A GASOLINE DEMAND MODEL FOR THE UNITED STATES LIGHT VEHICLE FLEET

by
DIANA A. REY

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Civil and Environmental Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Spring Term
2009

Major Professor: Haitham Al-Deek

ABSTRACT

The United States is the world's largest oil consumer demanding about twenty five percent of the total world oil production. Whenever there are difficulties to supply the increasing quantities of oil demanded by the market, the price of oil escalates leading to what is known as oil price spikes or oil price shocks. The last oil price shock which was the longest sustained oil price run up in history, began its course in year 2004, and ended in 2008. This last oil price shock initiated recognizable changes in transportation dynamics: transit operators realized that commuters switched to transit as a way to save gasoline costs, consumers began to search the market for more efficient vehicles leading car manufactures to close assembly plants producing low mileage vehicles, and the government enacted a new law entitled the Energy Independence Act of 2007, which called for the progressive improvement of the fuel efficiency indicator of the light vehicle fleet up to 35 miles per gallon in year 2020. The past trend of gasoline consumption will probably change; so in the context of the problem a gasoline consumption model was developed in this thesis to ascertain how some of the changes will impact future gasoline demand.

Gasoline demand was expressed in oil equivalent million barrels per day, in a two steps Ordinary Least Square (OLS) explanatory variable model. In the first step, vehicle miles traveled expressed in trillion vehicle miles was regressed on the independent variables: vehicles expressed in million vehicles, and price of oil expressed in dollars per barrel. In the second step, the fuel consumption in million barrels per day was regressed on vehicle miles traveled, and on the fuel efficiency indicator expressed in miles per gallon.

The explanatory model was run in EVIEWS that allows checking for normality, heteroskedasticty, and serial correlation. Serial correlation was addressed by inclusion of autoregressive or moving average error correction terms. Multicollinearity was solved by first differencing. The 36 year sample series set (1970-2006) was divided into a 30 years sub-period for calibration and a 6 year “hold-out” sub-period for validation. The Root Mean Square Error or RMSE criterion was adopted to select the “best model” among other possible choices, although other criteria were also recorded.

Three scenarios for the size of the light vehicle fleet in a forecasting period up to 2020 were created. These scenarios were equivalent to growth rates of 2.1, 1.28, and about 1 per cent per year. The last or more optimistic vehicle growth scenario, from the gasoline consumption perspective, appeared consistent with the theory of vehicle saturation. One scenario for the average miles per gallon indicator was created for each one of the size of fleet indicators by distributing the fleet every year assuming a 7 percent replacement rate. Three scenarios for the price of oil were also created: the first one used the average price of oil in the sample since 1970, the second was obtained by extending the price trend by exponential smoothing, and the third one used a longtime forecast supplied by the Energy Information Administration. The three scenarios created for the price of oil covered a range between a low of about 42 dollars per barrel to highs in the low 100’s.

The 1970-2006 light vehicle fleet gasoline consumption trend was extended to year 2020 by ARIMA Box-Jenkins time series analysis, leading to a gasoline consumption value of about 10 millions barrels per day in year 2020. This trend line was taken as the reference or baseline of gasoline consumption. The savings that resulted by application of the explanatory variable OLS model were measured against such a baseline of gasoline consumption.

Even on the most pessimistic scenario the savings obtained by the progressive improvement of the fuel efficiency indicator seem enough to offset the increase in consumption that otherwise would have occurred by extension of the trend, leaving consumption at the 2006 levels or about 9 million barrels per day.

The most optimistic scenario led to savings up to about 2 million barrels per day below the 2006 level or about 3 millions barrels per day below the baseline in 2020. The "expected" or average consumption in 2020 is about 8 million barrels per day, 2 million barrels below the baseline or 1 million below the 2006 consumption level. More savings are possible if technologies such as plug-in hybrids that have been already implemented in other countries take over soon, are efficiently promoted, or are given incentives or subsidies such as tax credits.

The savings in gasoline consumption may in the future contribute to stabilize the price of oil as worldwide demand is tamed by oil saving policy changes implemented in the United States.

ACKNOWLEDGMENTS

My special thanks to my advisor Dr. Haitham Al-Deek who properly guided me along the way. Besides proper guidance, Dr.Al-Deek supported me with patience, preoccupation, interest, dedication, and prompt response.

Special thanks also to Ravi Chilakamarri, Transportation Research Assistant, who together with Dr. Al-Deek provided continued assistance throughout the development of this work.

I express also thanks to the very good friends I made at UCF specially Kirolos Haleem Mesiha, Ana Maria Almonte, and Lamia Beroual.

Because of the valuable teachings the two transportation professors I had at UCF Dr. Mohamed Abdel-Aty, and Dr. Haitham Al-Deek deserve my gratitude. My appreciation is also acknowledged to Dr. Ahmed Khalafallah, who with Dr. Mohamed Abdel-Aty, and Dr. Haitham Al-Deek constitute the members of my committee.

I consider the support provided by my parents and my two brothers also extremely important since they provided me a lot of encouragement, and good advice along the years so that I could achieve the conclusion of my graduate studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS/ABBREVIATIONS

AC	Autocorrelation Coefficient
ACELA	Northeast United States High Speed Train Service
ADF	Augmented Dickey- Fuller Root Test
AIC	Akaike Information Criteria
AMTRAK	National Railroad Passenger Corporation
ANWR	United States Arctic National Wildlife Refuge Zone
API	American Petroleum Institute
APTA	American Public Transportation Association
AR(1)	First Order Autoregressive Process
AR(2)	Second Order Autoregressive Process
ARIMA	Autoregressive Integrated Moving Average Models
BP	British Petroleum Corporation
BTS	Bureau of Transportation Statistics
C1	Regression Coefficient Number One
C2	Regression Coefficient Number Two
C3	Regression Coefficient Number Three
CAFÉ	Corporate Average Fuel Standards
CALCARS	California Energy Commission Modeling Software
DLOGPRICE	First Difference of the Logarithm of PRICE
DMBD	First Difference of MBD
DMPG	First Difference of MPG

DMPGxVEH	First Difference of MPGxVEH
DOE	United States Department of Energy
DOT	United States Department of Transportation
DPRICE	First Difference of PRICE
DVEH	First Difference of VEH
DVMT	First Difference of VMT
DW	Durbin-Watson Statistics
e	Base of Natural Logarithms
EIA	United States Energy Information Administration
EVIEWS	Quantitative Micro Software LLC Statistical Software
F	F Statistics
Fc	Critical F Statistics
FFV	Flex Fuel Vehicles
GDPC	Gross Domestic Product Per Capita
GHG	Green House Gas Effect
IARR	Inverted Auto-Regressive Root
IMAR	Inverted Moving Average Root
JB	Jarque-Bera Statistics
LOG	Natural Logarithm
MA	Moving Average
MA(1)	First Order Moving Average Process
MA(2)	Second Order Moving Average Process
MAPE	Mean Absolute Percentage Error

MBD	Million Barrels per Day
MBD_111F	MBD OLS Regression Model on VMT_11F and MPG1
MBD_112F	MBD OLS Regression Model on VMT_12F and MPG1
MBD_113F	MBD OLS Regression Model on VMT_13F and MPG1
MBD_221F	MBD OLS Regression Model on VMT_21F and MPG2
MBD_222F	MBD OLS Regression Model on VMT_22F and MPG2
MBD_223F	MBD OLS Regression Model on VMT_23F and MPG2
MBD_331F	MBD OLS Regression Model on VMT_31F and MPG3
MBD_332F	MBD OLS Regression Model on VMT_32F and MPG3
MBD_333F	MBD OLS Regression Model on VMT_33F and MPG3
MBD_TESTBOT	MBD Model Lower 95 Percent Confidence Interval
MBD_TESTTOP	MBD Model Upper 95 Percent Confidence Interval
MBDF_TEST	MBD Model Performance in the Testing Period
MBDF_TRAIN	MBD Model Performance in the Training Period
MBDFA	MBD ARIMA Trend Model
MBDFA_BOT	MBD ARIMATrend Model Lower 95 percent Confidence Interval
MBDFA_TOP	MBD ARIMATrend Model Upper 95 percent Confidence Interval
MBDFG	MBD Constant Growth Rate Trend Model
MBDFL	MBD Simple Linear Trend Model
MBDFR	MBD Random Walk Trend Model
MILES	United States Total Lane Miles Length

MPAX	Million Passengers
MPG	Miles per Gallons
MPG1	MPG Scenario Number One
MPG2	MPG Scenario Number Two
MPG3	MPG Scenario Number Three
MPGxVEH	MPG times VEH
MPM	Million Passenger Miles
MTBE	Methyl Tertiary Buthyl Esther
NEMS	National Energy Modeling System
OLS	Ordinary Least Square
ORNL	Oak Ridge National Laboratory
p	p Statistics Coefficient
PAC	Partial Autocorrelation Coefficient
POP	United States Population
PRICE	Oil Price
PRICE1	PRICE Scenario Number One
PRICE2	PRICE Scenario Number Two
PRICE3	PRICE Scenario Number Three
PROB	Probability on the JB test
r	Constant Growth Rate Coefficient
R^2	Coefficient of Determination
RMSE	Root Mean Square Error
RWM	Random Walk Model

SCHWARTZ	Schwartz Criteria
SUV	Sport Utility Vehicle
SYSTAT	Systems and Statistical Software
VEH	United States Total Light Vehicle Fleet
VEH1	VEH Scenario Number One
VEH2	VEH Scenario Number Two
VEH3	VEH Scenario Number Three
VEHP	Vehicles over Population Ratio
VEHPF	VEHP Model
VMT	Vehicle Miles Traveled
VMT_11F	VMT OLS Regression Model on VEH1 and PRICE1
VMT_12F	VMT OLS Regression Model on VEH1 and PRICE2
VMT_13F	VMT OLS Regression Model on VEH1 and PRICE3
VMT_21F	VMT OLS Regression Model on VEH2 and PRICE1
VMT_22F	VMT OLS Regression model on VEH2 and PRICE2
VMT_23F	VMT OLS Regression Model on VEH2 and PRICE3
VMT_31F	VMT OLS Regression Model on VEH3 and PRICE1
VMT_32F	VMT OLS Regression Model on VEH3 and PRICE2
VMT_33F	VMT OLS Regression Model on VEH3 and PRICE3
VMT_TESTBOT	VMT Model Lower 95 Percent Confidence Interval
VMT_TESTTOP	VMT Model Upper 95 Percent Confidence Interval
VMTF_TEST	VMT Model in the Testing Period

VMTF_TRAIN	VMT Model in the Training Period
WHITE TEST	White Heteroskedasticity Test

CHAPTER 1. INTRODUCTION

1.1 Background

The importance of the transportation sector in the economy is enormous, and for the most part the transportation sector energy comes from oil. In 1970, the United States oil production peaked at 11.30 Million Barrels per Day (MBDs); that same year the U.S consumed 14.70 MBDs leading to a deficit of 3.40 MBDs. In 2006, the last year for which complete data is available, U.S oil production stood at 6.84 MBD, but the U.S. consumed 20.45 MBDs[1], therefore the deficit increased to 13.66 MBD. Consequently, in the period of 36 years the U.S has almost quadrupled the need for imported oil as the trade deficit has advanced at great pace (Jaffe 2008; Lucian et al. 2007).

One of the reasons why the U.S. trade deficit is so huge is related to the enormous amount of imported oil needed to sustain transportation and other economic activities (Emerson 2007). The reason for the need of such immense quantity of imported fuel is that as the rate of oil production has been decreasing since 1970, the rate of consumption has been increasing over time, as indicated above.

Almost 67 percent (or approximately 2/3) of the total oil consumed in the U.S in year 2006 was consumed by the transportation sector. The rest was allocated to the residential, commercial, and industrial sectors, including a small amount still needed for electric power generation. From the approximately 14 MBD consumed in transportation, almost 66 percent (or approximately 2/3) were consumed by cars and light trucks, making

[1] The data on oil consumption mentioned in this section is contained in Table 3.1 Section 3.1

the light vehicle fleet responsible for about half the total oil consumed daily in the United States .

Oil is the most important commodity. A rapid sudden increase in the price of oil is sometimes called an “oil spike” or an “oil price shock”. Since 1970, world energy markets had experienced five oil price shocks, and the last one began its course in year 2004 (Roubini and Setser 2004), and ended in 2008. Although the causes for the last oil price spike are complex and are currently being debated; they can be summarized as follows (Jaffe 2008):

I. In periods of high economic activity, some times there are difficulties for the oil supply to keep up with the world’s increasing oil demand;

II. The oil reserves of the world are being depleted at a greater pace; and

III. Speculations in the oil commodity markets.

The rapid increase in oil prices that began in the year 2004 is changing transportation dynamics and the interaction among different transportation modes. For example, during the past two years AMTRAK has seen substantial increases in both short haul commuters and long haul passenger travel. In Fiscal Year 2008 Amtrak increased ridership 11 percent over the figure for Fiscal Year 2007. The ACELA high speed train that serves the New York-Washington D.C. leg, and the Keystone Service that serves the route Harrisburg-Philadelphia-New York reported increases of up to approximately 20 percent during the last year. Also, large metropolitan areas of the United States have detected renewed interest in transit transport and many extensions and renewal projects are taking place across the nation (Bolte 2008). This indicates that some kind of transfer from personal car mode of transport is being shifted to the railroad mode of transportation.

The transportation dynamic changes that are occurring are affecting gasoline consumption. In addition to oil dependence, and oil price shocks the excessive gasoline consumption of oil is a factor in global warming (Green House Gas or GHG effect) and Environmental Pollution (Feng and Sauer 2004). It has also been recognized that the United States is presently in need of controlling and if possible diminishing gasoline demand, as clearly stated in the Energy Independence and Security Act of 2007 signed into law on December 18, 2007 (Emerson 2007).

1.2 Objectives of the Study

The United States light vehicle fleet consumes the largest chunk of the total amount of oil used in the United States. Therefore, it seems plausible to understand gasoline demand by examining in more detail the oil consumption patterns of the light vehicle transportation sector over the years. A gasoline demand model was proposed to be developed to that effect.

Gasoline demand models help identifying a convenient assortment of variables shaping gasoline demand, establishing fuel demand trends, and serving forecasting purposes.

The need to reduce gasoline consumption of oil in the United States has been widely acknowledged (Emerson 2006; Energy Policy Research Foundation 2008); and is now mandated by law. Therefore, the new light vehicle fleet standards required by the Energy Independence and Security Act of 2007 (The White House Press 2007), were input into the model developed to ascertain the possible impact upon gasoline consumption of the light vehicle fleet efficiency requirements of that law.

In summary the main three objectives of this thesis were:

I. To utilize time series methods and historical oil consumption data to evaluate the trend in gasoline consumption.

II. To develop a two steps explanatory variable model of gasoline consumption.

III. And, to evaluate possible effects of the new Corporate Average Fuel Standards (CAFÉ) on the United States light vehicle fleet gasoline consumption.

1.3 Methodology

To detect the trend of gasoline consumption, the time series for the gasoline consumption of oil expressed in equivalent million barrels per day was treated accordingly. The time series was processed by several methods like: simple linear trend, constant growth, random walk, and ARIMA Box-Jenkins (Vandaele 1983).

Once the gasoline demand trend was established, it was necessary to formulate an explanatory type variable model which in this work was dealt with by Ordinary Least Square (OLS) regression techniques (Gujarati 2003).

The variables to be manipulated were chosen among several transportation demand indicators identified by the Office of Energy Markets and End Use of the Energy Information Administration (Energy Information Administration 1995).

Some of the information needed has been compiled by the Oak Ridge National Laboratory (Davis and Diegel 2008), or is contained in transportation or energy statistics databases available through the United States Department of Energy (www.energy.gov), the Energy Information Administration (www.eia.gov), the American Petroleum Institute (www.api.gov), the British Petroleum Corporation (www.bp.com), the United States Bureau of Transportation Statistics (www.bts.org), the Federal Highway Administration (www.fhwa.gov), and the United States Census Bureau (www.uscensus.gov).

The statistical software tool used for data analysis was EVIEWS software provided by Quantitative Micro Software LLC (www.eviews.com) which has capabilities for time series processing including ARIMA Box-Jenkins and OLS Regression techniques manipulation. According to Studenmund (Studenmund 2000), EVIEWS is "the number one Windows based econometric software package in the world".

1.4 Significance of the Study

The United States daily oil consumption of approximately 20 MBDs represents roughly 25 percent of the world oil production creating the single most important source of demand pressure on the world oil commodity market (Medlock and Jaffe 2007). In addition, during the last ten years high rates of economic growth in China, India, other Asian countries, some of the former Soviet Union countries, and some of the Eastern European countries, have added unusual pressure on the demand for crude oil (Energy Policy Research Foundation 2008).

Oil consumption of gasoline by cars and light trucks also generates toxic emissions which constitute a health hazard and Green House Gas effects (GHG) that contribute to global warming (Fay and Golomb 2002). It also obligates the United States to import much of its oil requirements; contributing to the chronic U.S trade deficit and makes the country import oil dependent. Therefore, controlling future gasoline demand is of utmost importance to U.S policy makers and of interest to other transportation stakeholders.

A gasoline demand model helps understanding gasoline consumption patterns of the United States, and the effect of main factors and policies shaping future gasoline demand.

1.5 Thesis Organization

This work is organized as follows: the next chapter contains a literature review concerning gasoline demand models, Chapter Three complements some important statements mentioned briefly in the introduction concerning fuels for transportation, and defines the variables that are dealt with in later chapters. Chapter Four considers some time series models to establish the trend or base line of gasoline demand. Chapter Five develops the explanatory variable type model. Chapter Six hypothesizes some possible trend scenarios for the independent variables. Chapter Seven attempts to forecast gasoline demand up to year 2020. And, finally Chapter Eight draws and summarizes the conclusions of this thesis.

CHAPTER 2. LITERATURE REVIEW

According to Makridakis et al. (Makridakis et al. 1998), demand models may be classified in two types: time series models, and explanatory variable type models.

Uni-variate time series models try to present demand to establish trends based on the behavior of available past time series data and any statistical errors or stochastic variations that may occur. This type of demand models makes no consideration of any other factors affecting demand different from the past demand itself. Because of that they are also called uni-variate demand models (Wei 2006).

Explanatory variables type models look to express the demand or dependent variable as a function of some independent variables. Those independent variables, in the case of this thesis, were selected from a group of possible variables known as transportation demand indicators. These transportation demand indicators, as previously mentioned, were defined by the Office of Energy Markets and End Use of the United States Department of Energy (Energy Information Administration 1995).

Both types of models, time series uni-variate models and explanatory variable type models are found in the gasoline demand model literature. When the models refer to the individual household they are called disaggregate models and aggregate when they refer to a whole region, country or group of countries. Other distinction deals with the specification of the model: in general if the main objective is finding the price elasticity of demand (the relation between the increase in price and the decrease in the amount demanded) the model is specified as a log-log type of model, whereas if the objective is obtaining a demand projection or forecasting then the ordinary type of linear model is specified with no transformations. Also, there is a great variation in the group or assortment of dependent variables chosen to explain the dependent variable, each author choosing the group of variables as he or she best sees fit.

Pock (Pock 2005) based his gasoline demand model on an old previous model developed by Baltage and Griffin in 1968. Gasoline demand is presented in a logarithmic expression as a function of income, price, car ownership, and vehicle efficiency.

Data was processed in a Generalized Linear System model in STATA statistical software which allows checking for heteroskedasticity, normality of residuals, and correction for serial correlation. The model main interest was on the estimation of elasticities for income, price of energy, and car ownership per household. The findings were that income elasticity is positive, price elasticity is negative, and increased car ownership per household is negative on gasoline consumption.

Emerson (Emerson 2006) noted that gasoline demand can be expressed as a function of how many miles all cars travel or Vehicle Miles Traveled (VMT) , and the fuel economy or how far can one go on one gallon of gasoline or Miles Per Gallon (MPG). VMT may depend in turn on many variables such as fuel price, income, car registration, and even weather, and driving patterns. However, instead of formulating a regression equation on some of those variables, Emerson adopted an Energy Information Administration VMT forecast growth rate of about 2 percent per year, and then proceeded to formulate some hypothesis over possible improvements of the vehicle efficiency variable beyond 20 MPG.

Although the approach of modeling gasoline demand only by the existing relationship between VMT and MPG, thus implicitly holding all other variables constant is simple in nature, the procedure gave quantifiable indication of the impact and importance, the variable of vehicle efficiency or MPG indicator may have upon gasoline demand.

The gasoline demand model developed by Wiggins (Wiggins 2003) was focused on the selection of a group of variables that may explain gasoline consumption in the United States, and that may be used as predictors for the short-run future. The variables tested were: income, lane-miles, tax, drivers, mpg, crude price, war, and speed. An OLS regression model was created and run in SHAZAM statistical software. Wiggins found that the variables lane-miles, tax, war, and speed were statistically insignificant factors in gasoline consumption. Finally, the model was expressed in terms of income, the number of drivers registered, mpg, and crude oil price.

Hughes et al. (Hughes et al. 2008) focused their study on finding how the changes in oil price, and disposable income affected the consumption of gasoline during the periods 1975-1980 and 2001-2006; hence the objective was finding the corresponding elasticities for these periods. The calculated elasticity for price for the periods analyzed differs considerably from -0.034 to -0.077 during the period 2001-2006, versus -0.21 to -0.34 for the period 1975-1980. One important conclusion of the study was that U.S drivers have turned in later periods less responsive in adjusting to high gasoline price increases than in previous decades; the authors hypothesized that consumers are now more dependent on automobiles for daily transportation than during previous periods, therefore less able to reduce VMT as gasoline price increases, possibly because of the urban sprawl and less transit availability than in earlier decades.

The objective of the model put forth by Banazak et al. (Banaszak et al. 1999), was twofold: finding elasticities, and forecasting fuel demand. Fuel demand was expressed in a log-log model as a function of price of fuel, gross domestic product per capita, and the consumption of fuel in the previous year (the lag dependent variable used as independent variable). Both consumption of gasoline and consumption of fuel oil were considered for Korea and Taiwan. Forecasting proceeded thru several scenarios of gross domestic product per capita, and future price of fuel for Korea and Taiwan. The purpose was to make comparisons between the two countries and projections to help stakeholders plan the development of needed infrastructure such as new refineries needed to keep up with future gasoline and fuel oil demand.

As required by the California Senate (Page et al. 2007), the California Energy Commission regularly conducts forecasts of all aspects of energy consumption including fuel demand and fuel prices for transportation. The State of California utilizes a proprietary exclusive software called the "CALCARS Demand Model" which has capabilities to forecast vehicle ownership, VMT, gasoline and Diesel fuel demand, and the potential impact of government policies related to transportation. The input to the software consists of variables such as fuel prices, population, and vehicles by type and quantity. CALCARS also evaluates the impact of public policy on light vehicles petroleum demand, to develop strategies to reduce California's dependence of oil, to promote the use of alternate fuels, and to determine the effect upon gasoline consumption of new vehicle technologies. The software allows the State of California to forecast fuel demand under different scenarios of fuel concerning domestic oil production and imports.

The State of New York (New York State 2002) through the New York State Energy Planning Board (www.nyserda.org) modeled energy demand based upon regional projections developed by the U.S Department of Energy (DOE) Energy Information Administration (EIA) which in turn utilized the National Energy Modeling System or NEMS proprietary software (Energy Information Administration 2003).

The Middle Atlantic Region includes the states of Pennsylvania, New York, and New Jersey. Energy demand including coal, natural gas, fuel oil, jet fuel, and motor gasoline were taken for the State of New York as a fraction of the energy demand of the Middle Atlantic Region. Projections for the State of New York were made with the ARIMA time series module of the SYSTAT software.

Berkowitz et al. (Berkowitz et al. 1990) modeled gasoline demand at the household level as a function of the number and type of vehicle holdings, and vehicle usage (discretionary and non- discretionary) with the main objective of estimating the price elasticity of gasoline demand, and the elasticity of fuel efficiency. Cross sectional Canadian household data was collected thru a survey in Fall 1982 through a specially prepared questionnaire mailed to 2400 Canadian households, from which approximately 2000 answers were received and processed. Utility functions were extensively utilized in this approach to conclude that vehicle usage at the household level is insensitive to improvements in fuel efficiency.

The value of the independent variables in the models change rapidly over time; therefore model projections or conclusions become obsolete, making it necessary to update or redo the models as new data or new conditions become available or develop over time (Makridakis et al. 1998). Some of the models described above used proprietary exclusive software (CALCARS, NEMS) for interest only of specific audiences (California, New York Governments); although as noted by the examples above some models were set by way of commercially available software (SHAZAM, STATA, SYSTAT) making them available to larger audiences.

The explanatory variable model developed in this thesis utilized the last version of the widely known statistical software tool EVIEWS which is especially helpful to manipulate time series data and, as mentioned before, has ARIMA Box Jenkins processing capabilities. The thesis objective was to develop an explanatory variable OLS type model capable of projecting future gasoline demand and compare the forecast obtained with this model with a trend extension or reference line obtained by ARIMA.

CHAPTER 3. SOME ENERGY RELATED FACTORS IN LIGHT VEHICLE TRANSPORTATION

3.1 Oil Consumption for Transportation in the United States

The data for transportation energy consumption for the different transportation modes in the United States has been compiled by the Center for Transportation Analysis of the Oak Ridge National Laboratory (Davis and Diegel 2008) and it is presented in units of thousand barrels per day for the period 1970-2006 in Table 3-1.

Table 3-1: U.S. Oil Consumption and Production (Thousand Brrels/Day)

Year	Light Veh.	Bus & Trucks	Rail	Air	Water	Other	Transportation	Total	Production
1970	5,227	800	253	625	383	47	7,335	14,700	11,300
1971	5,534	833	260	630	391	49	7,697	15,570	11,100
1972	5,942	875	259	635	400	51	8,162	16,440	11,080
1973	6,201	918	258	640	408	53	8,478	17,310	10,840
1974	5,929	968	257	646	417	55	8,272	16,650	10,400
1975	6,081	1,010	249	651	425	57	8,473	16,320	10,010
1976	6,466	1,068	260	624	494	59	8,971	17,510	9,780
1977	6,617	1,179	265	655	536	62	9,314	18,430	9,890
1978	6,837	1,313	264	691	626	62	9,793	18,850	10,440
1979	6,591	1,367	270	723	721	55	9,727	18,510	10,180
1980	6,117	1,370	262	697	627	48	9,121	17,100	10,170
1981	6,054	1,398	253	706	724	43	9,178	16,060	10,180
1982	5,989	1,401	214	701	606	34	8,945	15,300	10,200
1983	6,149	1,426	212	699	562	30	9,078	15,230	10,250
1984	6,280	1,467	232	781	579	27	9,366	15,770	10,530
1985	6,450	1,469	216	814	579	25	9,553	15,720	10,580
1986	6,670	1,502	210	884	577	29	9,872	16,280	10,230
1987	6,778	1,546	213	920	588	27	10,072	16,670	9,950
1988	6,914	1,575	220	958	595	31	10,293	17,340	9,970
1989	6,992	1,613	221	960	611	32	10,429	17,410	9,160
1990	6,861	1,675	216	1,006	657	26	10,441	16,840	8,910
1991	6,689	1,713	202	940	692	24	10,260	17,030	9,080
1992	6,938	1,747	208	954	726	22	10,595	17,000	8,880
1993	7,169	1,797	215	961	654	24	10,820	17,440	8,590
1994	7,305	1,893	230	1,002	636	24	11,090	17,330	8,390
1995	7,415	1,968	239	1,036	669	20	11,347	17,900	8,320
1996	7,604	2,019	245	1,068	645	21	11,602	18,440	8,300
1997	7,781	2,040	246	1,114	575	22	11,778	18,470	8,270
1998	7,968	2,105	248	1,148	567	25	12,061	18,860	8,010
1999	8,228	2,308	257	1,196	626	25	12,640	19,460	7,730
2000	8,219	2,396	256	1,234	663	24	12,792	19,690	7,730
2001	8,290	2,388	257	1,167	547	24	12,673	19,570	7,670
2002	8,525	2,492	257	1,071	573	20	12,938	19,670	7,630
2003	8,829	2,424	263	1,073	497	22	13,108	19,910	7,400
2004	9,055	2,254	278	1,136	597	23	13,343	20,640	7,230
2005	8,890	2,519	281	1,199	626	22	13,537	20,630	6,900
2006	8,848	2,566	285	1,208	664	19	13,590	20,450	6,840

Source: David C. Stacy, S. Diegel, and R. G. Bounty, Transportation Energy Databook, Edition 27; Oak Ridge National Laboratory, National Transportation Research Center, Knoxville TN; 2008.

