.4	95.4	1,140.4	-949.6
1991-3	106.8	106.2	1,151.2	-938.9
1991-4	118.1	117.3	1,162.4	-927.7
1992-1	139.8	138.5	1,183.5	-906.6
1992-2	155.8	152.1	1,197.1	-893.0
1992-3	182.4	169.9	1,215.0	-875.1
1992-4	201.4	195.4	1,240.4	-849.7
1993-1	230.9	234.3	1,279.4	-810.7
1993-2	272.4	250.9	1,295.9	-794.1
1993-3	306.8	298.9	1,343.9	-746.1
1993-4	340.9	327.1	1,372.2	-717.9
1994-1	450.9	391.5	1,436.5	-653.6
1994-2	582.3	503.9	1,549.0	-541.1
1994-3	628.9	652.0	1,697.1	-393.0
1994-4	730.2	694.4	1,739.4	-350.6
1995-1	842.7	868.0	1,913.0	-177.1
1995-2	955.6	945.8	1,990.8	-99.2
1995-3	1,057.9	1,001.1	2,046.1	-44.0
1995-4	1,173.0	1,142.8	2,187.9	97.8
1996-1	1,414.5	1,347.3	2,392.3	302.2
1996-2	1,627.3	1,579.7	2,624.7	534.6
1996-3	1,907.0	1,727.0	2,772.0	682.0
1996-4	2,175.4	2,065.7	3,110.7	1,020.7
1997-1	2,659.1	2,548.4	3,593.4	1,503.4
1997-2	3,039.9	3,003.9	4,049.0	1,958.9
1997-3	3,679.7	3,305.2	4,350.2	2,260.2
1997-4	4,160.6	3,968.8	5,013.9	2,923.8
1998-1	4,906.3	4,888.7	5,933.8	3,843.7
1998-2	5,589.5	5,514.1	6,559.1	4,469.0
1998-3	6,240.8	6,162.7	7,207.7	5,117.6
1998-4	6,817.0	6,588.3	7,633.3	5,543.3

Table A5 Continued

Quarter	Actual	Predict	Upper 95%	Lower 95%
1999-1	7,734.4	7,757.7	8,802.7	6,712.6
1999-2	8,651.6	8,528.8	9,573.8	7,483.7
1999-3	9,748.5	9,363.9	10,408.9	8,318.9
1999-4	10,610.1	10,206.5	11,251.6	9,161.5
2000-1	11,836.6	11,917.1	12,962.2	10,872.1
2000-2	12,693.5	13,049.2	14,094.2	12,004.2
2000-3	13,647.3	13,711.0	14,756.0	12,666.0
2000-4	14,327.8	14,032.9	15,078.0	12,987.9
2001-1	16,864.1	15,558.0	16,603.0	14,512.9
2001-2	19,651.2	18,217.9	19,262.9	17,172.9
2001-3	22,112.3	21,547.4	22,592.4	20,502.4
2001-4	23,545.7	23,533.5	24,578.5	22,488.4
2002-1	25,111.8	26,624.8	27,669.8	25,579.7
2002-2	27,217.8	27,329.1	28,374.2	26,284.1
2002-3	28,896.9	28,830.2	29,875.3	27,785.2
2002-4	30,428.1	29,493.6	30,538.6	28,448.6
2003-1	32,811.0	32,884.6	33,929.7	31,839.6
2003-2	33,362.3	35,490.0	36,535.1	34,445.0
2003-3	34,197.9	34,984.8	36,029.8	33,939.7
2003-4	35,048.6	34,245.0	35,290.0	33,199.9

