High School Mathematics Competition in Modelling

Team 3455

Problem B - How Much Gas Should I Buy This Week?

Problem

The problem here is (just as the problem states) to determine how much gas to buy on any given week. Essentially, we develop a model to predict gasoline prices based on fundamental variables and a recursive formula.

Our Thoughts

We are given two cases, one where the consumer drives 100 miles per week and one where the consumer drives 200 miles per week. If we were to have the option of adding as much gas we would like, then there would be a large number of choices and complicated the model to reduce accuracy. Instead, we deemed accuracy to be of more importance and limited the choices of the consumer to binary form every week.

We divided our solution into 5 parts:

Part I) Establishing Preliminaries

Our Assumptions, Definitions, Variables, and General Analysis; these are constantly referred in the model

Part II) The General Pricing Model

Here we determine quantitative factors as well as the laws of supply and demand to create a model that could predict future gas prices. We create theoretical models and then find the effects of the actual data to create an accurate pricing model.

Component 1: Supply

- •Here we establish and proof a link between crude oil supply and gasoline prices. We also show that there is a response time such that we can use supply to anticipate gasoline prices

- •We create a theoretical model (created entirely based on observations we made and reasoning)

- •The model here uses current gasoline price, supply of crude oil, and a constant "C" (a natural result of one of the integrals), which are the combined effect of qualitative data (This constant is modeled in component 3)

Component 2: Individual Profiles

- This component is simply a basic analysis and profile of the top five oil exporting countries to the United States.

- Important factors are discussed and analyzed

Component 3: Qualitative Data

- We determined that the global weather and political stability to be the most important qualitative data that affects gasoline prices

- We categorize the different intensities of the two data and each category is given a quantitative value that affects the constant “C” and subsequently the price of gasoline in the future. Essentially, we are able to quantify qualitative data.

We now have a complete model that can predict future gasoline prices. The only unknown variable is β ,a constant value that is applied to calibrate the magnitude of the changes in supply to the changes in gas prices. We use the retail gasoline prices in 2011 to solve for β and the average value is used for the General Pricing Model.

In building the model, we accept that there exists uncertainties and near random movements. We express this concept and its effects by introducing a random variable ε and allow to act like a simplified Wiener Process on the model. This accounts for random movements that we cannot see.

Our next step is to test the model. We used the model starting in January of 2012 until the most recently available data. This is then compared to the actual prices.

Part III) The Binary Tree Model

In this part, we plan out and model the actual decision making possibilities of the consumer. Initially, we established that the consumer has a binary choice each time he has the chance to buy gas. We analyze each case given separately.

By looking at the outcomes of the binary trees, we create a table that the consumer can refer to. The table contains a column that calculates the average amount paid for every mile driven for very outcome. The consumer only needs to use the General Pricing Model from Part II to find the price of each week in the future and refer to the table in this part to find the outcome where he would be paying least per unit mile driven.

Part IV) The Boston Model

In this part, we take our General Pricing Model (which prices for the weekly United States Regular All Formulations Retail Gasoline Prices) and apply it to the city of Boston by adding in location-specific variables to recalculate the value “C” in the model.

This part is very similar to the testing of the General Pricing Model at the end of Part II.

Part V) The Letter

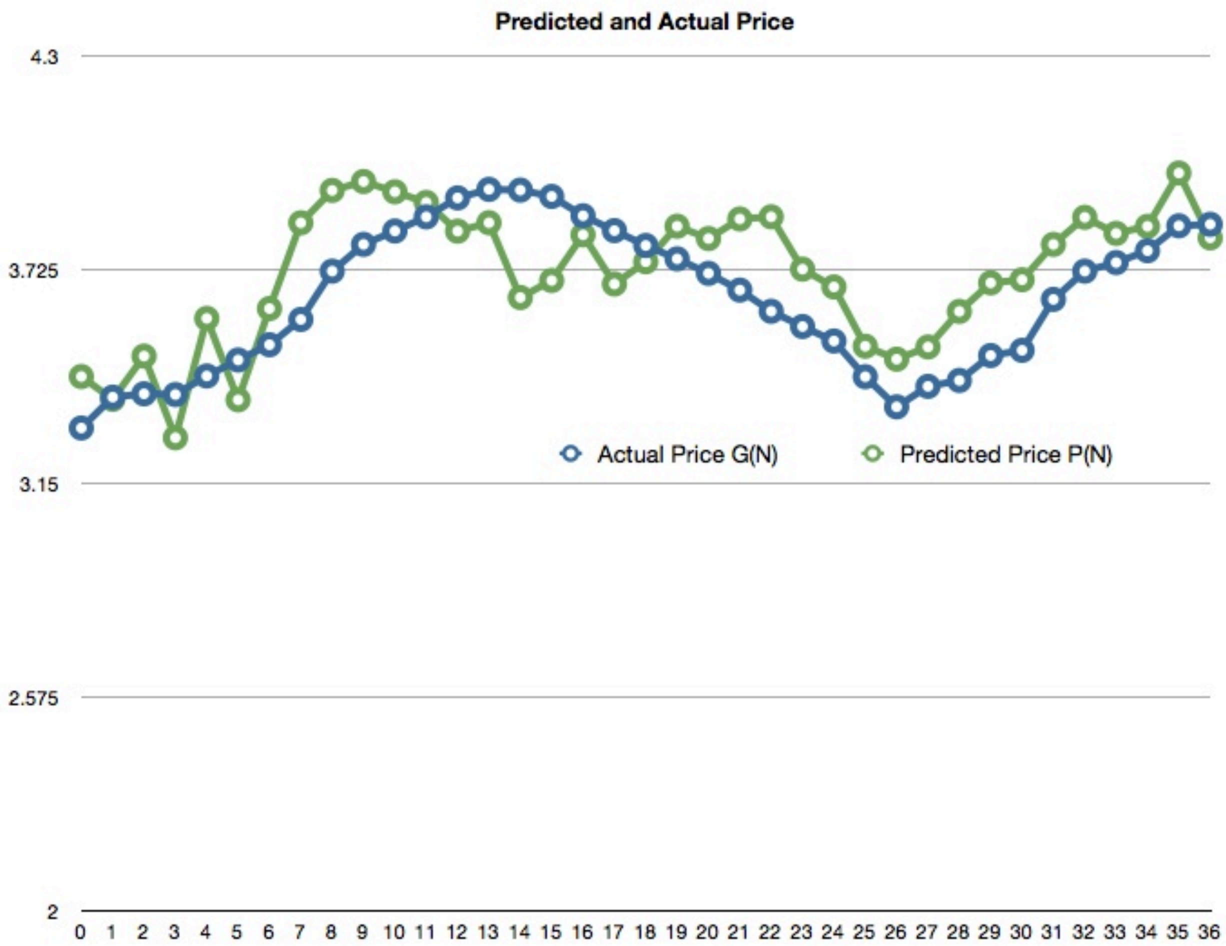
As requested, we wrote a letter to our local newspaper (The South China Morning Post) and presented them with an article detailing the basics of our model and how the average person can utilize it to save money.

This is written in a non-technical language and anyone is able to read it.

Results

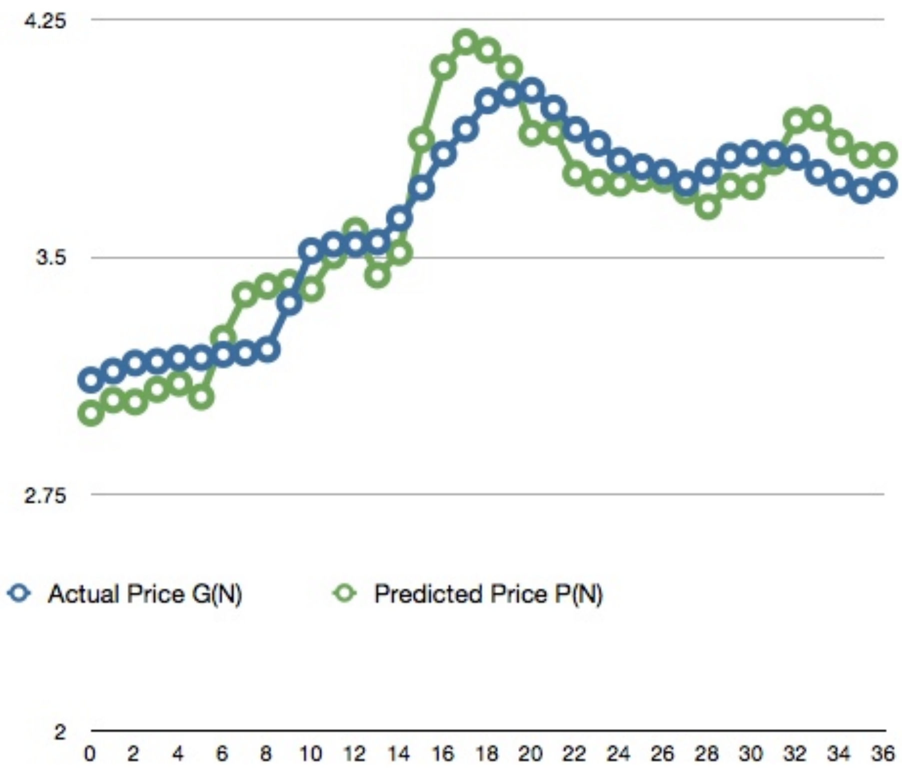
We obtained relatively positive results from our model. Here is a graph of our predicted price of 2012 and the actual price (from the testing of the General Pricing Model at the end of Part II):

Here is a a graph demonstrating our predicted price and the actual price:



Our percent error was 3.62% which is a very good result.

Here is a graph of the results from the Boston Model



Here, the correlation is slightly less but still maintains it's positive results.

HIGH SCHOOL MATHEMATICS COMPETITION IN MODELING

Team 3455

Problem B – How Much Gas Should I Buy This WEek

Prepared by: Team 3455

Date: November 2012

Problem Summary and Analysis:

The purpose of this paper is to develop a model that will attempt to predict and analyze retail gasoline prices so that a consumer could use it to determine how much gas to purchase on a given week.

A majority of the citizens in the United States depend upon gasoline to power their vehicles; however, the cost of gasoline is heavily volatile. Car owners need to be concerned with the constantly changing price of fuel. It would be of great convenience to a car owner if he were able to predict the price of gasoline, so as to be able to have an idea of how much gas to purchase on a given week. The objective of this model is to allow the car owner to input publicly available data and find a reasonable prediction for the following week's gasoline price. Knowing this, the car owners can make a more informed decision on how much gasoline to buy that week.

One part of this paper will create a generalized model for the United States while a latter part will apply the model to a specific city: Boston.

Summary of Solution Method:

To tackle this problem, we must first set definitions, assumptions, and variables that will be used. General economics theory dictates that changes in the price of gasoline are usually the result of changes in supply and demand of gasoline. We break down the factors determining both supply and demand to measurable segments. We analyze this through overall factors as well specific factors for the top five countries that export crude oil to the United States. This is done to simply the model to account for the time restraint but to minimize the influence of non-accounted data. This is the general pricing model. We then create a binary tree model based on the given assumptions and the results of the pricing model. This is the decision model. Next, we use the general pricing model and apply it to the city of Boston by adding location specific factors that further affect the price of gasoline. We call this the Boston Pricing Model. Lastly, we write a summary in the form of a one page non-technical letter to the local paper describing how the average person might use our model.

Part I) Assumptions, Definitions, Variables, and General Analysis

Part II) The General Pricing Model

- This model is meant to be a general indicator of the price of gasoline in the United States

Part III) The Decision Model

- This model takes the weekly price of gasoline for the next month based on the General Pricing Model and also the given assumptions on mileage and decision making to decide the best outcome for the driver

Part IV) The Boston Pricing Model

- This part takes the General Pricing Model from Part II and adds location-specific factors

Part V) The letter to the local paper

- This part is a non-technical summary that allows the average reader to understand our model and how to use it to their benefits

PART I

Assumptions:

1) A gas tank holds 16 gallons and the average mileage is 25 miles/gallon. This means that every full tank (Given)

2) The consumer has a choice to buy gas once a week (Given) (further addition to this assumption on assumption 4)

3) At the beginning of every week, the consumer will have the choice of filling the tank or buying a half tank. This is reasonable since half and full are naturally measurable choices and buying by other denominations would be unnatural.

4) The tank would always be half full or full after filling. It cannot be zero since the driver would have to drive the next week and thus must fill the tank. This also means that the driver can only fill the tank if he is empty, or half filled. If this was not true, then it is possible to have 3/4 or 1/4 filled tanks, thus complicating the problem beyond our capabilities in the given time. This assumption allows for a simplified binary tree model of the final decision on whether to fill full or half very time the opportunity arises.

5) The price elasticity of demand for gasoline is inelastic (meaning that there is a relatively low percentage change in the quantity demanded compared to percent change in price). This is reasonable to assume since in the short run (which is the main usage of our model), people are locked into their existing life and are willing to pay the extra amount for the necessity of gasoline. In the long run, people may switch to fuel-efficient vehicles or move to closer to their intended destination. However, that is beyond the range and objective of the model.

6) When deciding whether to fill up the tank, the consumers want to minimize cost per miles driven. This is reasonable as a basic economic assumption is that individuals will take opportunities to make themselves better off. Who wouldn't want to pay less for the same amount of miles driven? In other words, the objective of the consumer is to drive the most miles and pay-

ing the least possible amount. This also defines the objective of our model to allow the consumers to do this.

Definitions:

1) Crude Oil: the supply in crude oil is defined as the United States imports of Crude Oil and Petroleum Products
2) Retail Gasoline Prices are defined as Weekly United States Regular All Formulations Retail Gasoline Prices
3) Response time: the amount of time between when an independent variable changes and when the dependent variable changes due to the relationship between the two. In component 1 of Part II, we talk about the response time between gasoline prices and the supply of crude oil

Variables:

1) Crude Oil inputs: Crude oil is one of the fundamental components of gasoline. In Part II, component 1, we establish a powerful connection between the supply of crude oil and the price of gasoline.

2) Weather: The weather and temperature of the suppliers of crude oil as well as those who demand of gasoline will have an effect on the price of gasoline. For example, cold spells in Europe often lead to an increase demand for heating and thus an increase in gas prices. In Part II, we find out that the effect that the weather has on demand is a lot greater than the effect it has on supply.
3) Political Stability: The stability of an oil producing nation often drives the price of gasoline up and down. The tension and periodical violence in the Middle East (a heavy oil producing region) is measured and classified quantitatively in Part II component 3.

4) N : throughout this paper, N is generally used to denote the discrete time values whether it is month or week (this will be clarified). N=0 denotes the present time and N=1 is one time interval after the present, etc

5) G(N) is used throughout to denote the retail gasoline price at a time N. This is always in dollars per gallon

6) C(N) is used to denote the price of crude oil, the units are specified in the paper

7) G'(N) is used to denote the change of gasoline prices (sometimes represented only as G when used in a theoretical calculation)

8) C'(N) is used to denote the change of crude oil prices

9) M is used in PART II, Component 1 to indicate the time value of months

10) S(M) is used to denote the supply of crude oil (in thousand barrels) at a given month (sometimes represented only as S when used in a theoretical calculation)

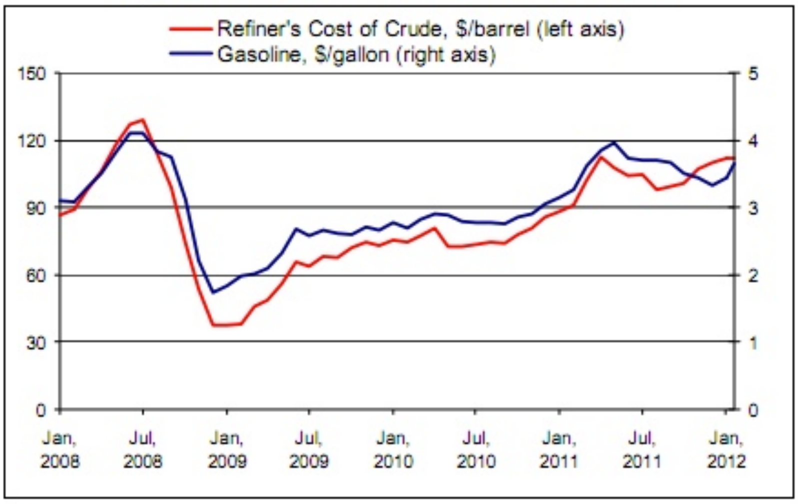
11) β is used in the General Pricing Model as a calibrating constant relating the change in S and the change in G.

