An Economic Analysis of Transportation and Energy Issues

by

Rob Tischer

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment of the requirement for the degree of

Doctor of Philosophy

Department of Economics

2012
This thesis entitled:

An Economic Analysis of Transportation and Energy Issues

By Rob Tischer

has been approved for the Department of Economics

-----------------------------------------------

Professor Jonathan Hughes, Chair

-----------------------------------------------

Professor Nicholas Flores

Date: _________________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.
Chapter one examines the nature of responses to market price changes in one market among other energy commodity markets continues to be uncertain. This paper looks at these energy commodity markets (oil futures, oil spot, gasoline futures, gasoline spot and gasoline retail) and explores the responses of market prices in all of those markets to an exogenous price shock in any one of the other markets. A four step estimation process is used to examine impacts in market prices from price and/or volatility shocks in another market, and a supply shock from an hurricane is analyzed to trace price impacts on these five markets. Results indicate that oil futures prices and price volatilities provide the most significant impacts in all markets.

Chapter two presents a flexible model that allows for the tracking of the private and social costs associated with traffic congestion and road quality deterioration. The structure of the model is similar to the class of models within the existing literature that have been used to examine transportation costs. The focus, however, is to quantify the economic impacts from a variety of different government policies designed to alleviate traffic congestion and road quality deterioration. Despite the widespread recognition of the efficacy associated with the use of congestion tolls, the reality is that congestion toll policies are rarely adopted. Instead government policies are often adopted that seek to decrease traffic congestion through the construction of new roadways. Additionally, given budget constraints, governments often adopt policies that do not provide for road maintenance at levels necessary to prevent the road surface quality from declining. Results indicate that governments can pursue a policy of little road maintenance for 7 to 10 years before the economy is adversely affected.

Chapter 3 examines a model of energy prices, possible affects of temporal aggregation associated with series that exhibit conditional heteroskedasticity. A seven equation recursive
vector error correction model is used as a vehicle for analyzing and forecasting retail gasoline prices using daily, weekly, and monthly data. Six of the seven series exhibit strong signs of conditional heteroskedasticity. The four main topics addressed are impacts from macroeconomic performance on retail gasoline prices, the impacts of oil and gas price volatilities on retail gasoline prices, the importance of oil futures prices for retail gasoline prices, and the affect of temporal aggregation on the implications and forecasting performance of the model. Evidence suggests that the presence of conditional heteroskedasticity does not inhibit the application of a VECM framework, that macroeconomic performance does impact retail gas prices, price volatilities are a source of variation in retail gas prices, oil futures prices have a significant positive effect on retail gasoline prices, and temporal aggregation can distort underlying information inherent in a low frequency data generating process.
Contents

Chapter

I. Hurricanes, Price Shocks and Price Impacts in Oil and Gas Markets........1

Introduction..............................................................................................................1

Literature..................................................................................................................2

Model.......................................................................................................................4

Data.........................................................................................................................11

   Endogenous Variables......................................................................................12

   Common Factor Variables..............................................................................19

   Macroeconomic Variables..............................................................................19

   Stock Market Variables..................................................................................22

   Interest Rate Variables...................................................................................23

   Oil Industry Variables.....................................................................................24

   Event Variables................................................................................................25

   Missing Data.......................................................................................................26

Results....................................................................................................................27

   Step 1................................................................................................................27

   Step 2................................................................................................................30

   Step 3................................................................................................................36

   Step 4................................................................................................................38

Conclusion..............................................................................................................62

II. An Investigation of Economic Costs Associated with Road Construction
   and Maintenance............................................................................................64

Introduction...........................................................................................................64

Literature...............................................................................................................66
Roadway Maintenance.................................................................66
Road Congestion...........................................................................67
Model..............................................................................................70
The Firm.........................................................................................71
Households.....................................................................................72
Transportation...............................................................................73
The General Equilibrium Model...................................................74
Data...............................................................................................76
Results............................................................................................79
Congestion Mitigation.................................................................80
Road Maintenance.......................................................................84
Implications....................................................................................89
Conclusion......................................................................................98

III. A Forecasting Model of Retail Gasoline Prices..........................100
Introduction...................................................................................100
Literature.........................................................................................101
Oil and Gas Pricing.................................................................101
Temporal Aggregation..............................................................103
Data.............................................................................................104
Model............................................................................................106
Results..........................................................................................108
Daily.............................................................................................108
Weekly.........................................................................................117
Monthly.......................................................................................125
Comparisons.................................................................................131
Impacts from Conditional Heteroskedasticity..............................132
Conclusion.....................................................................................137
Tables

Table

1. Tests for units roots and cointegration in price data series..............................................5
2. Data descriptions, variable name abbreviations, and model use...........................................12
3. Descriptive statistics for OILFUTP by year..........................................................................13
4. Descriptive statistics for gasoline future price series GASFUTP, GASFUTP+, and GASFUTP-.................................................................................................14
5. Coefficient estimates for GASFUTP series...........................................................................15
6. Descriptive statistics for crude oil spot prices, OILSPOTP..................................................18
7. Descriptive statistics for gasoline spot prices, GASSPOTP..................................................18
8. Descriptive statistics for gasoline retail prices, GASRETP....................................................19
9. Descriptive statistics for macroeconomic information variables.........................................20
10. Companies comprising the variable oilidx...........................................................................23
11. Parameter estimates from stage 1 OLS estimates (only estimates with p-value less than .001 are shown).....................................................................................27
12. Test results for unit root and cointegration for step 1 residuals.............................................31
13. VAR model significant coefficient estimates for oil futures price........................................33
14. VAR model significant coefficient estimates for gasoline futures price..............................34
15. VAR model significant coefficient estimates for oil spot price............................................34
16. VAR model significant coefficient estimates for gasoline spot price....................................35
17. VAR model significant coefficient estimates for gasoline retail price...................................36
18. Diagnostic test results for stage 2 residuals...........................................................................37
19. Univariate GARCH(1,1) estimates for $\eta_i$.................................................................38

20. Summaries for Step 4 VAR(21)................................................................................39

21. Threshold values for road congestion and maintenance............................................76

22. Descriptive statistics for daily data series.......................................................................110

23. Augmented Dickey-Fuller tests for unit roots. (*** indicate
 significant at 1% level)..................................................................................................110

24. Diagnostic tests for VAR(K) for daily data series......................................................107

25. Johansen cointegration tests for daily data with 38 lags............................................112

26. Descriptive statistics for weekly data series....................................................................118

27. Augmented Dickey-Fuller tests for unit roots. (*** indicate
 significant at 1% level)..................................................................................................119

28. Diagnostic tests for VAR(K) for weekly data series.........................................................119

29. Johansen cointegration tests for weekly data with 7 lags............................................120

30. Descriptive statistics for monthly data series...............................................................122

31. Augmented Dickey-Fuller tests for unit roots. (*** indicate
 significant at 1% level)..................................................................................................127

32. Diagnostic tests for VAR(K) for monthly data series......................................................127

33. Johansen cointegration tests for monthly data with 2 and 6 lags.................................128

34. Actual versus estimated coefficients for $\gamma_1$.................................................................133

35. Actual versus estimated coefficients for $\gamma_2$.................................................................134

36. Diagnostic tests for model with normal errors versus
 model with GARCH errors........................................................................................134
37. Forecast error variance decompositions for normal and GARCH errors.................136
38. Mean squared prediction errors for both error processes........................................137
Figures

Figure
1. Differences between created series, GASFUTP, and $GASFUTP^-$, OILFUTP, OILSPOTP, GASSPOTP, and GASRETP ...............................................................16
2. Scatter plots for GASFUTP and OILFUTP, and GASSPOTP and OILFUTP ..............17
3. Original versus created daily series for quarterly gross domestic product ..............21
4. Plots for Oil Stock index and S&P500 stock market series .........................................22
5. Plots of interest rate variables ................................................................................24
6. Plots of oil and gas industry data ..............................................................................25
7. Histogram plots for stage 1 residual series .............................................................32
8. Lag effects from prices on oil futures prices (OILFUTP) from Step 4 VAR(21) presented as elasticities at mean values ......................................................40
9. Price and volatility impacts on OILFUTP from individual market prices and price volatility impulses with 95 percent confidence bands ...............................41
10. Lag effects from price volatilities on oil futures prices (OILFUTP) from Step 4 VAR(21) presented as elasticities at mean values ........................................42
11. Lag effects from prices on gas futures prices (GASFUTP) from Step 4 VAR(21) presented as elasticities at mean values ......................................................43
12. Price and volatility impacts on GASFUTP from individual market prices and price volatility impulses with 95 percent confidence bands ...............................44
13. Lag effects from price volatilities on gasoline futures prices (GASFUTP) from Step 4 VAR(21) presented as elasticities at mean values ........................................45
14. Lag effects from prices on oil spot prices (OILSPOTP) from Step 4 VAR(21) presented as elasticities at mean values ......................................................46
15. Price and volatility impacts on OILSPOTP from individual market prices and price volatility impulses with 95 percent confidence bands ...............................47
16. Lag effects from price volatilities on oil spot prices (OILSPOTP) from Step 4 VAR(21) presented as elasticities at mean values ..........................................48
17. Lag effects from prices on gasoline spot prices (GASSPOTP) from Step 4 VAR(21) presented as elasticities at mean values........................................49

18. Price and volatility impacts on GASSPOTP from individual market prices and price volatility impulses with 95 percent confidence bands...........50

19. Lag effects from price volatilities on gasoline spot prices (GASSPOTP) from Step 4 VAR(21) presented as elasticities at mean values...........................51

20. Lag effects from prices on retail gasoline prices (GASRETP) from Step 4 VAR(21) presented as elasticities at mean values.................................52

21. Price and volatility impacts on GASRETP from individual market prices and price volatility impulses with 95 percent confidence bands.............53

22. Lag effects from price volatilities on retail gasoline prices (GASRETP) from Step 4 VAR(21) presented as elasticities at mean values...........................54

23. Change in commodity market prices from the presence of a hurricane in the Gulf of Mexico. Results obtained from step1.................................55

24. Daily patterns of changes in commodity market prices from the presence of a hurricane in the Gulf of Mexico...............................................................56

25. Commodity market price and price volatility responses to market price changes from a hurricane event including 95 percent confidence bands......57

26. Commodity market price and price volatility responses to only market price volatility changes from a hurricane event including 95 percent confidence bands...............................................................................................................................60

27. Commodity market price and price volatility responses to changes in market price and market price volatility from a hurricane event with 95 percent confidence bands...............................................................................................................................62

28. Average road congestion by city size for the period 1990 to 2009..................................................78

29. Road congestion levels occurring under different new road construction congestion thresholds............................................................................................81

30. Nominal road construction expenditures in the United States, 2001-2008..........................82

31. The impact on consumer welfare at different levels of congestion................................83
32. The impact on production output from different levels of congestion
33. Road quality in the RD for different levels of annual road maintenance
34. Consumer welfare for different road conditions absent significant increases in traffic congestion
35. Production output for different levels of road quality
36. The impact on production output levels as road quality deteriorates
37. An index of U.S. road conditions by road type from 1990 to 2008
38. A comparison of wage changes given different levels of annual road maintenance
39. Shows projected road conditions given current trends for the residential district and the central business district
40. Shows projections for traffic congestion from 2011 to 2019
41. Shows current and projected average vehicle speeds from 2011 through 2019 in the CBD
42. Changes in city density as a result of projected traffic congestion and road quality changes
43. Change in land prices faced by the firm in the central business district (CBD) over time
44. Change in land prices faced by households in the residential district (RD) over time
45. Projected impacts on production output
46. Projected impacts on labor supply and market wage
47. Plots of daily data for each data series
48. OLS-CUSUM of VAR(38). Shows potential structural changes in the data series
49. DGAS_RET squared residuals from VECM estimation
50. Responses of retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks
51. Responses of retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks ...............................................................115
52. Forecast error variance decomposition for daily retail gasoline prices (DGAS_RET) ........................................................................................................115
53. Actual versus fitted for retail gasoline prices for the last 100 observations .........116
54. Retail gasoline prices for January 2011 versus forecasted prices .........................117
55. Plots of weekly data for each data series ..................................................................118
56. OLS-CUSUM of VAR(7) for weekly data series .......................................................120
57. OLS-CUSUM of VAR(7) for weekly data series .......................................................121
58. Responses of weekly retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks ................................................122
59. Responses of weekly retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks ................................................123
60. Forecast error variance decomposition for weekly retail gasoline prices (WGAS RET) ........................................................................................................123
61. Actual versus fitted for retail gasoline prices for the last 6 weeks for proposed lags of 2, 5 and 7 weeks ........................................................................124
62. Weekly retail gasoline prices for January 2011 versus forecasted prices ..............125
63. Plots of monthly data for each data series .................................................................126
64. Actual versus fitted for retail gasoline prices for the last 12 months for proposed lags of 2, and 6 months .................................................................128
65. Responses of monthly retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks ..............................................129
66. Responses of monthly retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks ..............................................130
67. Forecast error variance decomposition for monthly retail gasoline prices (MGAS RET) ........................................................................................................130
68. Monthly retail gasoline prices for January and February 2011 versus forecasted prices..............................................................................................................131

69. 180-day forecast for daily VECM(38) model with 95% confidence interval ..........132

70. Impulse responses for 14 periods for both types of error terms.................................135

71. Impulse responses for 14 periods for both types of error terms.................................136
Chapter I

Introduction

The nature of responses to market price changes in one market among other energy commodity markets continues to be uncertain. This paper looks at five energy commodity markets (oil futures, oil spot, gasoline futures, gasoline spot and gasoline retail) and explores the responses of market prices in all of those markets from price and price volatility shocks that occur as a result of an hurricane in the Gulf of Mexico.