Table 3-1 also presents the total oil consumed in transportation activities (labeled "Transportation") as well as the total oil consumption (labeled "Total"). Total oil consumption includes besides transportation: residential, industrial, commercial, and a small amount still needed for electric power generation. Transportation oil consumption is divided into light vehicles (cars and light trucks) which use motor gasoline conventional or reformulated, trucks and buses which use diesel fuel, rail transportation (also diesel) , air which consumes aviation and jet fuel, and water transportation which consumes diesel fuel. The last classification in transportation fuel consumption, which is named 'Other" in the table, includes fuel for the operation of pipelines and gasoline consumed by motorcycles.

The table also includes data for oil production for the same period. Oil production is understood as coming from two sources which are: crude oil extracted directly from oil wells and natural gas plant liquids or condensates obtained in the processing of natural gas. The increased dependence of the United States from imported oil is clearly appreciated when the figures for total consumption and production for year 1970 are compared with those for year 2006. The proportions of oil consumed by light vehicles, buses and trucks, rail, aviation, and other consumption, to the total consumed in transportation are presented in Table 3-2. The variables in the table were called as follows:

Year: Designation of the Year the data was observed

LVTT: Light Vehicles/ Transportation Ratio

BTT: Buses and Trucks/Transportation Ratio

RTT: Rail/Transportation Ratio

ATT: Air/Transportation Ratio

WTT: Water/Transportation Ratio

OTT: Other/ Transportation Ratio

LVTALL: Light Vehicles/ Total Oil Consumption Ratio

TTTALL: Total Transportation/ Total Oil Consumption Ratio

TT: Transportation Oil Consumption

TALL: Transportation + Residential + Commercial + Power Generation

Table 3-2: U.S Transportation Oil Consumption (Percentages)

Year	LVTT	BTT	RTT	ATT	WTT	OTT	LVTALL	TTTALL
1970	71.26	10.91	3.45	8.52	5.22	0.64	35.56	49.90
1971	71.90	10.82	3.38	8.19	5.08	0.64	35.54	49.43
1972	72.80	10.72	3.17	7.78	4.90	0.62	36.14	49.65
1973	73.14	10.83	3.04	7.55	4.81	0.63	35.82	48.98
1974	71.68	11.70	3.11	7.81	5.04	0.66	35.61	49.68
1975	71.82	11.90	2.93	7.67	5.01	0.67	37.35	52.01
1976	72.14	11.88	2.89	6.94	5.49	0.66	37.04	51.35
1977	71.12	12.63	2.84	7.01	5.74	0.66	36.03	50.67
1978	69.90	13.37	2.69	7.04	6.37	0.63	36.42	52.10
1979	67.85	14.01	2.77	7.41	7.39	0.56	35.76	52.70
1980	67.17	14.97	2.86	7.62	6.85	0.52	35.93	53.50
1981	66.47	14.67	2.76	7.71	7.91	0.47	37.87	56.98
1982	67.02	15.63	2.39	7.82	6.76	0.38	39.26	58.58
1983	67.82	15.61	2.34	7.70	6.19	0.33	40.40	59.57
1984	67.05	15.66	2.48	8.34	6.18	0.29	39.82	59.39
1985	67.49	15.39	2.26	8.53	6.07	0.26	40.98	60.72
1986	67.54	15.23	2.13	8.96	5.85	0.29	40.92	60.59
1987	67.27	15.36	2.12	9.14	5.84	0.27	40.62	60.38
1988	67.15	15.31	2.14	9.31	5.78	0.30	39.83	59.32
1989	67.03	15.47	2.12	9.21	5.86	0.31	40.13	59.87
1990	65.70	16.05	2.07	9.64	6.30	0.25	40.72	61.98
1991	65.17	16.71	1.97	9.17	6.75	0.23	39.23	60.20
1992	65.47	16.50	1.96	9.01	6.86	0.21	40.78	62.30
1993	66.23	16.62	1.99	8.89	6.05	0.22	41.07	62.00
1994	65.85	17.08	2.08	9.04	5.74	0.22	42.11	63.95
1995	65.31	17.36	2.11	9.14	5.90	0.18	41.35	63.32
1996	65.51	17.42	2.11	9.21	5.56	0.18	41.18	62.86
1997	66.02	17.35	2.09	9.47	4.89	0.19	42.04	63.68
1998	66.02	17.48	2.06	9.53	4.71	0.21	42.17	63.87
1999	65.05	18.28	2.04	9.47	4.96	0.20	42.20	64.87
2000	64.19	18.76	2.00	9.66	5.19	0.19	41.63	64.85
2001	65.29	18.95	2.03	9.22	4.32	0.19	42.24	64.69
2002	65.80	19.31	1.99	8.30	4.44	0.15	43.17	65.61
2003	67.21	18.58	2.02	8.22	3.81	0.17	44.05	65.54
2004	67.66	17.00	2.10	8.57	4.50	0.17	43.47	64.24
2005	65.95	18.46	2.06	8.79	4.59	0.16	43.63	66.16
2006	65.64	18.59	2.07	8.75	4.81	0.14	44.29	67.48
Average	67.64	15.49	2.40	8.50	5,62	0.35	39.79	59.00
Maximun	73.14	19.31	3.45	9.64	7.91	0.67	44.29	67.48
Minimun	64.19	10.72	1.96	6.94	3.81	0.14	35.54	48.98

The averages are: approximately 68 percent for light vehicles; second is trucks and buses with 15.5 percent, third is aviation with 8.5 percent, water transportation has consumed on average 6 percent; and rail, the most efficient mode of transportation in terms of energy consumption, has taken on average only 2.5 percent of the total fuel spent in transportation activities.

Another interesting fact observed in Table 3-2 is that although light vehicles used to take about 71 percent of the oil spent in transportation in 1970, the proportion has been reduced to about 66 percent in 2006, while truck and buses have increased the proportion from 11 to 19 percent in the same period. When comparing the proportion of oil consumed by light vehicles to the total oil consumed, it is seen that the proportion has gone up from 36 percent in 1970 to 44 percent in year 2006 with an average of 40 percent in the 36 years period. Whatever the numbers and proportions are, by all accounts, light vehicle transportation stands as the most important mode of transportation in terms of fuel demand in the United States.

The data presented in Table 3-1 is also shown in Figure 3-1. To illustrate the relative importance of each transportation mode, the average percent of fuel consumption in Table 3-2 for each one of the series is also indicated in the same figure. The oil production data is also depicted in the same figure to illustrate the contribution of transportation to the increasing oil deficit along time.

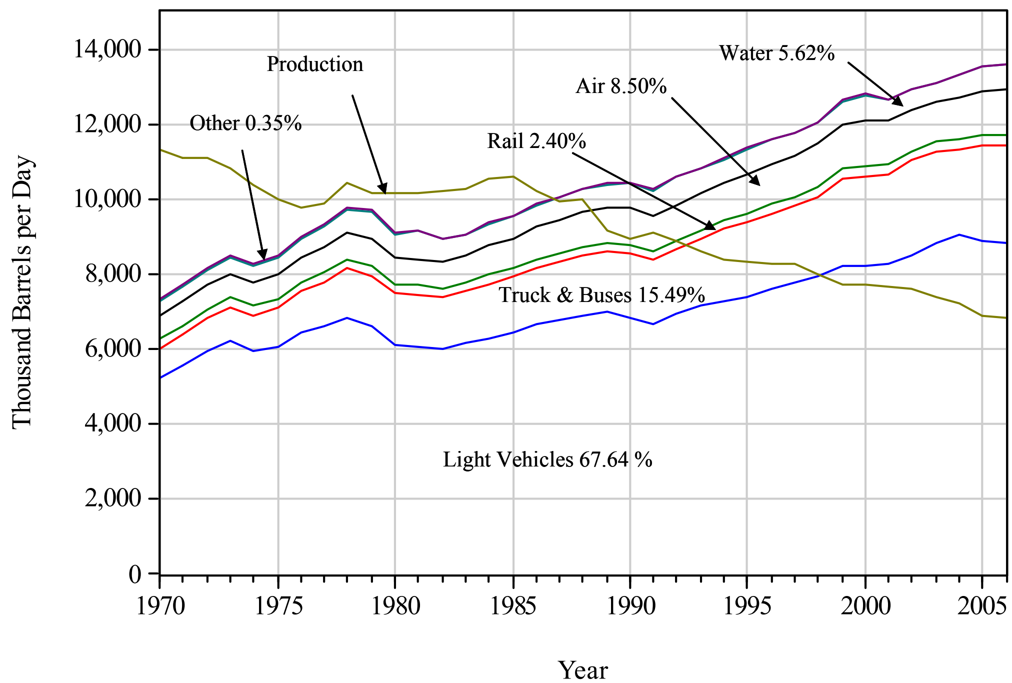


Figure 3-1: U.S Oil Production and Transportation Consumption

The year 1970 was the year when oil production peaked in the United States (British Petroleum Corporation 2008) and since has been slowly decaying as portrayed in the same figure. Besides the regions of current oil production, there are some untapped conventional oil reserves mostly offshore the Gulf of Mexico and in the zone known as the Alaska National Wildlife Refuge Zone or ANWR. These proved conventional oil reserves are, according to the American Petroleum Institute, capable of supporting 60 million cars for a period of 60 years, but efforts to obtain approval from the Unites States Congress to proceed with further exploration and production activities have failed so far (American Petroleum Institute 2008).

At the above mentioned levels of demand, the United States is the largest consumer of oil, but lately some other countries have experimented high rates of economic growth. For example: China, India, and other Asian countries, have been lately demanding as a consequence increasing amounts of oil.

Whenever there are difficulties to balance increasing demand with short supply, the price of oil increases sharply because the demand of oil is highly price inelastic (a large increase in price is necessary to cause a small quantity decrease in demand). Conversely, high sustained oil prices cause car vehicle owners trying to drive less, look for public transport, or acquire more efficient vehicles; consequently affecting gasoline demand.

3.2 Transportation Dynamics

Transportation stakeholders have recently experienced certain new changes in transportation dynamics. These changes affect the oil consumption patterns of the United States light vehicle fleet. Some of the observed changes are:

a. Passenger diversion to the railroad mode of transport.
b. Increased use of alternative fuels.
c. Incentives for bio-fuel producers and flex-fuel vehicle manufacturers.
d. Market competitiveness and development of other vehicle technologies.
e. The enactment of new CAFÉ standards as of December 2007.

Item “a” affects gasoline consumption because it reduces the number of vehicles on the road; “b” and “c” cause a direct reduction effect upon gasoline demand; and “d” and “e” look to improve the overall fuel efficiency of the light vehicle fleet consequently reducing gasoline requirements. Therefore, these items deserve further discussion.

3.2.1 Passenger Diversion to the Railroad Mode of Transport

The shift of riders to transit because of high oil prices has recently received a lot of attention and is well documented in anecdotic references. Many commuter routes are beyond capacity, new commuters have appeared at light rail stations and at crowded park and ride facilities; and transit operators have received more reports and complaints of crowded buses across the nation. These new happenings have been interpreted as if customers were switching to transit as a way to reduce gasoline consumption costs (American Public Transportation Association 2008).

The time series data in Table 3-3 has been reproduced from the American Public Transit Association database (www.apta.org) for two individual transit modes: light rail and commuter rail in million passenger miles (MPM). The table also contains the series data for all transit modes expressed in million passenger miles as well as in million passengers (MPAX). The last series run from 1970 to 2006 while the others are incomplete. Figure 3-2 is just the graphical depiction of the same data.

Table 3-3: U.S Transit Data Statistics (1970-2006)

Year	Light Rail (MPM)	Commuter Rail (MPM)	Total Transit (MPM)	Total Transit (MPAX)	Oil Price (Dollars/Barrel)
1970	-	-	-	7,332	9.63
1971	-	-	-	6,847	11.49
1972	-	-	-	6,567	12.28
1973	-	-	-	6,660	15.37
1974	-	-	-	6,935	48.66
1975	-	-	-	6,972	44.52
1976	-	-	-	7,081	46.72
1977	389	-	30,026	7,286	47.67
1978	392	-	31,664	7,616	44.65
1979	407	-	32,764	8,130	90.31
1980	381	6,516	39,854	8,567	92.77
1981	346	6,236	38,482	8,284	82.03
1982	379	6,027	37,124	8,052	70.90
1983	391	6,097	37,602	8,203	61.56
1984	416	6,207	39,424	8,829	57.45
1985	350	6,534	39,581	8,636	53.10
1986	361	6,723	40,204	8,777	27.28
1987	405	6,818	40,348	8,735	33.64
1988	477	6,964	40,580	8,666	26.14
1989	509	7,211	41,603	8,931	30.48
1990	571	7,082	41,143	8,799	37.66
1991	662	7,344	40,703	8,575	30.45
1992	701	7,320	40,241	8,501	28.54
1993	705	6,940	39,384	8,217	24.35
1994	833	7,996	39,585	7,949	22.12
1995	860	8,244	39,808	7,763	23.15
1996	957	8,351	41,378	7,948	27.30
1997	1,035	8,038	42,339	8,374	24.67
1998	1,128	8,704	44,128	8,750	16.18
1999	1,206	8,766	45,857	9,168	22.35
2000	1,356	9,402	47,666	9,363	34.29
2001	1,437	9,548	49,070	9,653	28.62
2002	1,432	9,504	48,324	9,623	28.83
2003	1,476	9,559	47,903	9,434	32.50
2004	1,576	9,719	49,073	9,575	42.00
2005	1,700	9,473	49,678	9,815	57.88
2006	1,866	10,361	52,154	10,017	67.02

Source: American Public Transportation Association; "2008 Public Transportation Fact Book", American Public Transportation Association; Washington, DC ; 2008. Also at www.apta.com. British Petroleum Corporation (www.bp.com).

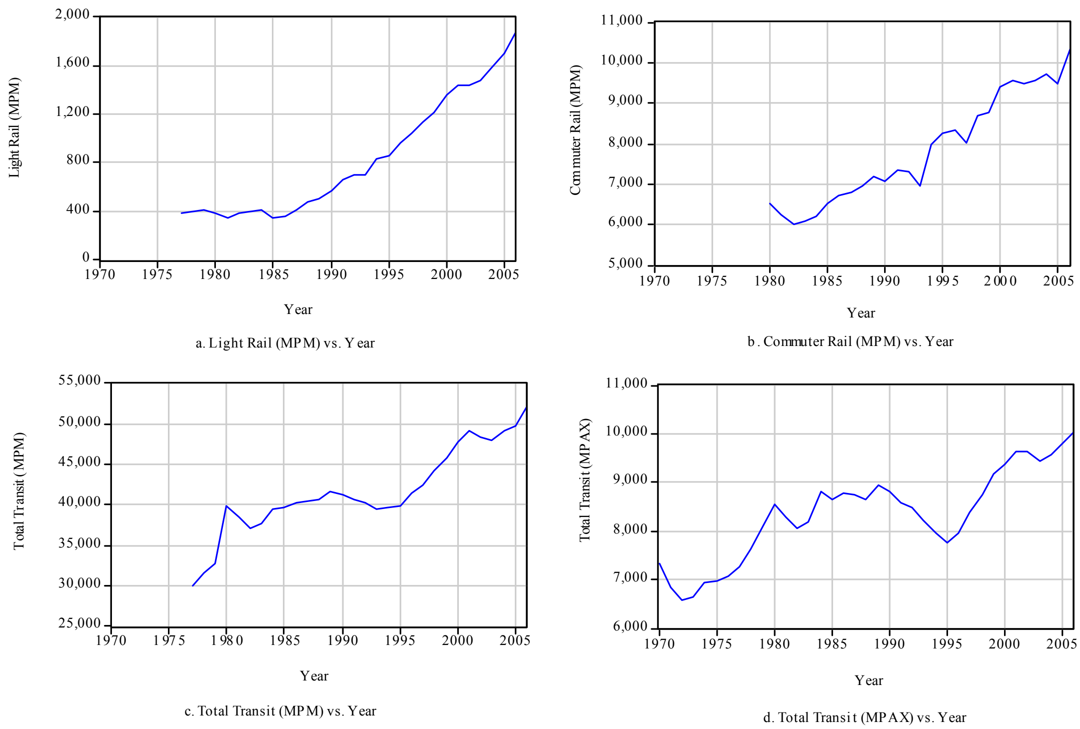


Figure 3-2: Passenger Data for some Transit Modes of Transportation

Light rail initiated a steady positive growth beginning about year 1986 (Figure 3-2a), commuter rail initiated a steady growth about the same time albeit interrupted by some small periods of negative growth (Figure 3-2b), and transit as a whole (all modes including heavy rail and buses; Figures 3-2c and 3-2d) experienced a turning point in 1995. Before 1995, transit as a whole stood at around 40,000 million passenger miles or less. Consumers also benefitted from a period of low oil prices between 1986 and 1995. So, it looks that perhaps because of corresponding low gasoline prices consumers were not very much interested in transit, or at least price might be one of the factors.

Table 3-3 also includes the time series for the price of oil (dollars per barrel) in order to correlate this with the variable transit as a whole (million passenger miles).The price of oil was obtained from the British Petroleum Corporation website (www.bp.com) and was referred to 2007 dollars as per the corresponding Consumer Price Index (Davis et al. 2008). The scatter plot of the total number of million passenger miles (MPM) versus the price of oil in dollars per barrel (PRICE), for the series from 1986 to 2006, is presented as Figure 3-3.

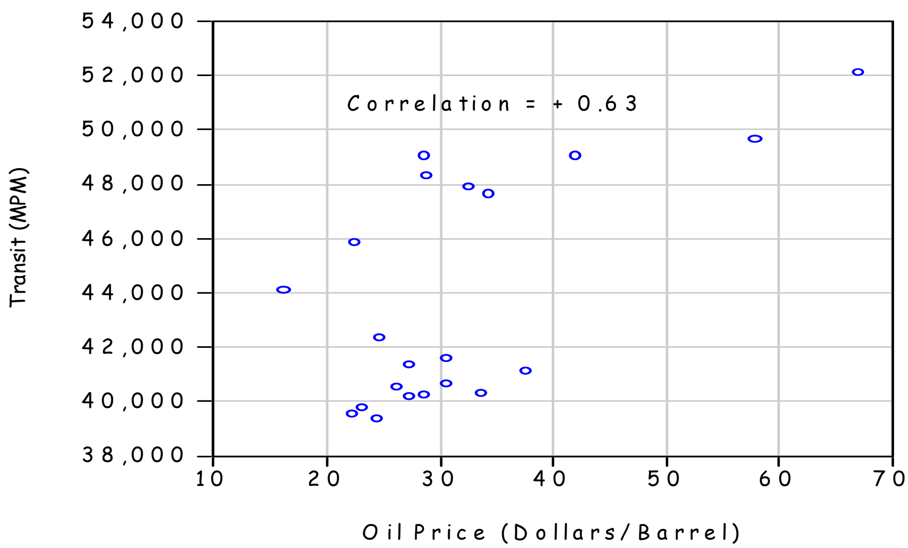


Figure 3-3: Correlation Transit (MPM) vs. Oil Price (Dollars/ Barrel)

This positive correlation statistically supports the evidence that transit ridership increases when oil price increases. However, it is important to realize that the transfer from passenger car travel to transit may only occur where good quality transit service is available as for example in some large metropolitan areas. Furthermore, the positive correlation between transit ridership and the price of oil also works in the opposite direction as some commuters may switch back to driving in times of low oil prices.

Accordingly, from the foregoing analysis the only hypothesis that can be formulated is that oil price (PRICE) may be a useful function to statistically express transit ridership (MPM), as it is so to express light vehicle fleet VMT and/or MBD.

3.2.2 Alternative Fuels

Oil refiners and gasoline blenders had been using a compound named Methyl Tertiary Buthyl Esther, better known as MTBE, to raise the octane number of gasoline. Because of environmental concerns, they are required now to substitute MTBE by ethanol (Medlock and Jaffe 2007). The percentage required to improve the octane number is about 3 percent ethanol by volume, but gasoline engines can tolerate up to 10 percent without any mechanical modifications required. The Energy Independence Act of 2007 (Bush 2007) also took advantage of this fact requiring gasoline refiners, importers, and blenders to progressively increase the percentage of ethanol up to the maximum 10 percent possible in year 2020. Therefore, ethanol producers are required also to progressively increase production up to 15 billion gallons in year 2020 (0.98 MBD) so that enough ethanol may be available (Dietert 2008).

The 15 billion gallons of ethanol mentioned above will be produced from corn, but ethanol can also be extracted from cellulose through an enzymatic process not quite yet established for industrial use. Nevertheless, it is expected that for year 2020 the industrial process to derive ethanol from cellulose will be already fully developed.

Ethanol producers will then be required to produce 15 additional billion gallons of cellulose derived ethanol by year 2020. This ethanol is expected to be demanded by specially manufactured flex-fuel vehicles (FFVs).

3.2.3 Incentives Related to Bio-fuels and FFV Vehicles

To help ethanol producers meet the targets demanded by the law, as well as to encourage the automotive industry to produce FFV vehicles, incentives in the form of tax credits were considered by the same above referred law.

3.2.4 Other Vehicle Technologies

The automotive industry is increasing the production of more efficient vehicles that the market is demanding, hybrids among them. Also, plug-in electric vehicles that are hybrids with batteries that can be charged externally are being tested in the laboratory and on the road (McManus 2006).

3.2.5 New CAFÉ Standards

The Energy Independence Act also requires car manufacturers to increase the overall light vehicle fleet efficiency from the 2007 approximate value of 21 MPG to 35 MPG in 2020. That is a 4 percent annual increase per year.

3.3 Transportation Demand Indicators

The transportation demand indicators listed in Table 3-5 (placed at the end of this chapter, after the variables are defined) were selected for analysis and considered later on in this thesis to develop the time series uni-variate and explanatory variable models.

The oil consumption of gasoline of the light vehicle fleet is expressed in million gallons per day, and the variable is identified as MBD. The price of oil is expressed in dollars per barrel, based on the purchasing power of year 2007, and is identified as

PRICE. The vehicle miles traveled indicator is expressed in trillion vehicle-miles, and is identified as VMT. The vehicle efficiency indicator is expressed in miles per gallon and is identified as MPG. The gross domestic product per capita is expressed in thousand dollars per capita, is based on the 2007 purchasing power (Williamson 2008), and is identified as GDPC. The lane mile indicator is expressed in million lane-miles (Bureau of Transportation Statistics 2007), and is identified as MILES. The size of the light vehicle fleet is expressed in million vehicles, and is identified as VEH. The population of the United States is expressed in million people (U.S. Census Bureau 2008), and is identified as POP. And finally the vehicle to population ratio, which is obtained by dividing VEH by POP, is identified as VEHP.

3.3.1 Oil Price Spikes and Gasoline Consumption

A rapid increase in oil price is often referred to as an oil spike or an oil price shock. Economists have often associated economic recessions or economic slowdowns to oil price shocks. Also, although gasoline is a very inelastic commodity, consumers have invariably responded to oil price spikes with cut downs in gasoline consumption (Energy Policy Research Foundation 2008).

Figure 3-4a, 3-4b, and 3-4c are the graphical representations of the time series PRICE, GDPC, and MBD. The periods of oil price increases are shaded in Figure 3-4a. The periods of economic slowdown or negative GDPC growth, are shaded in Figure 3-4b; and, the periods of decreasing demand for gasoline are shaded in Figure 3-4c.

Table 3-4 also identifies these periods. As seen in this table, the occurrence of oil price shocks corresponds to slowdowns in economic activity, as measured by the GDPC indicator, and decreases in gasoline consumption, as measured by the MBD indicator.

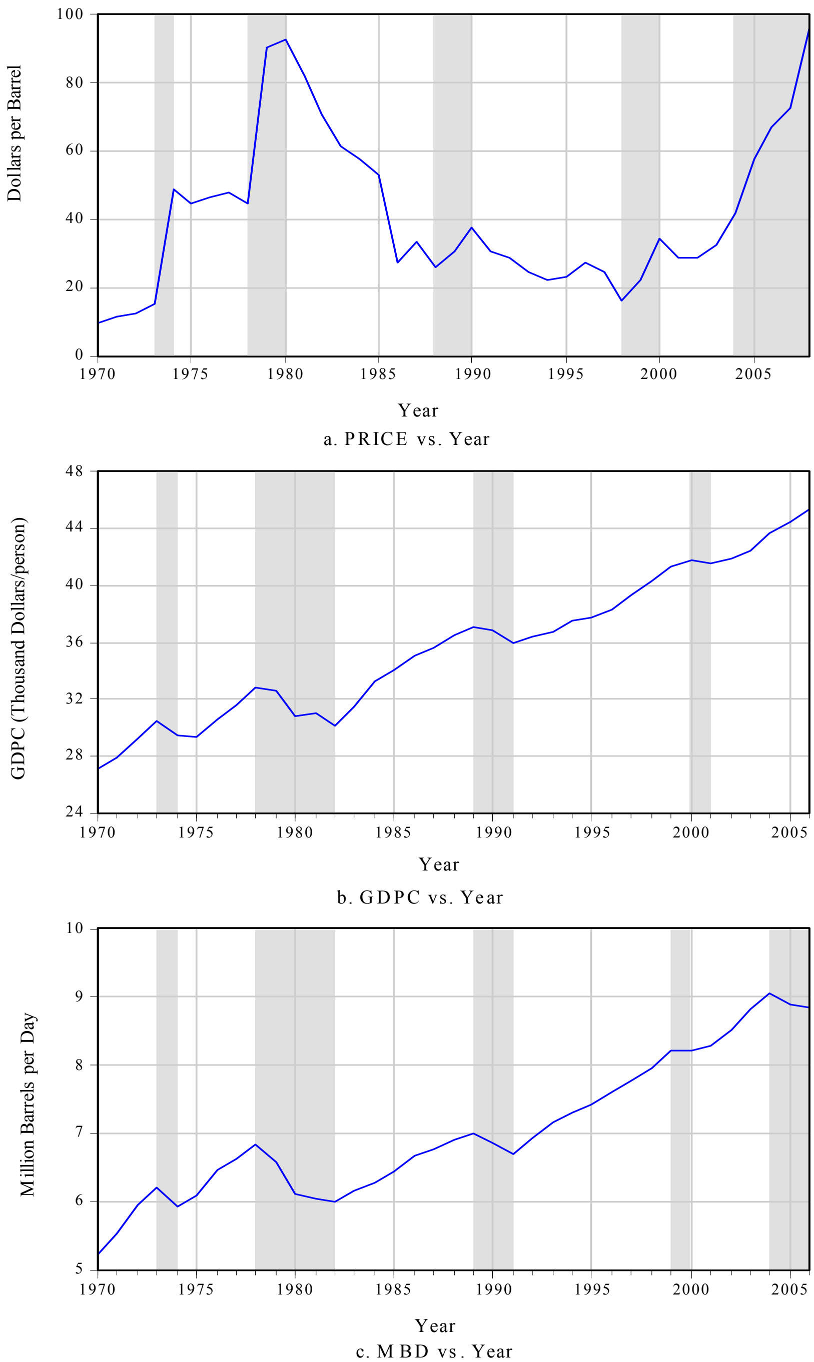


Figure 3-4: Time Series for PRICE, GDPC, and MBD

Table 3-4: Observations on Figure 3-4

PRICE	GDPC	MBD
Oil Price Shock: 1973-1974	Negative Growth :1973-1975	Decreased Consumption: 1974-1975
Oil Price Shock: 1978-1980	Negative Growth :1978-1981	Decreased Consumption: 1978-1982
Oil Price Shock: 1988-1990	Negative Growth :1989-1991	Decreased Consumption: 1989-1991
Oil Price Shock: 1998-2000	Negative Growth :2000-2001	Decreased Consumption: 1999-2000
Oil Price Shock: 2004-2007	Negative Growth :2006-2007	Decreased Consumption: 2004-2006

3.3.2 Vehicle Miles Traveled

The series for VMT is depicted graphically in Figure 3-5a. The periods of VMT negative growth rate or comparatively VMT slow positive growth rate are indicated in the figure. There is some correspondence of most of these periods with the periods of MBD negative growth, although this same association is less noticeable when it comes to PRICE or GDPC.