12) k is used to denote response time

General Analysis:

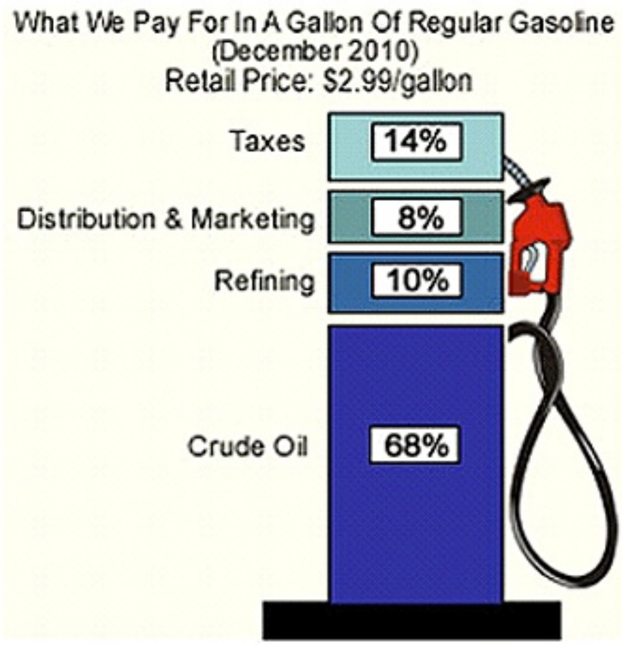
1) There is a general correlation between crude oil prices and retail gasoline prices as evidently shown in the graph below:

Figure 1. Crude Oil Costs and Gasoline Prices



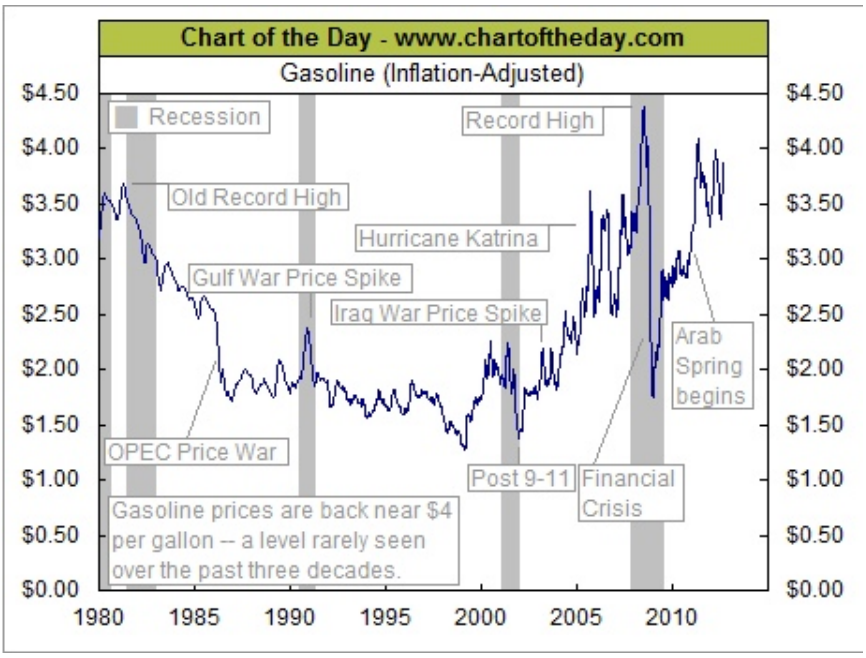
Source: EIA

2) According to the Energy Information Administration, for every dollar we pay for gasoline, roughly $.68 goes to paying for the crude oil, $.14 to tax, $.10 to the cost of refining, and 8% to the cost of distribution and marketing.



Source: Energy Information Administration

3) We believe that gasoline prices can be modelled and subsequently predicted by the combination of global factors.



As the graph demonstrates, factors and events such as the weather (Hurricane Katrina) and stability (Arab-Spring) and financial recessions and crisis have dramatic impact on gasoline prices.

PART II

The General Pricing Model

The General Pricing Model Breakdown:

The General Pricing Model is comprised of three components that uses various factors that affect the supply and demand of crude oil.

- Component 1 Supply: Analysis of the impact of crude oil imports and gasoline prices
  - There is an inverse relationship between crude oil imports and retail gasoline prices.
  - There is a time gap between a change in crude oil imports and the impact it has on the retail gasoline prices, thus creating an indicator for the gasoline prices
  - Because of the nature of oligopolies (which retail gas companies belong to), gasoline have sticky prices. This means there is a momentum effect (will be explained in component 1)

- Component 2 Individual Analysis: Analysis of the top 5 countries that import crude oil to the United States
  - Canada, Saudi Arabia, Mexico, Venezuela, and Colombia are analyzed in detail
    - The precise origin of the oil from the country is determined and the rainfall and temperature of that location is analyzed for correlation
  - Domestic crude oil production and refineries are analyzed

- Component 3 Qualitative: Qualitative factors are considered and quantified
  - Weather and Stability are two major qualitative factors that are analyzed and classified.
  - Temperature and political stability of the demand and supply side are considered

Component 1: Supply

Gasoline and Crude Oil

Based on the general analysis 1, there is a correlation between the price of crude oil and the retail gasoline prices. One may feel that since crude oil is an input of gasoline, that looking at current crude oil prices, we can determine future gasoline prices. However, this is not true. Looking at the data for crude oil prices and retail gasoline prices, we can see that (weekly) the movement of crude oil and the movement of retail gasoline occur at the same discrete time value:

•N = Date (In unit weeks) starting in the first week of January 2011

•G(N) = the U.S. Regular All Formulations Retail Gasoline Prices on week N (in dollars/gallon)

•C(N) = the Spot Price of WTI-Cushing, Oklahoma Crude Oil on week N

•G'(N) = the change in G(N) and discretely defined as G(N) - CGN-1)

•C'(N) = the change in C(N) and discretely defined as C(N) - C(N-1)

DATE (N)	G(N)	C(N) DOLLARS/ BARREL	C(N) DOLLARS/ GALLON	G'(N)	C'(N)
7 Jan 2011	3.089	89.54	2.8425396825397	0	0
14 Jan 2011	3.104	91.02	2.8895238095238	0.015	0.0469841269841
21 Jan 2011	3.11	89.75	2.8492063492064	0.006	-0.0403174603175
28 Jan 2011	3.101	86.11	2.7336507936508	-0.009	-0.1155555555556
4 Feb 2011	3.132	89.52	2.8419047619048	0.031	0.108253968254
11 Feb 2011	3.14	85.51	2.7146031746032	0.008	-0.1273015873016
18 Feb 2011	3.189	84.13	2.6707936507937	0.049	-0.0438095238095
25 Feb 2011	3.383	95.26	3.024126984127	0.194	0.3533333333333
4 Mar 2011	3.52	101.05	3.2079365079365	0.137	0.1838095238095
11 Mar 2011	3.567	103.74	3.2933333333333	0.047	0.0853968253968

DATE (N)	G(N)	C(N) DOLLARS/ BARREL	C(N) DOLLARS/ GALLON	G'(N)	C'(N)
18 Mar 2011	3.562	99.79	3.1679365079365	-0.005	-0.1253968253968
25 Mar 2011	3.596	104.41	3.3146031746032	0.034	0.1466666666667
1 Apr 2011	3.684	105.08	3.335873015873	0.088	0.0212698412698
8 Apr 2011	3.791	109.29	3.4695238095238	0.107	0.1336507936508
15 Apr 2011	3.844	107.75	3.4206349206349	0.053	-0.0488888888889
22 Apr 2011	3.879	109.11	3.4638095238095	0.035	0.0431746031746
29 Apr 2011	3.963	112.3	3.5650793650794	0.084	0.1012698412698
6 May 2011	3.965	105.84	3.36	0.002	-0.2050793650794
13 May 2011	3.96	99.87	3.1704761904762	-0.005	-0.1895238095238
20 May 2011	3.849	97.99	3.1107936507937	-0.111	-0.0596825396825
27 May 2011	3.794	99.55	3.1603174603175	-0.055	0.0495238095238
3 Jun 2011	3.781	100.92	3.2038095238095	-0.013	0.0434920634921
10 Jun 2011	3.713	100.05	3.1761904761905	-0.068	-0.027619047619
17 Jun 2011	3.652	95.87	3.0434920634921	-0.061	-0.1326984126984
24 Jun 2011	3.574	92.7	2.9428571428571	-0.078	-0.1006349206349
1 Jul 2011	3.579	93.7	2.9746031746032	0.005	0.031746031746
8 Jul 2011	3.641	97.12	3.0831746031746	0.062	0.1085714285714
15 Jul 2011	3.682	96.72	3.0704761904762	0.041	-0.0126984126984
22 Jul 2011	3.699	98.01	3.1114285714286	0.017	0.0409523809524
29 Jul 2011	3.711	97.83	3.1057142857143	0.012	-0.0057142857143
5 Aug 2011	3.674	90.85	2.884126984127	-0.037	-0.2215873015873
12 Aug 2011	3.604	82.86	2.6304761904762	-0.07	-0.2536507936508
19 Aug 2011	3.581	85.36	2.7098412698413	-0.023	0.0793650793651
26 Aug 2011	3.627	85.06	2.7003174603175	0.046	-0.0095238095238
2 Sep 2011	3.674	88.07	2.795873015873	0.047	0.0955555555556
9 Sep 2011	3.661	87.91	2.7907936507937	-0.013	-0.0050793650794
16 Sep 2011	3.601	88.93	2.8231746031746	-0.06	0.032380952381
23 Sep 2011	3.509	83.65	2.6555555555556	-0.092	-0.167619047619
30 Sep 2011	3.433	81.18	2.5771428571429	-0.076	-0.0784126984127

DATE (N)	G(N)	C(N) DOLLARS/ BARREL	C(N) DOLLARS/ GALLON	G'(N)	C'(N)
7 Oct 2011	3.417	79.43	2.5215873015873	-0.016	-0.0555555555556
14 Oct 2011	3.476	85.35	2.7095238095238	0.059	0.1879365079365
21 Oct 2011	3.462	86.82	2.7561904761905	-0.014	0.0466666666667
28 Oct 2011	3.452	92.32	2.9307936507937	-0.01	0.1746031746032
4 Nov 2011	3.424	93.24	2.96	-0.028	0.0292063492063
11 Nov 2011	3.436	96.97	3.0784126984127	0.012	0.1184126984127
18 Nov 2011	3.368	99.32	3.1530158730159	-0.068	0.0746031746032
25 Nov 2011	3.307	96.89	3.075873015873	-0.061	-0.0771428571429
2 Dec 2011	3.29	99.91	3.171746031746	-0.017	0.095873015873
9 Dec 2011	3.286	100.08	3.1771428571429	-0.004	0.0053968253968
16 Dec 2011	3.229	96.06	3.0495238095238	-0.057	-0.127619047619
23 Dec 2011	3.258	97.74	3.1028571428571	0.029	0.0533333333333
30 Dec 2011	3.299	99.81	3.1685714285714	0.041	0.0657142857143

The table above shows the retail price of gasoline and the price of crude oil. The last two columns shows the change in gasoline price and the change in crude oil. If we are able to use crude oil as an early indicator of retail gasoline prices, then we would be able to see a movement

$C'(N) \approx G'(N+k)$ for some value k over a consistent interval or discrete value N.

In a simpler explanation, it means that change in the price of crude oil in one direction will cause the price of gasoline to change in a similar direction and similar magnitude at a later time (hence the k value). However, based on this data, k = 0, thus there is no time difference between a change in crude oil and a similar change in gasoline price.

This is consistent with a form of the efficient market hypothesis, which states that the price of the commodity in question (gasoline) contains all public information concern the pricing of it (crude oil price). Therefore, although we can confirm the direct correlation between crude oil and gasoline prices, we cannot use it to predict it.

Supply of Crude Oil and Gasoline

Despite being unable to use present crude oil prices to predict future gasoline prices, we can use the laws of supply and demand of crude oil to make judgements on the price of crude oil and thus the price of gasoline (because of the connection in general analysis 1).

Given the commodity of crude oil, we can set a supply and demand to determine its price. Based on assumption 5 (the price inelasticity of demand) and general analysis 1 (relationship between crude oil and gasoline) we can conclude that the demand of crude oil will remain relatively the same in the short run while the main changes will occur in the supply. The supply of crude oil is the amount of crude oil that is imported by the United States.

The Relationship between The Supply of US Imports of Crude Oil and Petroleum Versus the Retail Prices of Gasoline:

M = Date (in unit Months) starting in the first week of January 2008

S(M) = total U.S. imports of Crude Oil and Petroleum Products in month M (in thousand barrels)

G(M) = the U.S. Regular All Formulations Retail Gasoline Prices in month M (in dollars/gallon)

M	S(M)	G(M)
1 Jan 2008	420616	3.043
1 Feb 2008	367143	3.028

M	S(M)	G(M)
1 Mar 2008	390528	3.244
1 Apr 2008	399944	3.458
1 May 2008	399957	3.766
1 Jun 2008	401942	4.054
1 Jul 2008	406854	4.062
1 Aug 2008	406644	3.779
1 Sep 2008	346859	3.703
1 Oct 2008	409269	3.051
1 Nov 2008	386421	2.147
1 Dec 2008	390817	1.687
1 Jan 2009	406925	1.788
1 Feb 2009	338661	1.923
1 Mar 2009	385841	1.959
1 Apr 2009	358846	2.049
1 May 2009	355774	2.266
1 Jun 2009	358083	2.631
1 Jul 2009	366727	2.527
1 Aug 2009	346658	2.616
1 Sep 2009	352675	2.554
1 Oct 2009	337212	2.551
1 Nov 2009	333152	2.651
1 Dec 2009	326556	2.607
1 Jan 2010	350315	2.715
1 Feb 2010	314452	2.644
1 Mar 2010	360262	2.772
1 Apr 2010	375769	2.848
1 May 2010	376383	2.836
1 Jun 2010	373317	2.732

M	S(M)	G(M)
1 Jul 2010	392915	2.729
1 Aug 2010	383045	2.73
1 Sep 2010	354677	2.705
1 Oct 2010	345415	2.801
1 Nov 2010	332884	2.859
1 Dec 2010	345099	2.993
1 Jan 2011	379676	3.095
1 Feb 2011	300672	3.211
1 Mar 2011	367341	3.561
1 Apr 2011	354229	3.8
1 May 2011	367839	3.906
1 Jun 2011	356304	3.68
1 Jul 2011	364469	3.65
1 Aug 2011	348033	3.639
1 Sep 2011	338087	3.611
1 Oct 2011	342638	3.448
1 Nov 2011	336523	3.384
1 Dec 2011	342995	3.266

This table shows both the total import (supply) of crude oil from other countries and the gasoline prices. Since we have implied that demand remains relatively constant in the short run, an increase in supply of crude oil results in a decrease in crude oil prices and a decrease in supply of crude oil results in an increase in crude oil prices.

This can be represented by $S(M) \propto 1/C(M)$

General analysis 1 and earlier discussion allows us to conclude that $C(M) \propto G(M)$

Therefore, we now conclude that $S(M) \propto 1/G(M)$.

One important aspect to note is that unlike the relationship between crude oil and gasoline (which is instantaneous), a change in the supply of crude oil does indeed happen at an earlier discrete time than the responding change in the price of gasoline. Therefore, we can actually use the supply of crude oil as an early indicator of the retail price of gasoline.

We now demonstrate this: **Observation 1**

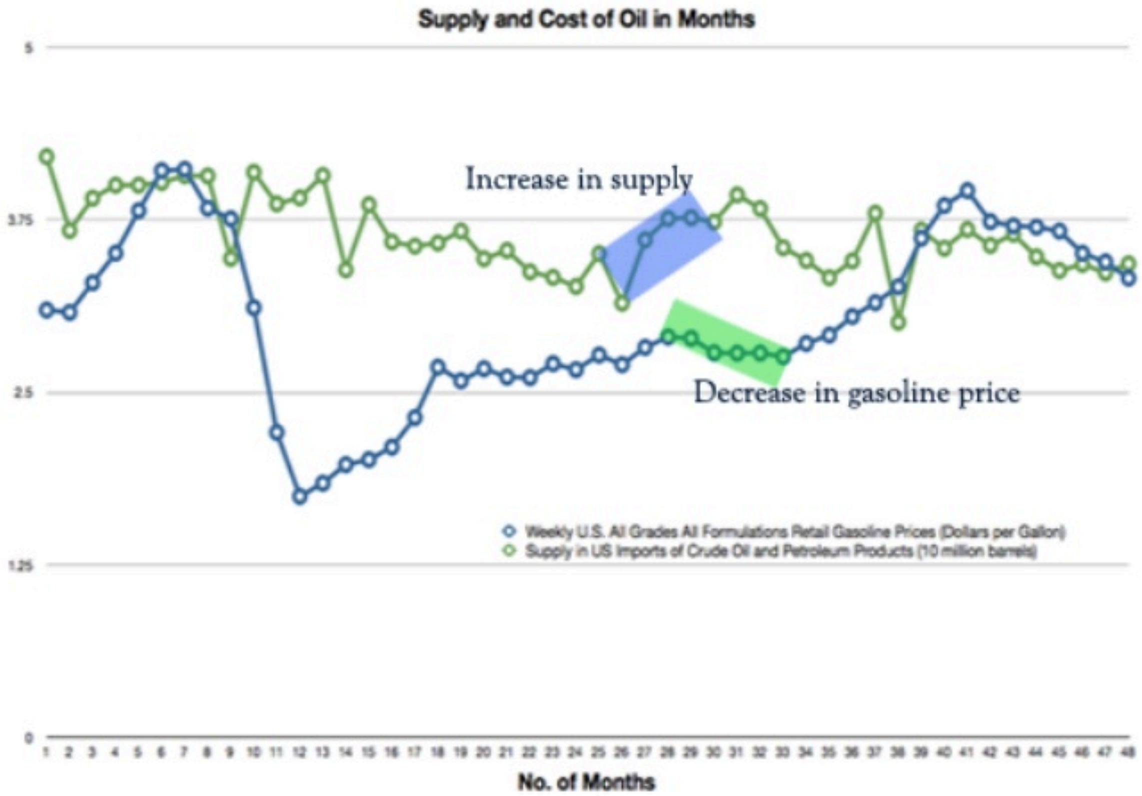


Figure 1.1

In this graph, we graph the retail price of gasoline in blue and the import of crude oil in green. Note that the actual value of the supply is in the hundred thousands, but the value graphed is S(M)/10,000 so that range of the graph could be seen closer together.