Table A6: Simple Average Models *Dave* and *Pave*

Quarter	BMI	Differenciated BMI	Percent Change BMI	Prediction by *Dave*	Prediction by *Pave*
1991-1	81,8				
1991-2	93,4	11,6	14,18%	767	92
1991-3	106,8	13,4	14,35%	779	105
1991-4	118,1	11,3	10,58%	792	120
1992-1	139,8	21,7	18,37%	804	133
1992-2	155,8	16,0	11,44%	825	158
1992-3	182,4	26,6	17,07%	841	176
1992-4	201,4	19,0	10,42%	868	206
1993-1	230,9	29,5	14,65%	887	227
1993-2	272,4	41,5	17,97%	917	260
1993-3	306,8	34,4	12,63%	958	307
1993-4	340,9	34,1	11,11%	992	346
1994-1	450,9	110,0	32,27%	1027	384
1994-2	582,3	131,4	29,14%	1137	508
1994-3	628,9	46,6	8,00%	1268	657
1994-4	730,2	101,3	16,11%	1315	709
1995-1	842,7	112,5	15,41%	1416	823
1995-2	955,6	112,9	13,40%	1528	950
1995-3	1.057,9	102,3	10,71%	1641	1078
1995-4	1.173,0	115,1	10,88%	1744	1193
1996-1	1.414,5	241,5	20,59%	1859	1323
1996-2	1.627,3	212,8	15,04%	2100	1595
1996-3	1.907,0	279,7	17,19%	2313	1835
1996-4	2.175,4	268,4	14,07%	2593	2151
1997-1	2.659,1	483,7	22,23%	2861	2453
1997-2	3.039,9	380,8	14,32%	3345	2999
1997-3	3.679,7	639,8	21,05%	3726	3428
1997-4	4.160,6	480,9	13,07%	4365	4150

Table A6: Continued

Quarter	BMI	Differenciated BMI	Percent Change BMI	Prediction by *Dave*	Prediction by *Pave*
1998-1	4.906,3	745,7	17,92%	4846	4692
1998-2	5.589,5	683,2	13,92%	5592	5533
1998-3	6.240,8	651,3	11,65%	6275	6303
1998-4	6.817,0	576,2	9,23%	6926	7038
1999-1	7.734,4	917,4	13,46%	7503	7688
1999-2	8.651,6	917,2	11,86%	8420	8722
1999-3	9.748,5	1.096,9	12,68%	9337	9757
1999-4	10.610,1	861,6	8,84%	10434	10994
2000-1	11.836,6	1.226,5	11,56%	11296	11965
2000-2	12.693,5	856,9	7,24%	12522	13348
2000-3	13.647,3	953,8	7,51%	13379	14315
2000-4	14.327,8	680,5	4,99%	14333	15390
2001-1	16.864,1	2.536,3	17,70%	15013	16158
2001-2	19.651,2	2.787,1	16,53%	17550	19018
2001-3	22.112,3	2.461,1	12,52%	20337	22161
2001-4	23.545,7	1.433,4	6,48%	22798	24936
2002-1	25.111,8	1.566,1	6,65%	24231	26553
2002-2	27.217,8	2.106,0	8,39%	25797	28319
2002-3	28.896,9	1.679,1	6,17%	27903	30694
2002-4	30.428,1	1.531,2	5,30%	29583	32587
2003-1	32.811,0	2.382,9	7,83%	31114	34314
2003-2	33.362,3	551,3	1,68%	33497	37001
2003-3	34.197,9	835,6	2,50%	34048	37623
2003-4	35.048,6	850,7	2,49%	34884	38566
	Average	***685,6***	***12,77%***		
		Dave	***Pave***		

