Why You Pay What You Pay at the Pump

Jason Orne

Abstract:

This paper investigates several factors that strongly influence the day to day consumer price of Gasoline in America. This study incorporates existing information on the most likely variables that should influence gasoline prices from the beginning of 2006 to the very end of 2009. The study looks at contributing factors that include the average price of gasoline from the week before (pennies per gallon), the current weeks average oil price (Dollars per barrel), the previous weeks average oil price (Dollars per barrel), and the US production, and consumption of gasoline (1000 barrels per day). Using macro-level data from the US Department of Energy, the results suggest that factors like past gas price, oil price, and gasoline usage had positive relationships with gas price, while the previous weeks oil price and gasoline output had negative ones.

JEL Classification: D12

Keywords: Gas prices, oil, gas consumption, gas usage

1.0 Introduction

The focus of this study is consumer gasoline prices in the United States. This study hopes to find the effects of several independent economic variables on the weekly average of gasoline prices throughout the United States. Gasoline prices are important to Americans because of our extreme reliance on the fossil fuel.

We use gasoline and other oil based products as a key part of our everyday lives. We rely on automobiles as our main and almost sole mode of transportation. Most other goods also undergo an increase in prices as shipping costs go up. With increased stability in oil prices its other uses will benefit as well. After all, things like plastics, medicines, compact disks, detergents, and lubricants are all made with some sort of oil byproduct.

This dependence makes price stability very important to us. Of course there are some supply shocks that will not be expected, but a formula to help foresee small fluctuations in gas price could be very helpful to determine exactly when it is the best time to buy.

Gasoline prices, like everything, goes up over time from inflation, but this tends to be a slow process and there is a lot of fluctuation. Of course, the inflation increasing is offset by the same increase in wages. Americans spend a large amount of money on gasoline, and they would like to get their money's worth and buy when prices are fluctuating down rather than up. By analyzing gas price against several easily determinable factors that can also be planned and projected will give us a better idea about next week's gasoline prices, which can influence our behavior today. This is why this study could prove useful as a practical tool or even just as a theoretical model to help influence future models.

The study's mission is to use national data collected by the Department of Energy to further test what factors have an effect (positive or negative) on the consumer price of gasoline. Gasoline prices are almost as important to us as food prices, due to our quantity and frequency of purchase of it. Gas prices are correlated with all of the factors listed above and this study sets out to find out by how much.

This study aims to gain perspective on the U.S. consumer price of gasoline and the impact of the variables used in this paper. This is important because gasoline prices are a key component in the economy's well being, as high gas prices lead to high prices in other goods as shipping becomes more expensive.

2.0 Literature review

There have been several regression studies done on gas prices in the past with a large range of variables and approaches. Adilov and Samavati (2009) studied the weekly change in gas prices against change in oil prices over several previous weeks. A 1% increase in the price of crude oil increases the price of gasoline by .196% this week and .288% next week and .348% and .438% in the weeks that follow. Their findings included things like in Texas, a 1% increase in the price of crude oil will cause a .603% increase in the price of gasoline three weeks later, and a 1% drop in the price of crude oil results in a decrease in gas prices by .481% by the third week after the crude oil price change.

Liddle (2009) took a very different approach to the determinants of gas price fluctuations. Specifically, he used the natural log of gasoline price as the dependant variable instead of the change in gas or the actual price of gas. Also, he studied its relationship with a totally different list of independent variables including: the natural logs of GDP per capita, gas use per capita, miles driven per capita, and registered vehicles per capita. He gave another abstract relationship that helped me to decide to attempt a direct relationship with gas price as the dependant variable.

Asplund et al. (1997) conducted a similar study on the price responses to changes in the input prices and market pressures, with the only difference being that their study was conducted on the Swedish gasoline market. They created a Ss-model for daily data covering January 1980 to December 1996. They found that gasoline prices appeared to follow cost movements closely. The independent variables they used were wholesale price paid by shell for gasoline, taxes, the exchange rate to US dollars, consumer price index, average manufacturing wage, and the number of days since the last change in gas prices. They provided a more worldly perspective to the gas issue this study examines.

Several papers, like the one from Borenstein et al. (1997), have found asymmetric price differences between oil up shocks and down shocks and the time before their effects on gasoline price are felt. They used their own estimation methodology for weekly price data. They concluded that there was a price asymmetry in oil price shocks effects on gas price. They found that oil up shocks take effect faster than down shocks. This would mean that prices would stay higher longer than necessary.