3.3.3 Miles per Gallon

The fuel efficiency of motor vehicles is evaluated in the United States by the miles per gallon indicator or MPG. The corresponding MPG series is depicted graphically in Figure 3-5b.

The standard for MPG were first set by the Corporate Average Fuel Standards, also known as the CAFE standards, in 1975 (Portney 2002). It is seen in Figure 3-5b that the MPG values improved considerably in the 15-year period extending between 1976 and 1991 reflecting the progressive effect of fleet substitution, as old light vehicles were gradually replaced by new more efficient ones.

Lower CAFE standards set by for SUV's than for passenger cars coupled with the relatively low oil prices of the period 1986-2004 caused the sales of SUVs, as proportion to light vehicles, to approximately double in the same period (McManus 2006). As a result the average MPG indicator has remained essentially flat or with little improvement since 1991. The so called "SUV loophole" has been closed in the new CAFE standards established as of December 2007.

3.3.4 Size of the Light Vehicle Fleet

The light vehicle fleet includes cars and light trucks; light trucks include vans, minivans, pick ups, and small utility vehicles or SUVs, and trucks with gross vehicle weight less than 8500 pounds (Energy Information Administration 2007). The light vehicle fleet VEH series is expressed in million vehicles, and is depicted graphically in Figure 3-5c. The light vehicle fleet grew very rapidly in earlier periods reaching the amount of 234.5 million vehicles in year 2006, although the corresponding growth rate, as shown later on in this work, has been decreasing over time.

The United States is the only country in the world, where there are more vehicles than drivers registered, although there is evidence that this tendency at least in very large cities, is beginning to change. In New York City for example, the registration of vehicles decreased 8.4 percent between years 2000 and 2003. It began to grow again in year 2004, but in 2007 the registration was still 6.2 percent lower than the peak registration that had been reached in year 2000 (Jeffrey 2006).

It appears also that in large cities, provided there is good public transportation, some people decide against car ownership on account of high taxes, parking difficulties, excessive fines for small violations, and high insurance premiums.

3.3.5 Population of the United States

The U.S. Census Bureau updates the population projections each year based on the 2000 census. The last update was released by the Public Information Office on August 14, 2008 (U.S. Census Bureau 2008). This data was used in this thesis.

3.3.6 Guideway

The guideway or transportation infrastructure needed by the light vehicle fleet to operate and perform its function is formed by the total length of lane miles across the country. It was expressed in million miles, was identified as MILES, and the corresponding time series is depicted graphically in Figure 3-5d.

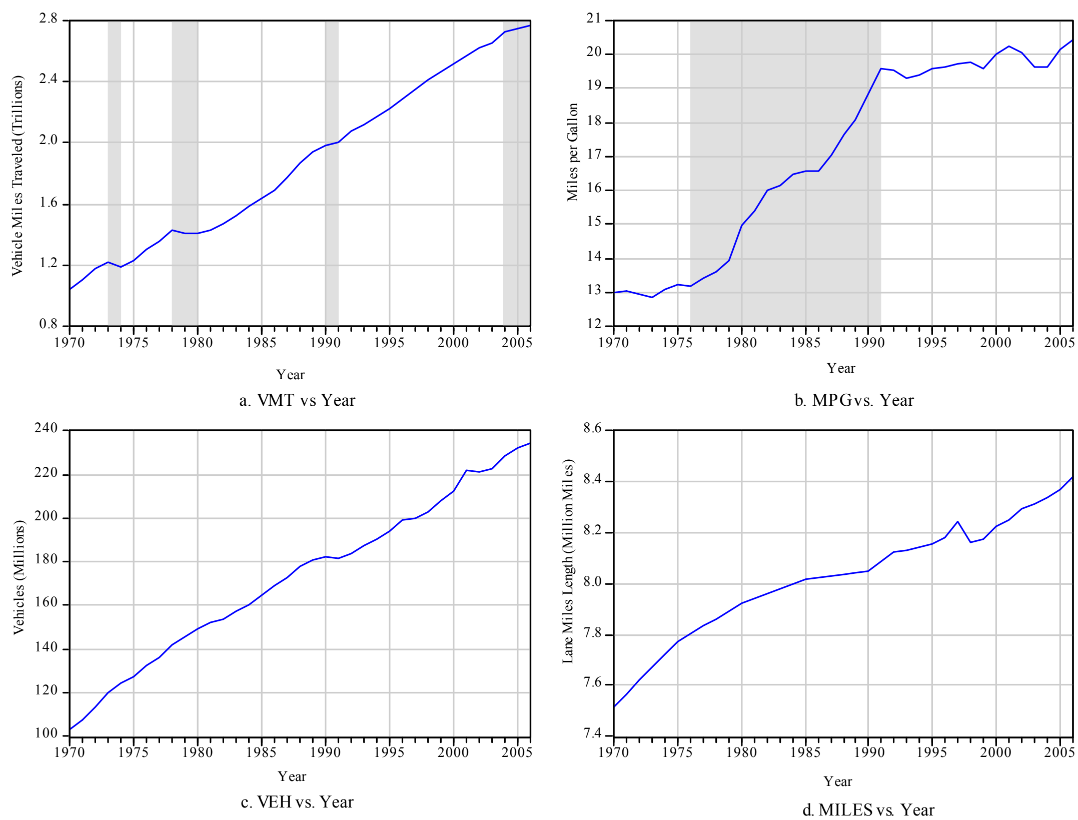


Figure 3-5: Time Series for VMT, MPG, VEH, and MILES

3.3.7 Vehicle over Population Ratio

The ratio vehicle over population coefficient was identified as VEHP. This ratio is useful to determine future projections of the light vehicle fleet size based on the projections of the population of the United States provided by the U.S. Census Bureau database (www.census.gov).

The series VEHP plotted in

Figure 3-6 has a parabolic trend. VEHP grew from 0.50 vehicles per person in 1970 to 0.79 in 2006.

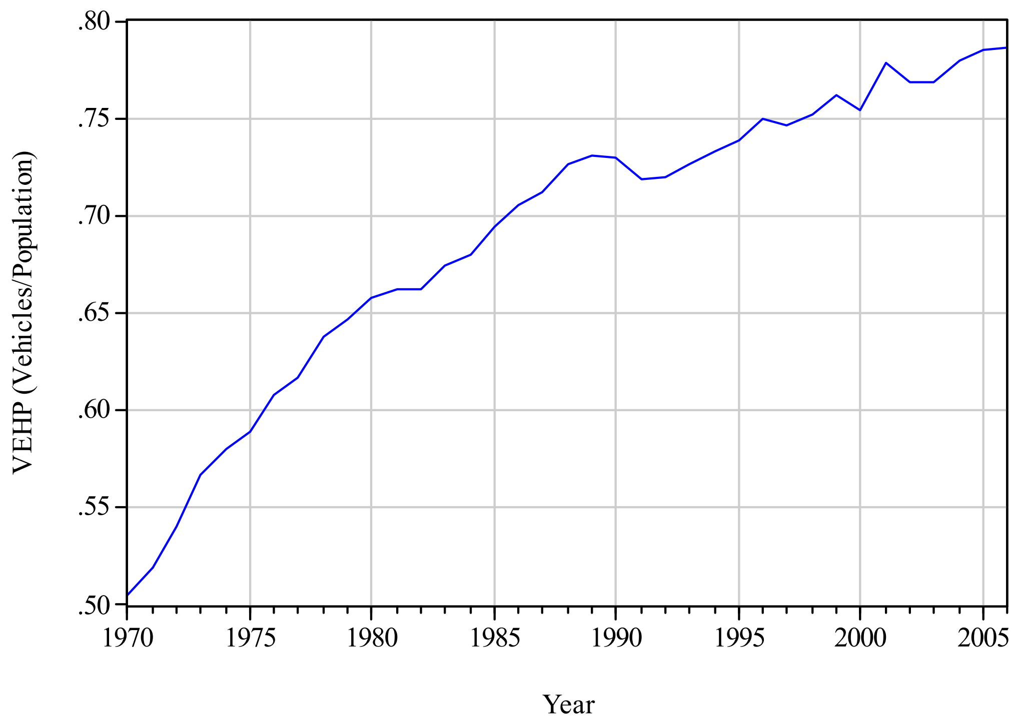


Figure 3-6: VEH/POP Ratio

Table 3-5: Transportation Demand Indicators

Year	MBD	VMT	MPG	GDPC	MILES	PRICE	VEH	POP	VEHP
1970	5.23	1.04	12.98	27.08	7.52	9.63	103.46	205.05	0.50
1971	5.53	1.10	13.02	27.83	7.57	11.49	107.90	207.66	0.52
1972	5.94	1.18	12.93	29.20	7.62	12.28	113.51	209.90	0.54
1973	6.20	1.22	12.86	30.49	7.67	15.37	120.07	211.91	0.57
1974	5.93	1.19	13.09	29.47	7.72	48.66	124.19	213.85	0.58
1975	6.08	1.23	13.24	29.29	7.77	44.52	127.12	215.97	0.59
1976	6.47	1.30	13.16	30.54	7.80	46.72	132.49	218.04	0.61
1977	6.62	1.36	13.40	31.57	7.83	47.67	135.91	220.24	0.62
1978	6.84	1.43	13.60	32.82	7.86	44.65	142.05	222.59	0.64
1979	6.59	1.41	13.91	32.53	7.89	90.31	145.45	225.06	0.65
1980	6.12	1.40	14.96	30.85	7.92	92.77	149.48	227.23	0.66
1981	6.05	1.43	15.40	31.05	7.94	82.03	152.03	229.47	0.66
1982	5.99	1.47	15.99	30.14	7.96	70.90	153.49	231.66	0.66
1983	6.15	1.52	16.15	31.44	7.98	61.56	157.66	233.79	0.67
1984	6.28	1.59	16.47	33.21	8.00	57.45	160.26	235.83	0.68
1985	6.45	1.64	16.56	34.09	8.02	53.10	165.10	237.92	0.69
1986	6.67	1.69	16.57	35.05	8.02	27.28	169.39	240.13	0.71
1987	6.78	1.77	17.06	35.61	8.03	33.64	172.59	242.29	0.71
1988	6.91	1.87	17.67	36.47	8.04	26.14	177.64	244.50	0.73
1989	6.99	1.94	18.08	37.07	8.04	30.48	180.50	246.82	0.73
1990	6.86	1.98	18.85	36.82	8.05	37.66	181.98	249.46	0.73
1991	6.69	2.01	19.58	36.00	8.09	30.45	181.33	252.15	0.72
1992	6.94	2.08	19.54	36.44	8.12	28.54	183.67	255.03	0.72
1993	7.17	2.12	19.29	36.70	8.13	24.35	187.32	257.78	0.73
1994	7.31	2.17	19.38	37.54	8.14	22.12	190.79	260.33	0.73
1995	7.42	2.23	19.60	37.75	8.16	23.15	194.13	262.80	0.74
1996	7.60	2.29	19.61	38.29	8.18	27.30	198.86	265.23	0.75
1997	7.78	2.35	19.73	39.31	8.24	24.67	199.97	267.78	0.75
1998	7.97	2.42	19.79	40.30	8.16	16.18	203.17	270.25	0.75
1999	8.23	2.47	19.58	41.27	8.18	22.35	207.79	272.69	0.76
2000	8.22	2.52	20.03	41.83	8.22	34.29	212.71	282.13	0.75
2001	8.29	2.57	20.24	41.56	8.25	28.62	221.82	284.80	0.78
2002	8.53	2.62	20.08	41.85	8.30	28.83	220.93	287.45	0.77
2003	8.83	2.66	19.62	42.47	8.32	32.50	222.86	290.12	0.77
2004	9.06	2.73	19.64	43.68	8.34	42.00	228.28	292.80	0.78
2005	8.89	2.75	20.17	44.49	8.37	57.88	231.91	295.51	0.78
2006	8.85	2.77	20.43	45.31	8.42	67.02	234.53	298.22	0.79

Source: S. Davis, S.W. Diegel, and R.G. Bounty; Transportation Energy Databook, Edition 27; Oak Ridge National Laboratory, Oak Ridge TN, 20008. S.Williamson and L.Officer, "Measuring Worth". www.measuringworth.com. Bureau of Transportation Statistics, "National Transportation Statistics", www.bts.gov. British Petroleum Corporation, "Statistical Review of World Energy 2008", www.bp.com. United States Census Bureau, "U.S. Census 2000 and 2008 Projections", www.census.gov.

CHAPTER 4. ESTABLISHING THE TREND OF GASOLINE DEMAND

The time series uni-variate methods of modeling demand may not be used for forecasting purposes in this particular study, because it is already known that a very important event occurred in the initial forecasting period: the new regulations concerning the fuel efficiency of vehicles, as per The Energy Independence Act of 2007, are expected to impinge on the future performance of the MBD series.

For forecasting purposes, an explanatory variable model is then needed. This model may take into account the fuel efficiency of vehicles variable, as well as other important variables that may be deemed necessary. Nevertheless, the time series uni-variate methods are extremely important and necessary here to determine a reference line based on the past trend of gasoline consumption. Any savings in fuel consumption may then be accrued against that reference or baseline.

4.1 Past Performance of the MBD Series

The MBD series was shown in Figure 3-4c. A general non-stationary long term increasing trend is observed in the picture. However, more clearly the series goes in a sig-saw pattern where there are several short periods of positive growth rate followed by shorter periods of almost zero or negative growth.

To establish the growth rate in different sub-periods of the MBD series 36 years total period, it is necessary to run the following regression for each one of the identified different sub-periods (Gujarati 2003) .

LOG(Yt) = C(1) + C(2) * T

Where:

$C(1) = LOG(Yo)$ and

$C(2) = LOG(Yt)$

Yt indicates the MBD time series at any time, Yo indicates the consumption where the MBD series began, and T is the time expressed in periods which takes the values 0,1,2,3, etc

The growth rate "r" searched for in such a procedure is contained in the constant C(2) and is given in percent by:

$r = (\exp(C(2)) - 1) * 100$

To determine the growth rate at different times of the MBD 1970-2006 series, it was partitioned in 8 pieces that comprised 4 sub-periods of positive growth and 4 sub-periods of almost zero or negative growth. Besides these central sub-periods, there were two additional pieces that were discarded, the first because there is no record as to when it began and the last because presumably it has not ended yet. A corresponding EVIEWS regression was run for each one of the 8 sub-periods with the results that are summarized in Table 4-1.

Table 4-1: MBD growth rates in several periods

Number	Year	Periods T	Number of Points	Slope C(2)	Growth Rate (r,%)	R^2	P
0	1970-2006	37	36	0.0125	1.26	0.96	0.00
1	1973-1974	1	1	-0.0449	-4.39	1.00	0.00
2	1974-1978	4	4	0.0370	3.77	0.97	0.00
3	1978-1982	4	4	-0.0350	-3.45	0.85	0.02
4	1982-1989	7	7	0.0230	2.33	0.98	0.00
5	1989-1991	2	2	-0.0220	-2.20	0.98	0.05
6	1991-1999	8	8	0.0240	2.44	0.99	0.00
7	1999-2001	2	2	0.0040	0.37	0.64	0.41
8	2001-2004	3	3	0.0300	3.04	0.99	0.00

Figure 4-1 also shows the 8 sub-periods. The 4 positive growth sub-periods went from 2.33 to 3.77 percent and lasted between 3 and 8 years, while the 4 almost zero or negative growth rate sub-periods went from almost zero to -4.39 percent and lasted between 1 and 4 years.

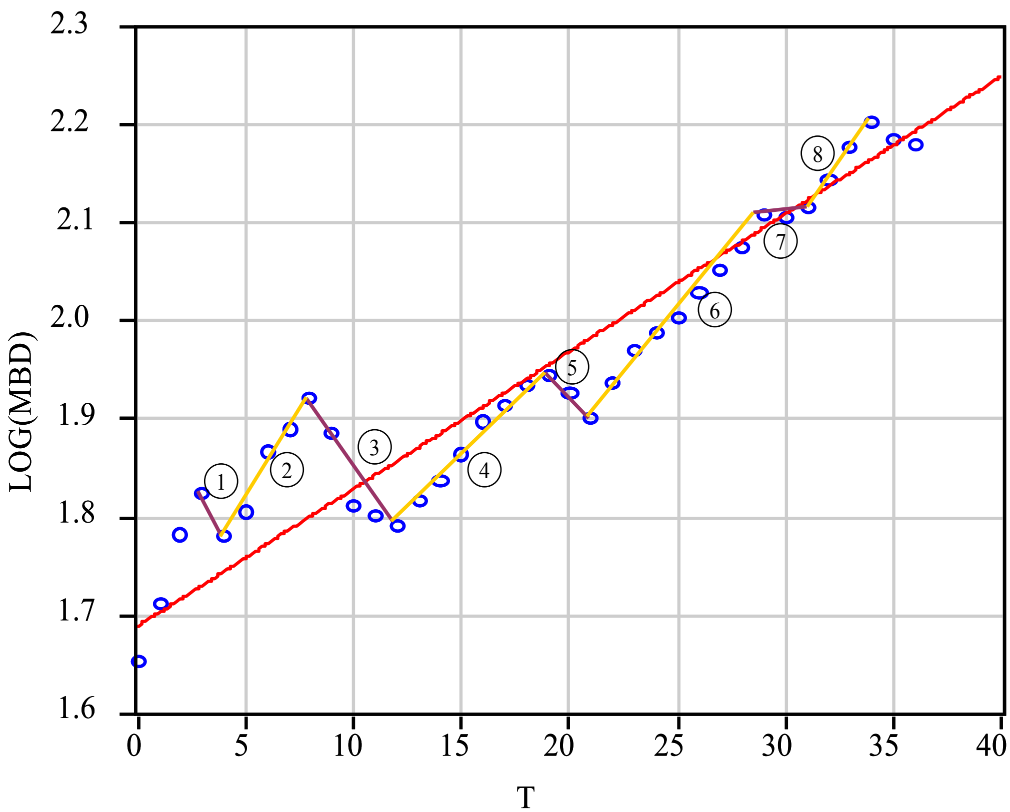


Figure 4-1: MBD Growth in Different Periods

The overall span of the MBD series was also considered in Table 4-1. It corresponds to the number 0 and gave an overall long term growth rate for the 36 years of 1.26 percent. The straight line in Figure 4-1 represents the corresponding growth.

4.2 The Constant Growth Rate Model

The constant growth rate model can also be used to extend the trend by using the EVIEWS forecasting option. The procedure is the same as outlined in section 4.1. The corresponding representation is as follows:

```
Estimation Command:
=========================
LS LOG(MBD) C @TREND

Estimation Equation:
=========================
LOG(MBD) = C(1) + C(2)*@TREND

Substituted Coefficients:
=========================
LOG(MBD) = 1.71642336297 + 0.0124953893586*@TREND
```

At the annual growth rate of 1.26 percent the demand for gasoline would grow from 8.85 MBD in year 2006 to 10.39 MBDs corresponding to about 20 percent increase in gasoline demand. A figure of 10.81 MBDs was reported by the Energy Information Administration for the same year (Energy Information Administration 2007). In the same report, the EIA estimated a constant growth rate of 1.4 percent per year for the MBD oil demand indicator of the United States light vehicle fleet.

The Root Mean Square Error or RMSE and the Mean Absolute Percentage Error or MAPE were selected and recorded for evaluation and comparison purposes. These statistics were recorded for two periods: a calibration or training period extending from 1970 to 2000 and a testing or validation period from 2001 to 2006. These values are shown in Table 4-2 at the end of this thesis section. The same table also presents values for the coefficient of determination and the AIC and Schwartz criteria.

4.3 The Linear Trend Model

Linear trend models are appealing because of their inherent simplicity, and hence they have been used extensively in the past for forecasting purposes (Wells and Weens 2004). Therefore, it is useful to construct a model of this type for comparison and evaluation purposes using EVIEWS.

The corresponding EVIEWS representation of this model is the following

```
Estimation Command:
=========================
LS MBD C @TREND

Estimation Equation:
=========================
MBD = C(1) + C(2)*@TREND

Substituted Coefficients:
=========================
MBD = 5.43549694183 + 0.089070457965*@TREND
```

By extension of the past trend with this model the gasoline demand would growth to 9.89 MBDs in year 2020.

The Root Mean Square Error or RMSE and the Mean Absolute Percentage Error or MAPE were selected and recorded for evaluation and comparison purposes. They are shown with those obtained by other methods in Table 4-2.

4.4 The Random Walk Model

In the random walk model or RWM the value of the dependent variable at any time is equal to its value one period before plus a random shock (Chatfield 2004). In the ARIMA language the random walk model is also known as ARIMA(0,1,0) model because the specification is equivalent to simple differencing of the dependent variable.

Perhaps the most important characteristic of random walk models is the persistence of random errors or random shocks. This persistence means that the impact of a particular shock never extinguishes itself, and because of that RWM processes are said to possess infinite memory. It is also said that RWM processes are able to remember the shocks forever (Gujarati 2003). As it was seen in part three of this thesis the MBD series has been affected by the periodic occurrence of oil price spikes or oil prices shocks. Therefore, it seems plausible to apply the RWM process for the MBD trend evaluation purpose.

In EVIEWS software, the representation of the random walk model is as follows:

```
Estimation Command:
=========================
LS MBD MBD(-1)

Estimation Equation:
=========================
MBD = C(1)*MBD(-1)

Substituted Coefficients:
=========================
MBD = 1.01374431477*MBD(-1)
```

The statistics selected for comparison and evaluation purposes are shown in Table 4-2. The RWM establishes an MBD trend value in year 2020 of 10.71 MBDs

4.5 ARIMA Type Models

Time series are often auto-correlated; therefore, when simple models such as linear trend or constant growth models are applied to them, the basic regression assumption requirement that the residuals should not be auto-correlated may be violated. Notwithstanding, simple models such as linear trend and constant growth models are widely used for forecasting because of its inherent simplicity, and because they often produce good forecasting performance when compared with other complex or more sophisticated models.

Box-Jenkins or ARIMA type models that strive to transform series into white noise were specially developed to process time series with trends. The process to verify whether a series is or has been transformed into white noise proceeds thru the checking-up of the correlogram. This checking-up or verification process of the correlogram constitutes the main statistical requirement to ascertain the correct applicability of Box-Jenkins or ARIMA type models.

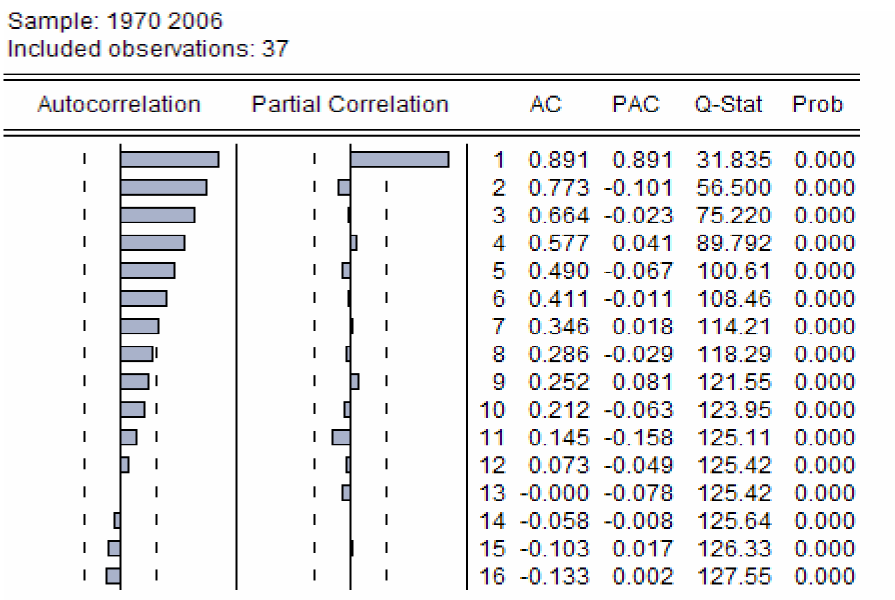
4.5.1 Solving the Stationarity Requirement

The application of the ARIMA time series modeling process, also known as Box-Jenkins methodology requires the time series to be stationary (Washington et al. 2004). As it is seen in Figure 3-4c the MBD series has an upward trend, and its mean is significantly increasing along time; therefore this time series may be considered non-stationary. Nevertheless, it is necessary to demonstrate the condition of non-stationarity.

The standard procedure is to check the correlogram (Makridakis et al. 1998). If the autocorrelation coefficients (ACs) tend quickly to zero as the lag increases, the variable is stationary, otherwise the variable is said to be non-stationary. If the variable turns out to be non-stationary, then it can be turned into stationary by simple differencing.

A second check-up may be obtained by looking at the Partial Correlation coefficients (PACs). Non-stationary series present significant PACs at lag one but insignificant PACs at lags greater than one (Gujarati 2003).

The AC and PAC EVIEWS output for the MBD series was reproduced in Figure 4-2. It is clearly seen that both conditions mentioned above are present in the MBD correlogram. Therefore, the MBD series is non-stationary and should be first differenced to turn it to stationary.



Sample: 1970 2006
Included observations: 37

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.891	0.891	31.835	0.000
		2	0.773	-0.101	56.500	0.000
		3	0.664	-0.023	75.220	0.000
		4	0.577	0.041	89.792	0.000
		5	0.490	-0.067	100.61	0.000
		6	0.411	-0.011	108.46	0.000
		7	0.346	0.018	114.21	0.000
		8	0.286	-0.029	118.29	0.000
		9	0.252	0.081	121.55	0.000
		10	0.212	-0.063	123.95	0.000
		11	0.145	-0.158	125.11	0.000
		12	0.073	-0.049	125.42	0.000
		13	-0.000	-0.078	125.42	0.000
		14	-0.058	-0.008	125.64	0.000
		15	-0.103	0.017	126.33	0.000
		16	-0.133	0.002	127.55	0.000

Figure 4-2: Correlogram for MBD

The Augmented Dickey-Fuller Unit Root Test or ADF test is a specially useful procedure where it is not immediately evident that the series in particular is not stationary (Quantitative Microsoftware 2007b). The null hypothesis for the ADF test is that the series is non-stationary.