The blue box shows an increase in supply and since $S(M) \propto 1/G(M)$, we know that the the retail price of gasoline should decrease. This is shown in the green box. The observation we want to note is that the green box occurs at a later interval than the blue box, meaning that the decrease in gasoline price happens a period of time after the increase in supply.

Here are some more examples of the observation:

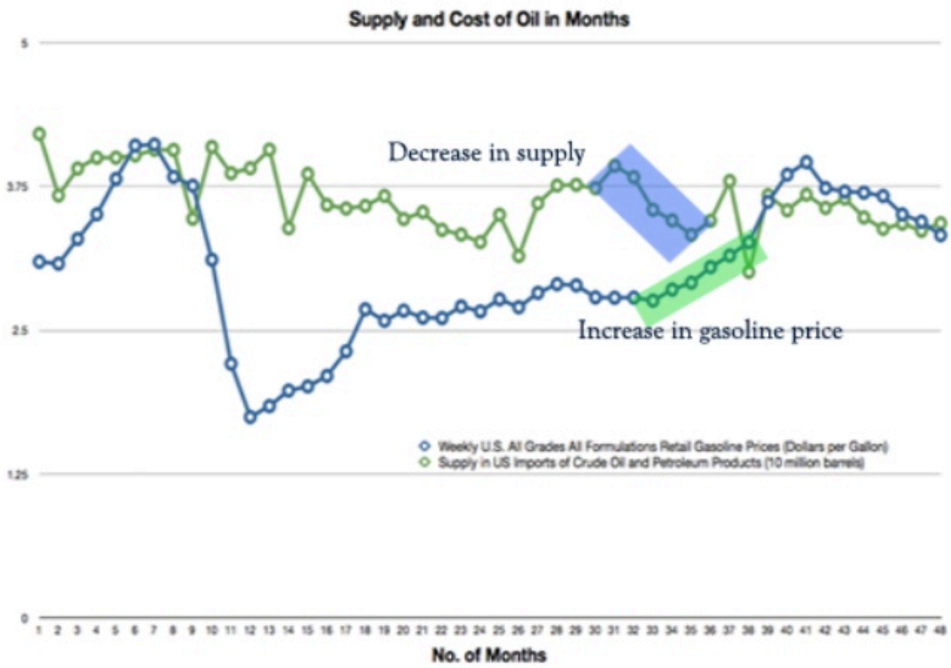


Figure 1.2

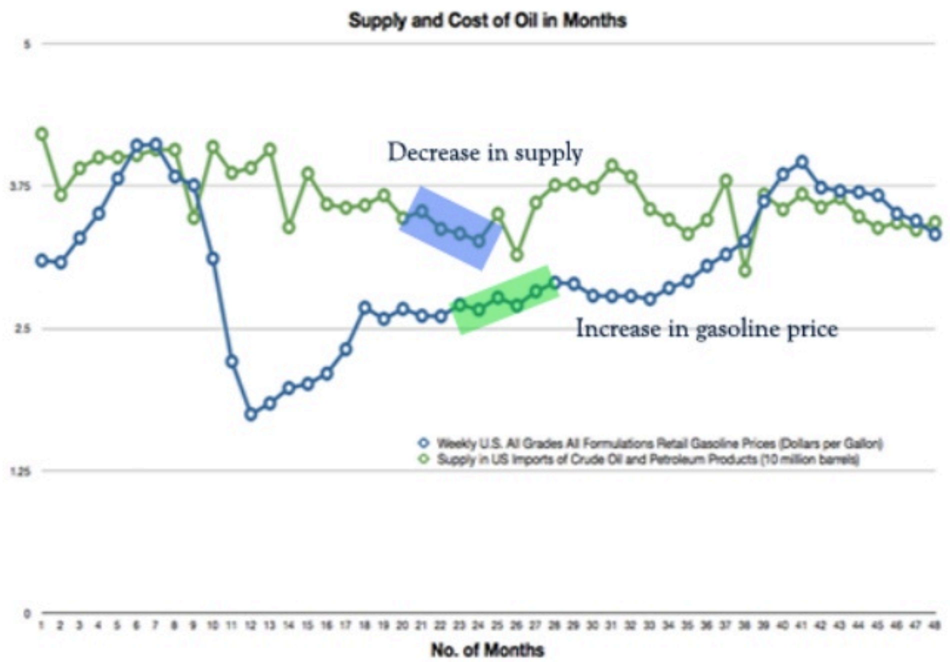


Figure 1.3

Based on this graph and others, we have concluded that $k = 1.857$ .This value is just an average of 8 observed response times.

Further explanation of details of the slope relationship:

$$G = \frac{\alpha}{S}$$

$$\frac{\partial G}{\partial t} = -\frac{\alpha}{S^2}\frac{\partial S}{\partial t}$$

G is the retail gasoline price function and S is the supply function.

The first line is just a representation of their inverse relationship for some constant α. Deriving both sides gives the second equation. The second equation says that any change in supply results in a change of the retail price of gasoline in the opposite direction.

Another observation that is key to this model is the concept of momentum. **Observation 2:** Generally momentum is considered as the quantity of motion of a moving body. In this case, we define it similarly as inertia, the tendency for a object to remain in its original state of movement. Here, we observed that although the retail price of gasoline moves inversely with the supply of crude oil, the price of gasoline seems to have a certain amount of momentum.

Simply put, if there was a large increase in supply, there would be a heavy drop in the price of gasoline. Here, the gasoline is moving downwards with a certain amount of momentum. Now, if there is a small decrease in supply, which would normally dictate a small increase in price, the momentum or tendency of the price is largely downwards and the small change in the other direction cannot reverse the trend. Instead it can only slow down the momentum (slow down the decrease in price). It is similar to how a large truck moving in one direction cannot be reversed in the other direction by simply throwing a ball at it, the most it can do is slow it down.

<u>Here is how we develop a mathematical relationship of this notion:</u>

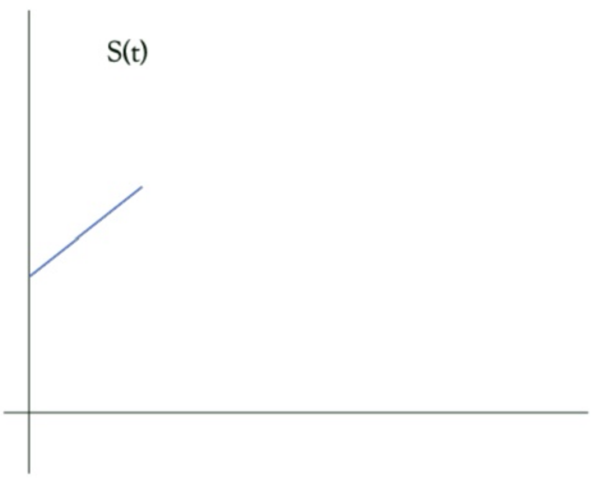
t = elapsed time

S(t) = the supply at a continuous value t

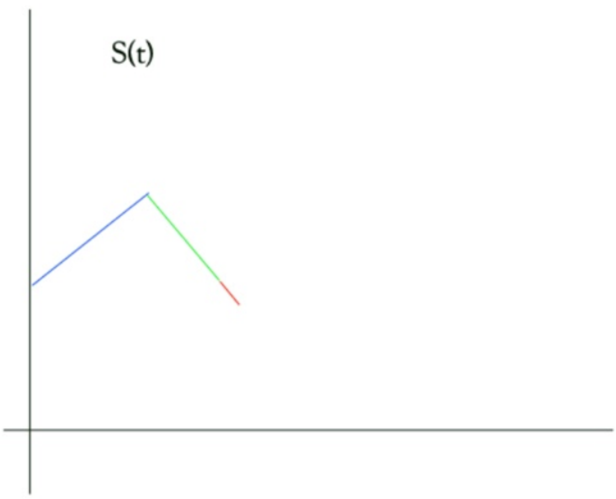
G(t) = the retail price of gasoline at a continuous value t

First, we establish that there is a relationship between the supply S(t) and retail gasoline price G(t). In general, an increase in S(t) will result in a decrease of G(t). However, by accounting for momentum, this will no longer be continuously valid after time t = 0 due to the introduction of the inertia.

Next, we start at time 0 and there is a simple movement of S(t), thus a movement of G(t+k) in the opposite direction, where k represents the delayed time for the price of gasoline to respond to the change in the supply as noted above in the first observation. Now, in order to reverse this trend, the movement of S(t) must overcome/retrace the amount it has experienced and then move into the opposite direction. Here is a visual reference to this notion:



(Figure 1.4)



(Figure 1.5)

Figure 1.4 shows an upwards movement of S(t); it is in blue. There is a corresponding movement in G(t+k). In order to reverse this trend, the movement S(t) must retrace back to its original position (green) and then continue in that direction (red).

Next, the slope of the retail price of gasoline with respect to time is in fact not dependent on the current slope of the supply function with respect to time but instead dependent on the accumulated slopes of the supply function in a set interval. In other words, the net change of S(t) is what determines the behaviour of the price function.

Therefore:

$$\beta \int S'(t)\,dt = \Delta S = -G'(t+k)$$

Equation 2.0

The price of gasoline G(t) has a negative sign because a positive movement of S(t) corresponds to a negative reaction of G(t) and vice versa, since a positive net change in supply results in a decrease of G(t); k is, as always, included in the price of the gasoline function because of the response time from Observation 1. There is a β in the equation to account for and calibrate the magnitude of the supply function. It is closely related to the constant α mentioned above but we deem the relationship between the two not important in the making of the model since it does not contribute as a real world ratio.

However, we do not want the relationship of the slopes, hence we must integrate once again to obtain:

$$\beta \int \left(\int S'(t)\,dt \right) dt = \int \Delta S\, dt = -G(t+k) + C$$

Equation 2.1

Although a double integration of a derivative may seem redundant, it is a necessary component due to the fact that we must consider only the original net change without it's own constant. By writing the integral of the derivative of S(t), we shift the function S(t) at t=0 to S(0) = 0.

This equation does not need any further explanation besides for the clarification of the "C" at the end of the equation. The constant is a result of the integration but in real world terms, C is the net result of the other factors that could come into play and of which we will analyze further below. It also include the current known value of gasoline price.

Equation 2.1 gives us the way to calculate retail gas prices as a continuous function crude oil imports in the United States. However, the values we are given are not continuous but discrete. Therefore, we must convert the equation into a discrete formula.

It is given here:

$$G(n+k) = G(n) + C - \beta n \sum_{i=1}^{n} (S(i) - S(i-1))$$

Equation 2.2

We first note that our model uses a given current retail gasoline price and the crude oil imports k months ago. We use the current value to predict the value of the future in one discrete interval at a time. This means that if we are in January, we use the data from January to find the price in February (One discrete month only).

We start with

$$\beta \int S'(t)dt = \Delta S = -G'(t+k)$$

Equation 2.3

and

$$\beta\int(\int S'(t)\,dt)\,dt = \int \Delta S\,dt = -G(t+k)+C$$

Equation 2.4

Since we observe changes by unit months. The change of the supply in discrete form is given by:

$S(i)-S(i-1)$ where i is a given month.

We then find the integral and thus result in Equation 2.2.

Next, our model is only useful if we are able to convert this formula to allow us to calculate gasoline prices based on weekly changes rather than monthly.

Note that we have values for two consecutive months (Let's say January and February). There are four weeks between the beginning of January and the beginning of February. This creates four intervals between the two points. We also have the supply of crude oil function at the discrete point of January and February. Let January be represented by time/date n = 0 and February as time/date n = 1, where n is the number of months after January. Now we can use a linear approximation to find the values of the beginning of each of the four individual weeks between the two months. This is given by

$$S(w) = (\frac{(w-1)}{4})(S(n)-S(n-1))+S(n-1)$$

Equation 2.5

where w is the w-th week of the month (in this case January) and S(w) denotes the supply at the beginning of the w-th week.

Again, we must mention that our model takes one supply value and the one discrete value after to determine gasoline prices. Therefore S(n) will always be S(1) and S(n-1) is S(0), where S(0) denotes the month we are looking at and w is the w-th week of that month.

Now that we have the value of supply in weeks, we can substitute this into Equation 2.2. Keep in mind the fact that our model only predicts one interval at a time.

It all simplifies to:

$$G(k) = G(0) + C - \beta\left(\frac{1}{4}(S(1) - S(0)\right)$$

Equation 2.6

This says that the gas price at any time k months in the future is the current value G(0) subtracting a magnitude time a quarter of the change of the monthly supply of crude oil and adding a constant value C determined by other factors mentioned in component 3 and for location-specific values in Part IV when we apply it to Boston. Keep in mind that the value that we have obtained for k above is 1.875 (this is from Observation 1). 1.875 months is also 7.5 weeks, which since using discrete values of week, we round to roughly 8 weeks.

In words, the price of gasoline one week from now is the price of gasoline 7 weeks ago subtract β multiplied by a quarter of the change in the supply of crude oil 3 weeks and and 7 weeks ago (Which is exactly a month and which is convenient since the data for crude oil is only given per month). Knowing this value, we then use the formula again to find the price of gasoline for the next week. One may mention that since our data for crude oil is in months, how we can move back 3 weeks or 7 weeks to find the supply at those points. We use

$$S(w) = (\frac{(w-1)}{4})(S(n) - S(n-1)) + S(n-1)$$

Equation 2.7.

Component 2: Individual Analysis

In this component, we take the top 5 countries who import crude oil to the United States. Based on the information given in the Energy Information Administration, the top imports of crude oil are from Canada, Saudi Arabia, Mexico, Colombia, and Kuwait.

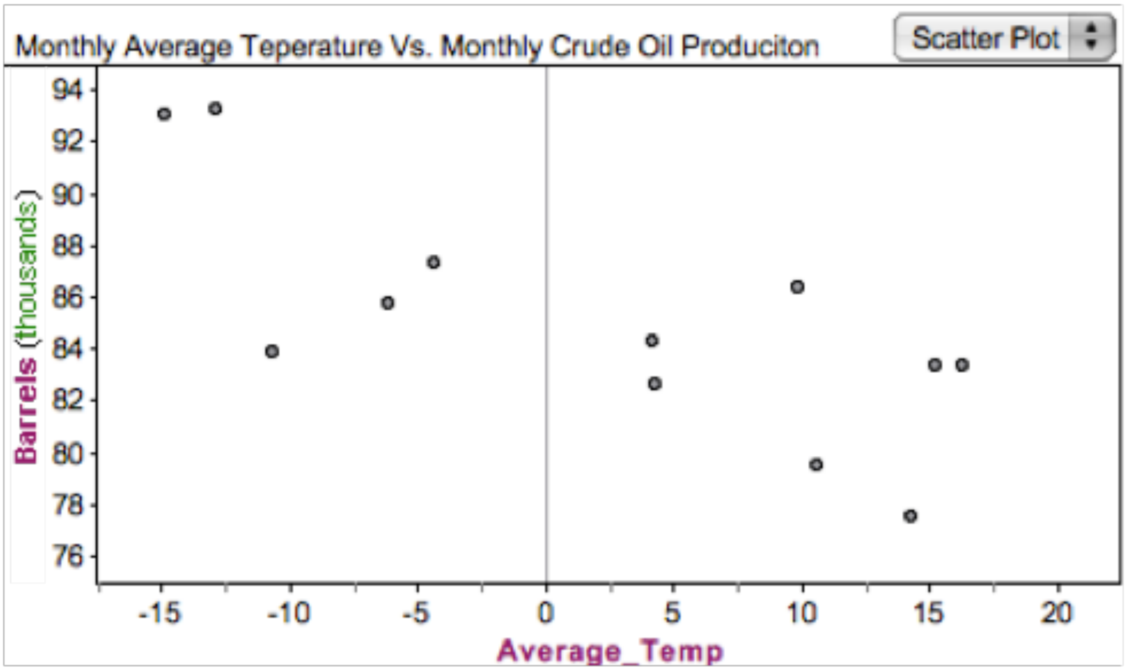
Here is a table showing the countries and their most recent export quantity to the United States:

COUNTRY	MARCH 2012	APRIL 2012	MAY 2012	JUNE 2012
Canada	90,876	87,943	93,562	91,533
Saudi Arabia	42,599	47,671	45,598	43,683
Mexico	31,130	30,054	30,873	27,447
Colombia	14,950	14,163	13,322	15,449
Kuwait	14,328	7,060	12,620	7,502
Total Sum of these 5 countries	193,883	186,891	195,970	185614
Total US Import of Crude Oil	328,923	319,005	345,098	341,782

Below is a brief profile of each of the major countries that were considered in the making of our model:

Canada

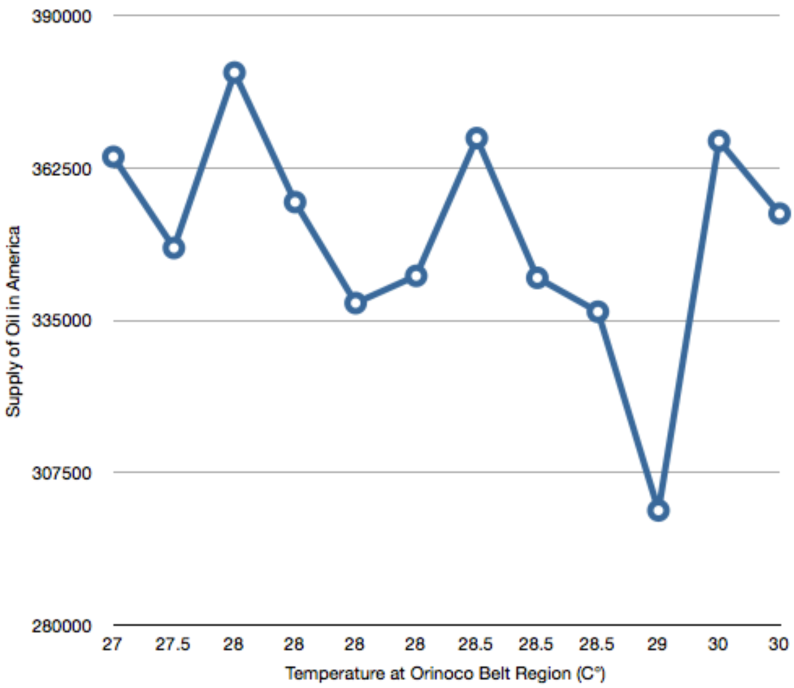
Canada is one of the world's largest crude oil producers on the planet today, and is the United State's largest crude oil supplier. Canada has exported about 900 million barrels per year within the last 4 years. 95% of Canada's Crude oil production is from Alberta which is located in Central Canada. Most of the Oil is obtained from the Athabasca Oil Sands which is located in Northeastern Alberta. Most of the oil extracted from this area is categorized as extremely crude oil and requires a series of refinements to produce gasoline. Most of the oil is transported to the United State by the pipelines. The climate of the region there can be quite different during summers and winters, ranging from sunny summers to snowy winters. Temperature in Alberta can range from -30 Celsius to 30 Celsius. Although temperature was initially a factor that would have considerable effect on the supply of crude oil, by graphing the data, there was little relationship and correlation between the two variables. Shown from the graph below. Hence it is concluded that the temperature cannot affect the change in supply by a significant factor. Shown below is the correlation between temperature and crude oil production.



Venezuela

Venezuela is the world's fifth largest oil-exporting country in the world. It is home to the one of the largest reserves of large heavy and light crude oil. By 2011, the reserves reached 297 billion barrels, surpassing that of Saudi Arabia. Venezuela is one of the first five countries of OPEC. A large percentage of Venezuelan oil is taken from the Orinoco Belt Region which is located south of the Eastern River Orinoco Basin. This belt region contains what is called the Orinoco Oil Sands, a very large deposit of extra heavy crude.

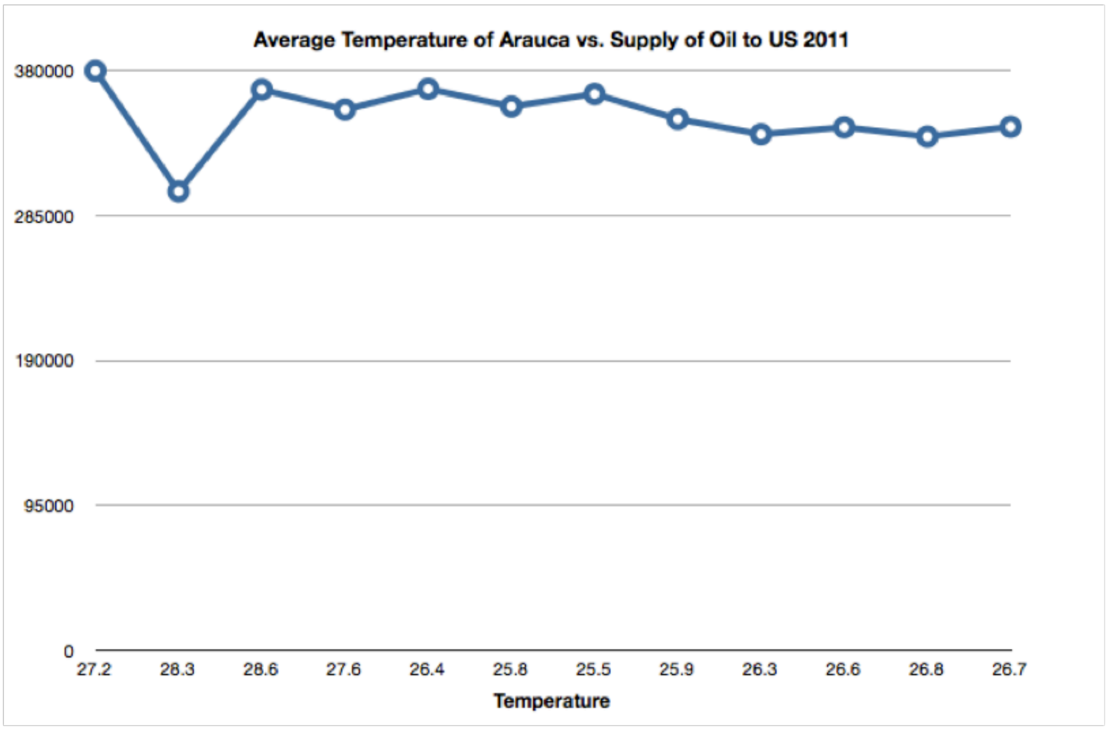
We wanted to see if the weather in the Orinoco Belt Region would affect Venezuelan supply of oil to the United States. We therefore plotted the total supply of oil to America against the temperature of Venezuela. On completion of this graph, we found no correlation between Venezuelan temperature and it's supply of oil to the United States. The graph of the relationship between Venezuelan temperature and American oil supply is shown below.



Columbia

In Columbia, South America, oil comes from the region Arauca Department of Colombia, located in the extreme north of the Orinoco part of Colombia, bordering Venezuela. Caño Limón oil fields located in Arauca territory accounts for most of Colombian oil output. The average temperature of the months of Arauca was found to see if there was any correlation between the weather of the region of one of the biggest exporters to the US and the supply of oil of the US.

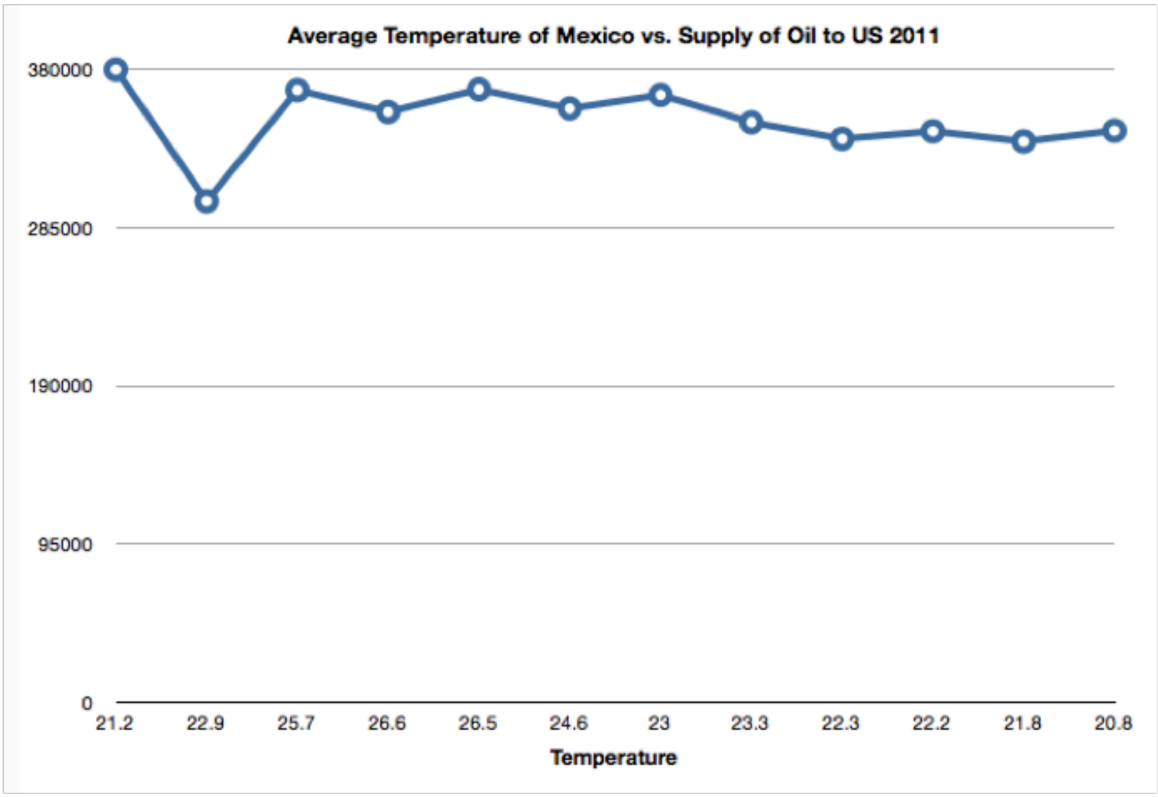
We wanted to see if the weather in the Arauca Department of Colombia would affect Columbia supply of oil to the United States. We therefore plotted the total supply of oil to America against the temperature of Aruca. On completion of this graph, we found no correlation between Aruca temperature and it's supply of oil to the United States. The graph of the relationship between Arauca, Columbia temperature and American oil supply is shown below.



Mexico

Oil is taken from Canterell Field production peaked at 2.1 million barrels per day (330,000 m^3/d) in 2003. antarell is located 80 kilometres (50 mi) off-shore in the Bay of Campeche. This complex comprises four major fields: Akal (by far the largest), Nohoch, Chac and Kutz. The average temperature of the months of Canterell Field was found to see if there was any correlation between the tempearture of Mexico and the supply of oil of the US.

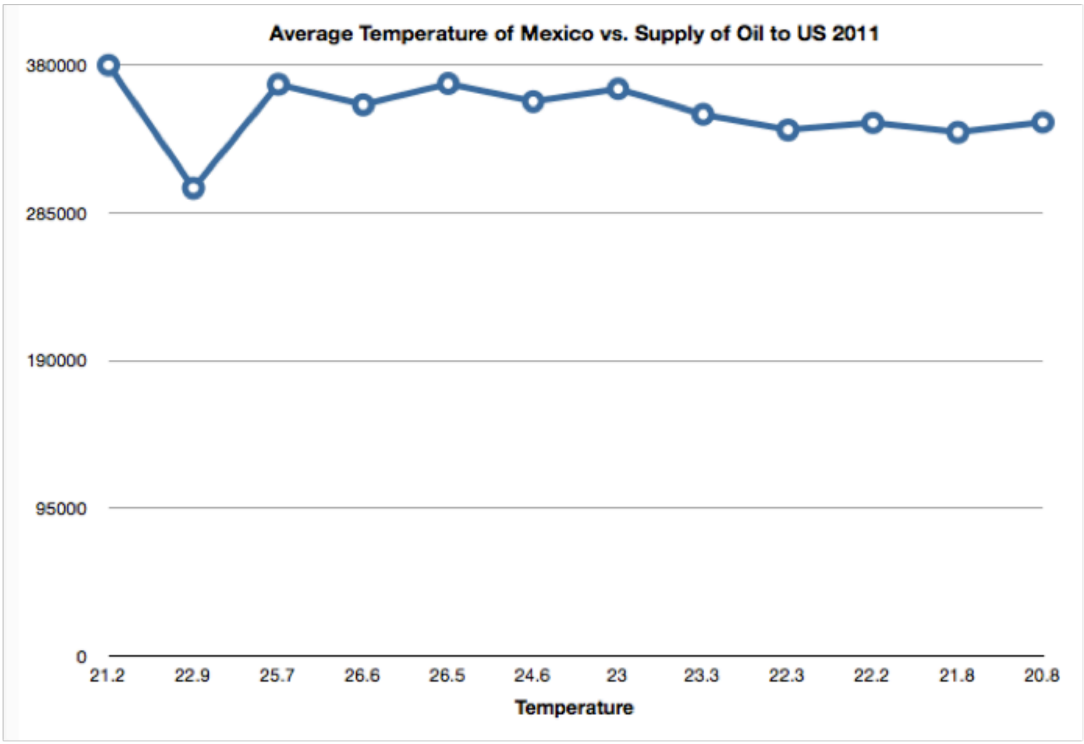
However, as the oil is taken offshore Mexico, the temperature of offshore is completely different than temperature of land. We therefore plotted the total supply of oil to America against the temperature of inshore Mexico. On completion of this graph, we found no correlation between Mexico temperature and it's supply of oil to the United States. The graph of the relationship between Mexico temperature and American oil supply is shown below.



Saudi Arabia

Oil is taken from Al-Ahsa Governorate (Ghawar Field) of the Eastern Province which is the largest province of Saudi Arabia, home to most of Saudi Arabia's oil production.

Similarly, we plotted the total supply of oil to America against the temperature of Ghawar Field. On completion of this graph, we found no correlation between Mexico temperature and it's supply of oil to the United States. The graph of the relationship between the temperature of Saudi Arabia and American oil supply is shown below.



Component 3: Qualitative

Components 1 Supply and part of Component 2 Individual Analysis dealt with the analysis of the supply portion of the model. Now, we must consider the demand portion of crude oil. Earlier, we had established that the short-term price elasticity of demand for crude oil is realtively low. Thus although supply may be shifting in the short-run, the general demand remains relatively stable. However, it is without doubt that world events may have a noticeable short-term effect on the demand. One example is that a harsher winter in February of 2012 in Europe led to a greater demand in heating and thus a slight increase in crude oil (and by general analysis 1, subsequently retail gasoline prices).

The Qualitative Data are presented as weather and stability. Although there are thousands qualitative factors so we will not be able to account for all of them. We deemed that weather and stability to be the ones that have the most effect on the price of gasoline.

Weather is the quality of the environment of specific areas that dictate demand. These are usually large important countries or regions. Political Stability indicates the level of confidence in the stability of a region. If the region heavily influences crude oil prices (such as the Middle East) then a shift in stability could potentially alter crude oil production and thus prices.

These qualitative values are the make up of the C values.

Weather:

The weather can be a major indicator of the demand of crude oil. A cold spell could potentially raise crude oil prices since people will need more heating in their homes. However, based on the analysis of component 2, minor changes in weather have almost no major impacts. We have also noticed that the actually quantitative values of temperature are unnecessary in calculation. Instead, we observe weather changes by categories.

<u>Categories </u>: If the temperature at any region deviates from the average by more than 5 degrees Celsius. The arbitrary value of 5 is used because it is not an extreme value but will create a noticeable difference.

If it is 5 degrees above the average temperature, there would be less heating required and thus a decrease in demand and a decrease in the price of gasoline.

If it is 5 degrees above the average temperature, there would be more heating required and thus an increase in demand and an increase in the price of gasoline.

The next interval is a change of 10 degrees Celsius. This is a largely noticeable change and thus will move gas prices.

The last is if there is a change of more than 20 degrees Celsius. This is an extreme movement in temperature and is very rare. It will have a dramatic affect on the retail gasoline prices. Here is a table showing the set effect of each category.

CATEGORY	TEMPERATURE CHANGE	EFFECT
Category 1	5 (C)	3%
Category 2	10 (C)	5%
Category 3	20 (C)	10%

The effect of whether is accounted for in the General Pricing Equation as a component of C. This to use this, we multiply G(0) or the initial gasoline price by the effect and this is the whether part of C. For example, if the price gasoline initially is G(0), and we experience a category 2 change in weather, we add or subtract .05*G(0). We mention again that a decrease in temperature indicates an increase in price and an increase in temperature indicates a decrease in price. Since temperatures and climate appear for a period of time, as long as the temperature is within the category parameters, the effect will be accounted for.