Bachmeier and Griffin (2002) wrote a paper that uses an error correction model with daily spot gasoline and crude oil price data from 1985-1998. Bachmeier and Griffin (2002) used two major differences than previous studies. The first was the use of the standard Engle-Granger two-step estimation procedure in comparison to the nonstandard methodology like Borenstein et al. (1997) used. The other difference is that they used daily data instead of weekly data. They found that using the nonstandard methodology with daily instead of weekly data there was little evidence of price asymmetry, with no evidence of asymmetry in gasoline prices with their standard method.

Radchenko (2004) did a study on the effect of the volatility of oil prices on how asymmetric the response in gasoline prices is to oil price increases and decreases. He used both time series to measure asymmetry and several measures of oil price volatility. His main finding was that asymmetry in gasoline prices declines as oil price volatility increases. His asymmetry results seem to support the oligopolistic coordination theory over the standard search theory with or without Bayesian updating. Oligopolistic coordination states that as input costs rise, sellers increase prices to avoid a decrease in profit margins. However, when the input decreases the sellers might want to sustain high prices until competitors cut their prices. This implies that sellers can use market power and decrease prices slowly as their implicit collusion breaks down and completion continues.

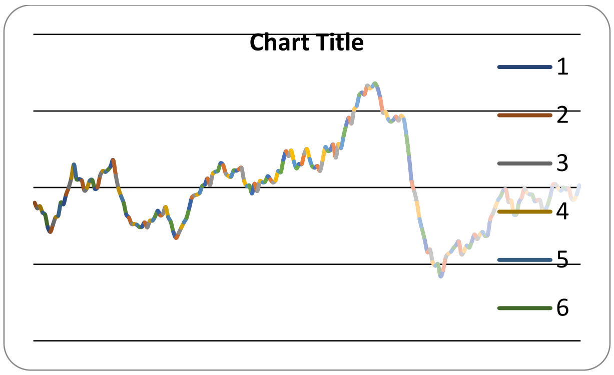
3.0 Trend

Below are graphs of the historical trends of each of the variables from 2006 to 2010. The first two variables: gas and oil both grow slowly with a great deal of fluctuation and a huge spike around June 2008. This works out well as it suggests that past gas, oil and past oil should all be strongly correlated with current weekly gas prices. Figures 1 and 2 show the weekly prices of gasoline and oil over the four year period in question. The last two figures show the historical levels of gasoline output and usage in the United States. Despite lots of fluctuation on both, output seems to grow over time and usage goes down over time with increases in technology. Over all, these figures show how each of the five variables chosen have behaved over the past four years.

Gas prices have various levels of correlation with the variables. Below are several scatter plots of each of the variables and their respective trend lines. The first three variables: past gas, oil, and past oil all have positive trend lines and the data points are all really close to the trend line. This means they are all strongly positively correlated with current weekly gas prices.

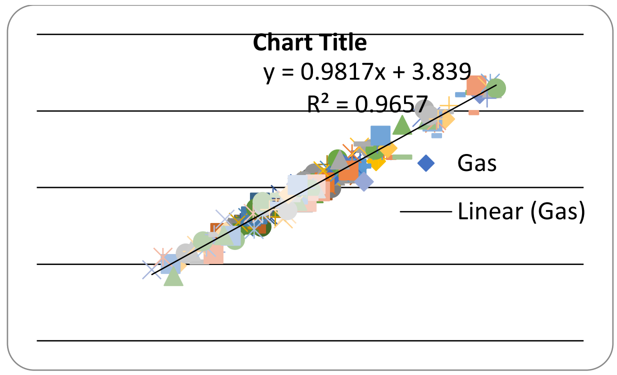
Figures 1, 2, and 3 show the three clearly more related variables out of the five. Figures 6 and 8 are scatter plots of gas price verses gasoline output and usage. Both of these have lower slopes, so they are less positively related. They are also further distributed from the trend line so they are slightly less reliable than the first three. Despite the fact that all of these variables are positively related to gas prices, some might act negatively in the end regression. Over all, these figures show how closely related the three variables are to current gas price.