The ADF procedure was applied to the MBD series as well as to the MBD differenced once series D(MBD), with the results shown in Figure 4-3 and Figure 4-4 respectively.

In Figure 4-3, it is accepted that the MBD series is non-stationary at the 5% level of significance.

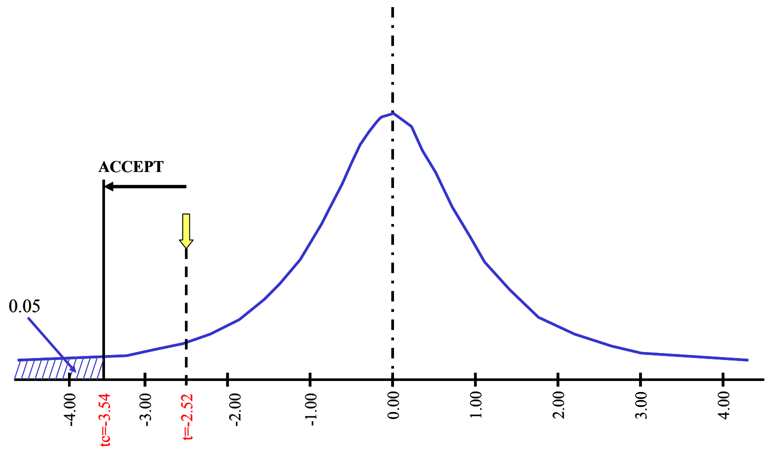


Figure 4-3: ADF Test for MBD

In Figure 4-4, it is rejected that the D(MBD) series is non-stationary at the 5 percent level of significance. It is therefore accepted by the ADF test that the D(MBD) series is stationary at the same level of significance.

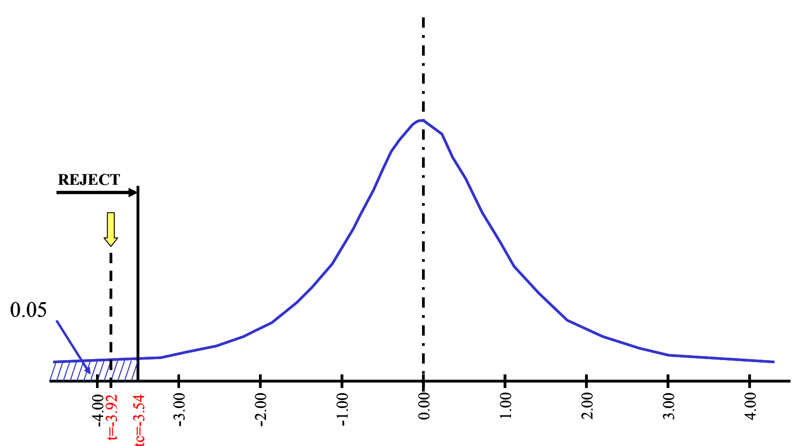


Figure 4-4: ADF Test for D(MBD)

The first difference of the MBD series or D(MBD) was plotted in Figure 4-5; unlike Figure 3-4c there is not any trend in this series. So the process of first differencing has turned the series into a stationary one, as was also proved by the above ADF test performed on the D(MBD) series.

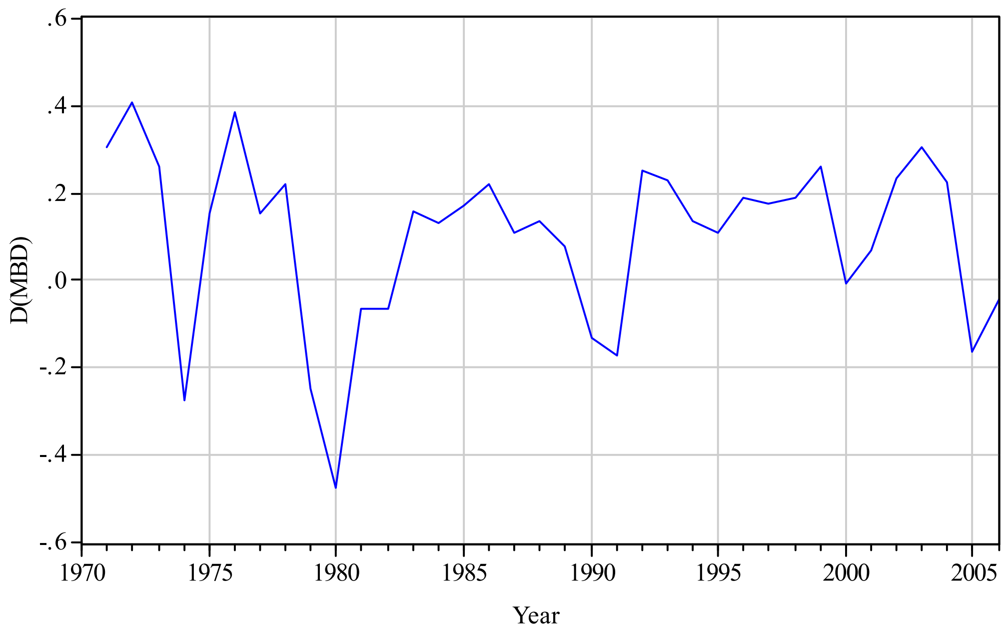


Figure 4-5: First difference for MBD

4.5.2 Identifying the ARIMA process

The correlogram of the D(MBD) is shown in Figure 4-6. As it is seen the pattern is completely different when compared to the correlogram for MBD in Figure 4-2. An autoregressive stationary AR process, or a stationary moving average MA process, and the corresponding order may both be identified by comparing the corresponding correlogram with standard correlograms of known processes (Pyndick and Rubinfeld 1991).

Sample: 1970 2006
Included observations: 36

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.354	0.354	4.9109	0.027
		2	-0.160	-0.327	5.9406	0.051
		3	-0.201	-0.013	7.6149	0.055
		4	-0.166	-0.152	8.7903	0.067
		5	0.084	0.196	9.0999	0.105
		6	0.037	-0.204	9.1638	0.165
		7	-0.143	-0.064	10.123	0.182
		8	-0.196	-0.171	12.000	0.151
		9	-0.080	0.066	12.327	0.195
		10	0.156	0.059	13.601	0.192
		11	0.146	-0.028	14.772	0.193
		12	-0.105	-0.186	15.397	0.220
		13	-0.189	-0.022	17.533	0.176
		14	-0.059	0.011	17.752	0.218
		15	0.026	-0.096	17.795	0.274
		16	0.090	0.011	18.343	0.304

Figure 4-6: Correlogram for D(MBD)

In this case, the fact that both AC and PAC at lag one are statistically different from zero suggests a first order autoregressive process AR(1). Also, the PAC at lag two (although in the limit) might suggest an AR(2) process. The AR(2) condition is confirmed by observation of the sinusoidal pattern in the AC and PAC correlograms. These sinusoidal patterns are typical of second order autoregressive processes (Pyndick and Rubinfeld 1991). There are no suggestions about any MA processes in the AC and PAC correlograms. The series is also integrated to grade 1 as it had to be differenced once to turn it stationary; therefore, the MBD series can be represented by an ARIMA (2,1,0) process.

4.5.3 The MBD ARIMA(2,1,0) PROCESS

The ARIMA(2,1,0) process was run in EVIEWS and the corresponding representation is as follows:

```
Estimation Command:
=========================
LS D(MBD) C AR(1) AR(2)

Estimation Equation:
=========================
D(MBD) = C(1) + [AR(1)=C(2),AR(2)=C(3)]

Substituted Coefficients:
=========================
D(MBD) = 0.0888045841203 + [AR(1)=0.430490161549,AR(2)=-0.318928409114]
```

The extension of the trend with the ARIMA(2,1,0) process led to a possible oil consumption of gasoline of 10.17 MBDs in year 2020. The statistics selected for evaluation and comparison purposes are also shown in Table 4-2.

4.6 Model Selection for the Baseline Definition

The criterion adopted to select a model to "best describe the trend" was the performance of the model in the testing period as measured by the Root Mean Square Error or RMSE statistic, although other statistics were also recorded in Table 4-2.

According to this criterion, from the four models previously presented in sections 4.2, 4.3, 4.4, and 4.5, the model that best described the trend turned out to be the ARIMA(2,1,0) model which for year 2020 led to a gasoline consumption of 10.17 million barrels per day.

Table 4-2: Some Recorded Statistics for the MBD Trend Models

Number	Model	R^2	AIC	Schwartz	F	Train		Test	
						RMSE	MAPE	RMSE	MAPE
1	Constant Growth	0.81	-3.18	-3.09	133	0.30	3.76	0.58	6.34
2	Linear Trend	0.81	0.63	0.79	133	0.31	3.97	0.67	7.40
3	RWM	0.92	-0.33	-0.29		0.20	2.36	0.17	1.72
4	ARIMA	0.18	-0.47	-0.33		0.17	2.04	0.14	1.39

Figure 4-7 shows the actual MBD series in the period 1970-2006 as well as the ARIMA trend line of gasoline consumption up to year 2020. This trend line was adopted as the reference or baseline of gasoline consumption. The ARIMA based 95 percent confidence interval is also shown in the same figure.

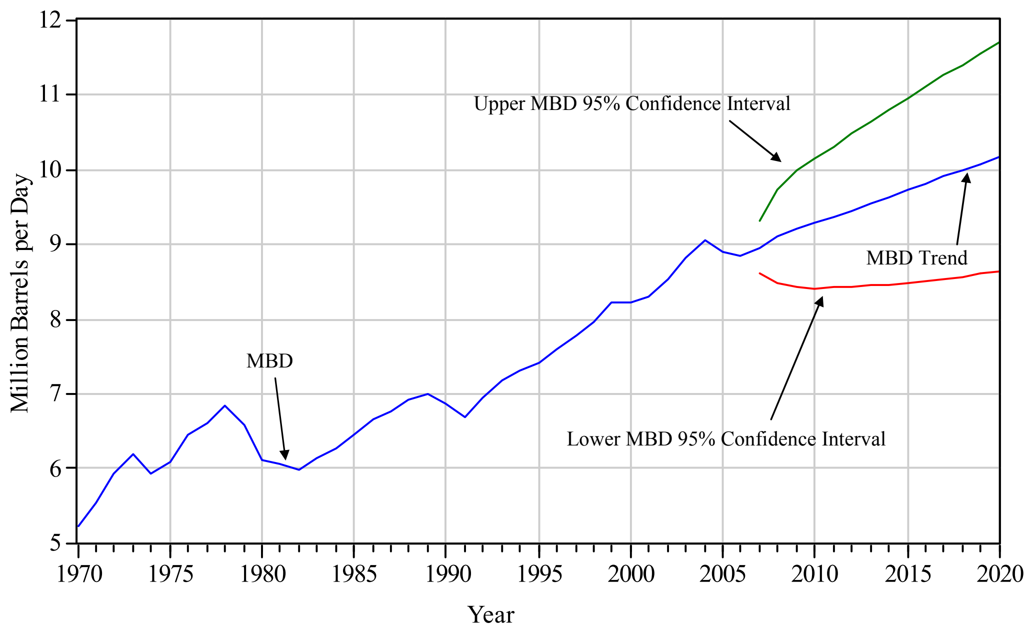


Figure 4-7: MBD Series and MBD Trend by ARIMA

The data corresponding to the various different approaches used to define the baseline of gasoline consumption are shown in Table 4-3. Values for the 95 percent confidence interval, defined by EVIEWS at two standard errors above and two standard errors below the average as per the ARIMA(2,1,0) procedure, are also presented in the same table. The identification of the variables is as follows:

MBD: observed, actual, or real data

MBDFG: trend obtained by constant growth rate model

MBDFL: simple linear trend model

MBDFR: trend obtained by random walk model

MBDFA: trend obtained by ARIMA model

MBDFA_TOP : upper limit 95 percent confidence interval obtained by ARIMA

MBDFA_BOT : lower limit 95 percent confidence interval obtained by ARIMA

Table 4-3: MBD Trend Evaluated by Several Methods

OBS	MBD	MBDFL	MBDFG	MBDFR	MBDFA	MBDFA_TOP	MBDFA_BOT
1970	5.23	5.44	5.56	NA	NA	NA	NA
1971	5.53	5.52	5.63	5.3	NA	NA	NA
1972	5.94	5.61	5.71	5.61	NA	NA	NA
1973	6.2	5.7	5.78	6.02	6.1	NA	NA
1974	5.93	5.79	5.85	6.29	6.26	NA	NA
1975	6.08	5.88	5.92	6.01	5.81	NA	NA
1976	6.47	5.97	6	6.16	6.31	NA	NA
1977	6.62	6.06	6.07	6.55	6.66	NA	NA
1978	6.84	6.15	6.15	6.71	6.64	NA	NA
1979	6.59	6.24	6.23	6.93	6.96	NA	NA
1980	6.12	6.33	6.31	6.68	6.49	NA	NA
1981	6.05	6.42	6.38	6.2	6.07	NA	NA
1982	5.99	6.5	6.46	6.14	6.26	NA	NA
1983	6.15	6.59	6.55	6.07	6.06	NA	NA
1984	6.28	6.68	6.63	6.23	6.32	NA	NA
1985	6.45	6.77	6.71	6.37	6.36	NA	NA
1986	6.67	6.86	6.8	6.54	6.56	NA	NA
1987	6.78	6.95	6.88	6.76	6.79	NA	NA
1988	6.91	7.04	6.97	6.87	6.83	NA	NA
1989	6.99	7.13	7.06	7.01	7.02	NA	NA
1990	6.86	7.22	7.14	7.09	7.06	NA	NA
1991	6.69	7.31	7.23	6.96	6.86	NA	NA
1992	6.94	7.4	7.33	6.78	6.73	NA	NA
1993	7.17	7.48	7.42	7.03	7.18	NA	NA
1994	7.31	7.57	7.51	7.27	7.27	NA	NA
1995	7.42	7.66	7.61	7.41	7.37	NA	NA
1996	7.6	7.75	7.7	7.52	7.5	NA	NA
1997	7.78	7.84	7.8	7.71	7.73	NA	NA
1998	7.97	7.93	7.9	7.89	7.88	NA	NA
1999	8.23	8.02	7.99	8.08	8.07	NA	NA
2000	8.22	8.11	8.1	8.34	8.36	NA	NA
2001	8.29	8.2	8.2	8.33	8.21	NA	NA
2002	8.53	8.29	8.3	8.4	8.4	NA	NA
2003	8.83	8.37	8.4	8.64	8.68	NA	NA
2004	9.06	8.46	8.51	8.95	8.96	NA	NA
2005	8.89	8.55	8.62	9.18	9.13	NA	NA
2006	8.85	8.64	8.73	8.54	8.83	NA	NA
2007	NA	8.73	8.84	8.97	8.96	9.33	8.6
2008	NA	8.82	8.95	9.09	9.1	9.73	8.47
2009	NA	8.91	9.06	9.22	9.21	9.99	8.43
2010	NA	9	9.17	9.34	9.28	10.16	8.41
2011	NA	9.09	9.29	9.47	9.36	10.31	8.42
2012	NA	9.18	9.4	9.6	9.45	10.48	8.43
2013	NA	9.27	9.52	9.74	9.54	10.64	8.44
2014	NA	9.35	9.64	9.87	9.63	10.81	8.46
2015	NA	9.44	9.76	10.01	9.72	10.96	8.49
2016	NA	9.53	9.89	10.14	9.81	11.11	8.51
2017	NA	9.62	10.01	10.28	9.9	11.26	8.54
2018	NA	9.71	10.14	10.42	9.99	11.4	8.57
2019	NA	9.8	10.26	10.57	10.08	11.55	8.61
2020	NA	9.89	10.39	10.71	10.17	11.69	8.64

CHAPTER 5. DEVELOPING AN EXPLANATORY VARIABLE MODEL OF GASOLINE DEMAND

5.1 Criteria for Model Build Up and Selection

There are several issues that may be considered in developing the gasoline demand regression model. In first place, a linear regression model makes basic assumptions that must be examined before the model is considered acceptable for forecasting purposes (Bowerman et al. 2005; Levine et al. 2001). These assumptions are:

I. The residuals must be independent, un-patterned, or un-correlated. Most of the times, this requirement constitutes a problem when dealing with non-stationary economic series (Wei 2006). However, EVIEWS allows for a special feature for serial correlation correction by way of inclusion of autoregressive or moving average error correction terms in the models (Startz 2007). Usually, autoregressive or moving average error correction terms of order 1 and 2, AR(1), AR(2), MA(1), MA(2), or a combination of these, is all what is required for the Durbin-Watson statistics to be around the 2.0 value range which defines the no serial correlation requirement (Wilson et al. 2002).

II. The residuals must be homoskedastic: no heteroskedasticity. The condition for constant variance of residuals can be examined by looking at scatter-plots of residuals (or of the squared residuals), against the dependent variable or against each one of the independent variables, to detect patterns indicative of increasing or decreasing error variances

(Bowerman et al. 2005). The other approach which is the procedure supported by EVIEWS (Quantitative Microsoftware 2007b) is to test for constant error variance; that is performing a test for heteroskedasticity. One of such tests that is widely used and that was applied here is the White F test of pure homoskedasticity.

III. Normality of residuals. The normality assumption may be checked by constructing a normality plot, a stem and leaf display, or a histogram (Bowerman et al. 2005); also, by the evaluation of the JB or Jarque-Bera statistic which is the procedure supported by EVIEWS. The JB statistic follows the Chi-Square distribution.

The second issue to be dealt with is the issue of multicollinearity. Multicollinearity exists when the independent variables relate to each other thus contributing redundant information to the description of the dependent variable. Upon this condition the coefficients determined for the independent variables may become unreliable. Bowerman et al. (Bowerman et al. 2005) consider multicollinearity to be a problem in a dataset if the correlation between two independent variables is greater than 0.90. Multicollinearity can be corrected by deleting the independent variables contributing the correlation, by forming combination of variables, or by differencing (Studenmund 2000).

Once a model or a set of possible models is identified, then a decision is needed to determine the accuracy of the model and a procedure to select the best model.

Some of the most widely accepted methods to evaluate the relative accuracy of forecasting regression models are the Root Mean Square Error measure of accuracy, or

RMSE, and the evaluation of the Mean Absolute Percentage Error or MAPE. Both of these statistics are provided by a variety of software statistics including EVIEWS.

To ascertain how the model may perform in actual forecasting, the existing observed data is divided in two periods. The first period is used for training or calibration, whereas the second period is used for testing or validation. The second period is also referred to in the forecasting literature as the “hold out period” (Wilson et al. 2002).

Other criteria used to select a model among a possible set of models are the standard coefficient of determination R^2, the Akaike Information Criteria or AIC, and the Schwartz Criteria (Quantitative Microsoftware 2007b) .

5.2 Selecting the Set of Explanatory Variables for the Model

The independent variables for the MBD model were chosen among the following transportation demand indicators: PRICE, MPG, GDPC, MILES, VEH, and VMT presented in Chapter 3 of this thesis. The variables PRICE and MPG were considered first for inclusion into the model because of special circumstances related to them. As of year 2004 the United States began experiencing the fifth oil price shock since 1970. The oil price reached the historical maximum of $147.00 in July 2008 and then began rapidly to decrease to about $40.00 in December (Energy Information Administration 2008a). The last oil price spike motivated the signing into law of the Energy Independence Act of 2007 that called for the progressive improvement of the fuel efficiency indicator MPG of the light vehicle fleet. The progressive improvement of MPG is expected to reduce the gasoline consumption of oil. Consumers to a certain degree also respond to oil price fluctuations by changing driving and behavioral patterns that affect oil consumption. Therefore, the variables PRICE and MPG were given priority for inclusion into the model.

The inclusion of more variables into the model requires examining the multicollinearity matrix shown in Table 5-1 which imposed the following constraints:

I. The regressors VEH, GDPC, VMT, and MILES are highly correlated among themselves; therefore, they contain redundant information to the expression of MBD.

II. MPG along with PRICE, as previously mentioned, were given priority for inclusion into the model, but since MPG is also highly correlated with VEH, GDPC, VMT, and MILES, the inclusion of any of them resulted in the problem of multicollinearity.

III. Multicollinearity can be solved by elimination of variables, combination of multicollinear variables, or transformation of the regression equation by first differencing.

Table 5-1: Multicollinearity Matrix

Variable	PRICE	MPG	VEH	GDPC	MILES	VMT
PRICE	1.000000	-0.265689	-0.148962	-0.297539	-0.009403	-0.310459
MPG	-0.265689	1.000000	0.965576	0.918024	0.932349	0.963251
VEH	-0.148962	0.965576	1.000000	0.965759	0.980284	0.977410
GDPC	-0.297539	0.918024	0.965759	1.000000	0.914354	0.979235
MILES	-0.009403	0.932349	0.980284	0.914354	1.000000	0.928913
VMT	-0.310459	0.963251	0.977410	0.979235	0.928913	1.000000

The following section illustrates the selection process of the "best model" among several possible choices. The criterion for final selection was how the model performed in the testing or "hold-out" period as indicated by the RMSE statistic.

5.3 Model Selection

In all the models described in Table 5-2 the independent variables passed at the 5 percent level of significance.

In Model 1 the dependent MBD variable as well as the independent variables were used with no transformation. However, an interaction new variable was created between the independent variables VEH and MPG to solve for the multicollinearity that existed between these two variables (Table 5-1). Serial correlation was solved by inclusion of AR(1) and MA(1) error correction terms in this model.

Model 2 was specified similar to Model 1, except that the dependent variable as well as the independent variables were first differenced to turn them stationary. The normality test showed some improvements, but the RMSE and MAPE statistics in the testing or hold-out period remained about the same as in Model 1.

In Model 3 the variables MPG, VEH, and PRICE, were used separately, but the dependent as well as the independent variables were first differenced to turn them stationary. Differencing also addressed the multicollinearity problem that existed between the independent variables VEH and MPG (Table 5-1). This model required only the AR(1) disturbance correction term for serial correlation. The RMSE statistics improved considerably as compared to Models 1 and 2.

In Model 4, the dependent variable as well as the independent variables were first differenced to turn them stationary; but the MBD model was specified in terms of MPG and VMT instead of MPG, VEH, and PRICE. Differencing also solved the multicollinearity problem that existed between the variables MPG and VMT (Table 5-1). This turned out to be the best model for MBD in terms of RMSE and MAPE. Therefore, this was the model finally chosen to describe MBD, but a separate specification was needed to solve for VMT. Model 4 corresponds then to the possibility of solving for MBD in two consecutives steps, a procedure also proposed by Emerson (Emerson 2006).

Models 5, 6, 7, and 8 were the models tried to solve in a second step for VMT. The regressors chosen to explain VMT were VEH and PRICE. There was no multicollinearity present since the variables VEH and PRICE are not highly correlated (Table 5-1). However, all the variables were first differenced to turn them stationary.

Model 5 required only the MA(1) error correction term for serial correlation.

Model 6 is similar to model 5, except that a logarithmic transformation used for the independent variable PRICE tended to improve the normality requirement in the model. It is seen also that the RMSE and MAPE statistics improved considerably but only in the testing period as compared to model 5. Serial correlation was addressed by inclusion of autoregressive error correction terms of order one and two.

Model 7 is similar to model 5, except that AR(1) instead of MA(1) was needed for serial correlation.

Model 8 is similar to Model 7, except that serial correlation was addressed by inclusion of both AR(1) and AR(2). Although AR(2) turned out to be statistically insignificant, it did improve the model not only by taking the DW criteria closest to 2 that defines no serial correlation but also improved the normality result. This turned out to be the best model for VMT in terms of the RMSE testing period statistic chosen to compare the models. It was therefore the model selected to describe VMT.

It is observed that the Akaike Information and the Schwartz Criteria were similar in al the VMT models. However, in the MBD models these criteria were better for Model 4 (the model finally selected to describe MBD) than for Models 1, 2, or 3.

All the models described in Table 5-2 passed the normality and constant error variance requirements. Normality was checked by the JB test, and no heteroskedasticity was checked by the White F test which requires $F < Fc$ (F critical). The Durbin-Watson (DW) statistics in all the models remained between the range 1.75 and 2.25 indicative of no serial correlation(Wilson et al. 2002).

Table 5-2 also shows the Inverted Autoregressive Root (IARR) and/or the Inverted Moving Average Root (IMAR) criteria, which both should be less than one for the ARMA processes to be stationary and invertible.

Table 5-2: Several VMT and MBD Models

MODEL No	DEPENDENT	INDEPENDENT VARIABLES			IARR	IMAR	ERROR CORRECTION			DW	NORMALITY TEST		WHITE TEST		R^2	F	AIC	SCHWARTZ	Training		Test	
							AR(1)	AR(2)	MA(1)		JB	PROB	F	Fc					RMSE	MAPE	RMSE	MAPE
1	MBD[1]	PRICE	VEH*MPG		1.00	-0.42	AR(1)		MA(1)	1.92	1.79	0.41	2.30	3.35	0.96	177	-0.93	-0.70	0.13	1.66	0.16	1.46
2	D(MBD)[1]	D(MPG*VEH)	D(PRICE)		0.08	-0.35	AR(1)		MA(1)	1.97	1.26	0.53	2.51	3.37	0.56	7.70	-0.90	-0.67	0.13	1.64	0.16	1.50
3	D(MBD)	D(MPG)	D(PRICE)	D(VEH)	0.84		AR(1)			2.10	0.86	0.65	0.18	3.37	0.70		-1.43	-1.29	0.11	1.31	0.10	0.90
4	D(MBD)	D(MPG)	D(VMT)			-0.68			MA(1)	1.85	0.04	0.98	1.18	3.35	0.98		-3.97	-3.83	0.03	0.36	0.03	0.30
5	D(VMT)	D(PRICE)	D(VEH)			-1.00			MA(1)	2.23	0.65	0.72	0.91	3.35	0.35		-4.59	-4.45	0.02	1.09	0.09	3.37
6	D(VMT)	D(LOG(PRICE))	D(VEH)		0.78 -0.13		AR(1)	AR(2)		1.98	0.38	0.83	0.67	3.39	0.32		-4.44	-4.25	0.03	1.16	0.04	1.12
7	D(VMT)	D(PRICE)	D(VEH)		0.75		AR(1)			2.12	0.54	0.76	0.77	3.37	0.33		-4.54	-4.40	0.02	1.11	0.04	1.16
8	D(VMT)	D(PRICE)	D(VEH)		0.80 -0.11		AR(1)	AR(2)		1.96	0.37	0.83	0.42	3.39	0.32		-4.44	-4.25	0.03	1.13	0.03	1.04

[1]This regression model required a constant term

No multicollinearity is observed among the variables used in the eight models discussed above as it is depicted in Table 5-3, where the corresponding variables were defined as follows:

PRICE: the price of oil in dollars per barrel;

DPRICE = D(PRICE): the first difference of PRICE;

DLOGPRICE = D(LOG(PRICE)): the first difference of the logarithm of PRICE;

MPGxVEH = MPG*VEH: interaction variable formed by MPG and VEH;

DMPGxVEH= D(MPG*VEH): the first difference of MPG*VEH;

DMPG = D(MPG): the first difference of MPG;

DVEH = D(VEH): the first difference of VEH;

DVMT = D(VMT): the first difference of VMT

Table 5-3: Correlation Matrix for Multicollinearity Checking

Variable	PRICE	DPRICE	DLOGPRICE	MPGxVEH	DMPGxVEH	DMPG	DVEH	DVMT
PRICE	1.000000	0.237950	0.169858	-0.338300	0.332382	0.515702	-0.160141	-0.594602
DPRICE	0.237950	1.000000	0.921774	-0.169417	0.080554	0.065455	0.117099	-0.579955
DLOGPRICE	0.169858	0.921774	1.000000	-0.222933	0.082274	0.043371	0.180490	-0.557413
MPGxVEH	-0.338300	-0.169417	-0.222933	1.000000	0.125521	0.020680	-0.361029	0.260129
DMPGxVEH	0.332382	0.080554	0.082274	0.125521	1.000000	0.860187	-0.020284	-0.073946
DMPG	0.515702	0.065455	0.043371	0.020680	0.860187	1.000000	-0.440274	-0.278759
DVEH	-0.160141	0.117099	0.180490	-0.361029	-0.020284	-0.440274	1.000000	0.178974
DVMT	-0.594602	-0.579955	-0.557413	0.260129	-0.073946	-0.278759	0.178974	1.000000

5.4 Model Description

The model selected to describe the MBD process is composed as mentioned before of two steps. In the first step a regression was set up to describe VMT, and in the second step a second regression was set up to define MBD. A similar two steps approach was also proposed by Emerson (Emerson 2006) as it was indicated in Chapter 2 of this thesis.