I employ a four step estimation process to examine responses to price and price volatility changes in oil and gasoline commodity markets from the occurrence of an hurricane in the Gulf of Mexico. This paper adds to the existing literature by estimating price and price volatility impacts from a discrete, short duration event that would be expected to impact both oil and finished motor gasoline production. I estimate direct impacts between prices and price volatilities, and I describe price and price volatility responses without imposing any *a priori* structural restrictions.

Our results indicate that commodity price responses to a seven-day hurricane in the Gulf of Mexico follow a predictable pattern. As the hurricane enters the Gulf, prices of oil and gasoline commodities increase over the first two days of the hurricane event. Commodity prices in most markets then tend to retreat somewhat, which is a direct effect of the increases in price volatilities in all markets. I find that price volatility in the oil futures market plays the most significant role in impacting prices in all commodity markets. Commodity prices tend to fall in the presence of higher price volatilities, *ceteris paribus*. Thus commodity prices would tend to increase more during the hurricane event, but the initial increases in price volatilities work to keep the overall price increases somewhat lower. After approximately ten days, three days after the hurricane event, declining price volatilities, particularly in the oil
futures market, work to increase commodity prices substantially\textsuperscript{1} over a period of one week to ten days. After that, prices tend to fall over the remaining week of the simulation period.\textsuperscript{2}

The remainder of the paper is organized as follows. Section 2 discusses some of the relevant literature, section 3 outlines the empirical model formulation, section 4 discusses the data employed, section 5 presents estimation results, and section 6 concludes.

Literature

Although this paper does not explore nor attempt to explain the causes of price shocks in the petroleum and gasoline commodity markets, there is a substantial literature that has examined these markets.


\textsuperscript{1}Prices increase around 12 cents per gallon in oil futures, oil spot, gasoline futures and gasoline spot markets, and about 4.5 cents per gallon in the retail gasoline market.

\textsuperscript{2}The simulation period was 28 days.
Herrera (2009) (menu costs, information, and strategic response) and Lewis (2009) (intermediate price shocks). Lutz Kilian, for one, has focused some effort on explaining gasoline price changes on demand and, or supply shocks in the petroleum market (Kilian (2008), Kilian and Park (2009), Kilian (2010), and Kilian and Murphy (2010)). Some other authors have also looked at this problem: Ghoddusi, Titman, and Tompaidis (2010), Houde (2010), and Kurita (2010).

Many studies have examined futures prices. Generally, futures markets are thought to be used for hedging strategies or speculation. Carter (1999) provides a general survey of research into commodity futures markets. Pindyck (2001) and Pindyck (2004) examined the short-run links between spot and futures prices as being related through production, inventory and price volatility. Commodity spot prices are often used to calculate futures prices. For example, Ribeiro and Hodges (2004) use a two-factor model of the spot price and convenience yield (inversely related to inventory levels as a measure of the value of immediate possession) to derive a formula for the futures price. Worthington and Higgs (2010) used a multivariate generalized autoregressive conditional heteroskedasticity model to examine the relationships between spot and futures prices for electricity. They found that futures price movements impacted spot price markets, but that impacts from changing volatilities complicated that simple relationship. Zagaglia (2010) investigated the oil futures market to try and determine impacts on futures prices from macroeconomic changes versus those from financial information changes. His analysis used principal components to extract uncorrelated information from a large set of macroeconomic and information factors in order to gauge their combined impacts on oil futures prices. Kogan, Lidvan and Yaron (2009) note an inverse relationship between futures price volatility and the slope of the forward curve (the sequence of futures contracts of different maturity dates), which they explain as a consequence of irreversible investment and investment constraints (imposing limitations on production adjustments).
Volatility has been extensively studied in the finance literature, and various methods have been established to examine price volatility. For example, Engle (1982) developed the autoregressive conditional heteroskedasticity (ARCH) model to examine and account for volatility in asset prices and returns. The multivariate, or MGARCH, framework was developed to examine the relationships between volatilities across multiple markets (see Engle, Granger and Kraft (1984), Bollerslev, Engle and Woolridge (1988), and Bollerslev (1990) for some early developments). Bauwens, Laurent and Rombouts (2006) provide a survey of MGARCH models useful for examining impacts from price volatility. Sadorsky (2006) examined volatility in futures markets using several different models. Sadorsky (1999) used VAR and GARCH models to examine stock market returns in light of oil price shocks, and oil price volatility.

Model

This paper employs a four-step process to ultimately explore the connection between price changes in five different commodity markets. The five markets are the petroleum futures market, OILFUT, the petroleum spot market, OILSPOT, the gasoline futures market, GASFUT, the gasoline spot market, GASSPOT, and the retail gasoline market, GASRET. Step one employs the notion of a set common factors that could be expected to impact these five markets and influence their behavior and, hence, price movements. The set of common factors includes publicly available information about economic conditions, investment opportunities, and some variables related to conditions in the oil and gasoline industries including events that impact the supply of oil and or gasoline, which include the presence of hurricanes in the Gulf of Mexico and conflicts that are believed to impact the supply of petroleum products.

If we examine the properties inherent in the raw price data, we observe that the presence of unit roots in each of the five series cannot be rejected at the one percent confidence
level. Further the null hypothesis that the series have four cointegrated variables cannot be rejected.

Table 1. Tests for units roots and cointegration in price data series.

Table 1 shows that at the one percent confidence level all five series have unit roots using the DF-GLS test. The Johansen Cointegration trace-test shows the presence of four cointegration relationships. Because of the possible presence of unit roots and the cointegration of the variables, the assumption of stability in the data series is not likely to hold. Therefore I can use the step 1 regressions to obtain a series that is stable and thus appropriate for further analysis of the relationships among these five markets. At the same time, by using common economic and industry factors, I can get some notion of how these factors influence prices in the five energy commodity markets.

These first stage regressions also draw motivation from the notion of risk factors explored in the financial econometrics literature. For example, Bailey and Chan (1993) hypothesized that the differences between commodity spot and futures prices could be explained by macroeconomic risk factors. Those authors defined the difference between spot and futures prices as a risk premium that could be explained by shocks to underlying economic vari-

\[ \text{Critical Values} \]

### DF-GLS Unit Root Test

<table>
<thead>
<tr>
<th></th>
<th>Lag=1</th>
<th>1pc</th>
<th>5pc</th>
<th>10pc</th>
</tr>
</thead>
<tbody>
<tr>
<td>OILFUTP</td>
<td>-2.6695</td>
<td>-3.48</td>
<td>-2.89</td>
<td>-2.57</td>
</tr>
<tr>
<td>GASFUTP</td>
<td>-2.8473</td>
<td>-3.48</td>
<td>-2.89</td>
<td>-2.57</td>
</tr>
<tr>
<td>OILSPOTP</td>
<td>-2.8022</td>
<td>-3.48</td>
<td>-2.89</td>
<td>-2.57</td>
</tr>
<tr>
<td>GASSPOTP</td>
<td>-3.2918</td>
<td>-3.48</td>
<td>-2.89</td>
<td>-2.57</td>
</tr>
<tr>
<td>GASRETP</td>
<td>-2.7844</td>
<td>-3.48</td>
<td>-2.89</td>
<td>-2.57</td>
</tr>
</tbody>
</table>

### Johansen Cointegration Test

<table>
<thead>
<tr>
<th></th>
<th>Lag=21</th>
<th>1pc</th>
<th>5pc</th>
<th>10pc</th>
</tr>
</thead>
<tbody>
<tr>
<td>r = 0</td>
<td>288.90</td>
<td>83.2</td>
<td>67.31</td>
<td>56.50</td>
</tr>
<tr>
<td>r = 1</td>
<td>146.59</td>
<td>59.14</td>
<td>62.59</td>
<td>70.95</td>
</tr>
<tr>
<td>r = 2</td>
<td>72.2</td>
<td>39.06</td>
<td>42.44</td>
<td>46.45</td>
</tr>
<tr>
<td>r = 3</td>
<td>36.13</td>
<td>22.76</td>
<td>25.32</td>
<td>30.45</td>
</tr>
<tr>
<td>r = 4</td>
<td>19.73</td>
<td>10.49</td>
<td>12.15</td>
<td>16.26</td>
</tr>
</tbody>
</table>

\[ ^3 \text{I am mindful that one could employ a properly specified vector error correction model to overcome these stability issues. However such a specification would require some additional information regarding the normalization of the error correction coefficients. This analysis seeks to avoid imposing any such a priori restrictions on the five commodity markets.} \]
ables that influence the spot market. Following typical asset pricing methods, the authors regressed for each commodity the quantity (futures price - spot price)/spot price on the information set in order to determine common bases for each commodity after accounting for the impacts of information.

The present analysis uses similar methods, here the set of common factors is larger and includes additional macroeconomics variables, stock prices, interest rates, and industry measures. The natural logarithms of each series of market prices is regressed on one-day lags of the natural logarithms of these common factors. The model can be written the following form where the superscript denotes step 1.

\[(1) \quad x_{1j,t} = \sum_{i=1}^{m} \beta_{i,j} z_{i,j,t-1} + \epsilon_{1j,t} \text{ for } j = \{OILFUTP, GASFUTP, OILSPOTP, OPPSPOTP, GASRETP\}\]

The \(z_i\) are the one-day lags of the natural logarithms of the so-called common factors. The residuals from this model are retained and used in the step 2 estimation procedure. This specification allows for the examination of the relationships among the oil and gas market price variables after accounting for the unique impacts from the common factors. The step 2 procedure examines the implications from a vector autoregression (VAR) model of the prices from the five oil and gasoline markets. The use of a vector autoregression (VAR) models to examine the joint evolution of oil and gasoline prices is not unusual. See for example Kilian (2009), Kilian and Murphy (2010) and Kilian (2010). The VAR proposed in this paper is designed to investigate short-run impacts from price changes within the five markets. The model is formulated as follows.

\[(2) \quad \hat{\epsilon}_{1t} = \hat{x}_{2t} = \sum_{i=1}^{p} A_i x_{2t-p} + \epsilon_{2t}\]

Equation (3.2) is the standard form for a VAR of order \(p\), where each dependent variable is modeled as a function of the \(p\) lags of all dependent variables and, again, the superscripts
denote the model step number. The value of \( p \) is chosen using the Forecast Prediction Error (FPE) over from one to thirty lags. One of the goals of this paper is to determine the relationships among the five markets, so no restrictions are placed on the VAR equations and a structural VAR model is not employed. The weakness of this is that it can make it more difficult to accurately identify the relationships between the markets. However, no \textit{a priori}, artificial restrictions are placed on the relationships in these five markets.

Financial data is often associated with volatility clustering (periods of high volatility followed by periods of low volatility). This typically manifests itself by the presence of heteroskedasticity in the conditional variance of the price series. To estimate the impacts from so-called conditional heteroskedasticity, step 3 employs a multivariate generalized autoregressive conditional heteroskedastic (MGARCH) model. There are several possible formulations for MGARCH modeling and the one chosen for step 3 is the PC-GARCH (principal components GARCH) model, also known as the O-GARCH (orthogonal GARCH) model from Alexander (2001). This PC-GARCH model is specified following Boswijk and Van der Weide (2006).

\begin{align*}
(3) \quad \hat{\epsilon}_t^2 & = \hat{\sigma}_t^3 = P \eta_t \\
(4) \quad H_t & = PA_tP' \\
(5) \quad \Lambda_t & = \begin{pmatrix}
\lambda_{1t} & 0 & 0 & 0 & 0 \\
0 & \lambda_{2t} & 0 & 0 & 0 \\
0 & 0 & \lambda_{3t} & 0 & 0 \\
0 & 0 & 0 & \lambda_{4t} & 0 \\
0 & 0 & 0 & 0 & \lambda_{5t}
\end{pmatrix} \\
(6) \quad \lambda_{it} & = \omega_i + \sum_{j=1}^{p} \alpha_{ij} \eta_{it-j}^2 + \sum_{k=1}^{q} \beta_{ik} \lambda_{it-k}
\end{align*}
Here $H_t$ is the conditional covariance matrix at time $t$, $\eta_t$ is a vector of independent or exogenous component processes at time $t$ and is assumed to be such that $\eta_t \sim N(0, H_t)$. Letting $V$ be the unconditional variance of the $x_t$, then $P$ can be estimated as $V = P \Lambda P'$ where $P$ is the orthonormal matrix of eigenvectors of $V$ and $\Lambda$ is the diagonal matrix of eigenvalues of $V$. As noted in Francq and Zakoian (2010), $\hat{H} = \frac{1}{T} \sum_{t=1}^{T} (x_t^3 - \bar{x}^3)(x_t^3 - \bar{x}^3)'$ provides an estimate for $H$ (where, again, the superscripts denote step 3 of the estimation process).