The Weather Variable is denoted as: ε

Here is a table showing the average temperatures around Europe and in China based on the months:

City	Climate	Average Monthly Temperature, Rainfall & Days of Rain											
		JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
Beijing	Average(°C)	-4.6	-2.2	4.5	13.1	19.8	24.0	25.8	24.4	19.4	12.4	4.1	-2.7
	Rainfall(mm)	3.0	7.4	8.6	19.4	33.1	77.8	192.5	212.3	57.0	24.0	6.6	2.6
	Days	2.0	3.1	4.1	4.6	5.9	9.7	14.1	13.2	6.8	5.0	3.7	1.6
Chengdu	Average(°C)	5.5	7.5	12.1	17.0	20.9	23.7	25.6	25.1	21.2	16.8	11.9	7.3
	Rainfall(mm)	5.9	10.9	21.4	50.7	88.6	111.3	235.5	234.1	118.0	46.4	18.4	5.8
	Days	6.0	7.7	10.9	13.8	16.1	15.9	17.4	15.5	16.2	15.2	8.3	5.7
Chongqing	Average(°C)	7.5	9.5	14.1	18.8	22.1	25.2	28.6	28.5	23.8	18.6	13.9	9.5
	Rainfall(mm)	19.7	19.5	39.3	89.7	157.8	166.4	142.4	138.4	136.3	97.3	47.8	25.0
	Days	9.5	9.3	11.2	13.4	17.7	15.5	11.1	10.8	13.9	16.4	13.5	10.1
Dalian	Average(°C)	-4.9	-3.4	2.1	9.1	18.5	19.4	23.0	20.9	20.6	13.6	5.8	-1.3
	Rainfall(mm)	7.6	7.7	12.5	35.8	43.9	86.1	175.6	153.0	68.5	35.6	21.6	10.8
	Days	3.7	3.0	4.1	5.8	6.7	9.4	13.1	10.6	7.2	5.5	5.3	4.1
Dunhuang	Average(°C)	-9.3	-4.1	4.5	12.4	18.3	22.7	24.7	23.5	17.0	8.7	0.2	-7.0
	Rainfall(mm)	0.8	1.6	1.2	2.9	1.6	6.7	12.1	5.3	1.8	1.0	1.1	0.7
	Days	1.1	1.7	0.8	1.2	1.2	2.6	3.9	2.6	0.8	0.7	1.0	1.4
Fuzhou	Average(°C)	10.5	10.7	13.4	18.2	22.1	25.5	28.5	28.2	26.0	21.7	17.5	13.1
	Rainfall(mm)	49.8	76.3	120.0	149.7	207.5	230.2	112.0	160.5	131.4	41.5	33.1	31.6
	Days	10.9	13.4	16.4	16.7	19.0	17.1	9.7	13.0	12.3	6.9	7.3	8.1
Guangzhou	Average(°C)	13.3	14.4	17.9	21.9	25.6	27.2	28.4	28.1	26.9	23.7	19.4	15.2
	Rainfall(mm)	36.9	54.5	80.7	125.0	293.8	287.8	212.7	232.5	189.3	69.2	37.0	24.7
	Days	8.0	10.7	14.2	15.1	18.7	20.0	16.2	16.4	13.0	6.5	5.6	5.8
Guilin	Average(°C)	7.9	9.1	13.2	18.4	23.0	26.2	28.3	27.7	25.6	20.7	15.1	10.2
	Rainfall(mm)	56.9	75.1	128.0	282.3	353.8	311.7	231.7	169.9	63.6	98.4	74.9	54.0
	Days	13.6	14.3	18.8	20.9	20.2	17.9	15.4	15.4	8.6	9.5	9.6	11.0
Haikou	Average(°C)	17.2	18.2	21.6	24.9	27.4	28.1	28.4	27.7	26.8	24.8	21.8	18.7
	Rainfall(mm)	23.6	30.4	52.0	92.8	182.6	241.2	206.7	239.5	302.8	172.4	97.6	38.0
	Days	9.2	10.1	10.2	11.0	16.3	16.3	14.3	16.4	16.1	12.2	9.6	8.6

(These two regions were chosen because they represent the most consumption of crude oil outside of the United States. The United States temperature isn't taken into consideration since PART IV will choose a specific city and it will be taken into account there)

Average temperature (°C) during the day

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
Minsk[20]	-2.7	-1.4	3.3	11.5	18.3	21.5	22.4	22.2	15.9	9.6	2.6	-1.0	10.1
Stockholm[21]	-0.7	-0.6	3.0	8.6	15.7	20.7	21.9	20.4	15.1	9.9	4.5	1.1	10.0
Budapest[22]	1.2	4.5	10.2	16.3	21.4	24.4	26.5	26.0	22.1	16.1	8.1	3.1	15.0
London[23]	7.9	8.2	10.9	13.3	17.2	20.2	22.8	22.6	19.3	15.2	10.9	8.8	14.8
Paris[24]	6.9	8.2	11.8	14.7	19.0	21.8	24.4	24.6	20.8	15.8	10.4	7.8	15.5
Barcelona[25][26]	13.4	14.6	15.9	17.6	20.5	24.2	27.5	28.0	25.5	21.5	17.0	14.3	20.0
Lisbon[27]	14.8	16.2	18.8	19.8	22.1	25.7	27.9	28.3	26.5	22.5	18.2	15.3	21.5
Valencia[28]	16.1	17.2	18.7	20.2	22.8	26.2	29.1	29.6	27.6	23.6	19.5	16.8	22.3
Athens[29]	12.9	13.6	16.0	20.3	25.3	29.8	32.6	32.3	28.9	23.1	18.6	14.7	22.3

Stability:

It is without doubt that the stability of nations around the world have a certain effect on the price of gasoline. Possible paths to war and violence in Northern Africa and the Middle East often drive prices up and down. Just like with weather, we will divide the stability factor into three categories.

Category 1: General Stability

When the middle east or another important region is in this category, there is relatively no unrest and generally peaceful. It creates a general decrease in gasoline prices as the supply of crude oil is stable and without disturbance.

Category 2: Protest and unrest

This category includes nations in general unrest. There are protests and or fear of violence in the nation. Category two instability may lead to violation of human rights and subsequent embargo on oil producing nations. This will lead to a fear that the future will have a lower supply of crude oil and thus an increase in gasoline prices.

Category 3: Social Violence and Military Action

A nation in this category is deemed very unstable and military action may be needed. An example may include the current state of Syria or the Arab Spring.

The table shows the categories and their assigned effect:

CATEGORY	EFFECT
Category 1	-2%
Category 2	+3%
Category 3	+5%

The usage of this table is similar to the usage of the categories of weather. When the state of an important oil producing nation is in one of the set category, the corresponding effect is multiplied by the effect and added to C. The Stability Variable is denoted as η.

Final C value

Thus, we now have a model to account for the two major qualitative factors: Weather and Stability.

The sum of the two factors is the C value in

$$G(k) = G(0) + C - \beta\left(\frac{1}{4}(S(1) - S(0)\right)$$

$$C = \varepsilon + \eta$$

Building the Model Based on 2011

The current model we developed is based on observations and theory. It takes into account current gasoline prices, supply and demand of crude oil (Which by General Analysis 1, we know to be related to gasoline prices), along with qualitative data. Although this model will work in theory, we are missing the β value to account for the conversion between the supply crude oil and factor magnitude. We will use the data in 2011 to solve for values of β and then find its average.

First, we have to simplify the model to be user friendly.

The current one is this:

$$\underline{G(k) = G(0) + C - \beta\left(\frac{1}{4}(S(1) - S(0)\right)}$$

This function gives us the price of gasoline 1.875 months (k value from observation 1) from now. In component 1, we rounded it to approximately 8 weeks. Therefore, we must shift the function by 7 weeks to allow it to predict the value one week from now:

$$G(n) = G(n-7) + C - \beta\left(\frac{1}{4}(S(-6) - S(-7)\right)$$

The problem now is that we only have the supply values in month. The original linear approximation used in component 1 only works to simplify the equation but is useless to actually contribute to the formula. There are 4 weeks in one month so S(0) = S(1) = S(2) = S(3) and S(4)= S(5) = S(6) = S(7). This means that the portion multiplied by β, which is the change of the supply using a linear approximation, is going to be the same.

To solve this, we realize that the supply of crude oil does not actually move in a linear motion across the time period of the four weeks from the price of one month to the price of the next as shown below:

Instead between month 1 and month 2, it moves in a motion we will assume to be nearly random (Brownian Motion). We add a random value ϕ to each S(1) and S(0) to account for and calculate these random changes. Now we have

$$\frac{1}{4}\Delta S = (S(1)+\phi)-(S(0)+\phi))$$

One may think that ϕ cancels out but it does not since not all ϕ are equal, they are random variables.

Formally expressed ϕ is a random variable following a normal distribution.

The mean of ϕ is 0.

The standard deviation of ϕ is 1.

On average,

$$\Delta S = (S(1) - S(0))$$

To account for the large value of the supply graph, instead of adding ϕ, we normalize the supply distribution and thus multiply ϕ by 125,000.

Now, our new user friendly equation is:

$$G(n) = G(n-7) + C - \beta \frac{(S(-6) + 125{,}000\phi) - (S(-7) + 125{,}000\phi)}{4}$$

Where n denotes the number of weeks from now and G(n) denotes the gasoline price at n.

This table gives the 2011 data where the week of January first 2011 is denoted by n=0

This shows the gasoline prices, the weather, and stability effects.

N	G(N)	G(N-7)	$C = \varepsilon + \eta$	ε (Weather Effect)	WEATHER CATEGORY	η (Stability Effect)	STABILITY CATEGORY
-7	2.892						
-6	2.876						
-5	2.856						
-4	2.958						
-3	2.98						
-2	2.982						
-1	3.052						
0	3.07	2.892	0.18312	0.09156	1	0.09156	2
1	3.089	2.876	0.1842	0.0921	1	0.0921	2

N	G(N)	G(N-7)	$C = \varepsilon + \eta$	ε (Weather Effect)	WEATHER CATE-GORY	η (Stability Effect)	STABILITY CATE-GORY
2	3.104	2.856	0.18534	0.09267	1	0.09267	2
3	3.11	2.958	0.18624	0.09312	1	0.09312	2
4	3.101	2.98	0.2488	0.0933	1	0.1555	3
5	3.132	2.982	0.15505	0	none	0.15505	3
6	3.14	3.052	0.1566	0	none	0.1566	3
7	3.189	3.07	0.157	0	none	0.157	3
8	3.383	3.089	0.15945	0	none	0.15945	3
9	3.52	3.104	0.16915	0	none	0.16915	3
10	3.567	3.11	0.176	0	none	0.176	3
11	3.562	3.101	0.17835	0	none	0.17835	3
12	3.596	3.132	0.1781	0	none	0.1781	3
13	3.684	3.14	0.1798	0	none	0.1798	3
14	3.791	3.189	0	-0.1842	2	0.1842	3
15	3.844	3.383	0	-0.18955	2	0.18955	3
16	3.879	3.52	-0.07688	-0.1922	2	0.11532	2
17	3.963	3.567	-0.07758	-0.19395	2	0.11637	2
18	3.965	3.562	0.11889	0	none	0.11889	2
19	3.96	3.596	0.11895	0	none	0.11895	2
20	3.849	3.684	0.198	0	none	0.198	3
21	3.794	3.791	0.19245	0	none	0.19245	3
22	3.781	3.844	0.3794	0.1897	2	0.1897	3
23	3.713	3.879	0.3781	0.18905	2	0.18905	3
24	3.652	3.963	0.3713	0.18565	2	0.18565	3
25	3.574	3.965	0.3652	0.1826	2	0.1826	3
26	3.579	3.96	0.28592	0.10722	1	0.1787	3
27	3.641	3.849	0.28632	0.10737	1	0.17895	3

N	G(N)	G(N-7)	$C=\varepsilon+\eta$	ε (Weather Effect)	WEATHER CATEGORY	η (Stability Effect)	STABILITY CATEGORY
28	3.682	3.794	0.03641	0.10923	1	-0.07282	1
29	3.699	3.781	0.03682	0.11046	1	-0.07364	1
30	3.711	3.713	0.03699	0.11097	1	-0.07398	1
31	3.674	3.652	0.03711	0.11133	1	-0.07422	1
32	3.604	3.574	0.22044	0.11022	1	0.11022	2
33	3.581	3.579	0.21624	0.10812	1	0.10812	2
34	3.627	3.641	0.21486	0.10743	1	0.10743	2
35	3.674	3.682	0.29016	0.18135	2	0.10881	2
36	3.661	3.699	0.29392	0.1837	2	0.11022	2
37	3.601	3.711	0.29288	0.18305	2	0.10983	2
38	3.509	3.674	0.28808	0.18005	2	0.10803	2
39	3.433	3.604	0.17545	0	none	0.17545	3
40	3.417	3.581	0.17165	0	none	0.17165	3
41	3.476	3.627	0.17085	0	none	0.17085	3
42	3.462	3.674	0.1738	0	none	0.1738	3
43	3.452	3.661	0.1731	0	none	0.1731	3
44	3.424	3.601	0.10356	0	none	0.10356	2
45	3.436	3.509	0.10272	0	none	0.10272	2
46	3.368	3.433	0.10308	0	none	0.10308	2
47	3.307	3.417	0.10104	0	none	0.10104	2
48	3.29	3.476	0.26456	0.09921	1	0.16535	3
49	3.286	3.462	0.2632	0.0987	1	0.1645	3
50	3.229	3.452	0.26288	0.09858	1	0.1643	3
51	3.258	3.424	0.25832	0.09687	1	0.16145	3

Now, we calculate the β values:

Keep in mind that we need the random values ϕ, which we found using a random number generator that uses a normal distribution with a mean of 0 and a standard deviation of 1.

N	G(N)	G(N-7)	S(N)	S(N-1)	C	β	FIRST ϕ	SECOND ϕ
-7	2.892							
-6	2.876							
-5	2.856							
-4	2.958							
-3	2.98							
-2	2.982							
-1	3.052							
0	3.07	2.892	379676	345099	0.18312	-0	-1.97	0.58
1	3.089	2.876	379676	345099	0.1842	-0.000001	0.9830	-0.59
2	3.104	2.856	379676	345099	0.18534	-0.000003	0.0493	-0.33
3	3.11	2.958	379676	345099	0.18624	-0.000001	-0.252	1.6
4	3.101	2.98	379676	379676	0.2488	0.000007	0.692	-0.539
5	3.132	2.982	300672	379676	0.15505	-0	-1.49	-0.00121
6	3.14	3.052	300672	379676	0.1566	0.000012	-1.61	-0.969
7	3.189	3.07	300672	379676	0.157	-0.000001	1.81	0.504
8	3.383	3.089	300672	379676	0.15945	0.000001	-0.415	2.33
9	3.52	3.104	367341	300672	0.16915	-0.000004	-0.106	-1.34
10	3.567	3.11	367341	300672	0.176	0.00001	1.73	1.64
11	3.562	3.101	367341	300672	0.17835	-0.000021	0.675	0.176
12	3.596	3.132	367341	300672	0.1781	-0.000029	0.094	0.224
13	3.684	3.14	354229	367341	0.1798	0.000096	-0.243	-0.0075
14	3.791	3.189	354229	367341	0	-0.000099	-0.25	-0.324
15	3.844	3.383	354229	367341	0	0.000011	0.925	1.31

N	G(N)	G(N-7)	S(N)	S(N-1)	C	β	FIRST ϕ	SECOND ϕ
16	3.879	3.52	354229	367341	-0.07688	-0.000038	-1.18	-0.594
17	3.963	3.567	367839	354229	-0.07758	0.000026	-0.0445	0.686
18	3.965	3.562	367839	354229	0.11889	0	-.0.694	0.701
19	3.96	3.596	367839	354229	0.11895	-0.000009	-2.01	-0.929
20	3.849	3.684	367839	354229	0.198	-0.000005	0.421	0.343
21	3.794	3.791	367839	367839	0.19245	0.000006	-0.888	-1.12
22	3.781	3.844	356304	367839	0.3794	0	-.1.13	-0.948
23	3.713	3.879	356304	367839	0.3781	0	-.0.701	-0.458
24	3.652	3.963	356304	367839	0.3713	0.000019	-0.39	-1.3
25	3.574	3.965	356304	367839	0.3652	-0.000014	-0.418	1.56
26	3.579	3.96	364469	356304	0.28592	-0.000079	0.521	0.389
27	3.641	3.849	364469	356304	0.28632	0.000281	-1.83	-0.174
28	3.682	3.794	364469	356304	0.03641	0.000009	1.06	-0.362
29	3.699	3.781	364469	356304	0.03682	-0.000006	-1.92	0.488
30	3.711	3.713	348033	364469	0.03699	-0.000004	1.61	0.321
31	3.674	3.652	348033	364469	0.03711	-0.000001	0.0641	0.239
32	3.604	3.574	348033	364469	0.22044	-0.000048	0.330	0.0293
33	3.581	3.579	348033	364469	0.21624	-0.000008	-1.48	0.54
34	3.627	3.641	348033	348033	0.21486	-0.000012	-1.1	0.51
35	3.674	3.682	338087	348033	0.29016	0.000005	0.742	-2
36	3.661	3.699	338087	348033	0.29392	-0.000014	-0.646	0.623
37	3.601	3.711	338087	348033	0.29288	-0.000016	0.596	0.769
38	3.509	3.674	338087	348033	0.28808	0.000097	1.1	-0.119
39	3.433	3.604	342638	338087	0.17545	-0.000049	0.742	0.335
40	3.417	3.581	342638	338087	0.17165	0.000016	-0.646	-0.691
41	3.476	3.627	342638	338087	0.17085	0.001491	0.596	0.0891
42	3.462	3.674	342638	338087	0.1738	0	-.1.4	-1.31

N	G(N)	G(N-7)	S(N)	S(N-1)	C	β	FIRST ϕ	SECOND ϕ
43	3.452	3.661	342638	342638	0.1731	-0.000013	-1.04	0.805
44	3.424	3.601	336523	342638	0.10356	-0.000008	0.312	1.04
45	3.436	3.509	336523	342638	0.10272	-0.000011	0.388	0.499
46	3.368	3.433	336523	342638	0.10308	0.000003	-1.97	-2.12
47	3.307	3.417	336523	342638	0.10104	-0.00003	-1.11	0.0644
48	3.29	3.476	342995	336523	0.26456	0.00001	1.02	-1.23
49	3.286	3.462	342995	336523	0.2632	0.000007	1.92	-1.74
50	3.229	3.452	342995	336523	0.26288	0.000519	-1.2	-0.0982
51	3.258	3.424	342995	336523	0.25832	0.000017	-0.307	-0.758

Based upon the above calculations, we have found many β values, each for every week of the 2011 data points. The average β value was 0.0000407. This is reasonable as the production values are large and thus a very small β value is needed to calibrate the values.