5.4.1 Defining VMT

The historical VMT data was divided in two periods, the first periods which is referred to as the training or calibration period went from 1970 to 2000; therefore, it lapsed 30 years and contained 31 points. The second period, which is referred to as the testing or validation period went from 2001 to 2006. Therefore, this period lapsed 5 years and contained 6 points.

The EVIEWS estimation output is the following.

Dependent Variable: D(VMT)
Method: Least Squares
Sample (adjusted): 1973 2000
Included observations: 28 after adjustments
Convergence achieved after 14 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(PRICE)	-0.001039	0.000303	-3.424630	0.0022
D(VEH)	0.006150	0.003034	2.027390	0.0539
AR(1)	0.685537	0.233201	2.939687	0.0072
AR(2)	0.088491	0.208586	0.424242	0.6752

R-squared	0.323256	Mean dependent var	0.048049
Adjusted R-squared	0.238663	S.D. dependent var	0.028180
S.E. of regression	0.024588	Akaike info criterion	-4.441513
Sum squared resid	0.014510	Schwarz criterion	-4.251198
Log likelihood	66.18118	Hannan-Quinn criter.	-4.383331
Durbin-Watson stat	1.960950		

Inverted AR Roots	.80	-.11

The model complies well with the basic assumptions, does not suffer from multicollinearity, and has good RMSE and MAPE values in the training as well as in the testing periods. Serial correlation was addressed by inclusion of the autoregressive terms AR(1) and AR(2) in the problem specification. Although the coefficient for AR(2) appears insignificant in the estimation output, it does contribute to a better result in the serial correlation DW test as well as in the Jarque-Bera test of normality as compared to the specification with only the autoregressive AR(1) term included.

The EVIEWS representation of this model is the following.

```
Estimation Command:
=========================
LS D(VMT) D(PRICE) D(VEH) AR(1) AR(2)

Estimation Equation:
=========================
D(VMT) = C(1)*D(PRICE) + C(2)*D(VEH) + [AR(1)=C(3),AR(2)=C(4)]

Substituted Coefficients:
=========================
D(VMT) = -0.00103855217842*D(PRICE) + 0.0061501240526*D(VEH) + [AR(1)=0.685537166453,AR(2)=0.0884910233406]
```

The sign for the independent variable PRICE coefficient is negative, as expected, because increases in this variable tend to decrease VMT. The sign for the independent VEH variable is positive, as expected, because the more vehicles on the road the more vehicle miles traveled.

Table 5-4, Figure **5-1**, and Figure 5-2 are the results of the VMT modeling process. The actual data for VMT in the period 1970-2006 were placed in the column named VMT. The modeling results in the calibration period were placed in the column named VMTF_TRAIN. As seen in the table and in Figure 5-1, the theoretical data follows reasonably well the actual data. The column named VMFT_TEST in the same table corresponds to the VMT modeling results for the validation period which look very similar to the actual VMT data. The columns named VMTF_TESTBOT and VMTF_TESTTOP depict the 95 percent confidence interval. As seen in Figure 5-2, at all times the actual VMT data fell within the 95 percent confidence interval of the modeling results.

Table 5-4: VMT Actual and Model Data (1970-2006)

Year	VMT	VMTF_TRAIN	VMTF_TEST	VMTF_TESTBOT	VMTF_TESTTOP
1970	1.040	-	-	-	-
1971	1.104	-	-	-	-
1972	1.178	-	-	-	-
1973	1.223	1.246	-	-	-
1974	1.190	1.222	-	-	-
1975	1.235	1.197	-	-	-
1976	1.304	1.279	-	-	-
1977	1.360	1.353	-	-	-
1978	1.426	1.429	-	-	-
1979	1.406	1.420	-	-	-
1980	1.403	1.434	-	-	-
1981	1.430	1.413	-	-	-
1982	1.468	1.448	-	-	-
1983	1.523	1.515	-	-	-
1984	1.585	1.558	-	-	-
1985	1.638	1.650	-	-	-
1986	1.694	1.707	-	-	-
1987	1.773	1.711	-	-	-
1988	1.872	1.857	-	-	-
1989	1.938	1.933	-	-	-
1990	1.983	1.980	-	-	-
1991	2.008	2.021	-	-	-
1992	2.078	2.042	-	-	-
1993	2.120	2.144	-	-	-
1994	2.171	2.159	-	-	-
1995	2.228	2.210	-	-	-
1996	2.286	2.282	-	-	-
1997	2.353	2.322	-	-	-
1998	2.418	2.424	-	-	-
1999	2.470	2.470	-	-	-
2000	2.523	2.512	-	-	-
2001	2.572	-	2.612	2.555	2.669
2002	2.625	-	2.560	2.490	2.629
2003	2.656	-	2.672	2.619	2.725
2004	2.727	-	2.701	2.646	2.756
2005	2.749	-	2.767	2.717	2.818
2006	2.772	-	2.772	2.721	2.822

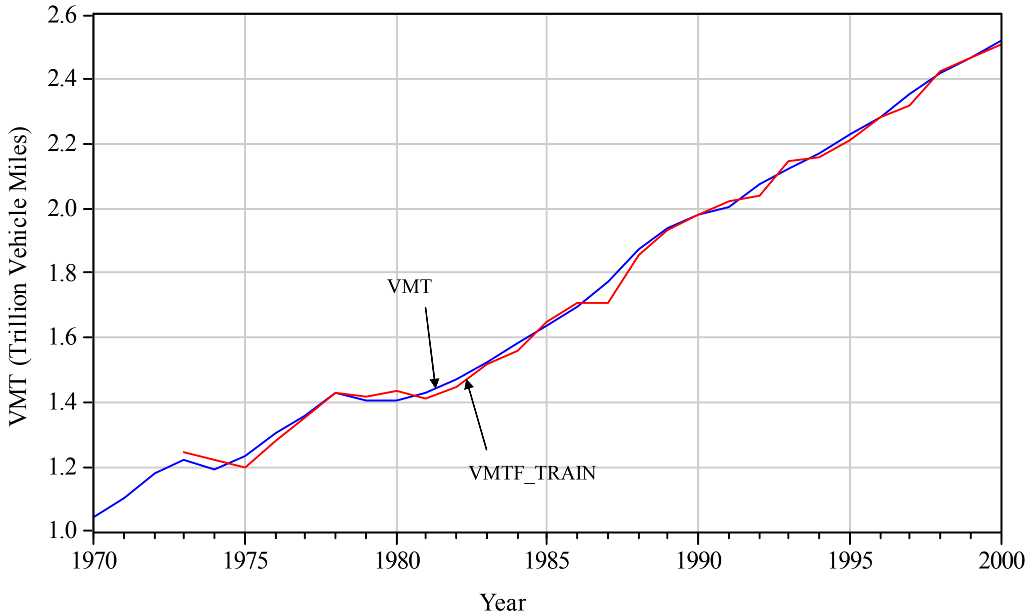


Figure 5-1: VMT Actual vs. VMT Model Data (1970-2000)

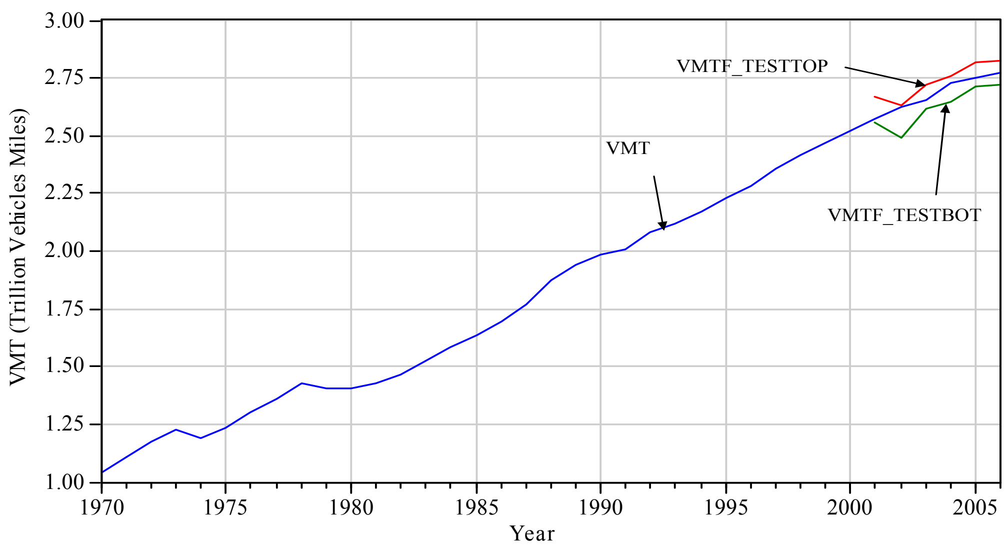


Figure 5-2: VMT Model 95 Percent Confidence Interval (2001-2006)

In view of the VMT modeling results, it was considered that this VMT model can be used to forecast the VMT dependent variable for the MBD model selected. This MBD model is described next.

5.4.2 Defining MBD

The historical MBD data was divided in two periods; the first period which was referred to as the training or calibration period went from 1970 to 2000; therefore, it lapsed 30 years and contained 31 points. The second period which was referred to as the testing or validation period went from 2001 to 2006 hence it lapsed 5 years and contained 6 points. The EVIEWS estimation output for this model is the following:

Dependent Variable: D(MBD)
Method: Least Squares
Sample (adjusted): 1971 2000
Included observations: 30 after adjustments
Convergence achieved after 6 iterations
MA Backcast: 1970

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(MPG)	-0.419708	0.019661	-21.34753	0.0000
D(VMT)	4.160165	0.154901	26.85689	0.0000
MA(1)	0.675050	0.141927	4.756301	0.0001

R-squared	0.976630	Mean dependent var	0.099741
Adjusted R-squared	0.974899	S.D. dependent var	0.199842
S.E. of regression	0.031662	Akaike info criterion	-3.972773
Sum squared resid	0.027067	Schwarz criterion	-3.832653
Log likelihood	62.59159	Hannan-Quinn criter.	-3.927947
Durbin-Watson stat	1.847264		

Inverted MA Roots	-.68

This model complies well with the basic regression assumptions and does not suffer from multicollinearity between the independent variables DMPG and DVMT, as seen in Table 5-2. It also has good RMSE and MAPE values in the training as well as in the testing periods as compared to the other MBD models studied in Table 5-1.

The corresponding EVIEWS representation is the following:

```
Estimation Command:
=========================
LS D(MBD) D(MPG) D(VMT) MA(1)

Estimation Equation:
=========================
D(MBD) = C(1)*D(MPG) + C(2)*D(VMT) + [MA(1)=C(3),BACKCAST=1971,ESTSMPL="1971 2000"]

Substituted Coefficients:
=========================
D(MBD) = -0.419707573798*D(MPG) + 4.16016458068*D(VMT) + [MA(1)=0.675049623828,BACKCAST=1971,ESTSMPL="1971 2000"]
```

Where BACKCAST refers to when the EVIEWS backward recursion algorithm was carried back in the MA error correction process, and ESTSAMPLE refers to the sample estimation period in the same process (Quantitative Microsoftware 2007b).

The sign for MPG is negative, as expected, as improved fuel efficiency should tend to decrease fuel consumption; and the sign for VMT is positive, as expected, as more vehicle miles traveled lead to more gasoline consumption.

Table 5-5, Figure 5-3, and Figure 5-4 are the results of the MBD modeling process. The actual data for MBD in the period 1970-2006 were placed in the column named MBD. The modeling results in the calibration period were placed in the column named MBDF_TRAIN. As seen in the table and in Figure 5-3, the theoretical data follows reasonably well the actual data. The column named MBDF_TEST in the same table corresponds to the MBD modeling results in the validation period, which look very similar to the actual data. The columns named MBDF_TESTBOT and MBDF_TESTTOP depict the 95% confidence interval. As seen in Figure 5-4 and also in the Table, at all times the actual MBD data fell within the 95% confidence interval of the modeling results.

In view of the MBD modeling results, this model can be used to forecast the MBD dependent variable. But before attempting that, it was necessary to create some needed scenarios for the independent variables VEH, PRICE, and MPG. This task was accomplished as shown in the next chapter.

Table 5-5: MBD Actual and Model Data (1970-2006)

Year	MBD	MBDF_TRAIN	MBDF_TEST	MBDF_TESTBOT	MBDF_TESTTOP
1970	5.227	-	-	-	-
1971	5.534	5.497	-	-	-
1972	5.942	5.901	-	-	-
1973	6.201	6.185	-	-	-
1974	5.929	5.979	-	-	-
1975	6.081	6.017	-	-	-
1976	6.466	6.450	-	-	-
1977	6.617	6.604	-	-	-
1978	6.837	6.817	-	-	-
1979	6.591	6.637	-	-	-
1980	6.117	6.108	-	-	-
1981	6.054	6.048	-	-	-
1982	5.989	5.973	-	-	-
1983	6.149	6.159	-	-	-
1984	6.280	6.271	-	-	-
1985	6.450	6.464	-	-	-
1986	6.670	6.673	-	-	-
1987	6.778	6.788	-	-	-
1988	6.914	6.932	-	-	-
1989	6.992	7.000	-	-	-
1990	6.861	6.849	-	-	-
1991	6.689	6.667	-	-	-
1992	6.938	7.014	-	-	-
1993	7.169	7.165	-	-	-
1994	7.305	7.343	-	-	-
1995	7.415	7.427	-	-	-
1996	7.604	7.644	-	-	-
1997	7.781	7.807	-	-	-
1998	7.968	8.005	-	-	-
1999	8.228	8.250	-	-	-
2000	8.219	8.249	-	-	-
2001	8.290	-	8.312	8.247	8.377
2002	8.525	-	8.560	8.493	8.626
2003	8.829	-	8.826	8.757	8.895
2004	9.055	-	9.118	9.051	9.185
2005	8.890	-	8.884	8.817	8.952
2006	8.848	-	8.878	8.814	8.942

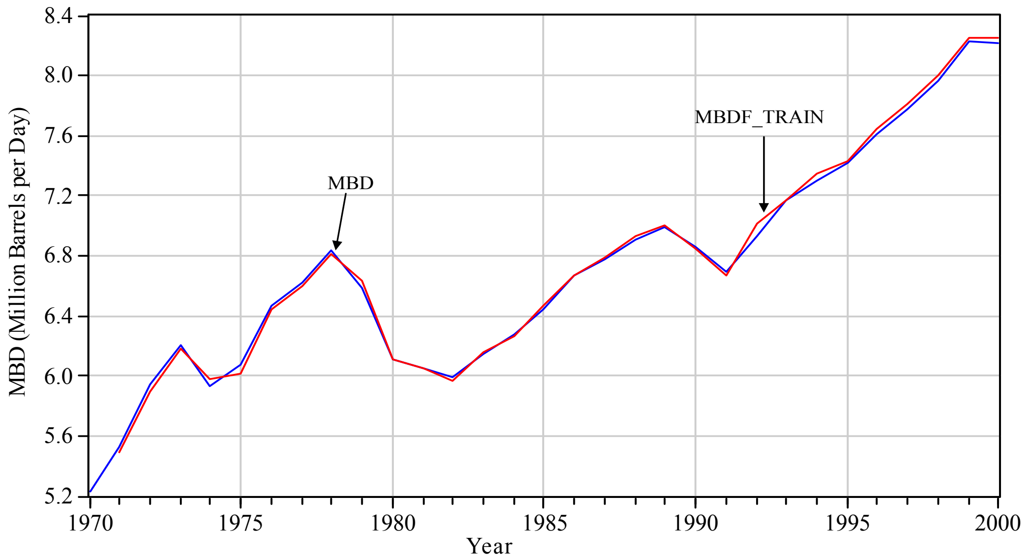


Figure 5-3: MBD Actual vs. MBD Model Data (1970-2000)

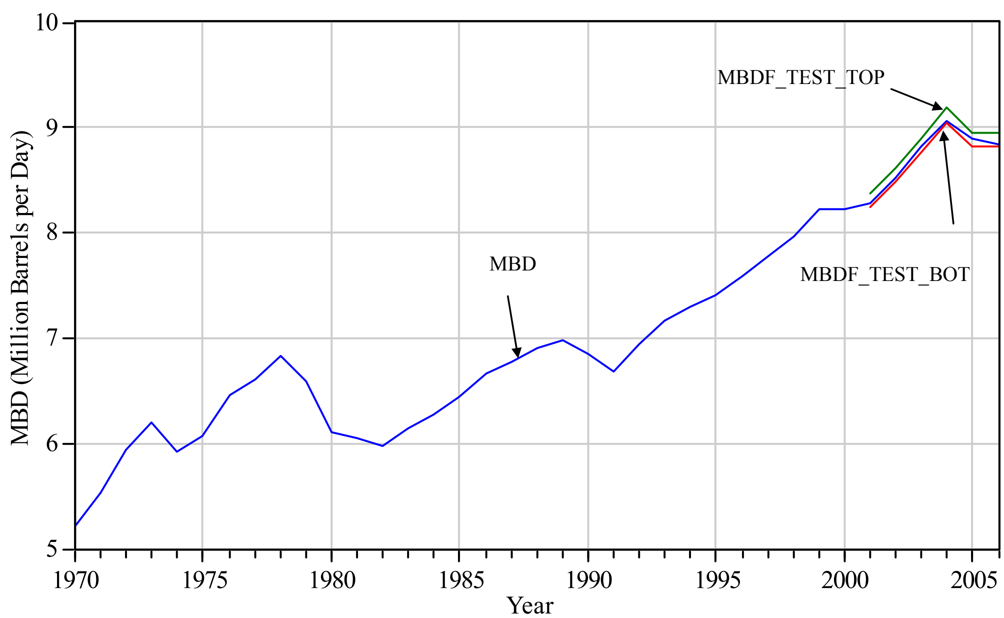


Figure 5-4: MBD Model 95 Percent Confidence Interval (2001-2006)

CHAPTER 6. FORECASTING SCENARIOS

In order to use the explanatory variable model developed in Chapter 5, it is first necessary to generate the data for the explanatory variables included in the model. That is, creating scenarios for the independent variables VEH, PRICE, and MPG.

6.1 Forecasting Scenarios for the VEH Variable

The growth rate of the light vehicle fleet slightly decreased in the 1970 to 2006 period of analysis. This fact led to several possible scenarios for the VEH variable, once the growth variation function was identified.

The MBD series was divided into several 5-year increments, or time intervals, for the purpose of evaluating the growth rate in each particular time interval. Corresponding regressions were run in EVIEWS with the results presented in Table 6-1. Also, a regression to evaluate the annual growth rate of the long term trend for the whole time interval 1970-2006 is shown in the same table.

Table 6-1: Growth Rate Estimation for the VEH Series

No	Period	C2	e^C2	r (percent)	R^2	p
0	1970-2006	0.0207	1.0209	2.09	0.97	0.0000
1	1971-1976	0.0400	1.0408	4.08	0.98	0.0001
2	1976-1981	0.0285	1.0289	2.89	0.98	0.0001
3	1981-1986	0.0222	1.0224	2.24	0.98	0.0001
4	1986-1991	0.0147	1.0148	1.48	0.85	0.0056
5	1991-1996	0.0185	1.0186	1.86	0.99	0.0000
6	1996-2001	0.0215	1.0218	2.18	0.92	0.0016
7	2001-2006	0.0127	1.0128	1.28	0.89	0.0031

The overall growth rate for the period 1970-2006 turned out to be about 2.1 percent. However, when the whole period was divided into seven five years sub-periods, it was observed that the growth trend decreased from a high of about 4.1 percent for the first period 1971-1976 to a low of about 1.28 percent for the last period 2001-2006 analyzed. Therefore, while the VEH variable had been increasing steadily along time its growth rate had been slowly decreasing. Therefore, there existed the possibility of representing the process by a decreasing growth rate model (Bowerman et al. 2005) which needed to be identified.

The forgoing analysis led to hypothesize three possible scenarios for the VEH variable. These scenarios were:

I. For the first VEH scenario or scenario VEH1 a constant annual growth rate of 2.1 percent was assumed for the 2007-2020 period (the same rate that occurred for the period 1970-2006). It is convenient to mention that this is a very improbable scenario as it is known that the growth rate of the VEH variable had been decreasing along time. This scenario may be labeled as the most pessimistic scenario for the VEH variable.

II. A second scenario was built by assuming that for the forecasting 2007-2020 period, a growth rate equal to the growth rate that was identified for the last five year sub-period (2001-2006) which turned out to be a growth rate of about 1.28 percent.

III. Although it is known that the growth rate of the VEH variable had been decreasing along time, this effect is more clearly seen when the variation is portrayed in terms of the VEHP or VEH/POP ratio variable. The behavior of VEHP along time was depicted graphically in Figure **3-6** of

Chapter 3 of this thesis. This turned out to be the most optimistic scenario in terms of the VEH variable.

6.1.1 Scenario VEH1

Scenario VEH1 was obtained by considering a light fleet growth rate of about 2.1 percent which is the long term growth trend of the VEH variable. The regressand in constant growth rate type models is the natural logarithm of the dependent variable, and the regressor is time or period which takes the value 0,1,2,3 and so on. The EVIEWS estimation output is the following:

Dependent Variable: LOG(VEH)
Method: Least Squares
Date: 12/08/08 Time: 23:52
Sample: 1970 2006
Included observations: 37

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.754331	0.012914	368.1394	0.0000
@TREND	0.020737	0.000617	33.60462	0.0000

R-squared	0.969938	Mean dependent var	5.127591
Adjusted R-squared	0.969079	S.D. dependent var	0.227913
S.E. of regression	0.040077	Akaike info criterion	-3.543499
Sum squared resid	0.056215	Schwarz criterion	-3.456422
Log likelihood	67.55473	Hannan-Quinn criter.	-3.512800
F-statistic	1129.270	Durbin-Watson stat	0.118505
Prob(F-statistic)	0.000000		

The corresponding EVIEWS representation is:

```
Estimation Command:
=========================
LS LOG(VEH) C @TREND

Estimation Equation:
=========================
LOG(VEH) = C(1) + C(2)*@TREND

Substituted Coefficients:
=========================
LOG(VEH) = 4.75433063346 + 0.0207366697724*@TREND
```

Table 6-2 presents the corresponding EVIEWS output. It is seen that the light vehicle fleet would grow to about 327 million vehicles in 2020 under this scenario.

Table 6-2: Scenario VEH1

Year	VEH1
2007	250.032
2008	255.271
2009	260.619
2010	266.080
2011	271.655
2012	277.347
2013	283.159
2014	289.092
2015	295.149
2016	301.333
2017	307.647
2018	314.094
2019	320.675
2020	327.394

6.1.2 Scenario VEH2

Considering the hypothesis about the light vehicle fleet continuously growing in the forecasting period at a ratio of approximately 1.28 percent (the growth rate in the period 2001-2006), the constant growth model in EVIEWS produced the following estimation output:

Dependent Variable: LOG(VEH)
Method: Least Squares
Date: 12/09/08 Time: 00:07
Sample: 2001 2006
Included observations: 6

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.994745	0.066924	74.63356	0.0000
@TREND	0.012797	0.001995	6.414208	0.0030

R-squared	0.911391	Mean dependent var	5.423450
Adjusted R-squared	0.889239	S.D. dependent var	0.025078
S.E. of regression	0.008346	Akaike info criterion	-6.472815
Sum squared resid	0.000279	Schwarz criterion	-6.542228
Log likelihood	21.41844	Hannan-Quinn criter.	-6.750683
F-statistic	41.14207	Durbin-Watson stat	1.568439
Prob(F-statistic)	0.003036		

The corresponding EVIEWS representation is the following:

```
Estimation Command:
=========================
LS LOG(VEH) C @TREND

Estimation Equation:
=========================
LOG(VEH) = C(1) + C(2)*@TREND

Substituted Coefficients:
=========================
LOG(VEH) = 4.9947451402 + 0.0127971671712*@TREND
```

Table 6-3 shows VEH2 would reach, under this scenario, a figure of almost 280 million vehicles.

Table 6-3: Scenario VEH2

Year	VEH2
2007	237.040
2008	240.100
2009	243.190
2010	246.320
2011	249.490
2012	252.710
2013	255.960
2014	259.260
2015	262.600
2016	265.980
2017	269.400
2018	272.870
2019	276.390
2020	279.950

6.1.3 Scenario VEH3

The series VEHP referred to before in

Figure **3-6** was modeled as a decreasing

growth trend function whose estimation output and EVIEWS representation are shown next:

Dependent Variable: VEHP
Method: Least Squares
Date: 12/09/08 Time: 00:37
Sample: 1970 2006
Included observations: 37

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.498321	0.004508	110.5500	0.0000
@TREND	2.919796	0.165148	17.67991	0.0000
@TREND^1.002	-2.891201	0.163919	-17.63803	0.0000

R-squared	0.991140	Mean dependent var	0.689573
Adjusted R-squared	0.990619	S.D. dependent var	0.078488
S.E. of regression	0.007602	Akaike info criterion	-6.843178
Sum squared resid	0.001965	Schwarz criterion	-6.712563
Log likelihood	129.5988	Hannan-Quinn criter.	-6.797130
F-statistic	1901.717	Durbin-Watson stat	1.037814
Prob(F-statistic)	0.000000		

```
Estimation Command:
=========================
LS VEHP C @TREND @TREND^1.002

Estimation Equation:
=========================
VEHP = C(1) + C(2)*@TREND + C(3)*@TREND^1.002

Substituted Coefficients:
=========================
VEHP = 0.498320676176 + 2.91979632152*@TREND - 2.89120112054
*@TREND^1.002
```

Figure 6-1 shows the actual VEHP series, as well as the theoretical VEHP model for the 1970-2006 observation period which is identified in EVIEWS as VEHPF.