Another possibility was to employ the generalized orthogonal GARCH (GO-GARCH) model discussed in Boswijk and Van der Weide (2006). However, their analysis suggests that if the data is skewed or not mesokurtic, then PC-GARCH may provide better estimates. These so-called factor MGARCH models are parsimonious specifications in that the parameters $\alpha_{ij}$ and $\beta_{ik}$ from equation (3.6) can be estimated by separate univariate GARCH models. Van der Weide (2002) shows that PC-GARCH is a special case of GO-GARCH, which are, in-turn, special cases of other MGARCH models.

Estimation steps are straightforward. Estimate $\hat{H} = \frac{1}{T} \sum_{t=1}^{T} (x_t^3 - \bar{x}^3)(x_t^3 - \bar{x}^3)'$ and determine $\hat{H} = \hat{P} \hat{\Lambda} \hat{P}'$. Use the elements of $\hat{\Lambda}$ as estimates for $\lambda_{1i}$. Calculate $\hat{\eta}_{i,t} = \hat{P}' \epsilon_{i,t}$ for all $i$ and all $t$. Estimate $\hat{\alpha}$ and $\hat{\beta}$ for use in Equation (3.6) using five univariate GARCH(p,q) models on $\eta_i$ where $p$ and $q$ are chosen by examining the appropriate extended autocorrelation function (eacf). Use Equation (3.6), the estimated $\hat{\alpha}$ and $\hat{\beta}$, and $\hat{\lambda}_{i,1}$ to obtain $\hat{\Lambda}_t$ for all $t$.

The results from step 3 are used to form volatilities from the relationship $\hat{\epsilon}_t^2 = x_t^3 = P\eta_t$ as the variance-covariance matrix of the $\epsilon_t^2$ (note that here the superscript denotes the step 2 residual, not the square of the residual), or $P\eta_t \eta_t'P'$. The variances for each were retained and used in the step 4 VAR. Thus, the step 4 VAR(p) is a ten equation system

\[ x_t^4 = \sum_{i=1}^{p} A_i x_{t-i}^2 + \epsilon_t^4 \]

where $\hat{x}_t^2$ includes the residuals from step 1 for the first five equations, and the volatilities from step 3 for the last five equations. This specification allows us to directly test the
impacts of prices on volatilities and volatilities on prices. Mahadevan and Suardi (2008) used a similar model when they employed quasi-maximum likelihood to estimate a trade model with trade volatilities by including the estimated variances of the error terms from a VECM using a BEKK framework in an MGARCH model. I use a two-step procedure because I am using PC-GARCH instead of the BEKK framework.

The estimates obtained from the VAR(p) are used to determine responses in all markets from price responses to an exogenous event that would be expected to cause price changes in all markets for oil and gasoline simultaneously, and responses from individual price changes that occur in each market separately. The event chosen is the presence of a seven-day hurricane\(^4\) in the Gulf of Mexico.

Recall from step 1 that I estimated the impact on oil and gasoline commodity prices from the presence of a Gulf hurricane. Such storms can impact oil production directly by interrupting drilling and shipping, and can also interrupt the production of gasoline by forcing the closure of refineries, thereby reducing the amount of finished motor gasoline being supplied. The price impacts from these interruptions are estimated in step 1 and applied to an impulse response analysis in step 4.

In order to examine how price shocks will impact the entire system over the following four weeks, I used the models to update from randomly drawn starting positions. This was done because traditional orthogonal impulse response functions do not usually give unambiguous pictures of these impacts. This is because the traditional orthogonal impulse response functions can vary greatly depending on how the variables are ordered in the specification of the VAR estimation. Essentially, the ordering of the variables implicitly sets up a structural vector autoregression where the first variable in the order is assumed to impact all variables, the second impacts all variables except the first and so on.

\(^4\)The other choice was the presence of a “conflict” impacting oil and/or gasoline production. The hurricane was chosen because those events tend to be more discrete (of shorter duration) and the overall price impacts are similar between the two events.
The method employed here depends, of course, on the starting point of the impulse response simulations. In order to overcome the reliance of the impulse responses on the selected range of the date, the impulse responses were calculated in the following manner. After the step 4 estimation is completed, $250^5$ periods are randomly selected. The periods selected are within the data range allowing for at least 21 periods prior to the selected period and 27 periods after for a total range of 49 days. Since daily data are used, we can then observe impacts from a change over the following 28 days by comparing the baseline of what actually happened to what the model will predict happens after the price impulses are implemented. What are observed then are short-run, intra-month responses over a four week period.

The output is 250 49-period series that represent the differences between the daily predicted prices and the baseline (actual) prices. The single impulse response for each day is then calculated as the mean of the 250 simulated differences. Finally, confidence intervals are calculated for the calculated differences between the predicted responses and the baseline prices for each of the five market prices and volatilities. The confidence intervals are formed as

$$
(\hat{\theta}^* - s_{\theta}^* t_{1-\alpha/2}, \hat{\theta}^* + s_{\theta}^* t_{1-\alpha/2})
$$

where $t_{1-\alpha/2}$ is the value of the t-distribution at the $\alpha$ level of significance, $\hat{\theta}^*$ is the mean estimated difference, $s_{\theta}^* = \left(\frac{1}{B-1} \sum_{j=1}^{B} (\hat{\theta}^*_j - \bar{\theta}^*)^2\right)^{\frac{1}{2}}$ is the standard error, where $B = 250$ is the number of periods selected and $\hat{\theta}^*_j$ and $\bar{\theta}^* = \frac{1}{B} \sum_{j=1}^{B} \hat{\theta}^*_j$ are the $j^{th}$ individual sample estimate and the mean of all sample estimates respectively. This method can also be used when calculating a bootstrap confidence interval. See Mackinnon (2002).

These price impulses are first examined by allowing prices in all markets to vary at once. From step 1 I noted that on a day the hurricane was in the Gulf of Mexico, prices in all five commodity markets are from two to seven cents per gallon higher than otherwise. In order to obtain a more realistic pattern, the step 1 models were re-estimated with separate dummy

---

5Some sensitivity analysis was conducted on the number of samples chosen. For example, the mean differences for a sample of 250 was within the 95 percent confidence interval for the sample of 400.

6The level chosen for this analysis is $\alpha = .05$. 

10
variables representing the first day of the hurricane, the second day, etc. I included seven such dummy variables representing the first seven days of a Gulf hurricane. It should be noted that there were only 21 observations for day 1 and 10 observations for day seven. The results from these regressions allow for the use of an approximate pattern of the aforementioned price increases over the duration of the hurricane event.

After examining the initial impacts from simultaneous changes in all markets, impacts from price changes in individual markets are calculated. The calculations are done in the manner described above and allow us to observe impacts in all five commodity markets from price changes in one market. These can then be compared to the pattern of price responses when prices are changing in all markets.

Data

The data consist of daily observations on five endogenous commodity market variables along with a set of lagged exogenous common factor variables. The final data used for estimation of the model parameters included 3762 daily observations from February 5, 1996 through December 31, 2010. Table 4.1 provides a list of variables used.
Table 2. Data descriptions, variable name abbreviations, and model use.

In Table 2, the petroleum price variables are used in every step of the estimation process and the other variables are used in every equation in step 1. The data for OILFUTP, GASFUTP, OILSPOTP, and GASSPOTP were obtained from the U.S. Energy Information Agency\(^7\), data for GASRETP were obtained from OPIS\(^8\).

**Endogenous Variables**

There are five variables considered endogenous to the model. The first, OILFUTP, is the Cushing, OK oil future price for the first contract. Contracts for oil are each for 1000 barrels (42 gallons per barrel) and contracts mature on the third business day prior to the 25th of the month if it is a business day, or the third business day prior to the first business day preceding the 25th if not. At that time, contract 2\(^9\) becomes the new contract 1. Table 3 presents some descriptive statistics for OILFUTP.

\(^7\)http://www.eia.doe.gov/petroleum/data.cfm

\(^8\)http://www.opisnet.com

\(^9\)A dummy variable was created to test for significant price changes on the day of the switch from contract 1 to contract 2, but the estimated coefficient on that dummy variable was not significantly different from zero.
Table 3. Descriptive statistics for OILFUTP by year.

From Table 3 we can note that oil future prices have generally increased over the sample period and that price volatility has increased (measured here by variance). We can also see that the annual distributions do not consistently skew left or right, nor are the distribution tails consistently fat or thin.

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean(OILFUTP)</th>
<th>Var(OILFUTP)</th>
<th>Skew(OILFUTP)</th>
<th>Kurtosis(OILFUTP)</th>
<th>Max(OILFUTP)</th>
<th>Min(OILFUTP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>0.625</td>
<td>0.003292</td>
<td>-0.112</td>
<td>-0.987</td>
<td>0.643</td>
<td>0.416</td>
</tr>
<tr>
<td>1997</td>
<td>0.690</td>
<td>0.000194</td>
<td>1.342</td>
<td>1.654</td>
<td>0.634</td>
<td>0.419</td>
</tr>
<tr>
<td>1998</td>
<td>0.342</td>
<td>0.001400</td>
<td>0.439</td>
<td>0.210</td>
<td>0.424</td>
<td>0.249</td>
</tr>
<tr>
<td>1999</td>
<td>0.460</td>
<td>0.001154</td>
<td>-0.122</td>
<td>-1.063</td>
<td>0.445</td>
<td>0.271</td>
</tr>
<tr>
<td>2000</td>
<td>0.721</td>
<td>0.000488</td>
<td>-0.034</td>
<td>-0.727</td>
<td>0.686</td>
<td>0.668</td>
</tr>
<tr>
<td>2001</td>
<td>0.617</td>
<td>0.000722</td>
<td>0.625</td>
<td>0.615</td>
<td>0.766</td>
<td>0.416</td>
</tr>
<tr>
<td>2002</td>
<td>0.623</td>
<td>0.000571</td>
<td>-0.634</td>
<td>-0.051</td>
<td>0.779</td>
<td>0.429</td>
</tr>
<tr>
<td>2003</td>
<td>0.738</td>
<td>0.000396</td>
<td>0.513</td>
<td>0.265</td>
<td>0.901</td>
<td>0.601</td>
</tr>
<tr>
<td>2004</td>
<td>0.688</td>
<td>0.001500</td>
<td>0.555</td>
<td>0.635</td>
<td>1.314</td>
<td>0.773</td>
</tr>
<tr>
<td>2005</td>
<td>1.351</td>
<td>0.000494</td>
<td>-0.132</td>
<td>-0.637</td>
<td>1.552</td>
<td>1.003</td>
</tr>
<tr>
<td>2006</td>
<td>1.578</td>
<td>0.001744</td>
<td>0.141</td>
<td>-1.219</td>
<td>1.834</td>
<td>1.329</td>
</tr>
<tr>
<td>2007</td>
<td>1.722</td>
<td>0.002349</td>
<td>0.485</td>
<td>-0.933</td>
<td>2.333</td>
<td>1.302</td>
</tr>
<tr>
<td>2008</td>
<td>2.575</td>
<td>0.004739</td>
<td>-0.925</td>
<td>-0.470</td>
<td>3.459</td>
<td>0.606</td>
</tr>
<tr>
<td>2009</td>
<td>1.479</td>
<td>0.001024</td>
<td>-0.462</td>
<td>-1.163</td>
<td>1.937</td>
<td>0.609</td>
</tr>
<tr>
<td>2010</td>
<td>1.696</td>
<td>0.001629</td>
<td>0.189</td>
<td>-0.666</td>
<td>2.179</td>
<td>1.619</td>
</tr>
</tbody>
</table>
Table 4. Descriptive statistics for gasoline future price series GASFUTP, GASFUTP+, and GASFUTP-.