Table A7 Quarterly Data for BCI, HS and MR

Quarter	BCI	HS	MR	Quarter	BCI	HS	MR
Jan-78	1,615	120.7	9.13	Jan-86	2,444	124.6	0.56
Apr-78	1,645	208.2	9.55	Apr-86	2,477	186.1	3.59
Jul-78	1,707	187.9	9.76	Jul-86	2,500	163.3	0.24
Oct-78	1,729	156.7	3.44	Oct-86	2,511	127.8	9.66
Jan-79	1,743	108.5	0.41	Jan-87	2,514	116.4	9.11
Apr-79	1,770	180.6	0.74	Apr-87	2,524	160.1	3.66
Jul-79	1,859	166.1	1.16	Jul-87	2,553	149.3	0.5
Oct-79	1,903	126.4	2.46	Oct-87	2,574	114.4	0.85
Jan-80	1,901	79.4	3.73	Jan-88	2,579	99.1	6.75
Apr-80	1,901	101.4	4.43	Apr-88	2,593	147.9	0.37
Jul-80	1,966	129.4	2.65	Jul-88	2,607	135	0.5
Oct-80	1,998	120.5	4.26	Oct-88	2,615	114.1	0.39
Jan-81	2,015	88.1	5.14	Jan-89	2,612	101.2	0.8
Apr-81	2,073	112.9	6.23	Apr-89	2,618	134.8	0.67
Jul-81	2,130	90.1	7.43	Jul-89	2,641	122.1	6.67
Oct-81	2,170	70.3	7.73	Oct-89	2,665	100.6	9.82
Jan-82	2,191	58.9	7.39	Jan-90	2,668	98.2	3.46
Apr-82	2,207	91.3	6.76	Apr-90	2,694	119.3	0.34
Jul-82	2,260	103.1	6.17	Jul-90	2,721	102.4	0.11
Oct-82	2,276	100.8	4.02	Oct-90	2,726	77.7	3.29
Jan-83	2,337	107.4	3.03	Jan-91	2,717	61.8	9.5
Apr-83	2,362	161.3	2.76	Apr-91	2,722	100.3	9.53
Jul-83	2,424	164.4	3.65	Jul-91	2,778	94.9	9.27
Oct-83	2,414	134.5	3.47	Oct-91	2,787	81	8.69
Jan-84	2,407	125.5	3.33	Jan-92	2,786	87.3	8.71
Apr-84	2,419	179.1	4	Apr-92	2,825	113.5	8.68
Jul-84	2,425	152.7	4.5	Jul-92	2,852	107.4	8.01
Oct-84	2,418	125.8	3.65	Oct-92	2,872	91.6	8.21
Jan-85	2,410	115.3	3.06	Jan-93	2,896	80.2	7.73
Apr-85	2,415	169.7	2.78	Apr-93	3,038	122.4	7.45
Jul-85	2,444	156.4	2.14	Jul-93	3,020	118.5	7.08
Oct-85	2,442	139.2	1.73	Oct-93	3,030	108.1	7.05

Table A7 Continued

Quarter	BCI	HS	MR	Quarter	BCI	HS	MR
Jan-94	3,098	98	7.3	Oct-00	3,545	119.2	7.64
Apr-94	3,122	140.9	8.44	Jan-01	3,541	115.9	7.01
Jul-94	3,111	132.6	8.59	Apr-01	3,553	153.5	7.13
Oct-94	3,112	114.2	9.1	Jul-01	3,609	143.1	6.97
Jan-95	3,109	90	8.81	Oct-01	3,592	121.8	6.78
Apr-95	3,097	123.6	7.95	Jan-02	3,586	123	6.97
Jul-95	3,115	129.1	7.7	Apr-02	3,606	158.2	6.82
Oct-95	3,125	108.7	7.35	Jul-02	3,652	152.8	6.29
Jan-96	3,131	100.9	7.24	Oct-02	3,648	134.3	6.08
Apr-96	3,162	142.8	8.11	Jan-03	3,651	124.9	5.84
Jul-96	3,220	136.8	8.16	Apr-03	3,663	163.6	5.51
Oct-96	3,300	111.8	7.71	Jul-03	3,704	170.3	6.01
Jan-97	3,329	99.1	7.79	Oct-03	3,756	157.1	5.92
Apr-97	3,379	139.7	7.92	Jan-04	3,809	141.6	5.61
Jul-97	3,385	133.4	7.47	Apr-04	3,953	179.8	6.13
Oct-97	3,364	119.1	7.2	Jul-04	4,047	177.3	5.89
Jan-98	3,368	108.3	7.05	Oct-04	4,127	153.2	5.73
Apr-98	3,376	149.3	7.09	Jan-05	4,118	149.4	5.76
Jul-98	3,396	148.3	6.86	Apr-05	4,184	191.8	5.72
Oct-98	3,422	133.1	6.77	Jul-05	4,216	189.2	5.76
Jan-99	3,418	121.4	6.88	Oct-05	4,302	159	6.22
Apr-99	3,425	149.1	7.21	Jan-06	4,334	154.7	6.24
Jul-99	3,479	148.6	7.8	Apr-06	4,335	173.6	6.6
Oct-99	3,500	128	7.83	Jul-06	4,364	152.6	6.56
Jan-00	3,521	119	8.26	Oct-06	4,445	119.4	6.25
Apr-00	3,548	149.6	8.32	Jan-07	4,424	107.2	6.22
Jul-00	3,543	135.1	8.03	Apr-07	4,438	139.3	6.22