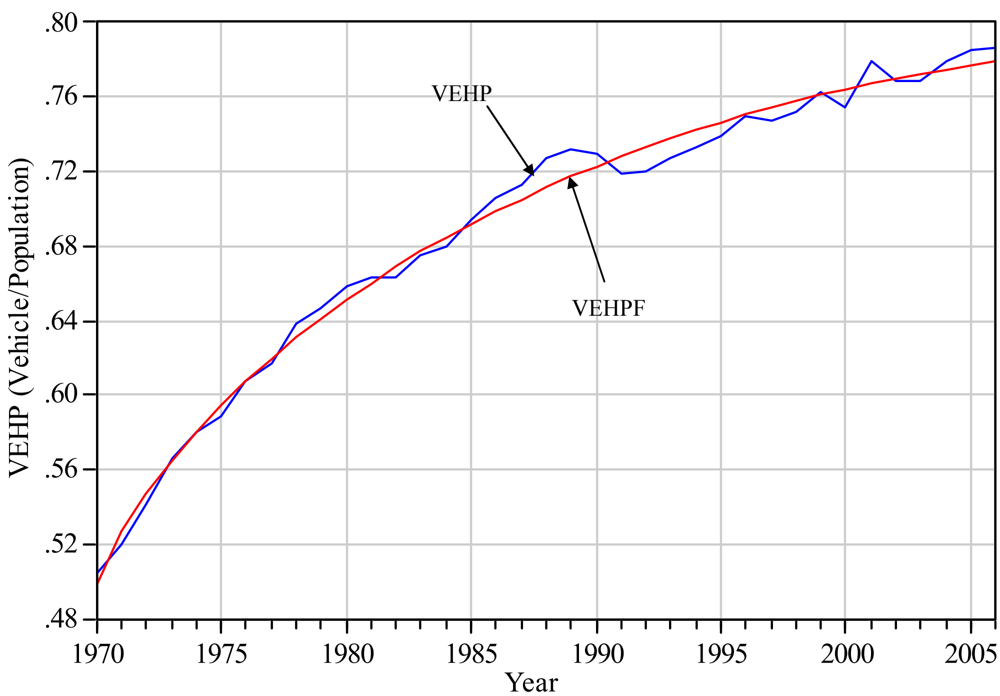


Figure 6-1: VEHP Actual and Decreasing Growth Trend Model

The extension of VEHPF model to the 2007-2020 forecasting period produced the results shown in Table 6-4. VEHPF increased from 0.780 in year 2007 to 0.790 in 2014 and 0.793 in 2020. The slow growth for the last periods is consistent with the theory of vehicle saturation (Ortuzar and Willumsen 2004). The projections for VEH were obtained by multiplying the VEHPF factors by the U.S Census Bureau population projections for the period 2007-2020.

Under this hypothesis, the number of vehicles would grow from 235 million in 2007 to about 266 million in year 2020 which results in the most optimistic growth trend for the VEH variable. This scenario is consistent with a constant VEH growth of only about one per cent per year which is the approximate growth rate of the U.S. population.

Table 6-4: VEHP Forecast by Decreasing Growth Trend Model

Year	VEHPF	POP	VEH3
2007	0.781	300.913	235.012
2008	0.783	303.598	237.636
2009	0.784	306.272	240.214
2010	0.786	308.936	242.747
2011	0.787	311.601	245.242
2012	0.788	314.281	247.711
2013	0.789	316.971	250.151
2014	0.790	319.668	252.557
2015	0.791	322.366	254.927
2016	0.791	325.063	257.258
2017	0.792	327.756	259.547
2018	0.792	330.444	261.793
2019	0.792	333.127	263.998
2020	0.793	335.805	266.159

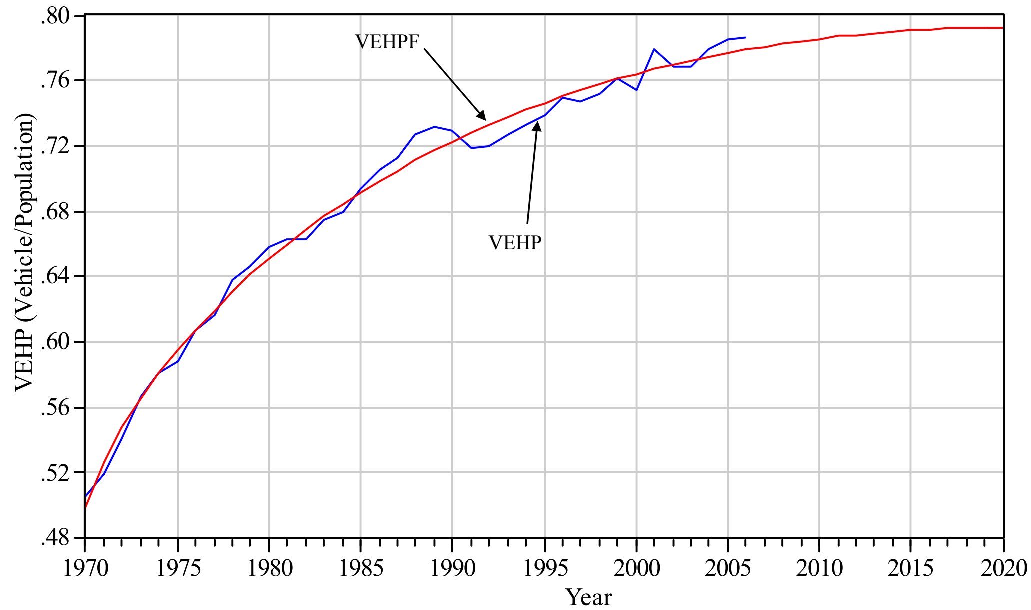


Figure 6-2 shows the extension of the VEHP decreasing growth trend parabolic model to the period 2007-2020 and the actual data in the observation 1970-2006 period.

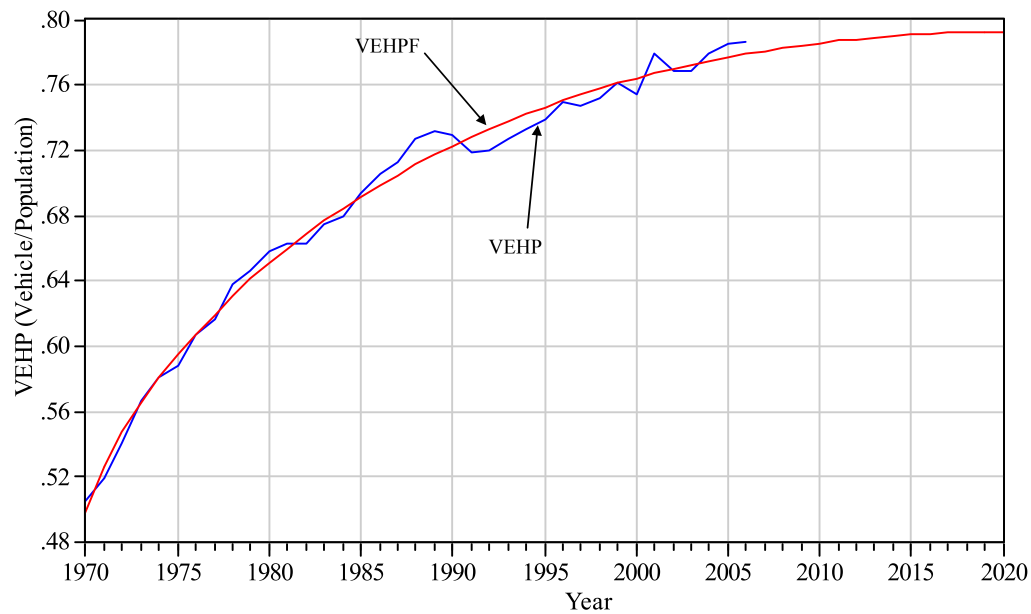


Figure 6-2: Actual VEHP Ratio and Decreasing Growth Rate Trend Model

6.2 Forecasting Scenarios for the MPG variable

The Energy Independent Act of 2007 called for an increase of the efficiency of new vehicles of 4 percent per year up to 35 MPG in year 2020 (The White House Press 2007). The replacement of the light vehicle fleet is around 7% per year (Canes 2007). The corresponding scenarios for the VEH variable were called as follows:

MPG1: Assume MPG as per the Energy Independent Act of 2007, fleet replacement of 7 percent per year, and VEH as per VEH1.

MPG2: Assume MPG as per the Energy Independent Act of 2007, fleet replacement of 7 percent per year, and VEH as per VEH2.

MPG3: Assume MPG as per the Energy Independent Act of 2007, fleet replacement of 7 percent per year, and VEH as per VEH3.

6.2.1 Scenario MPG1

The 7 percent per year replacement of the light vehicle fleet was incorporated into the VEH1 projections to determine the expected number of vehicles of each model year in circulation in every year up to 2020. Table 6-5 shows the expected light vehicle fleet size per year in million vehicles as well as its distribution per vehicle model year in every year up to year 2020.

The efficiency of the light vehicle fleet in circulation in year 2006 was estimated at approximately 20 MPG (Davis and Diegel 2008). The base average number for new 2007 model vehicles was estimated at around 21 MPG (Canes 2007). Increasing this number at a 4 percent annual growth rate beginning in year 2008 led to the 35 MPG in year 2020 as required by the Energy Independence Act of 2007. Table 6-6 is a compilation of the expected efficiency data of the new components of the light vehicle fleet up to year 2020.

Since the law calls for progressively increasing the MPG values to year 2020, and each year there are different amounts of models in circulation, then it is necessary to calculate an MPG average value for the light vehicle fleet every year up to year 2020. For example, the expected average MPG value for year 2009 is:

$$MPG_{2009} = \frac{(207.00 * 20.43) + (17.50 * 21.02) + (17.87 * 21.86) + (18.24 * 22.74)}{(207 + 17.50 + 17.87 + 18.24)} = 20.73 \text{ MPG}$$

The average expected MPG values for every year up to year 2020 were calculated in such a way as per indicated in Table 6-7.

Column 18 of Table 6-7 contains the sum of the products MPG*VEH (the MPG for each model year times the number of vehicles of each model year). Column 19 contains the expected average fleet efficiency data, which was obtained by dividing the sum of the products by the total number of vehicles in circulation in the corresponding year.

Table 6-8 contains the group VEH1-MPG1 expected data for every year up to year 2020.

Table 6-5: Fleet Distribution by Vehicle Model for VEH1

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Projections for Year	Total Number of Vehicles	Veh. Model Older than 2007	Veh. Model 2007	Veh. Model 2008	Veh. Model 2009	Veh. Model 2010	Veh. Model 2011	Veh. Model 2012	Veh. Model 2013	Veh. Model 2014	Veh. Model 2015	Veh. Model 2016	Veh. Model 2017	Veh. Model 2018	Veh. Model 2019	Veh. Model 2020
2007	250.03	232.53	17.50													
2008	255.27	219.90	17.50	17.87												
2009	260.62	207.00	17.50	17.87	18.24											
2010	266.08	193.84	17.50	17.87	18.24	18.63										
2011	271.66	180.40	17.50	17.87	18.24	18.63	19.02									
2012	277.35	166.68	17.50	17.87	18.24	18.63	19.02	19.41								
2013	283.16	152.67	17.50	17.87	18.24	18.63	19.02	19.41	19.82							
2014	289.09	138.36	17.50	17.87	18.24	18.63	19.02	19.41	19.82	20.24						
2015	295.15	123.76	17.50	17.87	18.24	18.63	19.02	19.41	19.82	20.24	20.66					
2016	301.33	108.85	17.50	17.87	18.24	18.63	19.02	19.41	19.82	20.24	20.66	21.09				
2017	307.65	93.63	17.50	17.87	18.24	18.63	19.02	19.41	19.82	20.24	20.66	21.09	21.54			
2018	314.09	78.09	17.50	17.87	18.24	18.63	19.02	19.41	19.82	20.24	20.66	21.09	21.54	21.99		
2019	320.67	62.22	17.50	17.87	18.24	18.63	19.02	19.41	19.82	20.24	20.66	21.09	21.54	21.99	22.45	
2020	327.39	46.03	17.50	17.87	18.24	18.63	19.02	19.41	19.82	20.24	20.66	21.09	21.54	21.99	22.45	22.92

Table 6-6: Projections for Light Vehicle Efficiency Data from 2008 to 2020 as per the Energy Independence Act of 2007

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Year	Veh. Model Older than 2007	Veh. Model 2007	Veh. Model 2008	Veh. Model 2009	Veh. Model 2010	Veh. Model 2011	Veh. Model 2012	Veh. Model 2013	Veh. Model 2014	Veh. Model 2015	Veh. Model 2016	Veh. Model 2017	Veh. Model 2018	Veh. Model 2019	Veh. Model 2020
2007	20.43	21.02													
2008	20.43	21.02	21.86												
2009	20.43	21.02	21.86	22.74											
2010	20.43	21.02	21.86	22.74	23.64										
2011	20.43	21.02	21.86	22.74	23.64	24.59									
2012	20.43	21.02	21.86	22.74	23.64	24.59	25.57								
2013	20.43	21.02	21.86	22.74	23.64	24.59	25.57	26.60							
2014	20.43	21.02	21.86	22.74	23.64	24.59	25.57	26.60	27.66						
2015	20.43	21.02	21.86	22.74	23.64	24.59	25.57	26.60	27.66	28.77					
2016	20.43	21.02	21.86	22.74	23.64	24.59	25.57	26.60	27.66	28.77	29.92				
2017	20.43	21.02	21.86	22.74	23.64	24.59	25.57	26.60	27.66	28.77	29.92	31.11			
2018	20.43	21.02	21.86	22.74	23.64	24.59	25.57	26.60	27.66	28.77	29.92	31.11	32.36		
2019	20.43	21.02	21.86	22.74	23.64	24.59	25.57	26.60	27.66	28.77	29.92	31.11	32.36	33.65	
2020	20.43	21.02	21.86	22.74	23.64	24.59	25.57	26.60	27.66	28.77	29.92	31.11	32.36	33.65	35.00

Table 6-7: Product of Number of Vehicles per Year by MPG for the Same Year

1 Year	2 Total Number of Vehicles	3 Veh. Model Older than 2007	4 Veh. Model 2007	5 Veh. Model 2008	6 Veh. Model 2009	7 Veh. Model 2010	8 Veh. Model 2011	9 Veh. Model 2012	10 Veh. Model 2013	11 Veh. Model 2014	12 Veh. Model 2015	13 Veh. Model 2016	14 Veh. Model 2017	15 Veh. Model 2018	16 Veh. Model 2019	17 Veh. Model 2020	18 Sum	19 Fleet Efficiency
2007	250.03	4750.58	367.90														5118.47	20.47
2008	255.27	4492.55	367.90	390.63													5251.07	20.57
2009	260.62	4229.11	367.90	390.63	414.77												5402.40	20.73
2010	266.08	3960.15	367.90	390.63	414.77	440.40											5573.84	20.95
2011	271.66	3685.56	367.90	390.63	414.77	440.40	467.61										5766.86	21.23
2012	277.35	3405.21	367.90	390.63	414.77	440.40	467.61	496.50									5983.01	21.57
2013	283.16	3118.99	367.90	390.63	414.77	440.40	467.61	496.50	527.18								6223.97	21.98
2014	289.09	2826.77	367.90	390.63	414.77	440.40	467.61	496.50	527.18	559.76							6491.51	22.45
2015	295.15	2528.43	367.90	390.63	414.77	440.40	467.61	496.50	527.18	559.76	594.35						6787.52	23.00
2016	301.33	2223.84	367.90	390.63	414.77	440.40	467.61	496.50	527.18	559.76	594.35	631.07					7114.00	23.61
2017	307.65	1912.87	367.90	390.63	414.77	440.40	467.61	496.50	527.18	559.76	594.35	631.07	670.07				7473.09	24.29
2018	314.09	1595.38	367.90	390.63	414.77	440.40	467.61	496.50	527.18	559.76	594.35	631.07	670.07	711.47			7867.07	25.05
2019	320.67	1271.24	367.90	390.63	414.77	440.40	467.61	496.50	527.18	559.76	594.35	631.07	670.07	711.47	755.43		8298.36	25.88
2020	327.39	940.30	367.90	390.63	414.77	440.40	467.61	496.50	527.18	559.76	594.35	631.07	670.07	711.47	755.43	802.11	8769.54	26.79

Table 6-8: Summary Results for VEH1-MPG1

Projection Year	VEH1	MPG1
2007	250.03	20.47
2008	255.27	20.57
2009	260.62	20.73
2010	266.08	20.95
2011	271.66	21.23
2012	277.35	21.57
2013	283.16	21.98
2014	289.09	22.45
2015	295.15	23.00
2016	301.33	23.61
2017	307.65	24.29
2018	314.09	25.05
2019	320.67	25.88
2020	327.39	26.79

6.2.2 Scenario MPG2

The 7 percent per year replacement of the light vehicle fleet was incorporated into the projections VEH2 to determine the expected number of vehicles of each model year in circulation in every year up to year 2020. The corresponding numbers are shown in Table 6-9. The figures for fleet distribution per model per year shown in Table 6-9 were then multiplied by the corresponding expected MPG data (Table 6-6), and then divided by the total number of vehicles in the fleet to determine the average MPG2 data for the light vehicle fleet up to year 2020, in the same way as it was performed to obtain MPG1. The corresponding calculations are shown in Table 6-10.

Column 19 in Table 6-10 contains the average MPG2 data for the fleet under scenario VEH2-MPG2, and Table 6-11 contains the group VEH2-MPG2 expected summary data for every year up to year 2020.

Table 6-9: Fleet Distribution by Vehicle Model for VEH2

1 Projections for Year	2 Total Number of Vehicles	3 Veh. Model Older than 2007	4 Veh. Model 2007	5 Veh. Model 2008	6 Veh. Model 2009	7 Veh. Model 2010	8 Veh. Model 2011	9 Veh. Model 2012	10 Veh. Model 2013	11 Veh. Model 2014	12 Veh. Model 2015	13 Veh. Model 2016	14 Veh. Model 2017	15 Veh. Model 2018	16 Veh. Model 2019	17 Veh. Model 2020
2007	237.04	220.45	16.59													
2008	240.10	206.70	16.59	16.81												
2009	243.19	192.77	16.59	16.81	17.02											
2010	246.32	178.65	16.59	16.81	17.02	17.24										
2011	249.49	164.36	16.59	16.81	17.02	17.24	17.46									
2012	252.71	149.89	16.59	16.81	17.02	17.24	17.46	17.69								
2013	255.96	135.22	16.59	16.81	17.02	17.24	17.46	17.69	17.92							
2014	259.26	120.38	16.59	16.81	17.02	17.24	17.46	17.69	17.92	18.15						
2015	262.60	105.33	16.59	16.81	17.02	17.24	17.46	17.69	17.92	18.15	18.38					
2016	265.98	90.09	16.59	16.81	17.02	17.24	17.46	17.69	17.92	18.15	18.38	18.62				
2017	269.40	74.66	16.59	16.81	17.02	17.24	17.46	17.69	17.92	18.15	18.38	18.62	18.86			
2018	272.87	59.03	16.59	16.81	17.02	17.24	17.46	17.69	17.92	18.15	18.38	18.62	18.86	19.10		
2019	276.39	43.20	16.59	16.81	17.02	17.24	17.46	17.69	17.92	18.15	18.38	18.62	18.86	19.10	19.35	
2020	279.95	27.16	16.59	16.81	17.02	17.24	17.46	17.69	17.92	18.15	18.38	18.62	18.86	19.10	19.35	19.60

Table 6-10: Product of Number of Vehicles per Year by MPG for the Same Year

1 Year	2 Total Number of Vehicles	3 Veh. Model Older than 2007	4 Veh. Model 2007	5 Veh. Model 2008	6 Veh. Model 2009	7 Veh. Model 2010	8 Veh. Model 2011	9 Veh. Model 2012	10 Veh. Model 2013	11 Veh. Model 2014	12 Veh. Model 2015	13 Veh. Model 2016	14 Veh. Model 2017	15 Veh. Model 2018	16 Veh. Model 2019	17 Veh. Model 2020	18 Sum	19 Fleet Efficiency
2007	237.04	4503.74	348.78														4852.52	20.47
2008	240.10	4222.89	348.78	367.41													4939.08	20.57
2009	243.19	3938.23	348.78	367.41	387.03												5041.45	20.73
2010	246.32	3649.91	348.78	367.41	387.03	407.69											5160.83	20.95
2011	249.49	3357.88	348.78	367.41	387.03	407.69	429.45										5298.25	21.24
2012	252.71	3062.26	348.78	367.41	387.03	407.69	429.45	452.40									5455.03	21.59
2013	255.96	2762.61	348.78	367.41	387.03	407.69	429.45	452.40	476.54								5631.92	22.00
2014	259.26	2459.26	348.78	367.41	387.03	407.69	429.45	452.40	476.54	502.00							5830.57	22.49
2015	262.60	2151.96	348.78	367.41	387.03	407.69	429.45	452.40	476.54	502.00	528.80						6052.06	23.05
2016	265.98	1840.63	348.78	367.41	387.03	407.69	429.45	452.40	476.54	502.00	528.80	557.03					6297.77	23.68
2017	269.40	1525.23	348.78	367.41	387.03	407.69	429.45	452.40	476.54	502.00	528.80	557.03	586.76				6569.13	24.38
2018	272.87	1205.89	348.78	367.41	387.03	407.69	429.45	452.40	476.54	502.00	528.80	557.03	586.76	618.09			6867.88	25.17
2019	276.39	882.54	348.78	367.41	387.03	407.69	429.45	452.40	476.54	502.00	528.80	557.03	586.76	618.09	651.11		7195.64	26.03
2020	279.95	554.92	348.78	367.41	387.03	407.69	429.45	452.40	476.54	502.00	528.80	557.03	586.76	618.09	651.11	685.87	7553.89	26.98

Table 6-11: Summary Results for VEH2-MPG2

Projection Year	VEH2	MPG2
2007	237.04	20.47
2008	240.10	20.57
2009	243.19	20.73
2010	246.32	20.95
2011	249.49	21.24
2012	252.71	21.59
2013	255.96	22.00
2014	259.26	22.49
2015	262.60	23.98
2016	265.98	23.68
2017	269.40	25.49
2018	272.87	25.17
2019	276.39	26.03
2020	279.95	26.98

6.2.3 Scenario MPG3

The 7 percent per year replacement of the light vehicle fleet was incorporated into the projections VEH3 to determine the expected number of vehicles of each model year in circulation in every year up to 2020. The resulting numbers are shown in Table 6-12.

The figures for fleet distribution per model per year shown in Table 6-12 were then multiplied by the corresponding expected MPG data (Table 6-6), and then divided by the total number of vehicles in the fleet to determine the average MPG3 data for the light vehicle fleet up to year 2020, in the same way as it was performed to obtain MPG2. The corresponding calculations are shown in Table 6-13.

Column 19 in Table 6-13 contains the average MPG3 data for the fleet under scenario VEH3-MPG3.

Table 6-14 contains the group VEH3-MPG3 expected summary data for every year up to year 2020.

Table 6-12: Fleet Distribution by Vehicle Model for VEH

1 Projections for Year	2 Total Number of Vehicles	3 Veh. Model Older than 2007	4 Veh. Model 2007	5 Veh. Model 2008	6 Veh. Model 2009	7 Veh. Model 2010	8 Veh. Model 2011	9 Veh. Model 2012	10 Veh. Model 2013	11 Veh. Model 2014	12 Veh. Model 2015	13 Veh. Model 2016	14 Veh. Model 2017	15 Veh. Model 2018	16 Veh. Model 2019	17 Veh. Model 2020
2007	235.01	218.56	16.45													
2008	237.64	204.55	16.45	16.63												
2009	240.21	190.31	16.45	16.63	16.81											
2010	242.75	175.86	16.45	16.63	16.81	16.99										
2011	245.24	161.18	16.45	16.63	16.81	16.99	17.17									
2012	247.71	146.31	16.45	16.63	16.81	16.99	17.17	17.34								
2013	250.15	131.24	16.45	16.63	16.81	16.99	17.17	17.34	17.51							
2014	252.56	115.97	16.45	16.63	16.81	16.99	17.17	17.34	17.51	17.68						
2015	254.93	100.50	16.45	16.63	16.81	16.99	17.17	17.34	17.51	17.68	17.85					
2016	257.26	84.82	16.45	16.63	16.81	16.99	17.17	17.34	17.51	17.68	17.85	18.01				
2017	259.55	68.94	16.45	16.63	16.81	16.99	17.17	17.34	17.51	17.68	17.85	18.01	18.17			
2018	261.79	52.85	16.45	16.63	16.81	16.99	17.17	17.34	17.51	17.68	17.85	18.01	18.17	18.33		
2019	264.00	36.58	16.45	16.63	16.81	16.99	17.17	17.34	17.51	17.68	17.85	18.01	18.17	18.33	18.48	
2020	266.16	20.11	16.45	16.63	16.81	16.99	17.17	17.34	17.51	17.68	17.85	18.01	18.17	18.33	18.48	18.63

Table 6-13: Product of Number of Vehicles per Year by MPG for the Same Year

1 Year	2 Total Number of Vehicles	3 Veh. Model Older than 2007	4 Veh. Model 2007	5 Veh. Model 2008	6 Veh. Model 2009	7 Veh. Model 2010	8 Veh. Model 2011	9 Veh. Model 2012	10 Veh. Model 2013	11 Veh. Model 2014	12 Veh. Model 2015	13 Veh. Model 2016	14 Veh. Model 2017	15 Veh. Model 2018	16 Veh. Model 2019	17 Veh. Model 2020	18 Sum	19 Fleet Efficiency
2007	235.01	4465.17	345.79														4810.96	20.47
2008	237.64	4179.05	345.79	363.65													4888.49	20.57
2009	240.21	3888.03	345.79	363.65	382.29												4979.76	20.73
2010	242.75	3592.76	345.79	363.65	382.29	401.78											5086.28	20.95
2011	245.24	3292.92	345.79	363.65	382.29	401.78	422.14										5208.57	21.24
2012	247.71	2989.13	345.79	363.65	382.29	401.78	422.14	443.45									5348.23	21.59
2013	250.15	2681.24	345.79	363.65	382.29	401.78	422.14	443.45	465.73								5506.06	22.01
2014	252.56	2369.29	345.79	363.65	382.29	401.78	422.14	443.45	465.73	489.02							5683.14	22.50
2015	254.93	2053.13	345.79	363.65	382.29	401.78	422.14	443.45	465.73	489.02	513.36						5880.33	23.07
2016	257.26	1732.83	345.79	363.65	382.29	401.78	422.14	443.45	465.73	489.02	513.36	538.77					6098.80	23.71
2017	259.55	1408.43	345.79	363.65	382.29	401.78	422.14	443.45	465.73	489.02	513.36	538.77	565.31				6339.71	24.43
2018	261.79	1079.81	345.79	363.65	382.29	401.78	422.14	443.45	465.73	489.02	513.36	538.77	565.31	592.99			6604.08	25.23
2019	264.00	747.41	345.79	363.65	382.29	401.78	422.14	443.45	465.73	489.02	513.36	538.77	565.31	592.99	621.92		6893.60	26.11
2020	266.16	410.90	345.79	363.65	382.29	401.78	422.14	443.45	465.73	489.02	513.36	538.77	565.31	592.99	621.92	652.09	7209.19	27.09

Table 6-14: Summary Results for VEH3-MPG3

Projection Year	VEH3	MPG3
2007	235.01	20.47
2008	237.64	20.57
2009	240.21	20.73
2010	242.75	20.95
2011	245.24	21.24
2012	247.71	21.59
2013	250.15	22.01
2014	252.56	22.50
2015	254.93	23.07
2016	257.26	23.71
2017	259.55	24.43
2018	261.79	25.23
2019	264.00	26.11
2020	266.16	27.09

6.3 Forecasting Scenarios for the PRICE Variable

Three scenarios were hypothesized for the PRICE variable. The first scenario used for the oil price forecast consisted in using the average annual oil price in the observation period. The second scenario consisted of extending the oil price trend by exponential smoothing, and the third scenario used the oil price forecast provided by the Energy Information Agency (Energy Information Administration 2008b). These three corresponding scenarios are named PRICE1, PRICE2, and PRICE3.