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean(GASFUTP)</th>
<th>Var(GASFUTP)</th>
<th>Skew(GASFUTP)</th>
<th>Kurtosis(GASFUTP)</th>
<th>Max(GASFUTP)</th>
<th>Min(GASFUTP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>0.660</td>
<td>0.00249</td>
<td>-0.29</td>
<td>0.30</td>
<td>0.779</td>
<td>0.548</td>
</tr>
<tr>
<td>1987</td>
<td>0.640</td>
<td>0.00167</td>
<td>-0.16</td>
<td>-0.56</td>
<td>0.747</td>
<td>0.547</td>
</tr>
<tr>
<td>1988</td>
<td>0.472</td>
<td>0.00298</td>
<td>-0.93</td>
<td>-0.64</td>
<td>0.550</td>
<td>0.348</td>
</tr>
<tr>
<td>1989</td>
<td>0.599</td>
<td>0.01622</td>
<td>-0.42</td>
<td>-0.97</td>
<td>0.769</td>
<td>0.396</td>
</tr>
<tr>
<td>2000</td>
<td>0.932</td>
<td>0.00817</td>
<td>-0.31</td>
<td>-0.92</td>
<td>1.114</td>
<td>0.707</td>
</tr>
<tr>
<td>2001</td>
<td>0.820</td>
<td>0.02392</td>
<td>-0.04</td>
<td>-0.93</td>
<td>1.154</td>
<td>0.518</td>
</tr>
<tr>
<td>2002</td>
<td>0.792</td>
<td>0.00739</td>
<td>-1.14</td>
<td>0.96</td>
<td>0.362</td>
<td>0.572</td>
</tr>
<tr>
<td>2003</td>
<td>0.938</td>
<td>0.00530</td>
<td>0.50</td>
<td>0.26</td>
<td>1.163</td>
<td>0.815</td>
</tr>
<tr>
<td>2004</td>
<td>1.052</td>
<td>0.01453</td>
<td>-0.13</td>
<td>-0.79</td>
<td>1.477</td>
<td>0.988</td>
</tr>
<tr>
<td>2005</td>
<td>1.656</td>
<td>0.05026</td>
<td>0.67</td>
<td>0.97</td>
<td>2.542</td>
<td>1.201</td>
</tr>
<tr>
<td>2006</td>
<td>1.934</td>
<td>0.10747</td>
<td>0.22</td>
<td>-1.49</td>
<td>2.498</td>
<td>1.431</td>
</tr>
<tr>
<td>2007</td>
<td>2.071</td>
<td>0.07677</td>
<td>-0.92</td>
<td>0.14</td>
<td>2.498</td>
<td>1.355</td>
</tr>
<tr>
<td>2008</td>
<td>2.488</td>
<td>0.59404</td>
<td>-0.72</td>
<td>-0.65</td>
<td>3.571</td>
<td>0.703</td>
</tr>
<tr>
<td>2009</td>
<td>1.934</td>
<td>0.09433</td>
<td>-0.61</td>
<td>-1.00</td>
<td>2.071</td>
<td>1.043</td>
</tr>
<tr>
<td>2010</td>
<td>2.125</td>
<td>0.02095</td>
<td>0.21</td>
<td>-0.92</td>
<td>2.453</td>
<td>1.649</td>
</tr>
</tbody>
</table>

The next variable, GASFUTP, is the New York Harbor regular reformatted RBOB gasoline future contract 1 price. This variable is partially constructed from two other variables. One is the New York Harbor reformulated regular gasoline future contract 1 price, \( GASFUTP^- \), which has data available only for the period January 2, 1985 through December 29, 2006. The other is the New York Harbor regular reformatted RBOB gasoline future contract price 1, \( GASFUTP^+ \), which is available from October 3, 2005 through the December 31, 2010 (data for that series continues to be available). The series was constructed for
the period February 5, 1996 through October 2, 2005. The goal was to take advantage of the apparent autoregressive structure of the existing $GASFUTP^-$ series, the overlapping data of the $GASFUTP^+$ and $GASFUTP^-$ series, the existing past observations for the $GASFUTP^-$ series, but to avoid “tethering” the series with the price series for the other four markets. Descriptive statistics for the three series are shown in Table 4.

The method used to construct the series was to re-order the two series, $GASFUTP^-$ and $GASFUTP^+$, in time descending order and use the re-ordered data to fit the following model.

$GASFUTP^+_t = \beta_1 GASFUTP^-_{t-1} + \beta_2 GASFUTP^+_{t-2} + \beta_3 GASFUTP^-_{t-3} + \beta_4 GASFUTP^-_t + \beta_5 GASFUTP^+_{t-2} \quad (9)$

The coefficient estimates were used to construct forecast values, which were then re-ordered into ascending time order. Table 5 shows the estimated parameters.

| Estimate | St. Err. | t-value | Pr(>|t|) |
|----------|---------|---------|---------|
| $\beta_1$ | 0.7985  | 0.0568  | 14.1    | $< e^{-16}$ |
| $\beta_2$ | 0.1391  | 0.0665  | 2.1     | 0.0372     |
| $\beta_3$ | -0.0166 | 0.0333  | -0.5    | 0.6198     |
| $\beta_4$ | 0.7758  | 0.0321  | 24.1    | $< e^{-16}$ |
| $\beta_5$ | -0.5398 | 0.0619  | -8.7    | $< e^{-16}$ |
| $\beta_6$ | -0.1540 | 0.0562  | -2.7    | 0.00064    |

Table 5. Coefficient estimates for $GASFUTP$ series.

It is not surprising that most coefficient estimates were significant, but more importantly is how the constructed series captures the nature of the existing $GASFUTP^-$ series and maintains the correct relative position with respect to the other four market price variables. Figure 1 shows some relationships between the series $GASFUTP$ and other series. The comparisons from left to right, top to bottom are with $GASFUTP$ and: $GASFUTP^-$, $OILFUTP$, $OILSPOTP$, $GASSPOTP$, and $GASRETP$. The vertical line in four of the graphs marks the start of the existing $GASFUT^+$ series. Most of the differences show some sort of trend and it is difficult to know whether the trend is due to some structural change in the
pricing relationships between the commodity markets, or a significant difference between the created portion of the GASFUTP series and the actual rbglf1 price series.

Figure 1: Differences between created series, GASFUTP, and $GASFUTP^-$, OILFUTP, OILSPOTP, GASSPOTP, and GASRETP.

Figure 2 may help answer that question. The graph on the right shows a scatter plot of OILFUTP versus GASFUTP with two straight lines. The line that is lower on the right is the line formed by the regression of OILFUTP on GASFUTP for the observations that were created as noted above. The other line is the same regression, but for observations on the actual series $GASFUTP^-$. Similarly, the graph on the left is the scatter plot of OILFUTP versus GASSPOTP. The two lines are regressions of OILFUTP on GASSPOTP for the same periods as for the figure on the right (the lower line on the right being the regression with the earlier observations). We note that the two graphs look quite similar and seem to possess similar characteristics, which provides some evidence that the series, GASFUTP, that was created is not obviously unreliable. Thus, the judgment has to be made between (1) using the created series and maintaining a much longer set of observations, (2) using only the
GASFUTP$^+$ series and losing thousands of observations, (3) using only the GASFUTP$^-$ series, losing a few hundred observations, but having a series for a market that no longer exists, or (4) dropping the notion of examining the gasoline future price market altogether.

Figure 2: Scatter plots for GASFUTP and OILFUTP, and GASSPOTP and OILFUTP.

For this paper, option (1) was chosen, and the analysis includes the series GASFUTP, which is designed to capture information about the regular reformatted RBOB gasoline futures prices. Looking back at the data descriptions for the GASFUTP series in Table 6, we can see that the year-to-year relationships are similar to those observed in oil futures market.

The next market variable being examined is the Cushing, OK crude oil spot market, OILSPOTP. Futures prices are related to the spot prices for their respective markets and the two prices do converge over time. It is quite easy to show that the futures price cannot be too different from the spot price or arbitrage opportunities will arise between the two markets. However, Alquist and Kilian (2010) have found that the futures price provides a poor forecast of spot prices. Table 4.5 shows the descriptive statistics for the OILSPOTP series.
Table 6. Descriptive statistics for crude oil spot prices, OILSPOTP.

Once again we observe that in 2008 oil spot prices increased (to as much as $3.46/gal), but more importantly, volatility as measured by the variance of the series increased to a level about 400 times the level of most other years.

The gasoline spot market is represented by New York Harbor regular reformatted gasoline, GASSPOTP. Table 7 shows the annual descriptive statistics for the GASSPOTP series.

Table 7. Descriptive statistics for gasoline spot prices, GASSPOTP.

Looking at Table 7, we can see that gasoline spot prices are about 8 to 10 cents per gallon more expensive than crude oil spot prices. Volatilities exhibit a pattern similar to the other variables, but gasoline spot prices tend to be more volatile than crude oil spot prices (with greater annual variances).
The final market is the retail gasoline market. The prices shown are the average U.S. retail prices for regular unleaded gasoline.

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean(GASRETP)</th>
<th>Var(GASRETP)</th>
<th>Skew(GASRETP)</th>
<th>Kurtosis(GASRETP)</th>
<th>Max(GASRETP)</th>
<th>Min(GASRETP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>1.235</td>
<td>0.00266</td>
<td>0.94</td>
<td>0.24</td>
<td>1.918</td>
<td>1.115</td>
</tr>
<tr>
<td>1997</td>
<td>1.232</td>
<td>0.00261</td>
<td>-0.62</td>
<td>0.26</td>
<td>1.922</td>
<td>1.139</td>
</tr>
<tr>
<td>1998</td>
<td>1.068</td>
<td>0.00134</td>
<td>-0.27</td>
<td>0.84</td>
<td>1.163</td>
<td>0.945</td>
</tr>
<tr>
<td>1999</td>
<td>1.159</td>
<td>0.01414</td>
<td>-0.63</td>
<td>-0.92</td>
<td>1.303</td>
<td>0.957</td>
</tr>
<tr>
<td>2000</td>
<td>1.691</td>
<td>0.00777</td>
<td>-0.74</td>
<td>0.26</td>
<td>1.826</td>
<td>1.295</td>
</tr>
<tr>
<td>2001</td>
<td>1.430</td>
<td>0.02874</td>
<td>-0.44</td>
<td>0.39</td>
<td>1.715</td>
<td>1.072</td>
</tr>
<tr>
<td>2002</td>
<td>1.351</td>
<td>0.01186</td>
<td>-1.36</td>
<td>0.34</td>
<td>1.471</td>
<td>1.111</td>
</tr>
<tr>
<td>2003</td>
<td>1.603</td>
<td>0.00698</td>
<td>0.73</td>
<td>0.70</td>
<td>1.734</td>
<td>1.453</td>
</tr>
<tr>
<td>2004</td>
<td>1.842</td>
<td>0.01650</td>
<td>-0.51</td>
<td>-0.51</td>
<td>2.048</td>
<td>1.506</td>
</tr>
<tr>
<td>2005</td>
<td>2.289</td>
<td>0.03065</td>
<td>0.76</td>
<td>0.69</td>
<td>3.070</td>
<td>1.757</td>
</tr>
<tr>
<td>2006</td>
<td>2.669</td>
<td>0.05986</td>
<td>0.21</td>
<td>-1.68</td>
<td>3.033</td>
<td>2.181</td>
</tr>
<tr>
<td>2007</td>
<td>2.786</td>
<td>0.03962</td>
<td>0.30</td>
<td>-0.18</td>
<td>3.223</td>
<td>2.125</td>
</tr>
<tr>
<td>2008</td>
<td>3.235</td>
<td>0.53151</td>
<td>-0.92</td>
<td>-0.17</td>
<td>4.106</td>
<td>1.695</td>
</tr>
<tr>
<td>2009</td>
<td>2.349</td>
<td>0.00566</td>
<td>-0.50</td>
<td>-1.32</td>
<td>2.067</td>
<td>1.698</td>
</tr>
<tr>
<td>2010</td>
<td>2.775</td>
<td>0.00943</td>
<td>0.78</td>
<td>0.27</td>
<td>3.086</td>
<td>2.593</td>
</tr>
</tbody>
</table>

Table 8. Descriptive statistics for gasoline retail prices, GASRETP.

From Table 8 we note that gasoline retail prices follow a pattern similar to that of gasoline spot prices. Retail prices tend to be about 70 cents per gallon more than spot prices, but, except in cases where price volatilities increase for all markets, retail prices tend to be somewhat less volatile than spot prices.

Common Factor Variables

The analysis used 18 separate variables that comprise the set of common factors for step 1 of the estimation. These can be roughly grouped into four different categories: macroeconomic, stock market, interest rate, and oil and gas industry.

Macroeconomic Variables

The macroeconomic variables included series measuring the M1 measure of the money supply, M1, seasonally adjusted nominal gross domestic product, NGDP, price indexes for producer prices, PPI, and consumer prices, CPI, employment, EMPTOT, and industrial production, INDPD1, INDPD3, and INDPD6. Except for gross domestic product,
which is available by quarter, these data are available in monthly series. Table 9 shows the descriptive statistics for these data series.