Table A8 Regression Model Fit Period Predictions

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
1979-1	1743	3224	5136	1313
1979-2	1770	2748	4545	951
1979-3	1859	2980	4848	1113
1979-4	1903	2718	4505	930
1980-1	1901	2211	3615	807
1980-2	1901	2212	3555	869
1980-3	1966	2223	3485	960
1980-4	1998	2189	3507	871
1981-1	2015	2369	3700	1039
1981-2	2073	2374	3661	1087
1981-3	2130	2528	3856	1199
1981-4	2170	2717	4099	1334
1982-1	2191	2478	3656	1300
1982-2	2207	2443	3605	1282
1982-3	2260	2508	3708	1308
1982-4	2276	2582	3870	1293
1983-1	2337	2269	3484	1053
1983-2	2362	2248	3524	972
1983-3	2424	2579	4086	1072
1983-4	2414	2788	4439	1138
1984-1	2407	2553	4085	1020
1984-2	2419	2537	4075	1000
1984-3	2425	2990	4753	1226
1984-4	2418	2989	4756	1222
1985-1	2410	2561	4086	1035
1985-2	2415	2548	4098	999
1985-3	2444	2741	4393	1088
1985-4	2442	2642	4306	979
1986-1	2444	2351	3848	854
1986-2	2477	2218	3713	723
1986-3	2500	2849	4556	1141
1986-4	2511	2427	4097	758

Table A8 Continued

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
1987-1	2514	3227	4880	1573
1987-2	2524	3360	5100	1621
1987-3	2553	2566	4114	1019
1987-4	2574	2738	4482	994
1988-1	2579	2446	3985	907
1988-2	2593	2796	4289	1303
1988-3	2607	1952	3272	631
1988-4	2615	1982	3332	632
1989-1	2612	2109	3469	750
1989-2	2618	1767	2965	568
1989-3	2641	1866	3113	619
1989-4	2665	2631	4051	1212
1990-1	2668	2840	4203	1477
1990-2	2694	2015	3206	823
1990-3	2721	2064	3382	746
1990-4	2726	2139	3491	787
1991-1	2717	1918	3015	820
1991-2	2722	2413	3512	1315
1991-3	2778	2514	3652	1376
1991-4	2787	2717	3981	1454
1992-1	2786	2875	4193	1556
1992-2	2825	3003	4408	1599
1992-3	2852	3151	4651	1651
1992-4	2872	3043	4531	1555
1993-1	2896	2971	4404	1538
1993-2	3038	2867	4263	1471
1993-3	3020	2979	4451	1508
1993-4	3030	3034	4581	1486
1994-1	3098	2876	4326	1427
1994-2	3122	2942	4434	1449

Table A8 Continued

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
1994-3	3111	3278	4909	1647
1994-4	3112	3381	5097	1666
1995-1	3109	3280	4893	1666
1995-2	3097	3248	4855	1641
1995-3	3115	3223	4833	1613
1995-4	3125	3326	5025	1628
1996-1	3131	3069	4623	1514
1996-2	3162	2991	4513	1468
1996-3	3220	3337	5024	1651
1996-4	3300	3380	5116	1644
1997-1	3329	3106	4690	1521
1997-2	3379	3146	4745	1546
1997-3	3385	3344	5042	1645
1997-4	3364	3290	5003	1577
1998-1	3368	3099	4697	1501
1998-2	3376	3111	4734	1489
1998-3	3396	3274	4985	1563
1998-4	3422	3335	5123	1547
1999-1	3418	3176	4866	1485
1999-2	3425	3209	4924	1494
1999-3	3479	3358	5138	1578
1999-4	3500	3484	5325	1643
2000-1	3521	3322	5056	1588
2000-2	3548	3362	5093	1631
2000-3	3543	3559	5394	1725
2000-4	3545	3482	5298	1666
2001-1	3541	3268	4953	1582
2001-2	3553	3216	4905	1526
2001-3	3609	3379	5153	1605
2001-4	3592	3363	5162	1564

Table A8 Continued

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
2002-1	3586	3111	4762	1460
2002-2	3606	3203	4902	1504
2002-3	3652	3357	5154	1560
2002-4	3648	3307	5120	1494
2003-1	3651	3159	4888	1431
2003-2	3663	3114	4833	1395
2003-3	3704	3218	5011	1425
2003-4	3756	3402	5304	1499
2004-1	3809	3285	5124	1446
2004-2	3953	3216	5041	1390