Testing our model on 2012 Retail Gas Prices

First, we must quantity the weather and stability effects as per the categories in component 3:

N denotes the weeks after January 2012

G(N) denotes

N	ε (Weather Effect)	WEATHER CATEGORY	η (Stability Effect	STABILITY CATEGORY	$C = \varepsilon + \eta$	G(N-1)
0	0.3258	3	0.09774	2	0.42354	3.258
1	0.3299	3	0.09897	2	0.42887	3.299
2	0.3382	3	0.10146	2	0.43966	3.382
3	0.3391	3	0.10173	2	0.44083	3.391
4	0.3389	3	-0.06778	1	0.27112	3.389
5	0.3439	3	-0.06878	1	0.27512	3.439
6	0.3482	3	-0.06964	1	0.27856	3.482
7	0.3523	3	-0.07046	1	0.28184	3.523
8	0.3591	3	-0.07182	1	0.28728	3.591
9	0.3721	3	0.11163	3	0.48373	3.721
10	0.3793	3	0.11379	3	0.49309	3.793
11	0.3829	3	0.11487	3	0.49777	3.829
12	0.3867	3	0.11601	3	0.50271	3.867
13	0.3918	3	0.11754	2	0.50934	3.918
14	0.3941	3	0.11823	2	0.51233	3.941
15	0.3939	3	0.11817	2	0.51207	3.939
16	0.3922	3	0.11766	2	0.50986	3.922
17	0.387	3	0.1161	2	0.5031	3.87
18	0.383	3	0.1149	2	0.4979	3.83
19	0.379	2	0.1137	2	0.4927	3.79
20	0.3754	2	0.11262	2	0.48802	3.754

N	ε (Weather Effect)	WEATHER CATEGORY	η (Stability Effect	STABILITY CATEGORY	$C = \varepsilon + \eta$	G(N-1)
21	0.3715	2	0.11145	2	0.48295	3.715
22	0.367	2	0.1101	2	0.4771	3.67
23	0.3613	1	0.10839	2	0.46969	3.613
24	0.3572	1	0.10716	2	0.46436	3.572
25	0.3533	1	0.10599	2	0.45929	3.533
26	0.3437	1	0.10311	2	0.44681	3.437
27	0.3356	2	0.10068	2	0.43628	3.356
28	0.3411	2	0.10233	2	0.44343	3.411
29	0.3427	2	0.10281	2	0.44551	3.427
30	0.3494	2	0.10482	2	0.45422	3.494
31	0.3508	2	0.10524	2	0.45604	3.508
32	0.3645	2	0.10935	2	0.47385	3.645
33	0.3721	2	0.11163	2	0.48373	3.721
34	0.3744	2	0.11232	2	0.48672	3.744
35	0.3776	2	0.11328	2	0.49088	3.776
36	0.3843	2	0.11529	2	0.49959	3.843
37	0.3847	2	0.11541	2	0.50011	3.847
38	0.3878	2	0.11634	2	0.50414	3.878
39	0.3826	2	0.11478	2	0.49738	3.826
40	0.3804	2	0.11412	2	0.49452	3.804
41	0.385	2	0.1155	2	0.5005	3.85
42	0.3819	2	0.11457	2	0.49647	3.819
43	0.3687	2	0.11061	2	0.47931	3.687
44	0.3568	2	0.10704	2	0.46384	3.568
45	0.3492	2	0.10476	2	0.45396	3.492
46	0.3449	2	0.10347	2	0.44837	3.449

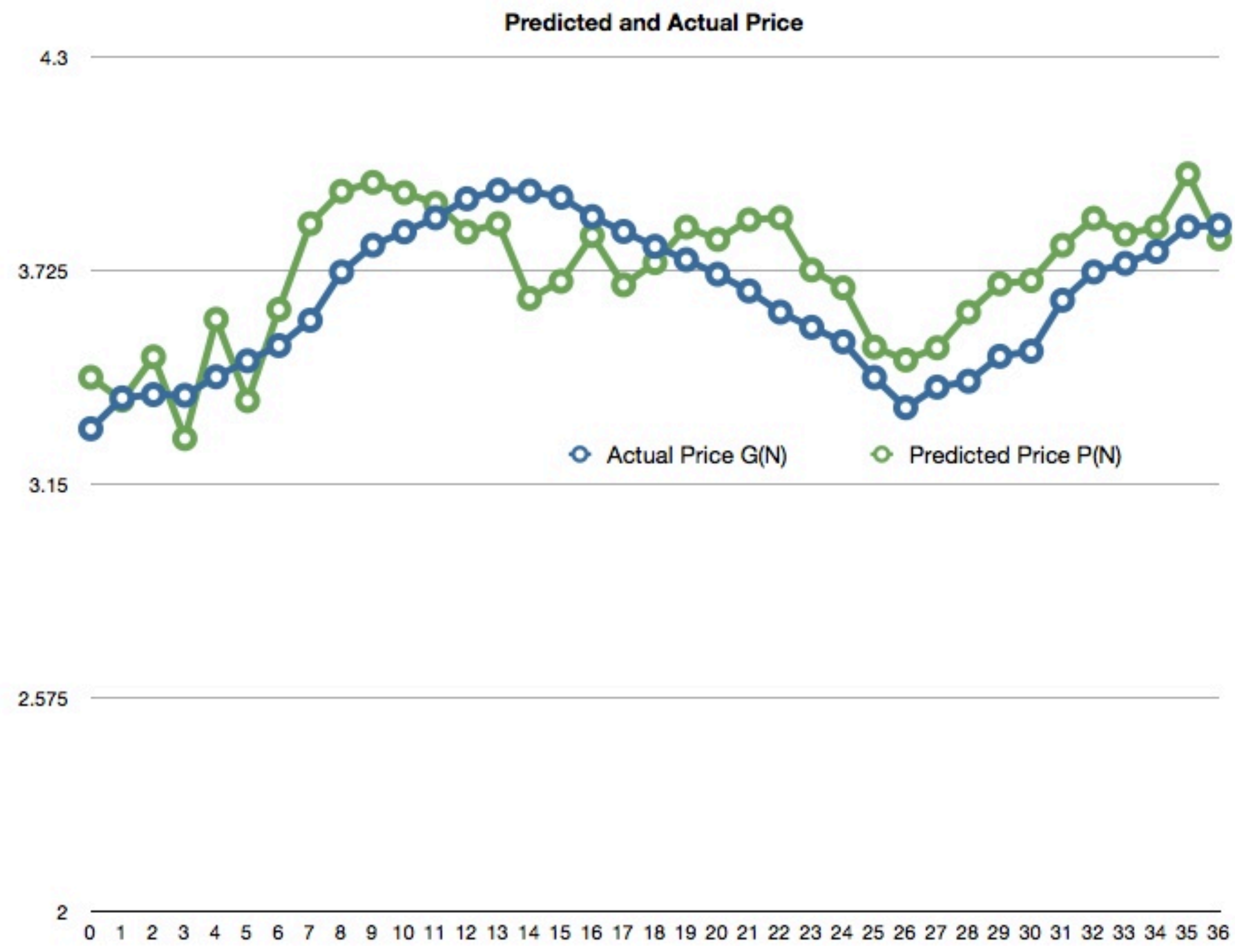
Now that we've determined the C values, we can determine our own price predictions and the actual price values given that β =.0000407

N	ACTUAL PRICE G(N)	PRE-DICTED PRICE P(N)	S(N)	S(N-1)	G(N-7)	C	FIRST ϕ	SECOND ϕ
-7	3.436		336523				-1.97	0.58
-6	3.368		336523				0.9830	-0.59
-5	3.307		336523				0.0493	-0.33
-4	3.29		342995				-0.252	1.6
-3	3.286		342995				0.692	-0.539
-2	3.229		342995				-1.49	-0.00121
-1	3.258		342995				-1.61	-0.969
0	3.299	3.437461	339255	336523	3.436	0.42354	1.01	0.7
1	3.382	3.376063	339255	336523	3.368	0.42887	-0.211	-0.520
2	3.391	3.492468	339255	336523	3.307	0.43966	1.002	0.824
3	3.389	3.272853	339255	342995	3.29	0.44083	1.73	1.34
4	3.439	3.593903	339255	342995	3.286	0.27112	-0.575	-0.576
5	3.482	3.374796	303455	342995	3.229	0.27512	0.194	-0.224
6	3.523	3.620275	303455	342995	3.258	0.27856	0.243	-0.0075
7	3.591	3.850986	303455	339255	3.299	0.28184	-0.25	-0.324
8	3.721	3.938154	303455	339255	3.382	0.28728	-0.835	-0.91
9	3.793	3.962052	328923	339255	3.391	0.48373	-1.18	-1.194
10	3.829	3.934435	328923	339255	3.389	0.49309	-0.5445	-0.586
11	3.867	3.905807	328923	339255	3.439	0.49777	-0.594	-0.701
12	3.918	3.828595	328923	303455	3.482	0.50271	-1.01	-0.929
13	3.941	3.851225	319005	303455	3.523	0.50934	0.361	0.343
14	3.939	3.650034	319005	303455	3.591	0.51233	-0.888	-1.12

N	ACTUAL PRICE G(N)	PRE-DICTED PRICE P(N)	S(N)	S(N-1)	G(N-7)	C	FIRST ϕ	SECOND ϕ
15	3.922	3.69583	319005	303455	3.721	0.51207	0.11	-0.188
16	3.87	3.818713	319005	328923	3.793	0.50986	0.402	-0.058
17	3.83	3.686425	319005	328923	3.829	0.5031	0.451	-0.136
18	3.79	3.74626	345098	328923	3.867	0.4979	-0.553	-0.91
19	3.754	3.841885	345098	328923	3.918	0.5527	0.311	-0.054
20	3.715	3.808671	345098	319005	3.941	0.48802	0.853	0.574
21	3.67	3.861379	345098	319005	3.939	0.48295	0.635	0.403
22	3.613	3.867182	341782	319005	3.922	0.4771	0.924	0.688
23	3.572	3.726372	341782	319005	3.87	0.46969	1.141	0.841
24	3.533	3.678498	341782	319005	3.83	0.46436	0.541	0.239
25	3.437	3.519015	341782	345098	3.79	0.45929	0.63	0.0293
26	3.356	3.484144	341782	345098	3.754	0.44681	0.88	0.29
27	3.411	3.517011	333184	345098	3.715	0.43628	0.825	0.231
28	3.427	3.612708	333184	345098	3.67	0.44343	0.642	0.153
29	3.494	3.689392	333184	341782	3.613	0.44551	0.036	-0.323
30	3.508	3.697802	333184	341782	3.572	0.45422	0.68	0.353
31	3.645	3.792897	333184	341782	3.533	0.45604	0.492	0.269
32	3.721	3.865744	337840	341782	3.437	0.47385	0.702	0.635
33	3.744	3.822605	337840	341782	3.356	0.48373	-0.646	-0.691
34	3.776	3.841569	337840	333184	3.411	0.48672	0.096	0.0891
35	3.843	3.984974	337840	333184	3.427	0.49088	-1.4	-1.31
36	3.847	3.810506	337840	333184	3.494	0.49959	-0.0453	-0.152
37	3.878			333184	3.508	0.50011	0.312	1.04
38	3.826			333184	3.645	0.50414	0.388	0.499
39	3.804			337840	3.721	0.49738	-1.97	-2.12
40	3.85			337840	3.744	0.49452	-1.11	0.0644

N	ACTUAL PRICE G(N)	PRE-DICTED PRICE P(N)	S(N)	S(N-1)	G(N-7)	C	FIRST ϕ	SECOND ϕ
41	3.819			337840	3.776	0.5005	1.02	-1.23
42	3.687			337840	3.843	0.49647	1.92	-1.74
43	3.568			337840	3.847	0.47931	-1.2	-0.0982
44	3.492				3.878	0.46384	-0.307	-0.758
45	3.449				3.826	0.45396	-.0.701	0.314

Here is a a graph demonstrating our predicted price and the actual price:



General Pricing Model Concluding Remark:

Based on the above construction of the model and the resulting graph above, we can say that our model is relatively accurate and significantly better than throwing a dart to predict prices. However, we shall analyze our data statistically to see how well our model did:

	ACTUAL PRICES	PREDICTED PRICES
MEAN	3.658	3.715
STANDARD DEVIATION	0.19	0.18

Based on this, our mean predicted price was slightly higher ($.057) than the actual price and the standard deviation is very close. Thus, we deem our model an accurate fit to the actual retail gasoline prices.

Lastly, we did a general error analysis. Our average percentage of error is 3.62% which is a reasonable space for error.

PART III

The Decision Model

The Decision Model tells the consumer what to do each week. It takes the predicted price of gasoline for the a month in the future and decides using a binary tree. The binary tree here branches off everytime the consumer is allowed a decision between 1) filling up full and half or 2)filling half or no gas. The problem gives two cases to consider: One where the consumer drives 100 miles per week and one where the consumer drives 200 miles per week.

As a reminder, the gas tank holds 16 gallons and the average mileage is 25 miles/gallon, therefore each full tank allows the consumer to drive 400 miles without having to refill.

Case 1:

In case 1, the consumer drives 100 miles per week. We start at Week (0), where the consumer has a choice of filling a full tank or a half tank. It then branches off into the two decisions. After one week, the tank could be 1/4 or 3/4 full depending on the original choice. By Assumption 4, the consumer would not make a decision and so continues on that week. By similar deduction, decisions are made on Week 0, Week (2), Week (4),...,Week (2n). The model below shows the branches that result from this, where f indicates filling the tank to full, h indicates adding half, 0 indicates adding no gas. w_0 , w_2 ,and w_4 represent Week (0), Week (2), and Week (4) respectively. p_0 , p_1 ,and p_2 represent the price of gasoline at the start of w_{2n}

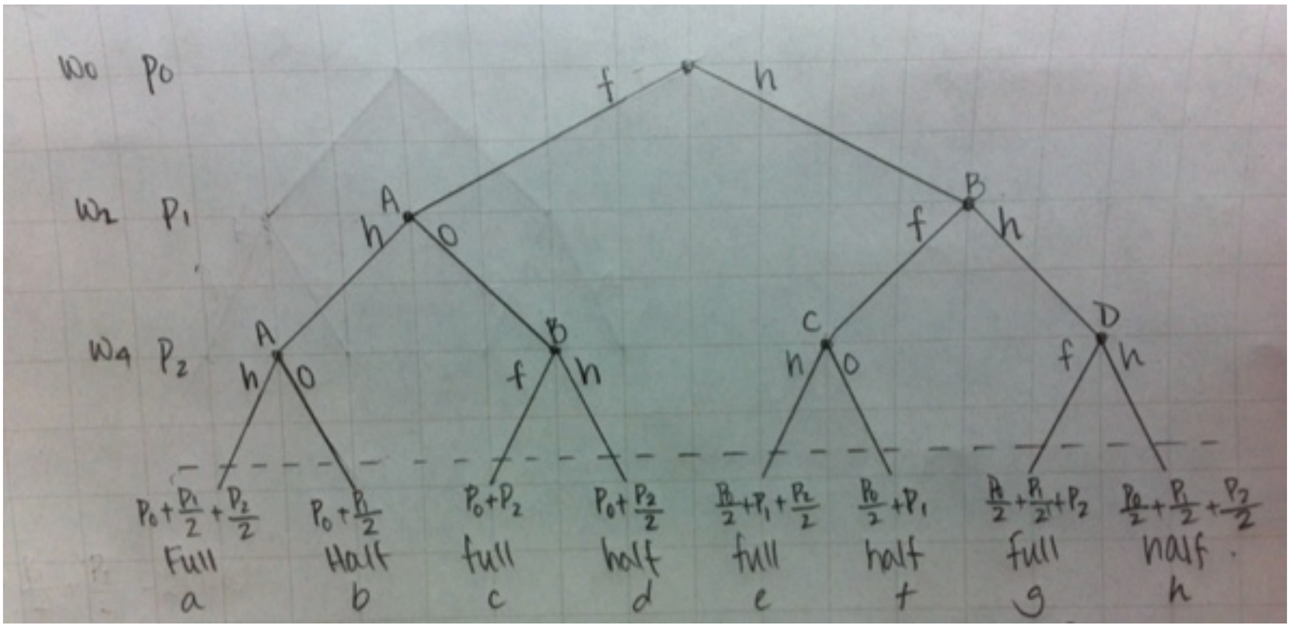


Diagram 3.1

This is a binary tree representation of what happens over the course of a month.