Table 9. Descriptive statistics for macroeconomic information variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Variance</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>1,200.36</td>
<td>41,406.92</td>
<td>0.98</td>
<td>-0.04</td>
<td>1,940.00</td>
<td>1,044.80</td>
</tr>
<tr>
<td>EMPTOT</td>
<td>130,693.44</td>
<td>23,566,330.00</td>
<td>-0.63</td>
<td>0.13</td>
<td>139,080.00</td>
<td>116,408.36</td>
</tr>
<tr>
<td>PPI</td>
<td>148.65</td>
<td>502.52</td>
<td>0.81</td>
<td>-0.99</td>
<td>205.50</td>
<td>122.50</td>
</tr>
<tr>
<td>CPI</td>
<td>186.75</td>
<td>413.67</td>
<td>0.15</td>
<td>-1.32</td>
<td>219.35</td>
<td>154.45</td>
</tr>
<tr>
<td>NGDP</td>
<td>11,392.32</td>
<td>5,112,245.00</td>
<td>-0.01</td>
<td>1.37</td>
<td>14,995.43</td>
<td>7,575.81</td>
</tr>
<tr>
<td>INDPROD</td>
<td>51.54</td>
<td>63.23</td>
<td>-0.62</td>
<td>0.50</td>
<td>71.20</td>
<td>21.50</td>
</tr>
<tr>
<td>INDPRODI</td>
<td>52.30</td>
<td>114.76</td>
<td>-0.85</td>
<td>0.53</td>
<td>71.50</td>
<td>17.30</td>
</tr>
<tr>
<td>INDPRODS</td>
<td>52.70</td>
<td>167.26</td>
<td>-0.82</td>
<td>0.21</td>
<td>70.00</td>
<td>13.60</td>
</tr>
</tbody>
</table>

There is nothing particularly noteworthy about the information shown in Table 9, but it should be noted that these statistics were generated after the monthly (or quarterly for the case of gross domestic product) data were converted to daily series. The problem being faced is that on any day these markets are subject to a variety of factors, but the data is not always available on a daily basis. What then can be done to make use of that information? Or, more to the point, how can the researcher fit the information into the structure of a statistical model given different data frequencies? Pavia-Miralles (2010) provides an extensive review of the various methods developed to take a low frequency data series and transform it into a high frequency data series. The methods range from fairly simple weighting schemes to more complex methods involving techniques such as Kalman filtering. One way, the method employed in this paper, is to take the non-daily series and convert them to daily series using the period-to-period growth rates of the low frequency series. Specifically, the method used is to fill-in the missing days using the rates of growth, by day, between two observations of the original series using the standard formula, \( x_t = (1 + r)^t \ x_0 \) for \( t \in [0, T] \) and \( x_T = (1 + r)^T \ x_0 \), where \( x_T \) is the current observation in the original, low frequency data series and \( x_0 \) is the preceding observation. This produces a trend that mimics the trend in the original series, an example of which is shown in Figure 3.
Figure 3:
Original versus created daily series for quarterly gross domestic product.

The strength of this approach is that the underlying trend in the data is preserved from the day of the current observation to the day of the next observation. This is important in order to obtain a stationary price series by removing the presence of unit roots in the price series.

Most of the series included in the macroeconomics part of the information set should be familiar. The price data, and employment were obtained from the Bureau of Labor Statistics, the data for the money supply and industrial productivity were obtained from the Federal Reserve Bank of St. Louis, and the gross domestic product data were obtained from the Bureau of Economic Analysis. The industrial productivity series, INDPROD1, INDPROD3 and INDPROD6 are defined as the proportion of industrial sectors whose productivity increased over the last 1, 3 and 6-months respectively.
Stock Market Variables

Data series designed to capture information from the stock market are the Standard and
Poor’s 500, SP500, which is contrasted with an index of oil producing and manufacturing
companies listed on the NYSE, NASDAQ or AMEX for the years 1983 through 2010. The
data on the S&P 500 were obtained from the Federal Reserve Bank of St. Louis and the data
for oilidx were obtained from CRSP data from Wharton Research Data Services\(^{10}\). Figure 4
shows plots of these two series.

![Figure 4: Plots for Oil Stock index and S&P500 stock market series.](http://wrds-web.wharton.upenn.edu/wrds/index.cfm)

The two stock market series follow very similar trends, but there are some easily noted
differences. Where the S&P 500 followed a downward trend, the index of oil manufacturers

\(^{10}\)http://wrds-web.wharton.upenn.edu/wrds/index.cfm
suffered a steep, one-day drop on July 23, 2001. The SP500 variable is intended to capture overall economic activity inherent in the stock market and we can observe that it is quite similar to the indirect measure of the popularity of stocks issued by petroleum producers.

The Oil Stock index series was created by identifying 38 companies (the companies, identified by CRSP permco are shown in Table 10) and tracking their share price and shares outstanding for each day over the sample period. These companies are all identified over the period 1983 through 2010 as having an NAICS code of 211111 (oil and natural gas extraction), an NAICS code of 324110 (petroleum refineries), a 3-digit sic code of 291 (petroleum refining), or a 3-digit SIC code of 131 (crude petroleum and natural gas extraction). The index was created by multiplying for each company, for each day the share price by the number of shares outstanding, then summing this product over all companies and dividing the sum by the total shares outstanding for all companies. Some of the companies' stocks do not trade every day, so CRSP reports the bid minus ask price. For these cases, the share price used was for the last share price for a day when trading occurred.

Table 10. Companies comprising the variable oilidx.

| ADAMS RESOURCES & ENERGY INC | HOLLY CORP |
| APACHE CORP | IMPERIAL OIL LTD |
| APOCO OIL & GAS INTERNATIONAL INC | MAGELLAN PETROLEUM CORP |
| ARABIAN AMERICAN DEVELOPMENT CO | MURPHY OIL CORP |
| B P PLC | NEXEN INC |
| BARNWELL INDUSTRIES INC | OCCIDENTAL PETROLEUM CORP |
| CHEVRON CORP NY | PETROLEUM DEVELOPMENT CORP |
| CONOCOPHILLIPS | PRIME ENERGY CORP |
| CRDEO PETROLEUM CORP | PYRAMID OIL CO |
| ESSO MOBIL CORP | RANGE RESOURCES CORP |
| FIELDPOINT PETROLEUM CORP | RESOURCE AMERICA INC |
| FRONTIER OIL CORP | SABOL LTD |
| GEORGEKINS INC | SOUTHWESTERN ENERGY CO |
| GEORGEKINS INC | SUNOCO INC |
| GOODROH PETROLEUM CORP | SWIFT ENERGY CO |
| H K N INC | T E L OFFSHORE TRUST |
| HILBORN & PAYNE INC | TESORO CORP |
| HESS CORP | TORNADO RESOURCES CORP |
| | UNIT CORP |

Interest Rate Variables

Data series for interest rates were obtained from the Federal Reserve Bank of St. Louis and include the 90-day treasury bill, TBILL90D, the 10-year, TBOND10Y, and 20-year
treasury bonds, TBOND20Y, Moody’s Aaa, AAA, and Baa, BAA, corporate bond yields. These interest rate data series are designed to capture the financial risk measures noted in the finance literature discussed in Section 2. It should also be noted that these data, unlike other factors in the data set, are available as daily series. Figure 5 shows the plots of these data over time.

Figure 5:
Plots of interest rate variables.

The interest variables were included in the stage 1 model in the forms BAA-AAA, TBILL90D, (TBOND10Y-TBOND20Y)/2-TBILL90D.

Oil Industry Variables
Three series were included to capture industry-specific changes that have occurred over the sample period in the petroleum and gasoline industries. Roughly, these series are designed to account for some possible changes in the supply of and demand for crude oil and gasoline in the United States. All of these variables were obtained from the U.S. Energy Information Agency and are available as weekly series. Figure 6 shows the graphs of these 3 series over time.

![Figure 6: Plots of oil and gas industry data.](image)

All of the series show considerable variability with fairly noticeable trends.

Event Variables
The analysis includes two variables to account for petroleum supply shocks. The first variable, HURCANE, is a dummy variable for the presence of an hurricane\textsuperscript{11} in the Gulf of Mexico on a particular day. No distinction was made for the presence of more than one hurricane on any given day. There were 167 days with the presence of a hurricane over the sample period from 27 named hurricanes.

The other variable is a dummy variable to indicate the existence on a particular day of a conflict deemed to have an impact on the supply of oil. There were three conflicts\textsuperscript{12} over the period sampled: the Iraq oil export suspension (June 3, 2001 through July 11, 2001), the oil strike in Venezuela (December 2, 2002 through March 30, 2003), and the Iraq war (March 19, 2003 through December 31, 2003). There were 299 days, identified in the sample, that were impacted by one or more of these conflicts.

Missing Data

Some of the daily data series had missing observations when daily data was unavailable. For example, the retail gas series, GASRETP, had 87 missing observations, all prior to 2001. The petroleum futures data series, OILFUTP, had 27 missing observations, most prior to 2007. There are a variety of strategies for handling missing data. One approach when a missing observation is encountered for a particular variable is to omit that observation for every variable. Another strategy is to use the last present observation to fill-in for the missing observation. A third strategy is to fill-in the missing observation using an algorithm. Algorithms can be complex, relying on information from outside the model, or simple such as using a rolling mean value, or linear interpolation. Because the number of missing observations was a small percentage of overall data, simple interpolation was used to generate values for the missing observations.

\textsuperscript{11}Source: http://www.nhc.noaa.gov
Results

Step 1

The results from the estimation of the step 1 regressions are shown in Table 11. The estimates were obtained from individual ordinary least squares regressions of the market price variables on the set of common factor variables (the natural logarithms of all variables were used in the estimation). No corrections were made for any possible violations of the assumptions that the error terms have zero mean and covariance matrix $\sigma^2 I$. This may call into question the reliability of the estimated standard errors presented in Table 11. Because of this, the focus is on estimates with estimated standard errors that provide for a significance level of at least 0.001.

Table 11. Parameter estimates from stage 1 OLS estimates (only estimates with p-value less than .001 are shown).

Most of the variables show that changes in the values of the common factors have significant impacts on market prices for all five markets. The estimated coefficient for money supply, M1, and consumer price index, CPI, tend to not be significantly different from zero.
The estimate for CPI does have a significant, positive impact on retail gas prices, which is not unexpected given that energy prices make up a large portion of the measure of CPI. It is surprising that increases in the CPI tend to decrease the futures price of oil. On its face this would imply an anticipated decline in the spot price of oil when the CPI increases (or increasing inflation rates bring an anticipation of lower oil spot prices in the future). One possible explanation is that an increase in consumer prices reduces demand for petroleum products resulting in lower oil prices.

Most economic variables show consistent impacts across all markets except for nominal gross domestic product, NGDP. NGDP has a positive impact on all commodity prices except retail gasoline prices. One could view an increasing GDP as resulting in an increase in the demand for gasoline and an increase in the supply, and the results here show that the supply increases are somewhat dominant. It is also possible that the relationship to be found is one that associates increasing retail gasoline prices with recession, and decreasing retail gasoline prices with economic expansion.

As expected, an increase in the producer price index leads to higher petroleum prices across all commodity markets, although impacts are much less in the market for gasoline futures, but are only somewhat less in the gasoline spot market. It may not be surprising that the futures market would react differently to the presence of inflation than the spot market, but that is not the case for oil. If we examine a gasoline futures contract as a financial asset, then it would appear that inflation risks are less important than they are for oil futures contracts.

There are different ways to view the results with respect to the Standard and Poor’s 500 stock market index, SP500. One way is to view it is a broad indicator of macroeconomic performance. One would expect that if the macroeconomy is expanding, the SP500 will increase. Another view is that SP500 is a measure of an expected or average return on an asset. Thus, when SP500 is increasing one might expect demand for other assets to decline. Still another way is to view SP500 as a measure of missing pricing information not included.
in other economic factors (see for example Chen, Roll and Ross (1986)). In this respect, the results indicate that the overall stock market performance does provide information in the pricing of petroleum-related commodities.

The industrial productivity “diffusion indexes are calculated as the percentage of series that increased over the indicated span (one, three, or six months) plus one-half the percentage that were unchanged.” These diffusion indexes show similar impacts in all commodity markets. As the one- and six-month indexes increase, commodity prices increase, but as the three-month index increases, commodity prices fall. One explanation is that productivity in sectors that demand petroleum products initially increases more rapidly than productivity in sectors that supply petroleum products, which causes petroleum commodity prices to rise. Over the three-month time frame sectors involved with petroleum supply show more rapid increases in productivity, which results in a decrease in petroleum commodity prices. Finally, over the six-month window, productivity increases more rapidly in heavy demand sectors, which, again, causes prices to increase.

The interest rate measures, the default spread, DEFSPREAD, which is Moody’s Baa corporate bond yield minus their Aaa rates corporate bond yield, the term rate measure, TERMRT, which is the average long-term treasury bond yield minus the 90-day treasury bill rate, and the 90-day treasure bill rate, TBILL90D. From the finance literature we anticipate that DEFSPREAD is an ex-ante measure of time varying risk premia reflecting systematic risk of underlying commodities related to long-term business cycle conditions. The TERMRT is a measure of storage costs and is related to short-term business cycles. The TBILL90D has been found to be negatively related to future stock market returns and is seen as a proxy for expected future economic activity. See Chordia and Shivakumar (2002) for an explanation.

Results indicate that as DEFSPREAD increases, which is associated with recessionary conditions, commodity prices fall and suggests that demand conditions are more relevant. Both short-run, TERMRT, and long-run, TBILL90D, business cycle measures are positively

\[ \text{Source: http://research.stlouisfed.org/fred2/} \]
related to commodity prices. Thus, both short-run and long-run economic expansions resulting in increased commodity prices appear to be associated with increases in demand for petroleum commodities across all five markets.