Table A9 ARIMA (2,1,0) (0,1,1)s No Intercept Model for Fit Period

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
1979-2	1770	1776	1820	1731
1979-3	1859	1836	1881	1791
1979-4	1903	1888	1935	1843
1980-1	1901	1925	1972	1878
1980-2	1901	1931	1973	1890
1980-3	1966	1977	2019	1935
1980-4	1998	1997	2040	1954
1981-1	2015	2004	2047	1961
1981-2	2073	2038	2080	1996
1981-3	2130	2167	2212	2123
1981-4	2170	2164	2208	2120
1982-1	2191	2181	2226	2137
1982-2	2207	2229	2274	2185
1982-3	2260	2283	2329	2237
1982-4	2276	2293	2340	2248
1983-1	2337	2284	2331	2239
1983-2	2362	2379	2426	2332
1983-3	2424	2440	2489	2392
1983-4	2414	2456	2505	2408
1984-1	2407	2433	2482	2385
1984-2	2419	2427	2475	2379
1984-3	2425	2487	2537	2438
1984-4	2418	2431	2479	2383
1985-1	2410	2427	2476	2380
1985-2	2415	2428	2476	2380
1985-3	2444	2464	2513	2416
1985-4	2442	2452	2500	2404
1986-1	2444	2449	2498	2401
1986-2	2477	2461	2510	2413
1986-3	2500	2529	2580	2480
1986-4	2511	2505	2555	2456

Table A9 Continued

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
1987-1	2514	2520	2570	2471
1987-2	2524	2537	2587	2487
1987-3	2553	2564	2615	2514
1987-4	2574	2560	2611	2510
1988-1	2579	2584	2635	2534
1988-2	2593	2600	2652	2549
1988-3	2607	2632	2684	2581
1988-4	2615	2614	2666	2563
1989-1	2612	2620	2672	2569
1989-2	2618	2629	2681	2577
1989-3	2641	2648	2701	2597
1989-4	2665	2650	2703	2599
1990-1	2668	2672	2725	2620
1990-2	2694	2685	2738	2632
1990-3	2721	2728	2782	2675
1990-4	2726	2737	2791	2683
1991-1	2717	2728	2782	2675
1991-2	2722	2732	2787	2679
1991-3	2778	2749	2803	2695
1991-4	2787	2797	2852	2742
1992-1	2786	2789	2844	2734
1992-2	2825	2801	2857	2747
1992-3	2852	2870	2926	2814
1992-4	2872	2863	2920	2807
1993-1	2896	2874	2931	2818
1993-2	3038	2925	2983	2868
1993-3	3020	3108	3169	3047
1993-4	3030	3029	3089	2970
1994-1	3098	3033	3093	2973
1994-2	3122	3170	3233	3108

Table A9 Continued

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
1994-3	3111	3142	3204	3081
1994-4	3112	3115	3177	3054
1995-1	3109	3130	3192	3069
1995-2	3097	3151	3213	3089
1995-3	3115	3096	3157	3036
1995-4	3125	3123	3185	3062
1996-1	3131	3142	3204	3080
1996-2	3162	3161	3224	3100
1996-3	3220	3177	3240	3115
1996-4	3300	3241	3305	3178
1997-1	3329	3334	3400	3269
1997-2	3379	3371	3438	3305
1997-3	3385	3413	3480	3346
1997-4	3364	3411	3479	3344
1998-1	3368	3369	3436	3304
1998-2	3376	3401	3468	3334
1998-3	3396	3391	3458	3325
1998-4	3422	3411	3478	3344
1999-1	3418	3439	3507	3372
1999-2	3425	3445	3514	3378
1999-3	3479	3441	3509	3374
1999-4	3500	3505	3575	3437
2000-1	3521	3513	3582	3444
2000-2	3548	3549	3620	3480
2000-3	3543	3581	3652	3511
2000-4	3545	3555	3625	3486
2001-1	3541	3552	3622	3483
2001-2	3553	3562	3633	3493
2001-3	3609	3571	3642	3501
2001-4	3592	3632	3704	3562
2002-1	3586	3595	3666	3525
2002-2	3606	3603	3675	3533