At the end of the month, we determine the state of the consumer after the 3rd time he makes a decision and before he drives any more miles after. Now, we take the sum of all the prices he paid. For example, if he paid for a full tank the 1st time, a half tank the 2nd time, and half tank the 3rd time, he would have paid the full price of the 1st time + 1/2 price of the 2nd time + 1/2 price of the 3rd and would have ended up with a full tank.

After 3 chances to fill up, there is obviously 2^3 or 8 different possibilities labelled a,b,c,d,e,f,g, and h. These are called outcomes.

It is in the best interest for the consumer to have paid the least amount of money per mile driven. So, after filling up the 3rd time at w_4 , the consumer would have driven 400 miles (4 weeks * 100 miles per week). We then need to consider the amount of gas left in the tank. This measures how much more the car could potentially travel. If there is a full tank left, we add the 400 miles to the total miles driven. If there is half a tank left, we add only 200 miles. This newly added value is called the potential amount and is determined by the state of the tank at each outcome.

Let:

$T_{outcome}$ denote the total miles traveled (400) + the additional potential amount for an outcome

$P_{outcome}$ denote the total amount paid after the 3rd chance for an outcome

$\frac{P_{outcome}}{T_{outcome}}$ denote the per mile amount paid for an outcome

Here is a table that shows what happens at the end of each outcome:

OUTCOME	$T_{outcome}$	$P_{outcome}$	$\frac{P_{outcome}}{T_{outcome}}$
a	800	$P_0 + \frac{P_1}{2} + \frac{P_2}{2}$	$(P_0 + \frac{P_1}{2} + \frac{P_2}{2}) / 800$
b	600	$P_0 + \frac{P_1}{2}$	$(P_0 + \frac{P_1}{2}) / 600$
c	800	$P_0 + P_2$	$(P_0 + P_2) / 800$
d	600	$P_0 + \frac{P_2}{2}$	$(P_0 + \frac{P_2}{2}) / 600$
e	800	$\frac{P_0}{2} + P_1 + \frac{P_2}{2}$	$(\frac{P_0}{2} + P_1 + \frac{P_2}{2}) / 800$
f	600	$\frac{P_0}{2} + P_1$	$(\frac{P_0}{2} + P_1) / 600$
g	800	$\frac{P_0}{2} + \frac{P_1}{2} + P_2$	$(\frac{P_0}{2} + \frac{P_1}{2} + P_2) / 800$
h	600	$\frac{P_0}{2} + \frac{P_1}{2} + \frac{P_2}{2}$	$(\frac{P_0}{2} + \frac{P_1}{2} + \frac{P_2}{2}) / 600$

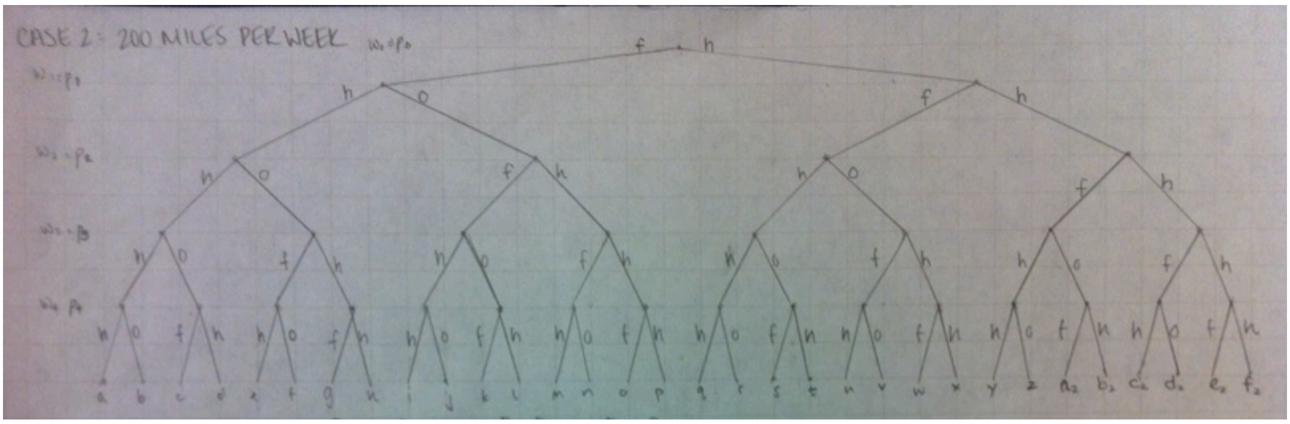
The consumer would utilize this model to make decisions on whether to buy a full tank or half tank or no gas. He or she is also assumed to be price-intelligent and will attempt to minimize the amount of money spent per mile driven (Assumption 6).

Knowing the retail price of gasoline today along with many other factors, we can use the General Pricing Model to approximate p_0 , p_1 ,and p_2 . Knowing the predicted values of each price, we can solve for $\frac{P_{outcome}}{T_{outcome}}$ in each outcome. The lowest $\frac{P_{outcome}}{T_{outcome}}$ value is the outcome that is preferred and the path that it represents in the binary tree is the set of decisions that the consumer should use.

Post-Creation Analysis:

Case 2:

In Case 2, the consumer drives 200 miles per week. According to assumption 4 and assumption 3, this means that at the beginning of every week, the driver has a choice of how much gasoline to fill in his tank. Again, we create a binary model similar to that of case 1. The difference is that since the driver drives more per week, he is given the chance to refill more often and thus, in the course of one month, he has a greater number of choices and thus a larger binary tree:



Like the one before, at every intersection, the consumer has a choice; h denotes filling half a tank, f denotes filling full, and 0 denotes no gasoline. At the end, the outcomes are labelled a-z, and a2-f2. Again, just like case 1, the tank will still have a certain amount of gas, and we know that a full tank is equivalent

The variables are the same:

Let

$T_{outcome}$ denote the total miles traveled (400) + the additional potential amount for an outcome

$P_{outcome}$ denote the total amount paid after the 3rd chance for an outcome

$\frac{P_{outcome}}{T_{outcome}}$ denotes the per mile amount paid for an outcome

OUTCOME	$T_{outcome}$	$P_{outcome}$	$\frac{P_{outcome}}{T_{outcome}}$
a	800	$P_0+\frac{P_1}{2}+\frac{P_2}{2}+\frac{P_3}{2}+\frac{P_4}{2}$	$(P_0+\frac{P_1}{2}+\frac{P_2}{2}+\frac{P_3}{2}+\frac{P_4}{2})/800$
b	600	$P_0+\frac{P_1}{2}+\frac{P_2}{2}+\frac{P_3}{2}$	$(P_0+\frac{P_1}{2}+\frac{P_2}{2}+\frac{P_3}{2})/600$
c	800	$P_0+\frac{P_1}{2}+\frac{P_2}{2}+P_4$	$(P_0+\frac{P_1}{2}+\frac{P_2}{2}+P_4)/800$
d	600	$P_0+\frac{P_1}{2}+\frac{P_2}{2}+\frac{P_4}{2}$	$(P_0+\frac{P_1}{2}+\frac{P_2}{2}+\frac{P_4}{2})/600$
e	800	$P_0+\frac{P_1}{2}+\frac{P_2}{2}+P_3+\frac{P_4}{2}$	$(P_0+\frac{P_1}{2}+\frac{P_2}{2}+P_3+\frac{P_4}{2})/800$
f	600	$P_0+\frac{P_1}{2}+P_3$	$(P_0+\frac{P_1}{2}+P_3)/600$
g	800	$P_0+\frac{P_1}{2}+\frac{P_3}{2}+P_4$	$(P_0+\frac{P_1}{2}+\frac{P_3}{2}+P_4)/800$
h	600	$P_0+\frac{P_1}{2}+\frac{P_3}{2}+\frac{P_4}{2}$	$(P_0+\frac{P_1}{2}+\frac{P_3}{2}+\frac{P_4}{2})/600$

OUTCOME	$T_{outcome}$	$P_{outcome}$	$\frac{P_{outcome}}{T_{outcome}}$
i	800	$P_0+P_2+\frac{P_3}{2}+\frac{P_4}{2}$	$(P_0+P_2+\frac{P_3}{2}+\frac{P_4}{2})/800$
j	600	$P_0+P_2+\frac{P_3}{2}$	$(P_0+P_2+\frac{P_3}{2})/600$
k	800	$P_0+P_2+P_4$	$(P_0+P_2+P_4)/800$
l	600	$P_0+P_2+\frac{P_4}{2}$	$(P_0+P_2+\frac{P_4}{2})/600$
m	800	$P_0+\frac{P_2}{2}+P_3+\frac{P_4}{2}$	$(P_0+P_2+\frac{P_4}{2})/800$
n	600	$P_0+\frac{P_2}{2}+P_3$	$(P_0+\frac{P_2}{2}+P_3)/600$
o	800	$P_0+\frac{P_2}{2}+\frac{P_3}{2}+P_4$	$(P_0+\frac{P_2}{2}+\frac{P_3}{2}+P_4)/800$
p	600	$P_0+\frac{P_2}{2}+\frac{P_3}{2}+\frac{P_4}{2}$	$(P_0+\frac{P_2}{2}+\frac{P_3}{2}+\frac{P_4}{2})/600$
q	800	$\frac{P_0}{2}+P_1+\frac{P_2}{2}+\frac{P_3}{2}+\frac{P_4}{2}$	$(\frac{P_0}{2}+P_1+\frac{P_2}{2}+\frac{P_3}{2}+\frac{P_4}{2})/800$
r	600	$\frac{P_0}{2}+P_1+\frac{P_2}{2}+\frac{P_3}{2}$	$(\frac{P_0}{2}+P_1+\frac{P_2}{2}+\frac{P_3}{2})/600$
s	800	$\frac{P_0}{2}+P_1+\frac{P_2}{2}+P_4$	$(\frac{P_0}{2}+P_1+\frac{P_2}{2}+P_4)/800$
t	800	$\frac{P_0}{2}+P_1+\frac{P_2}{2}+\frac{P_4}{2}$	$(\frac{P_0}{2}+P_1+\frac{P_2}{2}+\frac{P_4}{2})/600$
u	600	$\frac{P_0}{2}+P_1+P_3+P_4$	$(\frac{P_0}{2}+P_1+P_3+P_4)/800$
v	800	$\frac{P_0}{2}+P_1+P_3$	$(\frac{P_0}{2}+P_1+P_3)/600$
w	600	$\frac{P_0}{2}+P_1+\frac{P_3}{2}+P_4$	$(\frac{P_0}{2}+P_1+\frac{P_3}{2}+P_4)/800$
x	800	$\frac{P_0}{2}+P_1+\frac{P_3}{2}+\frac{P_4}{2}$	$(\frac{P_0}{2}+P_1+\frac{P_3}{2}+\frac{P_4}{2})/600$
y	600	$\frac{P_0}{2}+\frac{P_1}{2}+P_2+\frac{P_3}{2}+\frac{P_4}{2}$	$(\frac{P_0}{2}+\frac{P_1}{2}+P_2+\frac{P_3}{2}+\frac{P_4}{2})/800$

OUTCOME	$T_{outcome}$	$P_{outcome}$	$\frac{P_{outcome}}{T_{outcome}}$
z	800	$\frac{P_0}{2}+\frac{P_1}{2}+P_2+\frac{P_3}{2}$	$(\frac{P_0}{2}+\frac{P_1}{2}+P_2+\frac{P_3}{2})/600$
a2	600	$\frac{P_0}{2}+\frac{P_1}{2}+P_2+P_4$	$(\frac{P_0}{2}+\frac{P_1}{2}+P_2+P_4)/800$
b2	800	$\frac{P_0}{2}+\frac{P_1}{2}+P_2+\frac{P_4}{2}$	$(\frac{P_0}{2}+\frac{P_1}{2}+P_2+\frac{P_4}{2})/600$
c2	600	$\frac{P_0}{2}+\frac{P_1}{2}+\frac{P_2}{2}+P_3+\frac{P_4}{2}$	$(\frac{P_0}{2}+\frac{P_1}{2}+\frac{P_2}{2}+P_3+\frac{P_4}{2})/800$
d2	800	$\frac{P_0}{2}+\frac{P_1}{2}+\frac{P_2}{2}+P_3$	$(\frac{P_0}{2}+\frac{P_1}{2}+\frac{P_2}{2}+P_3)/600$
e2	600	$\frac{P_0}{2}+\frac{P_1}{2}+\frac{P_2}{2}+\frac{P_3}{2}+P_4$	$(\frac{P_0}{2}+\frac{P_1}{2}+\frac{P_2}{2}+\frac{P_3}{2}+P_4)/800$
f2	800	$\frac{P_0}{2}+\frac{P_1}{2}+\frac{P_2}{2}+\frac{P_3}{2}+\frac{P_4}{2}$	$(\frac{P_0}{2}+\frac{P_1}{2}+\frac{P_2}{2}+\frac{P_3}{2}+\frac{P_4}{2})/600$

This table shows the outcomes and their corresponding cost and cost per miles driven. The consumer will want to use this model along with the General Pricing Model to minimize $\frac{P_{outcome}}{T_{outcome}}$.

How to use this:

These two binary trees are very simple to use. A consumer would know the current price of the retail gasoline price. He would also know all the factors of input such as the supply of crude oil, the weather around the world, and the knowledge of world political stability. Using this, we can then determine the price values at each week in the future. Knowing this he then uses the table corresponding to the case he is taking (100 miles per week or 200 miles per week), and look at the last column, which tells the amount paid per miles driven and potential miles driven. The lowest of these values would be the optimal (for obvious reasons) and thus that outcome would dictate the decisions the consumer would use.

Is there an upper bound to mileage driven that changes the decision for buying weekly gasoline: Realistically, the answer should be no, no matter how much a driver drives per week, he would always want to obtain the least amount paid per miles driven and thus the binary tree model could be used with no upper bound to mileage driven.

However, the problem with this is that the data given to us is only in weekly (and monthly but it is then effectively converted to weeks). This means that we cannot tell the prices changes within a week. Using Assumption 5 (Which was assumed to create a slightly simpler model but a lot more effective), the driver can only have half a tank full or a full tank. By a corollary of Assumption 1 (Given), each tank has the potential to drive 400 miles. Therefore, the greatest amount that a driver could go a week is 200 miles. The next greatest that we can use our model for is 400 miles which would give the consumer no choice and thus not a good model. Anything above 400 miles is not possible because the maximum potential miles to be driven is only 400 and by Assumption 2, the driver can only fill up once a week and would thus be unable to drive his full attempted mileage. It appears that we have given up flexibility for accuracy in model, we believe to be an advantageous trade.

Post Binary Tree Model Analysis:

Using the General Pricing Model, we are able to generate future weekly prices. Using these prices, we use the Binary Tree Model table to find the optimal outcome and the consumer would follow that path. Here is an analysis on how well the consumer would do if he or she used our model. Based on the test of our model on 2012 retail gasoline prices, we found out that we accurately predicted the direction of the gasoline price movement 23 out of 37 times or 62.2% of the time.

We also know that two possible cases: Success in predicting the direction or failure.

We know that each prediction is independent since the prediction uses only present price values and that is constantly changing.

Lastly, for each Case (100 miles per week and 200 miles per week), we have a set number of decision making times.

These four conditions means that we have a binomial setting. This means the sample space X, which determines how well a consumer would do if he used our model, has a binomial distribution with n trials and probability of 62.2% of being correct each time.

We separate for each case:

Case 1:

There are two trials since we have to predict the change between week-0 to week 1 and week-1 to week-2. For any binomial setting, we have:

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

which means the probability of our model getting both predictions correct (and thus arriving at the most optimal decisions is 38.64%, 49.04% of getting a decent result, and only 14.32% of getting a unsatisfactory result.

Case 2:

Due to the complexity of case two, we only have a .8% chance of getting the worst result, and only 6% of getting four wrong predictions and thus a not so good result. If we define a reasonable and well modelled result as that predicting the majority of the price changes correctly, then our model achieves this more than 35% of the time. Hence, we the accuracy of our model gives it a powerful advantage to an average consumer trying to guess the prices with nothing but his or her gut feelings and a newspaper.