Factors that measure supply and demand for petroleum products include gross inputs into refineries, REFINPTQ, weekly net production of motor gasoline, GASSUPQ, and weekly product supplied of finished motor gasoline, GASDEMQ. REFINPTQ is a measure of gross inputs into refineries. Not surprisingly, inputs into refineries have no significant impact on oil market prices, but are positively related to gasoline commodity prices. This would indicate that refinery inputs are a reaction to increases in demand, and are thus associated with higher prices. One would naturally assume that gross inputs into refineries would be a variable that induced supply changes, but results do not agree. Instead, it appears that supply is chasing demand and even though REFINPTQ is ultimately responsible for increasing the supply of finished motor gasoline, it is associated with increasing gasoline prices. This interpretation is supported by the relative impacts in the three gasoline commodity markets. The gasoline futures market price is impacted much more than the spot and retail prices, which suggests a future increase in gas prices is associated with an increase in demand over supply.

GASSUPQ is typically viewed as a measure of gasoline supply. Results support this notion as all market prices are inversely related to changes in GASSUPQ. Not surprisingly, retail gasoline prices show the smallest impact from changes in gasoline supply as consumer demand for gasoline is quite price inelastic. GASDEMQ is typically viewed as a measure of the demand for gasoline. Results support this view as increases in the demand for gasoline cause increases in commodity prices in all markets.

Step 2

The residuals from step 1 were tested for the presence of unit roots and cointegration. As shown in Table 12, the Elliott-Rothenberg-Stock DF-GLS test shows no unit roots, and the
Johansen Cointegration trace test shows that we can reject the null hypothesis of four or fewer cointegrated series. However, separate Jarque-Bera tests demonstrate the null hypothesis that the series are normally distributed can be rejected.

Table 12. Test results for unit root and cointegration for step 1 residuals.

The series are all shown to be leptokurtic, which can be seen in the histograms shown in Figure 7.
Tables 13 through 17 show results from the step 2 VAR estimation. Results are significant at a p-level of at least 0.10. Table 13 gives the results for the oil futures price. We note a couple of things from Table 13: retail gasoline prices have little feedback into oil futures prices, and oil markets tend to have a much larger impact on the oil futures price than do gasoline markets.
Table 13. VAR model significant coefficient estimates for oil futures price

Step 2 uses a vector autoregression model (VAR) based upon the results of the Johansen cointegration test presented above. The VAR system will allow for the exploration of the interrelationships between the five commodity markets, and to provide a means to further analyze these markets in terms of price velocities (step 3). A VAR(21) system was estimated. The maximum lag value \( p = 21 \) was chosen based upon the final prediction error, FPE, where

\[
FPE = \min_p \left\{ \left[ \frac{T + Kp + 1}{T - Kp - 1} \right]^K \left| \bar{\Sigma}_e(p) \right| \right\}
\]

where \( \bar{\Sigma}_e \) is the maximum likelihood estimate of the variance-covariance matrix. Although the FPE is not an ideal selection criterion for unstable models, because of the stability properties in our data, the FPE was chosen. Another reason for choosing the FPE is because one goal of this paper is to determine responses to price shocks, which essentially involve a one-step ahead forecast, and because the use of daily data naturally suggests the possibility of significant impacts over a larger number of lags. Finally, the FPE, like the AIC, has been found to be consistent for large lag processes. See Lütkepohl (2006).
From Table 14, we can see that oil market prices have a significant influence on gasoline futures prices. Not surprisingly, the variable with the greatest impact on the gasoline futures price is the one-day lagged gasoline futures price. In Table 15, we observe the significant impacts on the spot price of oil.

**Table 14. VAR model significant coefficient estimates for gasoline futures price**

| GASFUTP | Lag | Estimate | Std. Error | Pr(b|t) |
|---------|-----|----------|------------|------|
| OILFUTP | 1   | 1.219    | 0.047      | 0.000 |
| OILFUTP | 3   | 1.76     | 0.054      | 0.012 |
| OILFUTP | 6   | 1.32     | 0.054      | 0.016 |
| OILFUTP | 8   | 1.16     | 0.054      | 0.023 |
| OILFUTP | 13  | 1.03     | 0.054      | 0.050 |
| GASFUTP | 1   | 0.99     | 0.035      | 0.000 |
| GASFUTP | 6   | 0.82     | 0.047      | 0.027 |
| GASFUTP | 20  | 0.66     | 0.047      | 0.036 |
| OILSPOT | 1   | 0.95     | 0.045      | 0.034 |
| OILSPOT | 2   | 0.96     | 0.050      | 0.037 |
| OILSPOT | 3   | 0.112    | 0.050      | 0.268 |
| OILSPOT | 5   | 0.08     | 0.050      | 0.004 |
| OILSPOT | 2   | 0.06     | 0.034      | 0.004 |
| OILSPOT | 9   | 0.06     | 0.038      | 0.034 |
| OILSPOT | 17  | 0.06     | 0.036      | 0.010 |
| OILSPOT | 18  | 0.01     | 0.036      | 0.001 |
| OILSPOT | 1   | 1.14     | 0.086      | 0.028 |
| OILSPOT | 11  | 1.73     | 0.073      | 0.036 |

**Table 15. VAR model significant coefficient estimates for oil spot price**

| OILSPOT | Lag | Estimate | Std. Error | Pr(b|t) |
|---------|-----|----------|------------|------|
| OILFUTP | 1   | 1.19     | 0.050      | 0.000 |
| OILFUTP | 3   | 0.32     | 0.050      | 0.000 |
| OILFUTP | 8   | 0.07     | 0.057      | 0.510 |
| OILFUTP | 13  | 2.14     | 0.057      | 0.001 |
| OILFUTP | 17  | 0.09     | 0.057      | 0.000 |
| OILFUTP | 20  | 1.26     | 0.057      | 0.000 |
| GASSPOT | 1   | 0.20     | 0.038      | 0.000 |
| GASSPOT | 2   | 1.21     | 0.030      | 0.167 |
| GASSPOT | 9   | 0.14     | 0.050      | 0.000 |
| GASSPOT | 20  | 0.08     | 0.049      | 0.753 |
| GASSPOT | 20  | 0.68     | 0.038      | 0.015 |
| OILSPOT | 1   | 0.33     | 0.048      | 0.000 |
| OILSPOT | 3   | 0.16     | 0.053      | 0.016 |
| OILSPOT | 5   | 0.10     | 0.053      | 0.049 |
| OILSPOT | 8   | 0.09     | 0.053      | 0.710 |
| OILSPOT | 13  | 0.11     | 0.052      | 0.010 |
| OILSPOT | 16  | 0.15     | 0.052      | 0.000 |
| GASSPOT | 8   | 0.06     | 0.037      | 0.000 |
| GASSPOT | 9   | 0.08     | 0.037      | 0.000 |
| GASSPOT | 11  | 0.06     | 0.037      | 0.000 |
| GASSPOT | 13  | 0.06     | 0.037      | 0.000 |
| GASRETP | 8   | 0.18     | 0.084      | 0.053 |
| GASRETP | 12  | 0.11     | 0.084      | 0.018 |
| GASRETP | 17  | 0.21     | 0.084      | 0.000 |
| GASRETP | 20  | 0.13     | 0.084      | 0.000 |

Table 15. VAR model significant coefficient estimates for oil spot price
The spot price of oil shows more impacts from oil markets than gasoline markets, but are influenced by gasoline markets than is the oil futures price. Gasoline spot and retail prices have significant impacts on the oil spot price only at longer lags. Table 16 shows significant coefficients for the spot price of gasoline. Here we observe that changes in the oil futures price shows the greatest number of significant lagged effects on the gasoline spot price. Additionally, we see that impacts from gasoline markets tend to be larger in magnitude than impacts from oil markets. As with the oil spot market, changes in gasoline retail prices influence the spot price of gasoline only at longer lags of more than one week.

| Lag | GASSPOTP Estimate | Std. Error | Pr(>|t|) |
|-----|------------------|------------|---------|
| OILFUTP | 1 | .280 | .056 | 0.0000 |
| OILFUTP | 3 | .174 | .065 | 0.0073 |
| OILFUTP | 8 | -.191 | .065 | 0.0029 |
| OILFUTP | 12 | .111 | .064 | 0.0347 |
| OILFUTP | 13 | .146 | .064 | 0.0295 |
| OILFUTP | 15 | .128 | .064 | 0.0316 |
| GASFUTP | 1 | .381 | .042 | 0.0000 |
| GASFUTP | 2 | .242 | .056 | 0.0002 |
| GASFUTP | 6 | -.101 | .056 | 0.1751 |
| GASFUTP | 13 | .117 | .055 | 0.0341 |
| OILSPO1P | 1 | .119 | .054 | 0.0202 |
| OILSPO1P | 3 | -.110 | .060 | 0.0646 |
| OILSPO1P | 8 | .142 | .059 | 0.0266 |
| OILSPO1P | 12 | -.106 | .059 | 0.2789 |
| GASSPOTP | 1 | .506 | .032 | 0.0000 |
| GASSPOTP | 2 | .139 | .041 | 0.0040 |
| GASSPOTP | 9 | .055 | .040 | 0.3100 |
| GASSRIP | 8 | .156 | .095 | 0.0899 |
| GASSRIP | 14 | .172 | .094 | 0.0606 |
| GASSRIP | 15 | -.212 | .094 | 0.0278 |

Table 16. VAR model significant coefficient estimates for gasoline spot price

Impacts observed in the retail gasoline market are found in Table 17. All markets’ prices have significant one-day lag effects on retail gasoline prices except for oil spot prices. Many of the markets show impacts at longer lags, which suggests that market effects must occur before they filter into the retail gasoline market. Thus, we observe that price changes in more “upstream” markets have a predictable immediate effect, which is followed some days later by secondary effects.
Table 17. VAR model significant coefficient estimates for gasoline retail price

Step 2 results show that the oil futures market has more impacts on other markets than other markets have on oil futures prices. Changes in retail gasoline prices do show some limited impacts in all markets, but generally at lags of a week or more, whereas future and spot markets tend to show more immediate impacts.

Step 3

As noted in Section 3, the residuals from stage 2 are examined using a variety of tests. The stage 2 residuals were tested for unit roots, cointegration, normality, and the presence of ARCH effects. Table 18 summarizes these tests. No unit roots are observed, the series is stable (the null hypothesis of four or less cointegrated series is rejected), the series are not normal, and ARCH effects are present.
Table 18. Diagnostic test results for stage 2 residuals.

The PC-GARCH model outlined in Section 3 above was estimated with the state 2 residuals. The extended autocorrelation function was examined for each series $\eta_i$ in stage 3. The results of the eacf analysis indicated that the most appropriate models were univariate GARCH(1,1) models. The coefficient estimates are shown in Table 19.
Table 19. Univariate GARCH(1,1) estimates for $\eta_i$

All coefficient estimates for the GARCH(1,1) models are significant. After estimating the step 3 PC-GARCH model, the estimated volatilities were formed as the variance of the step 2 residuals.

Step 4

Step 4 involves combining the residuals from step 1 with the estimated volatilities from step 3 into a ten-equation VAR(p) model. Summaries for the ten equations are shown in Table 20.
We note that the fit for most equations is quite good. Two exceptions are the equations for the volatilities of oil spot prices and retail gasoline prices. It is not surprising that volatility equations would be more difficult to fit, if the volatilities were less cyclical in nature. Recalling that GARCH processes are such that volatilities are cyclical (volatilities are high for some periods and then low for some periods), these results suggest that volatilities for the oil spot and retail gasoline markets do not follow such volatility patterns as closely.

<table>
<thead>
<tr>
<th></th>
<th>R-sq</th>
<th>Adj-R-sq</th>
<th>F(21,3519)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OILFUTP</td>
<td>0.9548</td>
<td>0.9521</td>
<td>352.3</td>
<td>0.0000</td>
</tr>
<tr>
<td>GASFUTP</td>
<td>0.9611</td>
<td>0.9588</td>
<td>413.2</td>
<td>0.0000</td>
</tr>
<tr>
<td>OILSPOTP</td>
<td>0.9535</td>
<td>0.9507</td>
<td>342.4</td>
<td>0.0000</td>
</tr>
<tr>
<td>GASSPOTP</td>
<td>0.944</td>
<td>0.9307</td>
<td>291.8</td>
<td>0.0000</td>
</tr>
<tr>
<td>GASRETP</td>
<td>0.9608</td>
<td>0.9712</td>
<td>855.5</td>
<td>0.0000</td>
</tr>
<tr>
<td>OILFUTV</td>
<td>0.9572</td>
<td>0.9546</td>
<td>373.7</td>
<td>0.0000</td>
</tr>
<tr>
<td>GASFUTV</td>
<td>0.9572</td>
<td>0.9546</td>
<td>373.7</td>
<td>0.0000</td>
</tr>
<tr>
<td>OILSPOTV</td>
<td>0.8889</td>
<td>0.8703</td>
<td>37.92</td>
<td>0.0000</td>
</tr>
<tr>
<td>GASSPOTV</td>
<td>0.9332</td>
<td>0.9392</td>
<td>233.5</td>
<td>0.0000</td>
</tr>
<tr>
<td>GASRETV</td>
<td>0.3386</td>
<td>0.2992</td>
<td>8.663</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 20. Summaries for Step 4 VAR(21).
Figure 8: Lag effects from prices on oil futures prices (OILFUTP) from Step 4 VAR(21) presented as elasticities at mean values.