Table A9 Continued

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
2002-3	3652	3636	3708	3565
2002-4	3648	3663	3735	3591
2003-1	3651	3651	3724	3580
2003-2	3663	3672	3745	3601
2003-3	3704	3696	3770	3624
2003-4	3756	3709	3783	3637
2004-1	3809	3773	3847	3699
2004-2	3953	3844	3920	3769

Table A10 Winters Additive Model for Fit Period

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
1978-1	1615	1615	1669	1561
1978-2	1645	1658	1711	1604
1978-3	1707	1686	1740	1633
1978-4	1729	1738	1791	1684
1979-1	1743	1747	1801	1694
1979-2	1770	1785	1839	1731
1979-3	1859	1810	1864	1756
1979-4	1903	1893	1947	1839
1980-1	1901	1928	1982	1874
1980-2	1901	1946	1999	1892
1980-3	1966	1939	1993	1885
1980-4	1998	1994	2048	1941
1981-1	2015	2016	2070	1963
1981-2	2073	2057	2111	2004
1981-3	2130	2119	2172	2065
1981-4	2170	2163	2217	2110
1982-1	2191	2194	2247	2140
1982-2	2207	2238	2292	2185
1982-3	2260	2250	2304	2196
1982-4	2276	2290	2344	2237
1983-1	2337	2293	2347	2240
1983-2	2362	2386	2439	2332
1983-3	2424	2408	2461	2354
1983-4	2414	2458	2512	2404
1984-1	2407	2430	2484	2377
1984-2	2419	2443	2497	2390
1984-3	2425	2452	2506	2399
1984-4	2418	2439	2493	2386
1985-1	2410	2418	2472	2365
1985-2	2415	2433	2487	2379
1985-3	2444	2436	2489	2382
1985-4	2442	2452	2505	2398

Table A10 Continued

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
1986-1	2444	2438	2491	2384
1986-2	2477	2465	2518	2411
1986-3	2500	2501	2554	2447
1986-4	2511	2509	2563	2456
1987-1	2514	2510	2564	2456
1987-2	2524	2538	2591	2484
1987-3	2553	2546	2600	2493
1987-4	2574	2562	2615	2508
1988-1	2579	2574	2628	2521
1988-2	2593	2604	2658	2551
1988-3	2607	2617	2671	2564
1988-4	2615	2615	2669	2562
1989-1	2612	2613	2666	2559
1989-2	2618	2634	2687	2580
1989-3	2641	2638	2692	2584
1989-4	2665	2647	2701	2593
1990-1	2668	2663	2717	2610
1990-2	2694	2691	2745	2638
1990-3	2721	2719	2772	2665
1990-4	2726	2732	2785	2678
1991-1	2717	2725	2779	2672
1991-2	2722	2739	2792	2685
1991-3	2778	2742	2796	2688
1991-4	2787	2789	2843	2736
1992-1	2786	2788	2841	2734
1992-2	2825	2810	2864	2757
1992-3	2852	2853	2906	2799
1992-4	2872	2865	2919	2811
1993-1	2896	2876	2929	2822
1993-2	3038	2927	2981	2874
1993-3	3020	3088	3142	3035
1993-4	3030	3045	3098	2991

Table A10 Continued

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
1994-1	3098	3042	3095	2988
1994-2	3122	3143	3197	3090
1994-3	3111	3164	3218	3111
1994-4	3112	3130	3184	3077
1995-1	3109	3118	3171	3064
1995-2	3097	3137	3191	3084
1995-3	3115	3119	3173	3066
1995-4	3125	3122	3176	3069
1996-1	3131	3122	3176	3069
1996-2	3162	3154	3207	3100
1996-3	3220	3187	3240	3133
1996-4	3300	3236	3290	3183
1997-1	3329	3316	3370	3263
1997-2	3379	3371	3425	3318
1997-3	3385	3423	3477	3370
1997-4	3364	3409	3462	3355
1998-1	3368	3370	3423	3316
1998-2	3376	3397	3451	3344
1998-3	3396	3403	3456	3349
1998-4	3422	3407	3461	3354
1999-1	3418	3425	3479	3372
1999-2	3425	3444	3498	3390
1999-3	3479	3449	3502	3395
1999-4	3500	3493	3547	3440
2000-1	3521	3505	3559	3451
2000-2	3548	3553	3606	3499
2000-3	3543	3580	3633	3526
2000-4	3545	3554	3608	3501
2001-1	3541	3544	3598	3491
2001-2	3553	3564	3617	3510
2001-3	3609	3575	3628	3521
2001-4	3592	3622	3676	3569