PART IV

The Boston Model

In order to apply the model, we have to find the factors that reside in Boston that could affect the gas price of a gallon of gasoline such as the monthly temperature. This will determine the C value in the model. The gasoline prices of each month of 2011 was found be averaging 5 days of month in price of dollars per gallon. The gasoline prices was graphed against the respective temperature of the month. From the graph below, we can conclude as the temperature increases, there is an increase in demand with a steady supply, leading to a increasing price of the average oil ($/g) in Boston in a linear progression model or trend.

The price is supposedly affected by adjustment of tax quarterly in the state of Massachusetts. However, the current tax of gasoline (0.21 dollars per gallon) and addition Federal tax/UST/Inspection fund tax stayed constant. Shipment similarly has no affect on the price as the gas price fluctuation is negligible in the current and previous year.

First, we must quantity the weather and stability effects as per the categories in component 3 as described above. N denotes the weeks after January 3 2012.

N	**ε** (Weather Effect)	WEATHER CATEGORY	**η** (Stability Effect	STABILITY CATEGORY	$C=\varepsilon+\eta$	G(N-1)
0	0.0622	1	0.0933	2	0.1555	3.11
1	0.06276	1	0.09414	2	0.1569	3.138
2	0.06328	1	0.09492	2	0.1582	3.164
3	0.06338	1	0.09507	2	0.15845	3.169

N	ε (Weather Effect)	WEATHER CATEGORY	η (Stability Effect	STABILITY CATEGORY	$C = \varepsilon + \eta$	G(N-1)
4	0.0636	1	-0.0636	1	0	3.18
5	0.06362	1	-0.06362	1	0	3.181
6	0.06382	1	-0.06382	1	0	3.191
7	0.06392	1	-0.06392	1	0	3.196
8	0.06414	1	-0.06414	1	0	3.207
9	0.0671	1	0.10065	3	0.16775	3.355
10	0.07038	1	0.10557	3	0.17595	3.519
11	0.0708	1	0.1062	3	0.177	3.54
12	0.0708	1	0.1062	3	0.177	3.54
13	0.07094	1	0.10641	2	0.17735	3.547
14	0.07244	1	0.10866	2	0.1811	3.622
15	0.07438	1	0.11157	2	0.18595	3.719
16	0.07652	1	0.11478	2	0.1913	3.826
17	0.07808	1	0.11712	2	0.1952	3.904
18	0.07988	1	0.11982	2	0.1997	3.994
19	0.08032	1	0.12048	2	0.2008	4.016
20	0.08054	1	0.12081	2	0.20135	4.027
21	0.07942	1	0.11913	2	0.19855	3.971
22	0.07806	1	0.11709	2	0.19515	3.903
23	0.07718	1	0.11577	2	0.19295	3.859
24	0.0761	1	0.11415	2	0.19025	3.805
25	0.0757	1	0.11355	2	0.18925	3.785
26	0.07536	1	0.11304	2	0.1884	3.768
27	0.07466	1	0.11199	2	0.18665	3.733
28	0.0754	1	0.1131	2	0.1885	3.77
29	0.07638	1	0.11457	2	0.19095	3.819

N	ε (Weather Effect)	WEATHER CATEGORY	η (Stability Effect	STABILITY CATEGORY	$C = \varepsilon + \eta$	G(N-1)
30	0.07658	1	0.11487	2	0.19145	3.829
31	0.07652	1	0.11478	2	0.1913	3.826
32	0.0763	1	0.11445	2	0.19075	3.815
33	0.07534	1	0.11301	2	0.18835	3.767
34	0.07472	1	0.11208	2	0.1868	3.736
35	0.07418	1	0.11127	2	0.18545	3.709
36	0.07458	1	0.11187	2	0.18645	3.729
37	0.07454	1	0.11181	2	0.18635	3.727
38	0.07354	1	0.11031	2	0.18385	3.677
39	0.07232	1	0.10848	2	0.1808	3.616
40	0.07126	1	0.10689	2	0.17815	3.563
41	0.07002	1	0.10503	2	0.17505	3.501
42	0.0708	1	0.1062	2	0.177	3.54
43	0.07042	1	0.10563	2	0.17605	3.521
44	0.06996	1	0.10494	2	0.1749	3.498
45	0.06966	1	0.10449	2	0.17415	3.483
46	0.06972	1	0.10458	2	0.1743	3.486
	0.0692				0.0692	3.46
	0.06864				0.06864	3.432
	0.06796				0.06796	3.398
	0.06712				0.06712	3.356
	0.06656				0.06656	3.328
	0.06648				0.06648	3.324

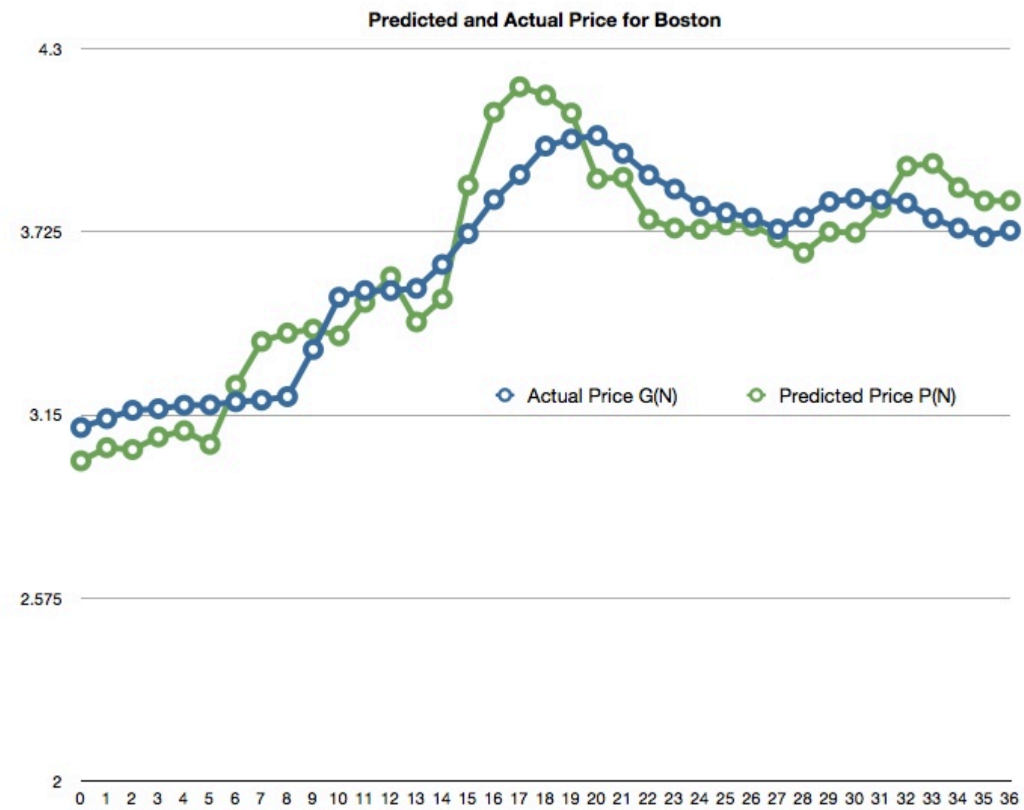
Now that we've determined the C values, we can determine our own price predictions for Boston and the actual price values given that β =.0000407

N	ACTUAL PRICE G(N)	PRE-DICTED PRICE P(N)	S(N)	S(N-1)	G(N-7)	C	FIRST ϕ	SECOND ϕ
-7	2.903		336523				-1.97	0.58
-6	2.966		336523				0.9830	-0.59
-5	3.012		336523				0.0493	-0.33
-4	3.006		342995				-0.252	1.6
-3	3.063		342995				0.692	-0.539
-2	3.068		342995				-1.49	-0.00121
-1	3.083		342995				-1.61	-0.969
0	3.11	3.005264	339255	336523	2.903	0.1555	0.72	0.7
1	3.138	3.046771	339255	336523	2.966	0.1569	-0.211	-0.249
2	3.164	3.040652	339255	336523	3.012	0.1582	0.92	0.84
3	3.169	3.080405	339255	342995	3.006	0.15845	1.02	0.924
4	3.18	3.099783	339255	342995	3.063	0	-0.575	-0.576
5	3.181	3.05696	303455	342995	3.068	0	0.101	-0.224
6	3.191	3.243027	303455	342995	3.083	0	0.183	-0.0075
7	3.196	3.380146	303455	339255	3.11	0	-0.25	-0.324
8	3.207	3.406874	303455	339255	3.138	0	-0.835	-0.91
9	3.355	3.419072	328923	339255	3.164	0.16775	-1.18	-1.194
10	3.519	3.397931	328923	339255	3.169	0.17595	-0.545	-0.586
11	3.54	3.502828	328923	339255	3.18	0.177	-0.594	-0.562
12	3.54	3.583447	328923	303455	3.181	0.177	-1.01	-0.629
13	3.547	3.44161	319005	303455	3.191	0.17735	0.361	0.543
14	3.622	3.513954	319005	303455	3.196	0.1811	-1.252	-1.02
15	3.719	3.870666	319005	303455	3.207	0.18595	0.025	0.525

N	ACTUAL PRICE G(N)	PRE-DICTED PRICE P(N)	S(N)	S(N-1)	G(N-7)	C	FIRST ϕ	SECOND ϕ
16	3.826	4.100003	319005	328923	3.355	0.1913	0.402	0.758
17	3.904	4.180144	319005	328923	3.519	0.1952	0.451	0.738
18	3.994	4.153823	345098	328923	3.54	0.1997	-0.453	0.002
19	4.016	4.097688	345098	328923	3.54	0.2008	0.111	0.521
20	4.027	3.891126	345098	319005	3.547	0.20135	0.853	1.174
21	3.971	3.895916	345098	319005	3.622	0.19855	0.335	0.603
22	3.903	3.763794	341782	319005	3.719	0.19515	0.824	0.888
23	3.859	3.736319	341782	319005	3.826	0.19295	-0.141	-0.181
24	3.805	3.732763	341782	319005	3.904	0.19025	0.541	0.439
25	3.785	3.746397	341782	345098	3.994	0.18925	1.3	0.93
26	3.768	3.743381	341782	345098	4.016	0.1884	1.008	0.619
27	3.733	3.706569	333184	345098	4.027	0.18665	0.825	0.331
28	3.77	3.658778	333184	345098	3.971	0.1885	0.642	0.153
29	3.819	3.724832	333184	341782	3.903	0.19095	0.036	-0.323
30	3.829	3.722032	333184	341782	3.859	0.19145	0.68	0.353
31	3.826	3.800157	333184	341782	3.805	0.1913	0.492	0.269
32	3.815	3.930644	337840	341782	3.785	0.19075	0.702	0.635
33	3.767	3.939225	337840	341782	3.768	0.18835	-0.646	-0.691
34	3.736	3.863649	337840	333184	3.733	0.1868	0.096	0.0891
35	3.709	3.821588	337840	333184	3.77	0.18545	-0.184	-0.252
36	3.729	3.822366	337840	333184	3.819	0.18645	-0.0453	-0.152
37	3.727			333184	3.829	0.18635	0.312	1.04
38	3.677			333184	3.826	0.18385	0.388	0.499
39	3.616			337840	3.815	0.1808	-1.97	-2.12
40	3.563			337840	3.767	0.17815	-1.11	0.0644
41	3.501			337840	3.736	0.17505	1.02	-1.23

N	ACTUAL PRICE G(N)	PRE-DICTED PRICE P(N)	S(N)	S(N-1)	G(N-7)	C	FIRST ϕ	SECOND ϕ
42	3.54			337840	3.709	0.177	1.92	-1.74
43	3.521			337840	3.729	0.17605	-1.2	-0.0982
44	3.498				3.727	0.1749	-0.307	-0.758
45	3.483				3.677	0.17415	-.0.701	0.314

Here is a a graph demonstrating our predicted price and the actual price:



Boston Pricing Model Concluding Remark:

Based on the above construction of the model and the resulting graph above, we can say that our model is relatively accurate. Based on our model, the mean is predicted to be $3.625, which is extremely close to that of the actual price ($3.611). We calculated a 3.23% error.

PART V

Letter to Local Newspaper

Dear Mark Clifford, Chief Editor of the South China Morning Post,

We are Team 3455 writing to you to present what can be developed to be a new article on gas prices.

Even an average teenager would be capable of telling you that one of the most important commodities in our industrialized society is fuel. As such, gasoline is without a doubt an extremely essential resource for almost the entire population of this community. From car owners to financial investors, many members of society are concerned with the rising gas prices that have been observed recently, and we are sure that many individuals would like to utilize a method of purchase that makes gasoline more affordable.

As car owners as well as consumers of fuel, it would be extremely helpful to be able to predict the fuel cost of the following week. Using this prediction, one would be able to judge how much money to spend on fuel each week. If gas prices were predicted to increase the following week, individuals should spend more on gasoline this week to reduce the amount of money spent on a more expensive gasoline the following week. Likewise, if gas prices were predicted to decrease the following week, individuals should reduce the number of times gasoline is purchased this week, and purchase more gas at a lower cost the following week.

The model used to predict future fuel costs involves mathematics, economic theory, and financial consideration. When building this model, we took into consideration what the price of gasoline is based upon. According to the Energy Information Administration, it is based on distribution, refining, taxes, and most prominently, the price of crude oil. As such, we looked further on crude oil and found that there is a response time, a period of time between a change in the supply and the change in the price of gas due to that change. Also, we have taken into account the momentum of the price of gasoline. Any changes to factors affecting the price of gasoline may be impacted by this momentum.

In order to use this model to predict the price of gas of the following week, one only needs to know the following:

- The price of gasoline seven weeks prior to today;
- The weather of the current location;
- The political stability of the crude oil's country of origin; and
- The supply of crude oil (in barrels)

This model has a mere 3.62% error, and we believe it to be relatively accurate and will be of great use to the citizens of the this community.

We hope you will consider publishing this in the next edition of the South China Morning Post as we feel that many readers would be interested in our research and work on the subject.

Sincerely,

Team 3455

BIBLIOGRAPHY

Calore, Paul. "So What Is The Profit-Margin For Refining Crude Oil Into Gasoline?" *Refinery Oil Prices*. What It Costs, n.d. Web. 17 Nov. 2012. <http://business.whatitcosts.com/>.

EIA. "Spot Prices." *Spot Prices for Crude Oil and Petroleum Products*. U.S. Energy Information Administration, 15 Nov. 2012. Web. 17 Nov. 2012. <http://www.eia.gov/>.

EIA. "U.S. Energy Information Administration - EIA - Independent Statistics and Analysis." *Energy & Financial Markets*. U.S. Energy Information Administration - EIA - Independent Statistics and Analysis, Oct. 2012. Web. 17 Nov. 2012. <http://www.eia.gov/>.

Flynn, Phil. "Oil Supply Risks Rise with Middle East Instability." *Oil Supply Risks Rise with Middle East Instability*. Future Mags, 23 Feb. 2012. Web. 18 Nov. 2012. <http://www.futuresmag.com/2012/02/23/oil-supply-risks-rise-with-middle-east-instability>.

Freudenrich, Craig, PhD. "How Oil Refining Works." *How Oil Refining Works*. HowStuffWorks, 4 Jan. 2001. Web. 17 Nov. 2012. <http://science.howstuffworks.com/>.

Goren, Arie. "Brent Crude Or WTI Crude Oil - Do We Know Which One Is A Better Investment?" *Brent Crude Or WTI Crude Oil - Do We Know Which One Is A Better Investment?* SeekingAlpha, 3 Aug. 2012. Web. 17 Nov. 2012. <http://seekingalpha.com/>.

Nerurkar, Neelesh, and Robert Pirog. "Rising Gasoline Prices 2012." *Rising Gasoline Prices 2012*. Congressional Research Service, 1 Mar. 2012. Web. 17 Nov. 2012. <http://www.fas.org/sgp/crs/misc/R42382.pdf>.

Poyzner, Mikhail. "Middle East, 2011 Timeline." *Middle East, 2011 Timeline*. World New Atlas, n.d. Web. 18 Nov. 2012. <http://www.mapreport.com/>.

"Retail Prices for Gasoline, All Grades." *Weekly Retail Gasoline and Diesel Prices*. Ed. EIA. U.S. Energy Information Administration, 13 Nov. 2012. Web. 17 Nov. 2012. <http://www.eia.gov/>.

Rothe, John. "Musings of a Money Manager." *Chart of the Day: Gas Prices Are Rising (Again)*. Musings of a Money Manager, 21 Sept. 2012. Web. 18 Nov. 2012. <http://www.johnrothe.com/2012/09/chart-of-the-day-gas-prices-are-rising-again/>.

Wagner, Hans. "What Determines Gas Prices?" *What Determines Gas Prices?* Investopedia, 15 Feb. 2011. Web. 17 Nov. 2012. <http://www.investopedia.com/>.

"Weather Statistics for Athens, Attica (Greece)." *Weather Statistics for Athens, Attica (Greece) â Yr.no*. Ed. Tom Erik Støwer. YrNo, Dec. 2011. Web. 18 Nov. 2012. <http://www.yr.no/place/Greece/Attica/Athens/statistics.html>.