Figure 1: Historic Gas Prices



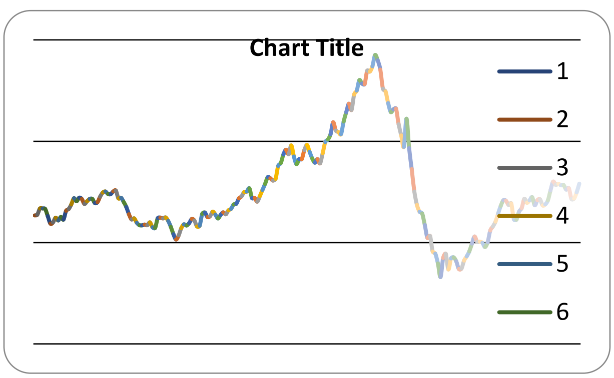
Author's Compilation

Figure 2: Graph of gas against past gas



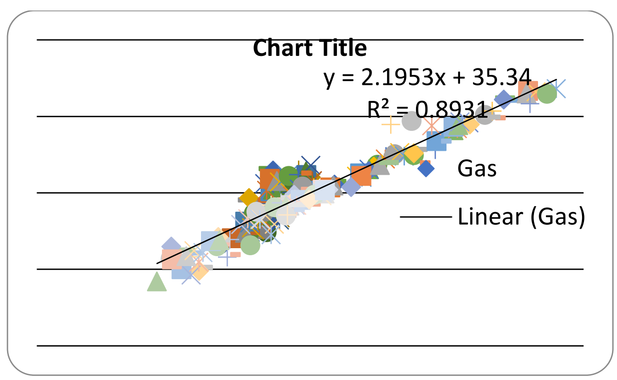
Authors Compilation

Figure 3: Historic Oil Prices



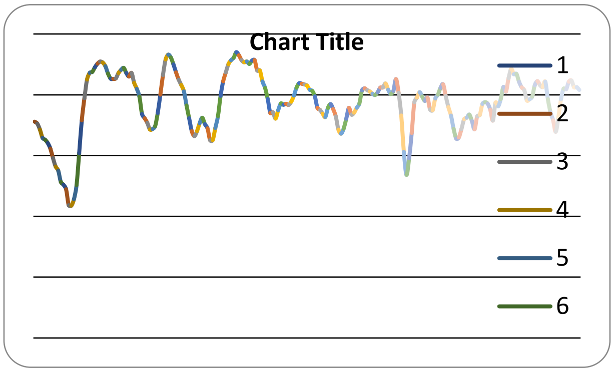
Authors Compilation

Figure 4: Graph of gas against oil



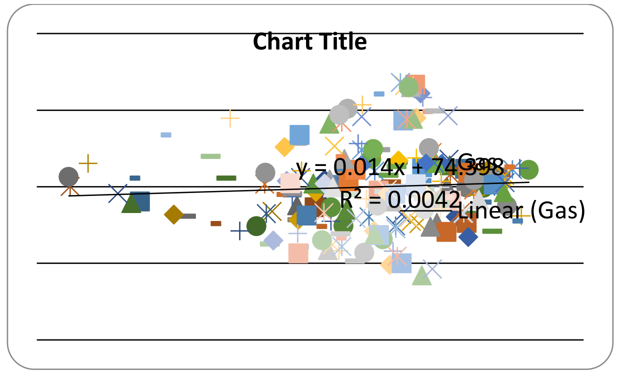
Authors Compilation

Figure 5: Historic Output



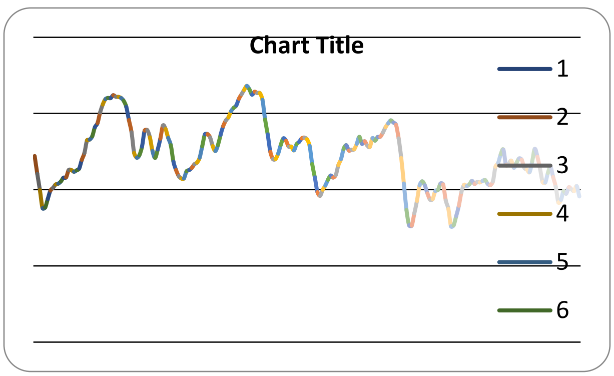
Authors Compilation

Figure 6: Graph of gas against output



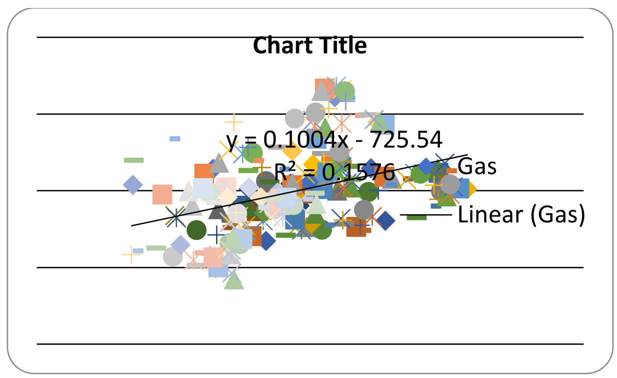
Authors Compilation

Figure 7: Historic Usage



Authors Compilation

Figure 8: Graph of gas against usage



Authors Compilation

4.0 Data and Methodology

This paper collected all of its data from the department of energy's web database. The data was limited to just the years 2006, 2007, 2008, and 2009. With the data gathered three linear regressions were conducted of gas price compared to the five variables chosen. Although all of the variables were used in each of the three regressions the difference came in the timeline covered by the regression. The first regression takes into account all of the data from January 1, 2006 to December 31, 2009 to examine the general trend. The second regression only uses data from the first year, 2006, to discover the relationship from before the crash. Finally, the third regression covers data from 2009 to 2010 to see if the regressed equation comes out significantly different during and after the financial crisis.

4.1 Definition of Variables:

$$GAS = \beta_1 \, PAST\_GAS + \beta_2 \, OIL + \beta_3 \, OUTPUT + \beta_4 \, USAGE + C$$

In this study there are five variables, the dependent and four independent variables. The dependant variable is GAS or gasoline price. In this case, gasoline price is the average price of gasoline over the United States over the course of a week, in terms of pennies per gallon. This allows it to be compared on a weekly basis to several other factors. The first independent variable is PAST_GAS which is the previous week's average gas price. It is found in the same manner as the dependant variable, and is also denoted in cents per gallon. This is included in the study to test whether trends continue to force prices up or down absent other change. The third variable, OIL, is the one week average of the oil prices across the United States, and is measured in dollars per barrel. Oil prices are important because they are a major input in the production of gasoline and costs of production are important. Next, PAST_OIL, is just the previous weeks average oil price, also in dollars per barrel. It is included for the same reasons as PAST_GAS, in addition to the expectations given by people's recognition of past prices. The variable OUTPUT gives the average daily US output of gasoline over the course of the week in terms of the number of thousands of barrels per day. This gives us a fair idea of supply of gasoline, which is a fair determinant of price. The final variable, USAGE, stands for the number of barrels, in thousands, of gasoline that US citizens use per day. This demand works in conjunction with the supply from the last variable to help determine price.