Table A10 Continued

Quarter	BCI	BCI Predict	Upper BCI	Lower BCI
2002-1	3586	3590	3643	3536
2002-2	3606	3607	3660	3553
2002-3	3652	3628	3681	3574
2002-4	3648	3663	3717	3610
2003-1	3651	3646	3700	3593
2003-2	3663	3674	3728	3620
2003-3	3704	3685	3739	3632
2003-4	3756	3715	3768	3661
2004-1	3809	3763	3817	3710
2004-2	3953	3848	3901	3794

Table A 11 Simple Average Models *Dave* and *Pave*

Quarter	BCI	Differenciated BCI	Percent Change BCI	Prediction by *Dave*	Prediction by *Pave*
1978-1	1.615				
1978-2	1.645	30,0	1,86%	1.637	1.629
1978-3	1.707	62,0	3,77%	1.667	1.659
1978-4	1.729	22,0	1,29%	1.729	1.722
1979-1	1.743	14,0	0,81%	1.751	1.744
1979-2	1.770	27,0	1,55%	1.765	1.758
1979-3	1.859	89,0	5,03%	1.792	1.785
1979-4	1.903	44,0	2,37%	1.881	1.875
1980-1	1.901	-2,0	-0,11%	1.925	1.919
1980-2	1.901	0,0	0,00%	1.923	1.917
1980-3	1.966	65,0	3,42%	1.923	1.917
1980-4	1.998	32,0	1,63%	1.988	1.983
1981-1	2.015	17,0	0,85%	2.020	2.015
1981-2	2.073	58,0	2,88%	2.037	2.032
1981-3	2.130	57,0	2,75%	2.095	2.091
1981-4	2.170	40,0	1,88%	2.152	2.148
1982-1	2.191	21,0	0,97%	2.192	2.189
1982-2	2.207	16,0	0,73%	2.213	2.210
1982-3	2.260	53,0	2,40%	2.229	2.226
1982-4	2.276	16,0	0,71%	2.282	2.279
1983-1	2.337	61,0	2,68%	2.298	2.296
1983-2	2.362	25,0	1,07%	2.359	2.357
1983-3	2.424	62,0	2,62%	2.384	2.382
1983-4	2.414	-10,0	-0,41%	2.446	2.445
1984-1	2.407	-7,0	-0,29%	2.436	2.435
1984-2	2.419	12,0	0,50%	2.429	2.428
1984-3	2.425	6,0	0,25%	2.441	2.440
1984-4	2.418	-7,0	-0,29%	2.447	2.446
1985-1	2.410	-8,0	-0,33%	2.440	2.439
1985-2	2.415	5,0	0,21%	2.432	2.431
1985-3	2.444	29,0	1,20%	2.437	2.436
1985-4	2.442	-2,0	-0,08%	2.466	2.465

Table A11 Continued

Quarter	BCI	Differenciated BCI	Percent Change BCI	Prediction by *Dave*	Prediction by *Pave*
1986-1	2.444	2,0	0,08%	2.464	2.463
1986-2	2.477	33,0	1,35%	2.466	2.465
1986-3	2.500	23,0	0,93%	2.499	2.498
1986-4	2.511	11,0	0,44%	2.522	2.522
1987-1	2.514	3,0	0,12%	2.533	2.533
1987-2	2.524	10,0	0,40%	2.536	2.536
1987-3	2.553	29,0	1,15%	2.546	2.546
1987-4	2.574	21,0	0,82%	2.575	2.575
1988-1	2.579	5,0	0,19%	2.596	2.596
1988-2	2.593	14,0	0,54%	2.601	2.601
1988-3	2.607	14,0	0,54%	2.615	2.615
1988-4	2.615	8,0	0,31%	2.629	2.629
1989-1	2.612	-3,0	-0,11%	2.637	2.638
1989-2	2.618	6,0	0,23%	2.634	2.635
1989-3	2.641	23,0	0,88%	2.640	2.641
1989-4	2.665	24,0	0,91%	2.663	2.664
1990-1	2.668	3,0	0,11%	2.687	2.688
1990-2	2.694	26,0	0,97%	2.690	2.691
1990-3	2.721	27,0	1,00%	2.716	2.717
1990-4	2.726	5,0	0,18%	2.743	2.744
1991-1	2.717	-9,0	-0,33%	2.748	2.750
1991-2	2.722	5,0	0,18%	2.739	2.740
1991-3	2.778	56,0	2,06%	2.744	2.745
1991-4	2.787	9,0	0,32%	2.800	2.802
1992-1	2.786	-1,0	-0,04%	2.809	2.811
1992-2	2.825	39,0	1,40%	2.808	2.810
1992-3	2.852	27,0	0,96%	2.847	2.849
1992-4	2.872	20,0	0,70%	2.874	2.877
1993-1	2.896	24,0	0,84%	2.894	2.897
1993-2	3.038	142,0	4,90%	2.918	2.921