6.3.1 Scenario PRICE1

The first scenario consisted of determining the average oil price in the observation period 1970-2008. Average oil prices were extracted from the Statistical Review provided by the British Petroleum Corporation. For year 2008 the average oil price was \$96.94 per barrel which was the historical maximum oil price ever.

The EVIEWS calculated statistics for the oil price time series are the following:

Sample: 1970 2008

	PRICE
Mean	41.63798
Median	33.64058
Maximum	96.94000
Minimum	9.625669
Std. Dev.	23.35691
Skewness	0.839238
Kurtosis	2.829612
Jarque-Bera	4.625264
Probability	0.099000
Sum	1623.881
Sum Sq. Dev.	20730.71
Observations	39

The average annual oil price used accordingly under scenario PRICE1 was 41.64 dollars per barrel. Figure 6-3 shows that the EVIEWS generated histogram for the oil price time series follows approximately a Gamma distribution. It is also observed a low probability for annual average oil prices beyond one hundred dollars per barrel.

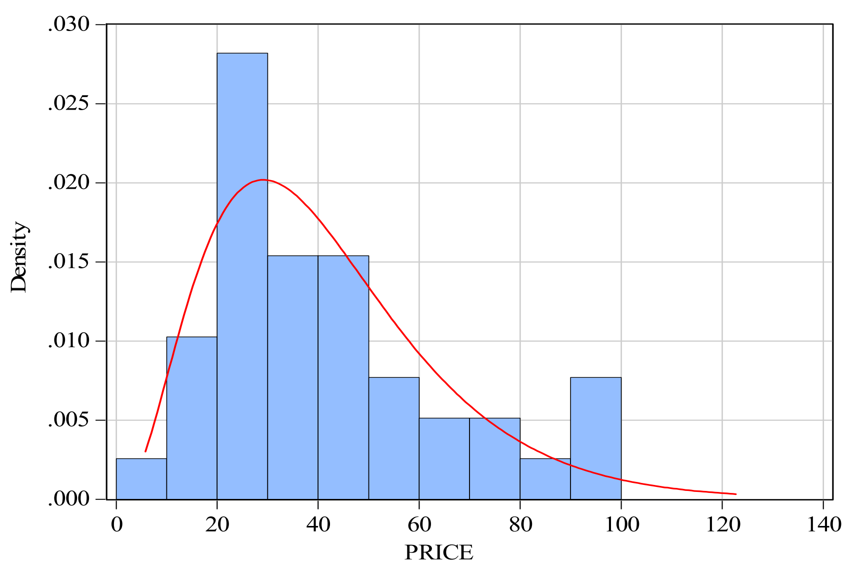


Figure 6-3: Histogram and Theoretical Distribution for the PRICE Variable

6.3.2 Scenario PRICE2

The exponential smoothing model output in EVIEWS in the observation period 1970-2008 for scenario PRICE 2 is shown next:

Sample: 1970 2008
Included observations: 39
Method: Holt-Winters No Seasonal
Original Series: PRICE
Forecast Series: PRICESM

Parameters:	Alpha	1.0000
	Beta	0.0000
Sum of Squared Residuals		5754.345
Root Mean Squared Error		12.14691
End of Period Levels:	Mean	96.94000
	Trend	1.097515

Table 6-15 contains the corresponding trend extension obtained by this method. As observed, oil prices were assumed to go from about 98 dollars per barrel in year 2009 to about 110 dollars per barrel in 2020. This scenario turned out to be the high price scenario as compared to the other two.

Table 6-15:Oil Price Trend by Exponential Smoothing

Year	PRICE2
2009	98.04
2010	99.14
2011	100.23
2012	101.33
2013	102.43
2014	103.53
2015	104.62
2016	105.72
2017	106.82
2018	107.92
2019	109.01
2020	110.11

6.3.3 Scenario PRICE3

The International Energy Agency, as mentioned before, has provided a long term forecast for the oil price from which the values for the period 2009-2020 were extracted. The corresponding information is shown in Table 6-16. Oil prices obtained by this method went from about 68 for year 2009 to about 52 for year 2020. As seen, this scenario placed between the other two at the beginning of the forecast period but at the end of this period it tended to look more like the low oil price scenario.

Table 6-16:Table 6-16 Oil Price Trend from the EIA

Year	PRICE3
2009	68.32
2010	65.18
2011	62.67
2012	60.06
2013	57.36
2014	54.67
2015	52.03
2016	49.37
2017	49.50
2018	50.15
2019	50.92
2020	51.55

CHAPTER 7. MBD FORECASTING

Three scenarios were defined for the VEH1, VEH2, and VEH3 independent variables, corresponding to approximate light vehicle fleet annual growth rates of 2.1, 1.28, and 1 percent respectevly. After combining these VEH data with the 7 percent fleet replacement historical parameter, and the 4 percent annual increase in MPG mandated by the Energy Independence Act of 2007, three scenarios for the MPG variables corresponding to each one of the VEH variables were found. Therefore, on one side there were three combinations of VEH-MPG: VEH1-MPG1, VEH2-MPG2, and VEH3-MPG3; and on the other side, there were three scenarios created for the PRICE variable: PRICE1, PRICE2, and PRICE3; then there were nine and only nine possibilities to combine MBD. This so happened, because the MBD variable was ultimately made dependent upon the variables VEH, PRICE, and MPG.

To obtain MBD it is necessary first to obtain VMT, because MBD is a function of VMT and MPG, but VMT is a function of VEH and PRICE. Then, because there were three scenarios for VEH and three scenarios for PRICE then there were also nine and only nine possibilities to combine VMT.

The coding selected to make a distinction among the nine VMT regression possibilities, as well as the nine MBD regression possibilities are described next.

7.1 VMT Coding

Figure 7-1 indicates the coding selected to differentiate the nine VMT scenarios. The first digit indicates the scenario for VEH, the second digit indicates the scenario for PRICE, and the letter F is readily assigned by EVIEWS to distinguish between the actual, or observed values, and the model theoretical, or calculated values.

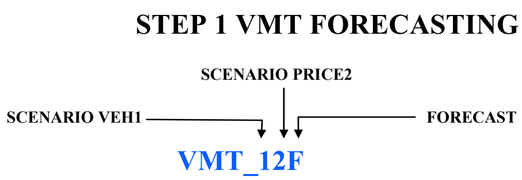


Figure 7-1: VMT Coding

With the above defined coding the nine scenarios for VMT are:

VMT_11F : VEH1 & PRICE1
VMT_12F : VEH1 & PRICE2
VMT_13F : VEH1 & PRICE3

VMT_21F : VEH2 & PRICE1
VMT_22F : VEH2 & PRICE2
VMT_23F : VEH2 & PRICE3

VMT_31F : VEH3 & PRICE1
VMT_32F : VEH3 & PRICE2
VMT_33F : VEH3 & PRICE3

7.2 MBD Coding

The MBD coding is depicted in Figure 7-2. The first digit indicates the corresponding scenario for the VEH variable, the second digit indicates the scenario for the MPG variable, the third digit indicates the scenario for the PRICE variable, and the letter F is, as mentioned before, the EVIEWS designation for forecasting or modeling.

STEP 2 MBD FORECASTING

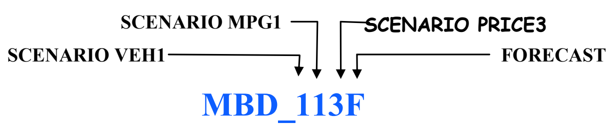


Figure 7-2: MBD Coding

With the above referred coding the nine possible scenarios for the MBD variable are:

MBD_111F:	VMT_11F & MPG1
MBD_112F:	VMT_12F & MPG1
MBD_113F:	VMT_13F & MPG1
MBD_221F:	VMT_21F & MPG2
MBD_222F:	VMT_22F & MPG2
MBD_223F:	VMT_23F & MPG2
MBD_331F:	VMT_31F & MPG3
MBD_332F:	VMT_32F & MPG3
MBD_333F:	VMT_33F & MPG3

7.3 MBD Forecasting Set: MBD_111F, MBD_112F, and MBD_113F

In step 1 VEH1 and PRICE1 were input into the VMT model described in section 5.4.1 to generate VMT_11F; in the same way VEH1 and PRICE2 were input into the same model to generate VMT_12F; and, in the same way VEH1 and PRICE3 were input into the model to generate VMT_13F.

The corresponding output is shown in Table 7-1. It is seen that VMT under scenarios VMT_11F, VMT_12F, and VMT_13F will grow to values between 3.36 trillion vehicles miles and 3.43 trillion vehicles miles in year 2020.

Table 7-1: VMT Scenario Set: VMT_11F, VMT_12F, and VMT_13F

Year	VMT_11F	VMT_12F	VMT_13F
2007	2.874	2.874	2.874
2008	2.890	2.890	2.890
2009	2.988	2.930	2.960
2010	3.028	2.968	3.003
2011	3.067	3.006	3.045
2012	3.106	3.044	3.087
2013	3.145	3.082	3.129
2014	3.184	3.120	3.170
2015	3.223	3.158	3.212
2016	3.263	3.196	3.255
2017	3.303	3.235	3.295
2018	3.343	3.275	3.335
2019	3.385	3.315	3.375
2020	3.427	3.356	3.416

VMT_11F, VMT_12F and VMT_13F were combined each one with MPG1, in step2 as per the model defined and described in section 5.4.2 to obtain MBD_111F, MBD_112F, and MBD_113F. The corresponding MBD output is shown in Table 7-2.

It is seen that under Forecast Set: MBD_111F, MBD_112F, and MBD_113F, fuel consumption would reach a value between 8.59 and 8.89 MBDs in year 2020. This Forecast Set (MBD_111F, MBD_112F, and MBD_113F) contains the most pessimistic forecast for gasoline consumption at 8.89 MBDs in year 2020. Since gasoline consumption in 2006 stood at 8.85 MBDs, even in the most pessimistic scenario, the fuel efficiency improvements under the Energy Independence Act of 2007 will be approximately able to offset the additional consumption projected by the trend in year 2020.

Table 7-2: MBD Forecast Set: MBD_111F, MBD_112F, and MBD_113F

Year	MBD_111F	MBD_112F	MBD_113F
2007	9.237	9.237	9.237
2008	9.264	9.264	9.264
2009	9.604	9.361	9.489
2010	9.677	9.429	9.576
2011	9.723	9.470	9.632
2012	9.742	9.484	9.663
2013	9.732	9.469	9.664
2014	9.697	9.429	9.640
2015	9.629	9.357	9.584
2016	9.538	9.261	9.504
2017	9.419	9.137	9.385
2018	9.269	8.983	9.232
2019	9.093	8.801	9.052
2020	8.885	8.589	8.842

The MBD Forecast Set described here is portrayed graphically in Figure 7-3. Looking vertically at the figure in the forecasting region or horizontally in Table 7-2, it is possible to illustrate the effect of oil price increase on gasoline consumption. The increasing price from scenario PRICE1 (oil price in the low 40's) to PRICE2 (oil price in the low 100's) according to this analysis would cause a decrease in consumption of about 0.30 MBDs

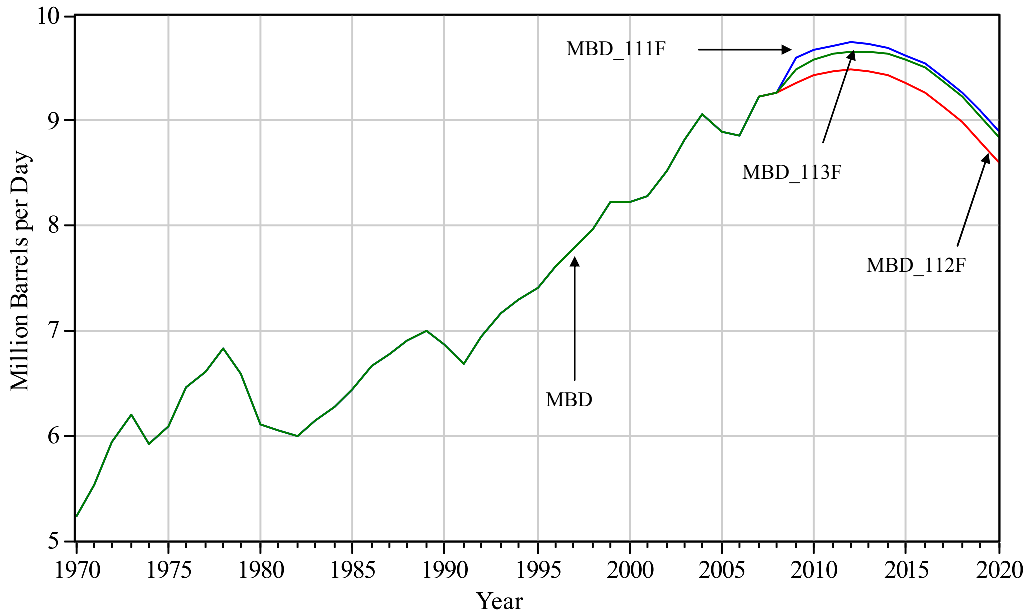


Figure 7-3: MBD Forecast Set: MBD_111F, MBD_112F and MBD_113F

Forecast MBD_113F that placed in the middle (median) of the three forecasts is illustrated in Figure 7-4. The actual MBD series as well as its ARIMA extended MBD series trend (the baseline) are also illustrated in the same Figure. MBD_113F gasoline consumption in 2020 (8.84 MBDs) is about the same as the actual or observed consumption in year 2006 (8.85 MBDs) but 1.33 MBDs below what would be consumed in year 2020 (10.17 MBDs).

It is also seen in Figure 7-4 that the forecast exceeds the trend in the initial forecasting period. This event seems rather improbable. This fact can be interpreted as the VEH1 hypothesis, in which this forecast is based, being too much on the pessimistic side; therefore, leading to extreme high MBD forecasting values.

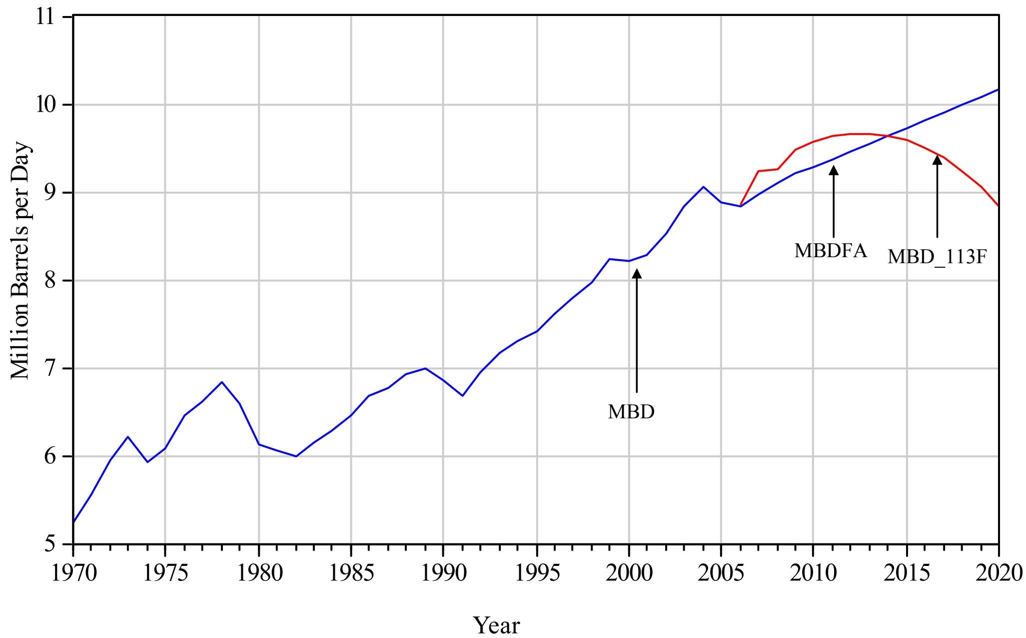


Figure 7-4: Baseline (MBDFA) Comparison with Forecast MBD_113F

7.4 MBD Forecasting Set: MBD_221F, MBD_222F, and MBD_223F

In step 1 VEH2 and PRICE1 were input into the VMT model described in section 5.4.1 to generate VMT_21F; in the same way VEH2 and PRICE2 were input into the same model to generate VMT_22F; and, in the same way VEH2 and PRICE3 were input into the model to generate VMT_23F.

The corresponding output is shown in Table 7-3. It is seen that VMT under this scenario set will grow to values between 3.06 trillion vehicles miles and 3.13 trillion vehicles miles.

Table 7-3: VMT Scenario Set: VMT_21F, VMT_22F, and VMT_23F

Year	VMT_21F	VMT_22F	VMT_23F
2007	2.794	2.794	2.794
2008	2.797	2.797	2.797
2009	2.881	2.822	2.853
2010	2.906	2.847	2.882
2011	2.931	2.870	2.909
2012	2.954	2.892	2.935
2013	2.978	2.914	2.961
2014	3.000	2.936	2.987
2015	3.023	2.957	3.012
2016	3.045	2.979	3.037
2017	3.068	3.000	3.059
2018	3.090	3.021	3.081
2019	3.112	3.042	3.103
2020	3.135	3.064	3.125

VMT_21F, VMT_22F, and VMT_23F were combined each one with MPG2, in step 2, as per the model defined in section 5.4.2 to obtain MBD_221F, MBD_222F, and MBD_223F. The corresponding MBD output is shown in Table 7-4.

It is seen that under this scenario set, the light vehicle fleet fuel consumption will be between 7.30 and 7.59 MBDs in year 2020. This scenario set which is illustrated in Figure **7-5** turned out to be the intermediate MBD forecast set.

Table 7-4: MBD Forecast Set: MBD_221F, MBD_222F, MBD_223F

Year	MBD_221F	MBD_222F	MBD_223F
2007	8.905	8.905	8.905
2008	8.875	8.875	8.875
2009	9.158	8.915	9.043
2010	9.172	8.923	9.070
2011	9.152	8.899	9.061
2012	9.103	8.845	9.024
2013	9.027	8.765	8.960
2014	8.916	8.649	8.860
2015	8.775	8.503	8.730
2016	8.604	8.327	8.570
2017	8.403	8.121	8.369
2018	8.164	7.878	8.127
2019	7.896	7.605	7.856
2020	7.591	7.296	7.549

Looking at the figure vertically in the forecasting region or horizontally in the corresponding table, it is possible to appreciate the effect of oil price increase. The increasing price from scenario PRICE1 (oil price in the low 40's) to PRICE2 (oil price in the low 100's) would lead to decrease in oil consumption up to about 0.30 MBDs.

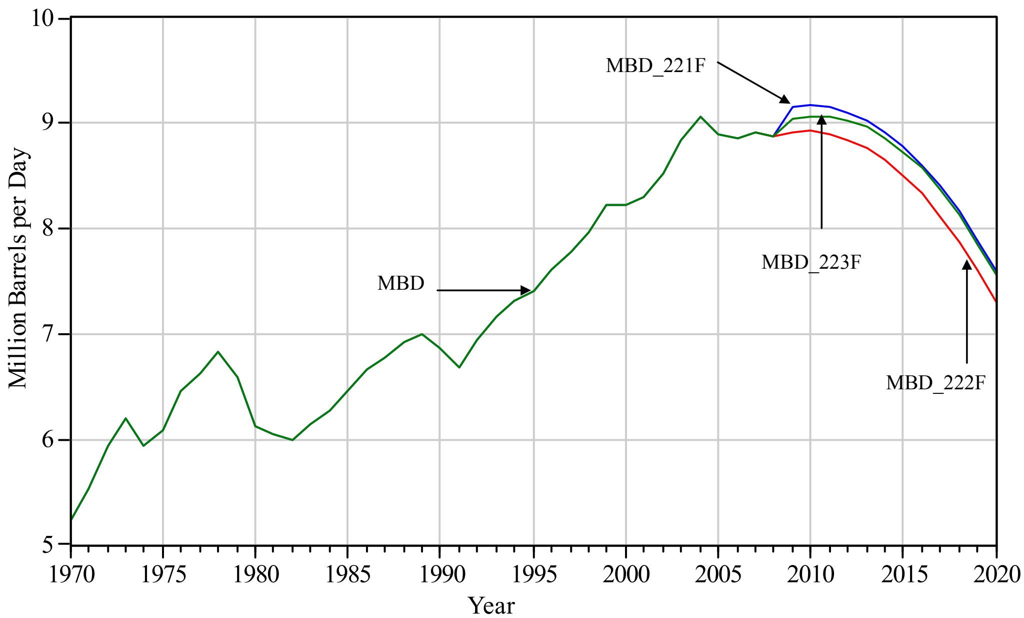


Figure 7-5: MBD Forecast Set: MBD_221F, MBD_222F and MBD_223F

Forecast MBD_223F that placed in the middle (median) of the three forecasts is illustrated in Figure 7-6. The actual MBD series as well as its ARIMA extended MBD series trend (the baseline) are also illustrated in the same Figure. MBD_223F gasoline consumption in 2020 (7.55 MBDs) is 1.30 MBDs below the actual consumption in year 2006 (8.85 MBDs) and 2.62 MBDs below what would be consumed in 2020 (10.17 MBDs).

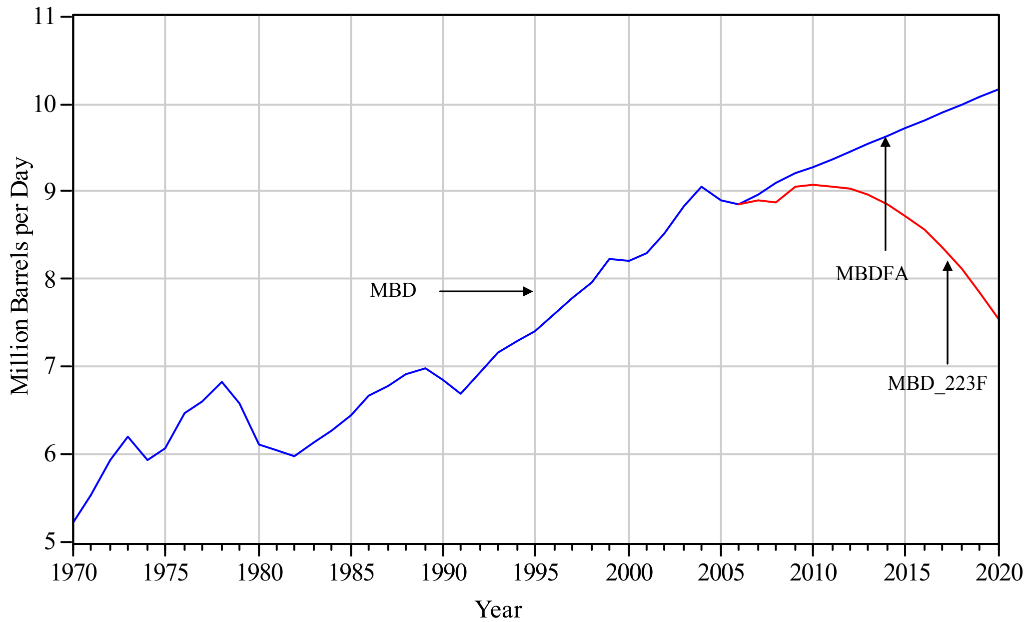


Figure 7-6: MBDFA (Baseline) Comparison with MBD_223F

7.5 MBD Forecasting Set: MBD_331F, MBD_332F, and MBD_333F

In step 1 VEH3 and PRICE1 were input into the VMT model described in section 5.4.1 to generate VMT_31F; in the same way, VEH3 and PRICE2 were input into the same model to generate VMT_32F; and in the same way VEH3 and PRICE3 were input into the model to generate VMT_33F.

The corresponding output is shown in Table 7-5. It is seen that VMT under this scenario set will grow to values between 2.98 trillion and 3.05 trillion vehicles miles in year 2020.

Table 7-5: Table 7-5 VMT Scenario Set: VMT_31F, VMT_32F, and VMT_33F

Year	VMT_31F	VMT_32F	VMT_33F
2007	2.781	2.781	2.781
2008	2.782	2.782	2.782
2009	2.863	2.804	2.835
2010	2.884	2.825	2.860
2011	2.905	2.844	2.883
2012	2.924	2.862	2.905
2013	2.942	2.879	2.926
2014	2.959	2.895	2.946
2015	2.976	2.910	2.965
2016	2.992	2.925	2.984
2017	3.007	2.939	2.999
2018	3.022	2.953	3.013
2019	3.036	2.966	3.026
2020	3.050	2.979	3.040

VMT_31F, VMT_32F, and VMT_33F were combined each one with MPG3, in step2, to define MBD_331F, MBD_332F, and MBD_333F. The corresponding MBD output is shown in Table 7-6.

It is seen that under this scenario set, which is illustrated in

Figure 7-7, the light

vehicle fleet fuel consumption will reach between 6.90 and 7.19 MBDs in year 2020. This forecast set contains the most optimistic MBD forecast at 6.90 MBDs in year 2020.

Table 7-6: MBD Forecast Set: MBD_331F, MBD_332F, and MBD_333F

Year	MBD_331F	MBD_332F	MBD_333F
2007	8.853	8.853	8.853
2008	8.812	8.812	8.812
2009	9.082	8.839	8.967
2010	9.080	8.832	8.979
2011	9.043	8.790	8.952
2012	8.976	8.718	8.896
2013	8.875	8.612	8.807
2014	8.741	8.473	8.685
2015	8.570	8.298	8.526
2016	8.368	8.091	8.335
2017	8.130	7.848	8.096
2018	7.856	7.569	7.819
2019	7.546	7.255	7.506
2020	7.193	6.897	7.150

The decrease in consumption because of oil price increase in this forecast set is consistent with similar findings in the other two cases.

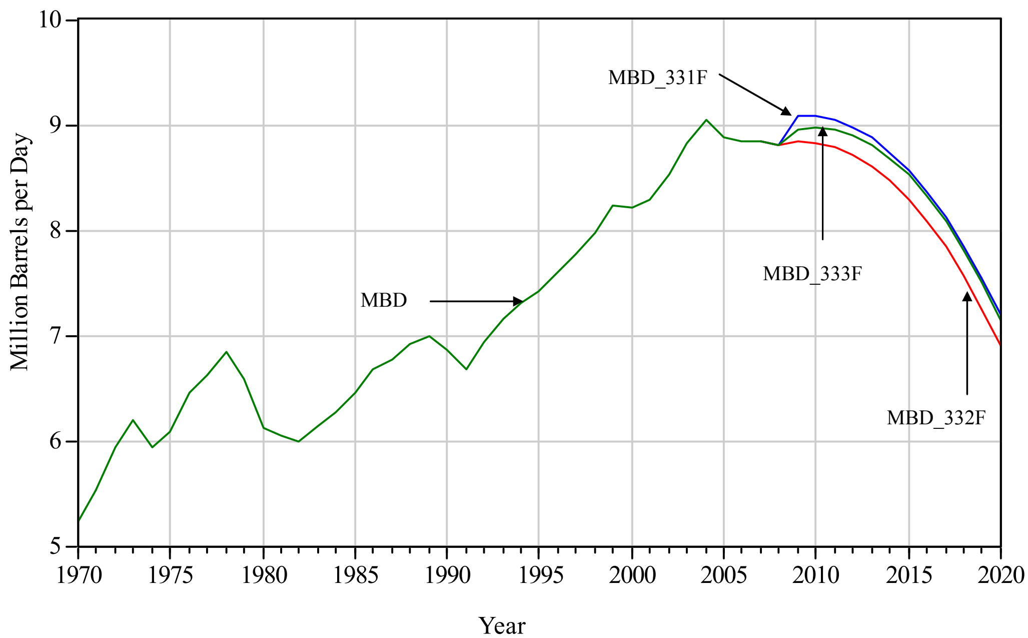


Figure 7-7: MBD Forecast Set: MBD_331F, MBD_332F and MBD_333F

Forecast MBD_333F, which placed in the middle (median) of the three forecasts, is illustrated in Figure 7-8. The actual MBD series as well as its ARIMA extended MBD series trend (the baseline) are also illustrated in the same Figure. MBD_333F gasoline consumption in 2020 (7.15 MBDs) is 1.70 MBDs below the actual or observed consumption in year 2006 (8.85 MBDs) but 3.02 MBDs below what would be consumed in year 2020 (10.17 MBDs).