5.0 Empirical Results:

The data set had 209 weekly observations spread over four years on each of the five variables. After starting the regressions there seemed to be a strange coefficient in the PAST_OIL variable which was unnerving. The conclusion was that it was most likely multicollinearity. This study remedied the situation by dropping PAST_OIL from the regressions and came up with 3 new regressions. These seem to be much more accurate now that the multicollinearity is absent from the study. The basic results of the three linear regressions in terms of the coefficients of the variables, the constant, the R^2, and the adjusted R^2 are all in Table 1 below. As far as the significance of the three regressions, all of the R^2's are above .9 and the adjusted R^2's drop less than one hundredth, so they are all very good.

Table 1: Regression Output

Years	PAST_GAS	OIL	OUTPUT	USAGE	C	R2	adj R2
2006-2010	0.730431	0.552249	-0.014742	0.021907	-57.53318	0.973193	0.972667
2006	0.53267	1.950899	-0.018074	0.020365	-72.69223	0.90882	0.90106
2009	0.397277	1.339661	-0.014995	0.020618	-34.3894	0.97002	0.967469

This table reveals several things that were expected. First of all, the coefficient of each PAST_GAS are positive which makes sense as inflation is constant and as prices rise, they will rise faster and growth slows as shocks cause them to fall. OIL is also positive as the higher the input price the higher the product price. OUTPUT came out negative as excessive supply drives down prices. USAGE is positive in all three cases, which is fitting because increasing demand tends to push prices up. The constant happens to be negative in each case, but that is probably a coincidence because it is generated to help better fit the equation.

Table 2: Regression results from 2006 to 2010

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PAST_GAS	0.730431	0.039067	18.69706	0
OIL	0.552249	0.086098	6.414164	0
OUTPUT	-0.01474	0.003051	-4.83173	0
USAGE	0.021907	0.004075	5.375811	0
C	-57.5332	30.24425	-1.90229	0.0585
R^2	0.973193	Adj R^2	0.972667	

The values of these coefficients all have slightly different meanings. The meanings of the first regressions coefficients are as follows. For instance, the first coefficient, .7304, means that for every penny per gallon of last week's price there is a .7304 cent increase in current price. Similarly, for every dollar per barrel price of oil there is a .5523 cent per gallon increase in the price of gasoline. Although the seemingly less significant supply and demand variable have lower coefficients, the scale is much grander. For every 1000 barrels of gasoline produced per day there is a .01474 cent drop in price and for every 1000 barrels of gasoline used per day there is an increase in price by .021907 cents. These coefficients are of course deceiving because of the scale of the variables, in that millions of barrels are produced and used, while prices are only in hundreds, so the net effects are similar, making all of the variables significant.