Table A11 Continued

Quarter	BCI	Differenciated BCI	Percent Change BCI	Prediction by *Dave*	Prediction by *Pave*
1993-3	3.020	-18,0	-0,59%	3.060	3.064
1993-4	3.030	10,0	0,33%	3.042	3.046
1994-1	3.098	68,0	2,24%	3.052	3.056
1994-2	3.122	24,0	0,77%	3.120	3.125
1994-3	3.111	-11,0	-0,35%	3.144	3.149
1994-4	3.112	1,0	0,03%	3.133	3.138
1995-1	3.109	-3,0	-0,10%	3.134	3.139
1995-2	3.097	-12,0	-0,39%	3.131	3.136
1995-3	3.115	18,0	0,58%	3.119	3.124
1995-4	3.125	10,0	0,32%	3.137	3.142
1996-1	3.131	6,0	0,19%	3.147	3.152
1996-2	3.162	31,0	0,99%	3.153	3.158
1996-3	3.220	58,0	1,83%	3.184	3.189
1996-4	3.300	80,0	2,48%	3.242	3.248
1997-1	3.329	29,0	0,88%	3.322	3.328
1997-2	3.379	50,0	1,50%	3.351	3.358
1997-3	3.385	6,0	0,18%	3.401	3.408
1997-4	3.364	-21,0	-0,62%	3.407	3.414
1998-1	3.368	4,0	0,12%	3.386	3.393
1998-2	3.376	8,0	0,24%	3.390	3.397
1998-3	3.396	20,0	0,59%	3.398	3.405
1998-4	3.422	26,0	0,77%	3.418	3.425
1999-1	3.418	-4,0	-0,12%	3.444	3.452
1999-2	3.425	7,0	0,20%	3.440	3.447
1999-3	3.479	54,0	1,58%	3.447	3.455
1999-4	3.500	21,0	0,60%	3.501	3.509
2000-1	3.521	21,0	0,60%	3.522	3.530
2000-2	3.548	27,0	0,77%	3.543	3.551
2000-3	3.543	-5,0	-0,14%	3.570	3.579
2000-4	3.545	2,0	0,06%	3.565	3.574

Table A11 Continued

Quarter	BCI	Differenciated BCI	Percent Change BCI	Prediction by *Dave*	Prediction by *Pave*
2001-1	3.541	-4,0	-0,11%	3.567	3.576
2001-2	3.553	12,0	0,34%	3.563	3.572
2001-3	3.609	56,0	1,58%	3.575	3.584
2001-4	3.592	-17,0	-0,47%	3.631	3.640
2002-1	3.586	-6,0	-0,17%	3.614	3.623
2002-2	3.606	20,0	0,56%	3.608	3.617
2002-3	3.652	46,0	1,28%	3.628	3.637
2002-4	3.648	-4,0	-0,11%	3.674	3.683
2003-1	3.651	3,0	0,08%	3.670	3.679
2003-2	3.663	12,0	0,33%	3.673	3.682
2003-3	3.704	41,0	1,12%	3.685	3.695
2003-4	3.756	52,0	1,40%	3.726	3.736
2004-1	3.809	53,0	1,41%	3.778	3.788
2004-2	3.953	144,0	3,78%	3.831	3.842