Figure 8 shows impacts on the oil futures price from the estimated VAR(21). This figure shows the impact on the oil futures price at period zero from the prices in each of the five markets in the 21 prior periods (where impacts are shown as elasticities $\frac{dP_i}{dP_j} \frac{P_j}{P_i}$ for $i = OILFUTP$ and $j = \{OILFUTP, GASFUTP, OILSPOTP, GASSPOTP, GASRETP\}$ at the mean values). A couple of points observed are that oil market prices tend to dominate and are counter-cyclical with one another, and when oil past futures prices are increasing the current oil futures price, the contemporaneous past oil spot prices are decreasing the current oil futures price. For gasoline market prices, the futures price of gasoline has the largest impact on the oil futures price while the impacts from retail ad spot gasoline markets are relatively small. The impulse response function is shown in Figure 9.
Figure 9:
Price and volatility impacts on OILFUTP from individual market prices and price volatility impulses with 95 percent confidence bands.

Figure 9 shows price responses from individual price and volatility shocks. We can see that, among price changes, only the changes in the oil futures price, the gasoline futures price and the oil spot price causes a significant change in the oil futures price.
Figure 10: Lag effects from price volatilities on oil futures prices (OILFUTP) from Step 4 VAR(21) presented as elasticities at mean values.

Figure 10 shows the estimated impacts on the current oil futures price from past values of price volatilities. The oil futures price volatility clearly displays the largest impact on the current oil futures price. The only other two market volatilities that impact the oil futures price are the price volatilities of gas futures and gas spots. The impulse response functions shown in Figure 10 indicate that only the oil futures price volatility has a significant impact on the oil futures price. Figure 11 shows the estimated impacts on the gas futures price from lagged market prices.
Figure 11: Lag effects from prices on gas futures prices (GASFUTP) from Step 4 VAR(21) presented as elasticities at mean values.

We observe from Figure 11 that the largest impact on the current gas futures price comes from the immediate prior period price. Other than that, we note that the past oil futures prices and oil spot prices tend to dominate even past gas futures prices. Thus it appears that the market for gasoline futures is somewhat heavily influenced by prices in the oil markets. Figure 12 shows the impulse response functions for gasoline futures prices. Apart from some marginally significant impacts of oil futures prices early and oil spot prices later, the only significant price impacts come from gasoline futures prices.
Figure 12:
Price and volatility impacts on GASFUTP from individual market prices and price volatility impulses with 95 percent confidence bands.

Figure 12 also shows that oil futures price volatility has a large significant impact on gasoline futures prices. The increase in the volatility of oil futures prices that occur when a hurricane event is present in the Gulf of Mexico eventually results in an increase in gasoline futures prices of about seven cents per gallon.
Figure 13 shows price volatility impacts on the current period gas futures price. Interestingly, the volatility of the oil futures price has the largest impact on the current gas futures price, and the impacts are largest from the volatilities more than 12 days past. Gas futures price volatility impacts tend to be pro-cyclical with oil futures price volatility and the only other relatively large impact come from the gas spot price volatility, which appears to be counter-cyclical with the futures market price volatilities.
Figure 14:
Lag effects from prices on oil spot prices (OILSPOTP) from Step 4 VAR(21) presented as elasticities at mean values.

Figure 14 shows lagged price effects on the current oil spot price. Both oil and gasoline futures prices show larger impacts at more current lags (although the largest impact comes from the first period lag of the oil spot price), and prices in both futures markets tend to show the largest effects in oil spot prices over later lags also. Interestingly, the effects from lagged oil spot prices and lagged oil futures prices appear to counter each other. Figure 15, which shows the impulse responses from price and volatility changes in only one market, confirm these effects.
Figure 15:
Price and volatility impacts on OILSPOTP from individual market prices and price volatility impulses with 95 percent confidence bands.

Oil spot price impacts from price volatilities in all five commodity markets are shown in Figure 16. As with the futures markets for oil and gasoline, the largest impact on the current period oil spot price from price volatility comes from the volatility of oil futures prices. The volatilities of gasoline spot and futures prices also affect the oil spot price to a much greater degree than the oil spot price volatility itself, but both still have much smaller impacts than oil futures price volatility. Impulse responses shown in Figure 16 show that only oil futures price volatility has a significant impact on the oil spot price.
Figure 16:
Lag effects from price volatilities on oil spot prices (OILSPOTP) from Step 4 VAR(21) presented as elasticities at mean values.

The price impacts on the current period gasoline spot price are shown in Figure 17. The largest impacts occur in the previous two lag periods and come from the gasoline spot price, the gasoline futures price, and the oil futures price. The futures prices have impacts as large as those of the one-lagged gasoline price itself. Compared to other commodity markets, the gasoline spot market appears to be more broadly affected by prices in other markets.
Figure 17: Lag effects from prices on gasoline spot prices (GASSPOTP) from Step 4 VAR(21) presented as elasticities at mean values.

Figure 18 shows the impulse responses for gasoline spot prices from individual changes in commodity prices and volatilities. We observe that significant price impacts come from changes in gasoline futures prices and gasoline spot prices. Oil futures prices create significant impacts, but only very early. Oil spot prices impact gasoline spot prices, but only after approximately 24 days.
Figure 18: 
Price and volatility impacts on GASSPOTP from individual market prices and price volatility impulses with 95 percent confidence bands.

Figure 19 gives the impacts from lagged volatilities on the current period gasoline spot price. Once again we observe that the greatest impact comes from the oil futures price volatility. All impacts appear to dampen as impacts at longer lags provide greater impacts. Gasoline futures price volatilities appear to be as important in affecting gasoline spot prices as are gasoline spot price volatilities. As with other markets, Figure 19 shows that oil futures volatility causes gasoline spot prices to eventually increase around seven cents per gallon. Other volatility changes have little effect on gasoline spot prices.
Figure 19:
Lag effects from price volatilities on gasoline spot prices (GASSPOTP) from Step 4 VAR(21) presented as elasticities at mean values.

Impacts on current period retail gasoline prices from lagged prices in all five commodity markets are found in Figure 20. Over the 21-day lag period the oil futures price is prominent until the earliest five lags, the spot price of oil is somewhat more consistent over all lags, and the futures price of gasoline does not appear to have any particularly large impacts. The most significant impact comes from the first lag of the retail gasoline price.
Figure 20: Lag effects from prices on retail gasoline prices (GASRETP) from Step 4 VAR(21) presented as elasticities at mean values.

Figure 21 shows impacts on retail gasoline prices from changes in prices in each market separately. Oil futures prices tend to push retail gas prices slightly lower over the 28-day period. Gasoline spot prices and gasoline futures prices tend to push retail gasoline prices slightly higher by 1 to 1.5 cents per gallon. The biggest impact on retail gasoline prices comes from the price response of lagged gasoline retail prices.
Figure 21:
Price and volatility impacts on GASRETP from individual market prices and price volatility impulses with 95 percent confidence bands.

The impacts from price volatilities on the retail gas price are shown in Figure 22. Once again the price volatility of oil futures provides the greatest impact. Other gasoline market price volatilities provide some steady impacts on retail gasoline prices, but retail gasoline price volatility has little effect on retail gasoline prices. Figure 22 confirms the responsive to oil futures price volatility, which push retail gasoline prices lower during the hurricane event, but afterward tend to increase retail gasoline prices by as much as three cents per gallon.
As noted above, the results from step 4 are used to compute impacts from a Gulf of Mexico hurricane on oil and gas commodity markets. Initially prices in all markets are allowed to change as a result of the hurricane. The data from step 1 were used to compute the price changes in each market over a seven day hurricane event. The results obtained in step 1 show that days when a hurricane is present have the following price impacts.

Figure 22:
Lag effects from price volatilities on retail gasoline prices (GASRETP) from Step 4 VAR(21) presented as elasticities at mean values.
Figure 23:
Change in commodity market prices from the presence of a hurricane in the Gulf of Mexico. Results obtained from step 1.

We can observe in Figure 23 that on a day when an hurricane is present in the Gulf of Mexico, prices in all five commodity markets are higher. Gasoline spot prices increase the most (almost seven cents per gallon), while retail gasoline prices increase the least (less than three cents per gallon). Also as noted above, it would be expected that price increases would vary by day over the life of the hurricane. The daily price patterns estimated in separate step 1 regressions are shown in Figure 24.
In Figure 24 we see how oil and gasoline commodity market prices evolve over time during a Gulf of Mexico hurricane event. The general pattern of price changes is similar in all five markets. On day one prices in all markets increase by at least fifty percent of the averages generated in the step 1 regression models. These initial price increases do not change much until day 5 of the hurricane event. This probably reflects the average time it takes for a hurricane that forms in or enters the Gulf of Mexico to move close enough to oil producing and refining areas to more directly impact supplies. Prices remain high over the next two or three days, but then drop off fairly rapidly (again most likely as the hurricane moves away from oil producing and refining areas).

The price responses in all markets to the price impacts from a seven-day hurricane when prices in all five markets are allowed to change are shown in Figure 25.
Figure 25:
Commodity market price and price volatility responses to market price changes from a hurricane event including 95 percent confidence bands.

The impulse responses shown in Figure 25, as noted in section three, are calculated as mean responses from 250 randomly drawn starting days. The upper and lower 95-percent confidence bounds are shown as well. Looking first at the price responses to simultaneous changes in only market prices in the five commodity markets, we observe that the oil futures price increases over six cents per gallon over the first couple of days (these prices are usually reported as dollars per barrel, which would be an increase of over $2.50 per barrel) before dropping somewhat over days three and four (this is almost universally true in all five...
markets). Prices increase again over days five and six, increasing to a high of almost nine cents per gallon more than the pre-hurricane price. The nine cent increase is more than the maximum increase attributable to the hurricane itself, which is estimated to be about 7.6 cents per gallon. Over the remainder of the 28 day simulation period, oil futures prices steadily decrease to be about four cents per gallon more than they were before the hurricane.

Gasoline spot prices show a pattern similar to that of prices in the oil futures market. Prices increase as the hurricane enters the Gulf, drop somewhat more sharply, then increase sharply to a maximum of about 7.3 cents per gallon over pre-hurricane levels (the maximum increase estimated from the hurricane itself is about 7.2 cents per gallon) and then drop steadily over the next three weeks end up about 2.7 cents per gallon more than they were prior to the hurricane event. Oil spot prices increase over the first six days of the hurricane about 9.5 cents per gallon, which is almost two cents per gallon more than the amount of increase attributable to the hurricane event (7.75 cents per gallon). Oil spot prices also then drop steadily to be about four cents per gallon more than they were before the hurricane event.

Prices in the gasoline spot market increase about 7.5 cents a gallon in the first two days of the hurricane event, but drop about two cents per gallon over the next two days. Over days five and six, prices again increase to a maximum of about 9.3 cents per gallon, which is less than the estimated maximum price increase attributable to the hurricane of almost 10 cents per gallon. Prices then fall over the next three weeks to be about 3.5 cents per gallon more than they were before the hurricane event.

Retail gasoline prices have the smallest estimated direct price impacts from hurricane events (a first day increase of just over two cents per gallon up to a maximum increase of about 4.26 cents per gallon on day six of the event), which is not surprising given that there are likely to be sufficient amounts of gasoline inventories available and our prices are average prices for the entire United States. However, Figure 25 clearly shows that price increases in the other four commodity markets have significant effects on retail gasoline prices. Retail
gasoline prices increase about 4.5 cents per gallon over the first two days of the hurricane event before falling about one-half cent per gallon over the next three days in the usual pattern. On days six and seven, retail gas prices again increase, rising to be about six cents per gallon more than they were prior to the hurricane event. Retail gasoline prices then fall modestly (about one-half cent per gallon) over the next seven to ten days before falling more rapidly through day 28, ending up about four cents per gallon higher than they were before the hurricane.

Price volatilities, for the most part, do not change significantly as a result of market price increases. Volatilities tend to increase initially in most markets before falling back to be not significantly different from the observed baseline volatilities. Volatilities in gasoline markets tend to increase more than volatilities in oil markets. One should keep in mind when examining Figure 25 that only estimated price changes have been imposed. We also examine impacts from estimated changes in price volatilities associated with a simulated seven-day hurricane event.
Figure 26: Commodity market price and price volatility responses to only market price volatility changes from a hurricane event including 95 percent confidence bands.