Table 3: Regression of 2006 data

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PAST_GAS	0.53267	0.134765	3.952582	0.0003
OIL	1.950899	0.629639	3.098442	0.0033
OUTPUT	-0.01807	0.005373	-3.36363	0.0015
USAGE	0.020365	0.010032	2.029912	0.048
C	-72.6922	62.51831	-1.16274	0.2508
R^2	0.90882	Adj R^2	0.90106	

There is of course a better way to determine the significance of each variable in the three regressions. That is of course the t statistic and p value of each variable. In the first regression all variables are significant at 99% confidence level as shown in Table 2. The second is not quite as good as its variables have three in the 99% and one only in 95% confidence levels, shown in Table 3. The last is the worst only having the first two in the 99% confidence level and the supply and demand variable in the 95% confidence level which is still very good (Table4). Overall, the three regressions are all reliable and serve as a good basis of comparison against each other.

Table 4: Regression of 2009 data

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PAST_GAS	0.397277	0.077067	5.15493	0
OIL	1.339661	0.178786	7.493115	0
OUTPUT	-0.015	0.006941	-2.16037	0.0359
USAGE	0.020618	0.008179	2.520972	0.0152
C	-34.3894	74.56405	-0.46121	0.6468
R^2	0.97002	Adj R^2	0.967469	

There were not too many major differences between the models of gas price in 2006 and 2009. The coefficients of each of the variables share the same signs. The scale of the coefficients of PAST_GAS and OIL were a little off with 2006 being higher in each case. The coefficients of OUTPUT and USAGE were almost identical, with OUTPUT just a little higher in 2006. The only other difference is that OUTPUT and USAGE were highly significant in 2006 but had slightly lower t-scores in 2009, which seemed a little strange. This means that gas and oil prices held more sway in 2006 than in 2009, which makes sense as people were concerned with the mortgage crisis in 2009, and the havoc it reeked on the economy.

There are a few options for how to regulate gasoline prices and cut down on the price shocks. The most obvious is to create forward contracts on oil so that input prices are steadier, which will make the variable OIL constant. This will in turn lower the fluctuations of gas, and thus PAST_GAS which will lead to lowering fluctuations of gas prices again. Demand is too hard to control so it will hold its small swing over price. Supply is easier to control as you could create a production schedule to increase, decrease, or stay the same. These would be the easiest and most effective ways to control gasoline prices.

6.0 Conclusion:

Gasoline prices are highly important to us because of our huge reliance on fossil fuels. However gas prices seem hard to predict or control, with shocks and all, but this study shows regressions of gasoline prices over three timeframes. Unfortunately no attention was given to the differences in prices between states or even sections of the United States and their effects. This paper looks at gas prices for the greatest effected population in question, the whole United States. The results show that oil prices, past gas, output and usage are all significant factors with regards to gasoline price. The strongest effect goes to OIL which makes sense as it is the only input cost used. Then PAST_GAS seems to hold the next most sway over gasoline prices. OUTPUT and USAGE both had small scales which makes sense as they are so big and how long supply and demand take to change prices over an input cost change.

Bibliography

Adilov, N, & Samavati, H. (2008). Pump prices and oil prices: a tale of two directions. *International Atlantic Economic Society*, 37, 51-64

Asplund, M, Eriksson, R, & Friberg, R. (1997). Price adjustments by a gasoline retail chain. *Working Paper Series in Economics and Finance*, *194*

Bachmeier, L, & Griffin, J. (2002). New evidence on asymmetric gasoline price responses. *Texas A&M University, College Station*, *77843*

Borenstein, S, Cameron, C & Gilbert, R. (1997). Do Gasoline Prices Respond
Asymmetrically to Crude Oil Price Changes? *Quarterly Journal of Economics* 112 305-339.

Liddle, B. (2009). Long-run relationship among transport demand, income, and gasoline price for the US. Elsevier, 14, 73-82

Radchenko, S. (2004). Oil price volatility and the asymmetric response of. *University of North Carolina at Charlotte*

US Department of Energy. *http://www.eia.doe.gov*