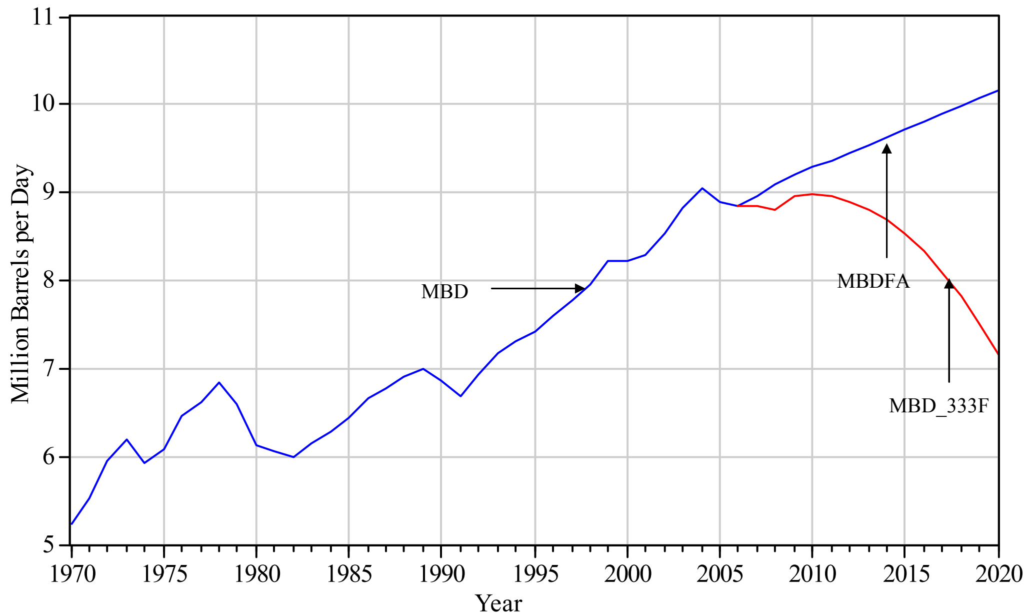


Figure 7-8: MBDFA (Baseline) Comparison with MBD_333F

7.6 Forecasting Summary

Table 7-7 summarizes the forecasting work for the light vehicle fleet gasoline consumption, as follows:

a. The column named MBDFA contains the ARIMA extension of the actual MBD 1970-2006 series. This is the most likely consumption that would have occurred in absence of the Energy Independence Act of 2007. This line was considered the baseline or reference line of gasoline consumption. The consumption in year 2020 would have reached 10.17 MBDs under the trend found by the ARIMA hypothesis.

b. The average (the most commonly used measure of central tendency) can be applied to most sets of data, and forecasting data in this respect are no different. The column named MBD_AVGF contains the average of the nine MBD forecasts considered at 7.78 MBDs in year 2020.

c. The median of the nine forecasts performed, which was named MBD_MEDF corresponds to forecast MBD_223F at 7.55 MBDs in year 2020. The median is in this case a better central tendency predictor than the average because the median, as opposed to the average, is unaffected by extreme values in any set of data.

d. The most optimistic forecast named MBD_OPTF in the table corresponds to forecast MBD_332F at 6.90 MBDs in year 2020.

e. The most pessimistic forecast for MBD named MBD_PESF corresponds to forecast MBD_111F at 8.89 MBDs in year 2020.

Table 7-7: Forecasting Summary

Year	MBDFA	MBD_AVGF	MBD_MEDF	MBD_OPTF	MBD_PESF
2007	8.96	9.00	8.90	8.85	9.24
2008	9.10	8.98	8.88	8.81	9.26
2009	9.21	9.16	9.04	8.84	9.60
2010	9.28	9.19	9.07	8.83	9.68
2011	9.36	9.19	9.06	8.79	9.72
2012	9.45	9.16	9.02	8.72	9.74
2013	9.54	9.10	8.96	8.61	9.73
2014	9.63	9.01	8.86	8.47	9.70
2015	9.72	8.89	8.73	8.30	9.63
2016	9.81	8.73	8.57	8.09	9.54
2017	9.90	8.55	8.37	7.85	9.42
2018	9.99	8.32	8.13	7.57	9.27
2019	10.08	8.07	7.86	7.25	9.09
2020	10.17	7.78	7.55	6.90	8.89

Figure 7-9 shows the actual 1970-2006 MBD series, its ARIMA extension to year 2020 or baseline of gasoline consumption, and the gasoline consumption as per forecast MBD_223F (the median forecast). Savings, because of the Energy Independence Act of 2007, under this MBD scenario will amount to 1.30 MBDs below the consumption of year 2006 or 2.62 MBDs below what would be consumed in year 2020.

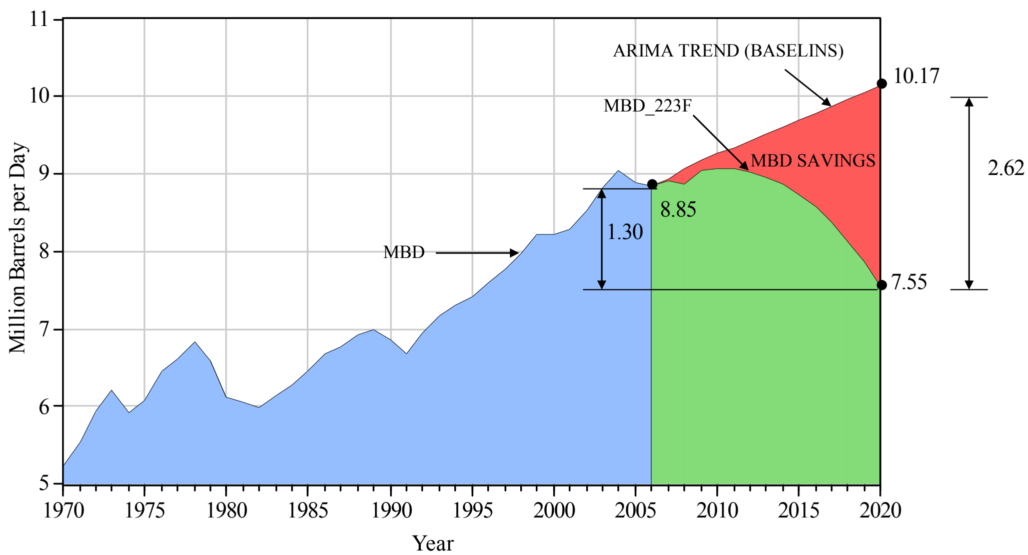


Figure 7-9: MBD Savings under Forecast Scenario MBD_223F

CHAPTER 8. CONCLUSIONS

Although humans by nature can not foresee the future, transportation planners, politicians, and other transportation stakeholders need, using available tools, draw possible scenarios over how current events may affect future outcomes. This is particularly true in the case of the U.S. light fleet fuel consumption, as the need to control, or if possible diminish fuel demand, has been widely acknowledged. In the same context former President Bush affirmed that “the United States has become addicted to oil” (British Broadcasting Corporation 2006). At least three major deleterious effects or negative consequences of the extremely high demand of this commodity deserve special consideration. These effects are: oil dependence, its contribution to the trade imbalance, and its contribution to the GHG pollution effect.

Developing a gasoline demand model, to obtain information on how the fuel demand will shape in the immediate future, leads to the formulation of two types of conclusions: on one side there are conclusions about the strategies used or needed to obtain a working gasoline consumption model, and on the other side there are conclusions related to the results once this model was put into function. It is important also to determine how the lessons learnt may be used for further studies.

8.1 Obtaining a Working Model

No matter how sophisticated time series uni-variate demand models may be, what they do is just the extension of a trend based on the past demand itself. Although models of this type have been used for forecasting purposes (New York State 2002), in the present situation they could not, because the demand trend is expected to be affected by the variations forced upon it by the new light fleet efficiency requirements mandated by law (Bush 2007). To forecast gasoline demand, an explanatory variable model had to be developed in this case. However, uni-variate models were needed to define the trend and establish a reference or baseline upon which possible future savings in gasoline demand can be measured.

The reference or baseline of gasoline consumption was defined by an ARIMA (2,1,0) model after comparing it with other three models: simple linear trend, constant growth, and RWM.

A two steps OLS explanatory variable model was purportedly developed to perform light vehicle fleet fuel consumption forecasting. The transportation demand indicators finally selected were MBD, VMT, VEH, PRICE, and MPG.

The models were checked to satisfy the basic OLS assumptions: normality, heteroskedasticity, and serial correlation (Levine et al. 2001). As time series are for the most usually autocorrelated (Wei 2006), procedures were needed to solve for this problem. To that respect EVIEWs is particularly powerful as it permits the use of autoregressive and moving average error correction terms (Quantitative Microsoftware 2007a). AR(1), AR(2), and MA(1) terms were applied to solve for serial correlation.

Gasoline consumption, the dependent variable, was expressed in oil equivalent million barrels per day or MBD. The regressor price of oil in dollars per barrel or PRICE and the fuel efficiency indicator in miles per gallon or MPG were given priority for inclusion into the OLS model. Other regressors considered were the size of the light vehicle fleet expressed in million vehicles or VEH, the vehicle miles traveled expressed in trillion miles or VMT, the gross domestic product per capita expressed in thousand dollars per person or GDPC, and the length of the infrastructure guide-way expressed in million lane miles or MILES.

Most of the regressors turned out to be highly correlated among themselves, thus contributing redundant information to the expression of MBD. The multicollinearity problem was solved by expressing the MBD regressand in a two steps model in first difference form (Studenmund 2000). MBD was expressed as function of VMT and MPG while VMT was expressed as function of VEH and PRICE.

For model developing the 1970-2000 series were partitioned in two sets. The first set used for calibration or training lapsed the first 30 years (1970-2000) thus contained 31 points, while the second set used for testing or validation lapsed 5 years (2001-2006) thus contained six points. The RMSE statistic performance in the testing or validation period was used as criterion for model selection among several possible models (Wilson et al. 2002), although other criteria were also recorded

8.2 Forecasting MBD

As ultimately MBD was made dependent on VEH, MPG, and PRICE, to forecast MBD scenarios were needed for them.

Three scenarios were hypothesized for VEH. The first scenario called VEH1 identified the constant growth rate trend in the observation period (1970-2006) and applied this growth to the forecasting period (2007-2020). This growth rate was about 2.1 percent.

It was noted that the VEH growth rate had been in general decreasing along time to a minimum of 1.28 percent in the last five year sub-period of observation; then, a second scenario called VEH2 was created by extending this trend to the forecasting period.

The decreasing VEH growth rate trend characteristic was best observed when instead of plotting VEH along time, the ratio VEH/POP called VEHP was plotted instead. Then, a standard decreasing growth rate model (Bowerman et al. 2005) was applied to VEHP. Once the model was identified, the corresponding trend was extended to the forecasting period. The model seemed consistent with the theory of vehicle saturation (Ortuzar and Willumsen 2004). The corresponding VEH forecast called VEH3 was obtained by multiplying VEHP in the forecasting period by the corresponding United States population projections for the same period (U.S. Census Bureau 2008).

The light fleet average annual MPG values in the forecasting period were obtained by applying the MPG new mandated efficiency values to the new portion of the fleet. Those mandated MPG values must increase a 4 percent each year up to 35 MPG in 2020 (The White House Press 2007). A fleet replacement of 7 percent was used to determine the new portion of the fleet (Canes 2007).

A set of average MPG values were obtained for each one of the three VEH scenarios. There were therefore three scenarios identified for MPG which were called MPG1, MPG2, and MPG3.

The oil price range to be applied to the forecasting period was established between a low given by the mean historical of the series in the observation period, and a high obtained by extending the historical trend by exponential smoothing. The low turned out to be about 42 dollars per barrel while the high turned out to be around 100 dollars per barrel. It is assumed that annual averages oil prices below 42 dollars per barrel will have low future probability in a growing economy, while annual average values above 100 dollars per barrel have proved to be difficult to sustain. A third scenario for PRICE in between was adopted by retrieving a long time forecast for the price of oil existing in the Energy Information Agency database (Energy Information Administration 2008b). The corresponding three scenarios were called PRICE1, PRICE2, and PRICE3.

Scenario VEH1 led to the most pessimistic scenario set for MBD, but even on it the MPG increase in fuel efficiency seems to be able to offset the MBD increase that otherwise would have occurred in year 2020 by extension of the trend. Under this MBD scenario set the 2020 gasoline consumption will remain at the 2006 levels of about 9 million barrels per day. However, since the trend would have taken gasoline consumption in year 2020 to about 10 MBD, then there will be net savings of about 1 MBD in year 2020.

Scenario VEH1 looks less likely than VEH2 or VEH3 because it seems somewhat difficult for VEH to grow at a sustained annual rate of 2.1 percent when it is already known that the VEH growth rate has been in general decreasing along time.

The most optimistic MBD scenario set was obtained for VEH3 in the hypothesis of vehicle saturation. Under this hypothesis the savings in fuel consumption in year 2020 will reach a maximum of almost 2 millions barrels per day based on the 2006 gasoline consumption level. However, since the trend would have taken gasoline consumption in year 2020 to about 10 MBDs, the net savings under VEH3 will be in the order of 3

MBDs. This scenario set for MBD having been construed upon the vehicle saturation hypothesis seems rather difficult to evaluate.

The average and median forecasts will lead to gasoline consumption levels of less than 8 MBD in year 2020 or more than 1 MBD below the 2006 level. However, since the trend would have taken gasoline consumption to about 10 MBD in year 2020, expected net savings well beyond 2 MBD are very probable.

Gasoline consumption below the 7 to 8 MBD figures in year 2020 are possible to achieve if technologies like plug-in hybrids already implemented in other countries (Day 2008) and already in trial in the United States take over fast enough, are strongly promoted, or are given incentives in the form of subsidies like tax credits; or if the Federal Government, under President Obama administration, extends a waiver to California and other States to require carmakers a fuel efficiency of 42.5 MPG in year 2020 as proposed by the State of California.

8.3 Recommendations for Further Research

The light vehicle fleet consumed 66 percent of the total oil consumed in transportation in year 2006 placing itself as the most important mode of transport in terms of energy consumption, second was trucks and buses which consumed 19 percent, and third was air transportation that consumed about 9 percent of the total oil demand for transportation activities (Davis and Diegel 2008). It seems also important to develop similar fuel consumption models for these other modes of transportation to recognize and interpret trends of energy consumption as well as energy efficiency. These fuel consumption models may also be of interest to transportation stakeholders.

The variable PRICE intervenes with negative sign in light vehicle fleet MBD models because as oil and gasoline price increase, drivers try to decrease vehicle miles traveled; but the correlation of PRICE with transit ridership is positive because as oil price increases some drivers switch to transit wherever this mode is available. Transit ridership models may be developed to understand how this and other variables may interact in the evolution of this very important transportation mode in the United States.

Further research is needed to explore the possibility of applying other econometric techniques, such as transfer functions and simultaneous equations, to the problem of modeling transportation energy demand.

The economic recession that began to unfold in 2008 and have continued its course in year 2009 has led so far to low oil prices consistent with scenario PRICE1, and low light vehicle growth best represented with the minimum VEH growth scenario VEH3. Therefore, the forecast for MBD that best represents the current situation is MBD_331F.

LIST OF REFERENCES

American Petroleum Institute. (2008). "The Truth About Oil and Gasoline: An API Primer." American Petroleum Institute, Washington, DC. http://www.api.org/aboutoilgas/upload/truth_primer.pdf Accessed: Aug. 15, 2008

American Public Transportation Association. (2008). *2008 Public Transportation Fact Book*, American Public Transportation Association, Washington, DC.

Banaszak, S., Ckravorty, U., and Leung, P. (1999). "Demand for Ground Transportation Fuel and Pricing Policy in Asian Tigers: A Comparative Study of Korea and Taiwan." *Energy Journal*, 20(2), pp 45-65.

Berkowitz, M., Gallini, N., Miller, E., and Wolfe, R. (1990). "Disagreggregate Analysis of the Demand for Gasoline." *Canadian Journal of Economics*, 23(2), pp 253-275.

Bolte, S. (2008). "Another Successful Year for Amtrak." Progressive Railroading, Trade Press Publishing Corporation, Milwaukee, WI., No.11 pg 13.

Bowerman, B. L., O'Connell, R., and Koehler, A. (2005). *Forecasting Time Series and Regression*, Thomson Brooks and Cole, Belmont, CA.

British Broadcasting Corporation. (2006). "February 1, 2006 Address to the Nation: President Bush Urges End to Oil Addiction." Electronic Article, British Broadcasting Corporation. http://news.bbc.co.uk/2/hi/americas/4665758.stm Accessed: June 16, 2008

British Petroleum Corporation. (2008). "Statistical Review of World Energy 2008." British Petroleum Corporation, London, United Kingdom. http://www.bp.com/productlanding.do?categoryId=6929&contentId=7044622 Accessed: November 10, 2008

Bureau of Transportation Statistics. (2007). "National Transportation Statistics." Research and Innovative Technologies Administration (RITA), Washington, DC. http://www.bts.gov/publications/national_transportation_statistics/ Accessed: May 30, 2008

Bush, G. W. (2007). "Public Law 110-140: Energy Independence and Security Act of 2007." United States Congress, Washington, DC. http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_public_laws&docid=f:publ140.110 Accessed: July 20, 2008

Canes, M. (2007). "Where is the United States Motor Fuels Market Headed?" Research Report, Energy Policy Research Incorporated, Washington, DC, pp 1-16. http://www.eprinc.org/pdf/TenYearOutlookFuelsJuly.pdf Accessed: September 18, 2008

Chatfield, C. (2004). *The Analysis of Time Series, An Introduction*, Chapman & Hall CRC Company, Boca Raton, FL.

Davis, S., and Diegel, S. W. (2008). *Transportation Energy Data Book: Edition 27*, Oak Ridge National Laboratory (ORNL), Oak Ridge, TN.

Day, P. (2008). "Revving China's Auto Industry." Electronic Article, BBC News, Dec. 13, 2008, British Broadcasting Corporation. http://news.bbc.co.uk/2/hi/americas/4665758.stm Accessed: December 15, 2008

Dietert, J. (2008). "Ethanol Supply and Demand Outlook." Ethanol Supply and Demand Outlook, Research Report, Simmons & Company International, Aberdeen, Scotland, United Kingdom, pp 1-14. http://www.eprinc.org/pdf/dietert.pdf Accessed: April 1, 2008

Emerson, S. A. (2006). "Does the United States Need a Gasoline Strategy." Technical Report, Energy Security Analysis Inc., Wakefield, MA. http://www.esai.com/pdf/Emerson-Repsol-KSG-gasolinefor%20ESAI.pdf Accessed: March 09, 2008

Emerson, S. A. (2007). "The Future of the Global Oil Market." Technical Report, Energy Security Analysis, Inc.; Boston, Massachusetts, pp 1-15. http://www.esai.com/pdf/FutureGlobalOilMarket.pdf Accessed: August 23, 2008

Energy Information Administration. (1995). "Energy Efficiency Report. Measuring Energy Efficiency in the United States. Chapter 5: Transportation Demand Indicators." Office of Energy Markets and End Use, Energy Information Administration, Washington, DC. http://www.eia.doe.gov/emeu/efficiency/ee_ch5.htm Accessed: May 15, 2008

Energy Information Administration. (2007). *Annual Energy Outlook 2007 with Projections to 2030*, Energy Information Administration, Washington, DC.

Energy Information Administration. (2008a). "Energy Information Administration: Monthly Report on Crude Oil Prices." Energy Information Administration, Washington, DC. http://tonto.eia.doe.gov/dnav/pet/hist/rbrteM.htm Accessed: January 12, 2009

Energy Information Administration. (2008b). "Official Statistics from the U.S. Government: Petroleum Product Price Long Term Forecast." International Energy Agency; Washinton, DC. http://www.eia.doe.gov/oiaf/aeo/excel/aeotab_12.xls Accessed: January 12, 2009

Energy Policy Research Foundation. (2008). "Why Do Oil Prices Jump So High When Supply Glitches Occur?" Research Report, Energy Policy Research Foundation, Inc., Washington, DC, pp 1-5. http://www.pirinc.org/download/whydooilprices.pdf Accessed: March 3, 2008

Fay, J., and Golomb, D. (2002). *Energy and the Environment*, Oxford University Press, New York, NY.

Feng, A., and Sauer, A. (2004). "Comparison of Passenger Vehicle Fuel Economy and GHG Emission Standards Around the World." Research Report, The Pew Center on Global Climate Change, Arlington, VA, pp 1-33. http://www.pewclimate.org/global-warming-in-depth/all_reports/fuel_economy Accessed: December 15, 2008

Gujarati, D. N. (2003). *Basic Econometrics*, McGraw Hill, New York, NY.

Hughes, J. E., Knittel, C. R., and Sperling, D. (2008). "Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand." *Energy Journal*, 20, pp 113-134.

Jaffe, A. M. "Oil Speculative Bubble and Long Term Supply Issues." *Testimony to the Selected Committee on Energy Independence and Global Warming of the United*

States, Presented at the U.S. House of Representatives, Conference Proceedings, Washington, DC., June 11, 2008, pp 1-4.
Jeffrey, O. (2006). "Planning and Economic Development in New York: Parking Puzzles." The Urban Prospect, Citizens Housing and Planning Councill, New York, NY pp 1-4. http://www.chpcny.org/pubs/UP_Parking.pdf Accessed: July 27, 2008
Levine, D. M., Ramsey, P. P., and Smidt, R. K. (2001). *Applied Statistics for Engineers and Scientists*, Prentice Hall Inc., Upper Saddle River, NJ.
Lucian, P., Goldstein, L., and Cumins, L. (2007). "Where is the United States Oil Market Headed?" Energy Policy Research Foundation Publications and Podcasts, Research Report, Energy Policy Research Foundation Inc., Washington, DC. http://www.pirinc.org/contact.html Accessed: August 18, 2008
Makridakis, S., Wheelright, S. C., and Hyndman, R. J. (1998). *Forecasting Methods and Applications*, John Wiley and Sons, New York NY.
McManus, W. (2006). "Can Proactive Fuel Economy Strategies Help Automakers Mitigate Fuel Price Risks." Research Report, University of Michigan Transportation Research Institute, Ann Arbor, MI. http://www.umtri.umich.edu/content.php?id=1111 Accessed: September 15, 2008
Medlock, K. B., and Jaffe, A. M. (2007). "United States Gasoline Markets and Oil Import Dependence." Technical Report, James A. Baker III Institute for Public Policy of Rice University, Houston , Texas, pp 1-16. http://www.bakerinstitute.org/publications/Gas_FAQ_Transcript.pdf Accessed: November 12, 2007
New York State. (2002). "New York State Energy Plan." Technical Report, New York State Energy Research and Development Authority, Albany, NY, pp 1-20. http://www.nyserda.org/Energy_Information/energy_state_plan Accessed: August 27, 2008
Ortuzar, J., and Willumsen, G. (2004). *Modeling Transport*, John Wiley and Sons, Hoboken, NJ.
Page, J., Weng-Gutierrez, M., and Schremp, G. (2007). "Overview of Proposed Transportation Energy Analysis for the 2007 Integrated Energy Policy Report." California Energy Commission, Sacramento, CA. http://www.energy.ca.gov/2007publications/CEC-600-2007-005/CEC-600-2007-005.PDF Accessed: August 20, 2008
Pock, M. "Estimation of Gasoline Demand Function." *University of Vienna 2005 Seminar in Panel Data*, Conference Proceedings, Vienna, Austria, 2005, pp 1-10.
Portney, P. (2002). "Effectiveness and Impact of Corporate (CAFE) Standards." *Report ISBN-13:978-0-309-09037-7, Transportation Research Board of the National Academies Press*, Washington, DC.
Pyndick, R., and Rubinfeld, D. (1991). *Econometric Models and Econometric Forecasts*, McGraw Hill Book Company, New York, NY.
Quantitative Microsoftware. (2007a). *EVIEWS User's Guide I*, Quantitative Microsoftware LLC, Irvine CA.
Quantitative Microsoftware. (2007b). *EVIEWS User's Guide II*, Quantitative Microsoftware LLC, Irvine, CA.

Roubini, N., and Setser, B. (2004). "The Effects of the Recent Oil Price Schock on the United States and Global Economy." Research Report, Stern School of Business, New York University, New York, NY. http://pages.stern.nyu.edu/~nroubini/papers/OilShockRoubiniSetser.pdf Accessed: September 10, 2008

Startz, R. (2007). *EVIEWS Illustrated for Version VI*, Quantitative Microsoftware LLC, Irvine, CA.

Studenmund, A. H. (2000). *Using Econometrics - A Practical Guide*, Addison Wesley Longman, Boston, MA.

The White House Press. (2007). "President Bush Signs H.R.6, the Energy Independence and Security Act of 2007." Electronic Article, The White House Press, Washington, D.C., pp 1-3. http://www.whitehouse.gov/news/releases/2007/12/20071219-6.html Accessed: December 19,2007

U.S. Census Bureau. (2008). "United States Census 2000 and 2008 National Projections." United States Census Bureau, Philadelphia, PA. http://www.census.gov/ Accessed: September 29, 2008

Vandaele, W. (1983). *Applied Time Series and Box-Jenkins Models*, Academic Press Inc., San Diego, CA.

Washington, S. P., Karlaftis, M. G., Mannering, F. L., and Cullinane, K. (2004). "Statistical and Econometric Methods for Transportation Data Analysis." *Maritime Economics & Logistics*, 6, pp 187-189.

Wei, W. (2006). *Time Series Analysis Univariate and Multivariate Methods*, Pearson Adison Wesley Inc., Boston, MA.

Wells, A., and Weens, J. G. (2004). *Air Transportation: A Management Perspective*, Books Code Thompson Learning Inc., Belmont, CA.

Wiggins, N. (2003). "United States Gasoline Consumption: An Econometric Model." eReview of Economics, Working Paper, University of Missouri St. Louis, Missouri. http://www.umsl.edu/~nabe/wiggins.pdf Accessed: August 12, 2008

Williamson, S. (2008). "Measuring Worth." Lawrence Officer and Samuel Williamson Gross National Product Statistics. http://www.measuringworth.com/aboutus.html Accessed: January 18, 2009

Wilson, J. H., Keating, B., and Galt, J. (2002). *Business Forecasting With Accompanying Excel-Based Forecast XTM Software*, McGraw Hill Higher Education, New York, NY.