Figure 26 shows price and volatility responses from changes in volatility associated with a typical seven-day hurricane event in the Gulf of Mexico. The volatility impulses were calculated by running separate individual linear regressions of days one through seven of the presence of a hurricane in the Gulf of Mexico on commodity market price volatilities used in step 4. This provided separate market impacts on price volatilities over the seven days of the simulated hurricane event. We see that in all markets the increases in price volatilities initially work to drive market prices lower. After seven to ten days, market prices begin to
recover and in all markets except retail gasoline, prices end up between five and ten cents per
gallon more than before the hurricane. Retail gasoline prices also end up higher, but their
increase is only between two and three cents per gallon. Volatilities in all markets increase
and remain higher over the duration of the study period.

Figure 27 shows the evolution of prices and price volatilities when both vary simulta-
neously. Once again we see the price effects of a seven-day hurricane event in the Gulf of
Mexico. Prices in each market are allowed to increase as well as price volatilities. In all
markets after the immediate price effects from the hurricane event, prices are pushed lower
as price volatilities\(^{14}\) increase. Prices in the oil futures, oil spot and gasoline spot markets
then drop back to near pre-hurricane levels through day ten. After the tenth day, prices
begin to increase as price volatilities continue to decline. In the gasoline futures market,
prices see-saw for the first ten days at an average of about two cents per gallon higher than
they were in the baseline. After the tenth day gasoline futures prices increase rapidly until
they are nearly 15 cents per gallon higher than before the hurricane. Over the last ten
days of the study period gasoline futures prices fall to be about 8.5 cents per gallon above
pre-hurricane levels. Retail gasoline prices respond by increasing over the first two days and
then decreasing for the most part until the tenth day. After the tenth day, retail gasoline
prices increase again over the next week until they are about eight cents per gallon more
than they were prior to the hurricane. Over the final week or so of the study period, retail
gasoline prices decline about 2 cents per gallon to end up almost six cents per gallon higher
than they were before the hurricane event.

\(^{14}\)I observed that the presence of a hurricane caused price volatilities to increase in all markets although
the increase in the retail gasoline market is initially much smaller than in the other four markets.
Figure 27:
Commodity market price and price volatility responses to changes in market price and market price volatility from a hurricane event with 95 percent confidence bands.

Overall when I account for price and volatility impacts we observe that after nearly two weeks market prices start to rise well above what we would expect from a seven-day hurricane event. When comparing price response patterns between price effects and volatility effects, it appears that volatility effects have a significant influence in the overall evolution of market prices reacting to a Gulf of Mexico hurricane event.

Conclusion
This paper has estimated price and price volatility responses between five commodity markets in the face of a Gulf of Mexico hurricane. An hurricane event was chosen because hurricanes tend to be of short, well-defined duration and would be expected to impact both oil production, shipping and gasoline production. The analysis was done in four steps. Step 1 used ordinary least squares regressions of market prices on several macroeconomic, interest rate, stock market, and industry-specific variables in order to gauge market price impacts from common factors and to generate stationary price series for step 2 of the model. The second step employed a VAR(21) system to determine influences of price shocks throughout the five markets. The presence of conditional heteroskedasticity in the residuals from step 2 made it appropriate to employ an MGARCH model for step 3 in order to obtain estimates of the price volatilities associated with the five commodity markets. Estimated price volatilities from step 3 were combined with residuals from step 1 to estimate a ten-equation VAR(21) model.

The results show several conclusions. I find that price lags of up to 21 days are significant. Not surprisingly, the single biggest impact on daily commodity prices comes from own price lagged one day. Price responses to the presence of an hurricane in the Gulf of Mexico tend to follow a pattern more associated with responses to changes in price volatilities than with direct changes in prices. Without accounting for changing price volatilities, average price changes during the presence of a Gulf hurricane are as high as nearly seven cents per gallon (gasoline spot market), to a low of just under three cents per gallon (retail gasoline market). Price volatility in the oil futures market tends to have large significant impacts on prices in all commodity markets. Retail gasoline prices show the smallest responses to hurricane events, but tend to be influenced by prices in several different markets. Price volatilities do not appear to be significantly impacted by changes in market prices.

Our results further show that a seven-day hurricane event in the Gulf of Mexico will result in increased commodity prices in all five markets over a 28-day period. Oil futures
prices are from 1 cent to nearly 15 cents per gallon higher. Gasoline futures prices are from one-half cent to over 14 cents per gallon higher. Oil spot prices are from one to nearly 14 cents per gallon higher and tend to remain at higher levels longer than prices in other markets. Gasoline spot prices increase about seven cents per gallon during the hurricane, and eventually increase to be about 13 cents per gallon higher before ending up about 8 cents per gallon higher at the end of the 28-day period. Finally, retail gasoline prices are from 3 cents to 7.5 cents per gallon more over the 28 days, ending up about 5.5 cents per gallon higher.

These results provide evidence of the relationships between oil and gas commodities. They also show that models of commodity prices must take into account changes in price volatilities as well as price changes in other markets. Finally, our results confirm the notion that prices in commodity markets respond more to changes in some commodity markets than others.

Chapter II

Introduction

Problems associated with traffic congestion continue to increase, but politically viable solutions continue to evade implementation. The typical government response to congestion problems is to increase roadway capacity by building new roads. However, Duranton and Turner (2011) recently showed that increasing the supply of roads does not eliminate traffic congestion. In this paper I propose a simple model of transportation that conforms to the findings of Duranton and Turner (2011) and comports with the requirements of the fundamental law of traffic congestion noted by Downs (1960 and 1992).
While there is a fairly wide body of literature that has focused on traffic congestion and associated private and social costs. Another separate but related topic, the examination of the economic impacts of road maintenance, has received much less attention. The model employed in this paper also examines impacts on welfare, output, employment and land consumption that arise from deteriorating road quality and government expenditures for road maintenance in the face of traffic congestion and mitigation efforts. These topics are explored in the context of an applied or computable general equilibrium model.

The examination of traffic congestion began in earnest in the 1960’s. These efforts built upon spatial models of urban areas which noted the existence of what has been called the bid-rent curve (See Alonzo (1964)). The bid-rent curve is a description of land prices at various distances from the city center. The typical bid-rent curve shows decreasing land prices as one moves farther from the city center. Much of the subsequent research into traffic congestion has focused on the shape of the bid-rent curve in the face of increasing transportation costs that arise as a result of congestion, and the use of congestion tolls to privatize the social costs of traffic congestion.

The purpose of this paper is to present a flexible model that allows for the tracking of the private and social costs associated with road quality deterioration and government efforts to maintain roads. The structure of the model is similar to the class of models within the existing literature that have been used to examine transportation costs. The focus, however, is to quantify the economic impacts from a variety of different government policies designed to alleviate road quality deterioration and, to some extent, traffic congestion.

Despite the widespread recognition of the efficacy associated with the use of congestion tolls, the reality is that congestion toll policies are rarely adopted. Instead government policies are often adopted that seek to decrease traffic congestion through the construction of new roadways. Additionally, given budget constraints, governments often adopt policies that do not provide for road maintenance at levels necessary to prevent the road surface quality from declining.
This paper presents a closed general equilibrium model of a monocentric city in which land is allocated to production, housing and roads. Although the term “monocentric city” is adopted to describe the import of the model, the model actually describes an urban system where households and firms incur transportation costs in a way that influences their economic decisions. The model is used to answer the following questions. What is the cost in terms of output and welfare from road quality deterioration in light of traffic congestion? What impacts on land prices, labor supply, market wage, production, and welfare arise as a result of government policies related to roadway construction and maintenance? What are the effects of various government policies to reduce transportation costs, and are their outcomes similar or different? What are the economic impacts if the government does not fully address road maintenance?

The paper is organized as follows. Section 2 discusses some of the relevant literature, section 3 outlines the empirical model formulation, section 4 discusses the data employed, section 5 presents model results, and section 6 concludes.

Literature

Roadway Maintenance

Fallah et al. (2010) developed a dynamic model of road deterioration by examining 17 miles of roadway in Virginia from 2002 to 2007. The first step of that model was designed to incorporate mechanical models of road deterioration in order to develop and calibrate a simulation of road deterioration over time. Their analysis focused on the relationship between load cycles and pavement cracking. Load cycles are determined by the amount and type of traffic, and the impacts of load cycles are determined relative to the number of allowed load cycles governed by the physical characteristics of the pavement. They estimated the amount of months from the completion of road maintenance through three qualitative
stages of road deterioration. This road cycle is described as fresh->not severe->severe->very severe, and the time from stage-to-stage is approximately fresh->not severe, 24 months; not severe->severe 12 months; and severe->very severe 18 months.

Chasey et al. (2002) employed a simulation model to examine the impacts from deferred road maintenance. The authors used a so-called Level of Operation Index to describe the physical condition of the roadway. Additionally, the authors employed a Level of Availability index to measure road congestion. By combining these two measures, the authors were able to provide some simulated estimates of the impact on drivers from deferring routine maintenance in favor of additional reconstructive maintenance.

In a much earlier paper, Abelson and Flowerdew (1975), examined road maintenance programs in Jamaica in order to design a model of road maintenance that would optimize the balance between road maintenance costs and the costs to drivers of road deterioration. In that study, the authors noted as many as thirty separate activities that could be considered under the umbrella of road maintenance. Further, the authors noted that in predicting the future physical state of a road, one would consider the current road state and apply a “do nothing” baseline. Empirical observations indicated that road deterioration varies widely based on the amount of traffic and the beginning state of the road section.

Karan et al. (1976) estimated the relationship between road quality and average vehicle speed. The authors examined 72 sites in southern Ontario that exhibited varying degrees of pavement quality. Pavement quality was measured by the riding comfort index (RCI), which measures the roughness of the road surface and generates a value on a ten-point scale, ten being the smoothest and zero being the roughest. The authors obtained data on RCI, vehicle to capacity ratio, and speed limit and fit four different functional specifications to relate average vehicle speed to the three other variables.

Road Congestion
Research into the costs of congestion has been ongoing for many years. One early effort, Vickrey (1963), noted the significant cost of congestion suggested by even very simple models, and related the cost of congestion to land values. Smeed (1968), in the published version of an inaugural address as Professor of Traffic Studies, showed the impact of traffic congestion (measured as vehicles per hour per foot of roadway width) on average vehicle speed.

Vickrey (1969) attempted to relate transportation investment to the nature of the underlying congestion. The author examined several types of congestion that he labeled simple interaction, multiple interaction, bottleneck, triggerneck, network and control, and general density. When examining what he termed interaction delay, Vickrey (1969) provided a simple model of the added delay to a traffic system that comes from the addition of a single extra vehicle. Additionally, the author examined the trade-offs possible for eliminating congestion due to a traffic bottleneck (a situation where a portion of the roadway does not have the capacity to accommodate all traffic at some particular time interval). The analysis showed, given simplified assumptions, that appropriate congestion tolls may be necessary in order to improve congestion impacted traffic flows.

The Vickrey (1969) model has been extended in a variety of ways. Arnott et al. (1993) extended the Vickrey (1969) congestion model to account for price elastic trip demand, to treat the rush hour period as a single time interval, to provide more refined estimates of efficiency gains from congestion toll pricing, and to examine implications of self-financing road construction. These authors found previously unobserved efficiency gains arising from congestion toll pricing because of effects from redistributing drivers’ departure times.

Another avenue of research is to relate the impacts from congestion costs to land rents. Solow (1973) is an early example of this literature that employed the notion of a monocentric city with households that commute to the city center. The author added a congestion cost to the cost of transportation as the ratio of the population to the area of the road. By combining the total rent paid at any location with notion that the marginal rent paid must equal the marginal transportation cost, the author ended up with two first-order differential
equations that were redefined into one second-order differential equation that was solved given certain boundary conditions, which were used to solve for various parameter values. Results showed that the rent gradient, the value of land rents as a function of travel distance, are significantly impacted when congestion costs are included in the model.

Many of the model assumptions made in Solow (1973) have been challenged. Wheaton (2004) rejected the monocentric model in favor of a model where firms and households are interspersed. The author explored agglomeration technologies that limit dispersion of firms, and the amount of land consumed was assumed to be related to the price of land. In this model, the author sought to maximize worker productivity (as a function of location) times the amount of land devoted to production less the demand for travel (as a function of location) subject to constraints on firms size, travel demand, and boundary constraints. The primary conclusions made by the author included that as congestion costs decline and firms and households move farther apart, firm dispersion depends significantly on agglomeration.

Langer et al. (2008) examined congestion by constructing an hedonic model of housing prices that included factors measuring highway congestion. The authors found significant benefits realizable from instituting congestion tolls. Their results also indicated the existence of heterogeneous responses to congestion tolls, which allow for some consumers to move farther away from their workplace, while in general city population densities increase and housing prices decrease. Hymel (2009) looked at the impacts from congestion on employment growth. The author used a panel data model with fixed effects and planned highway miles as an instrument for congestion, and found that initial congestion reduced employment growth.

Parry (2008) and Lindsey (2006) reviewed the literature with respect to congestion pricing. Lindsey (2006) examined the literature on congestion pricing to determine if a consensus existed among researchers. The author found a general consensus among researchers supporting road pricing. However, many researchers are skeptical that road pricing can overcome political opposition. Parry (2008) looked at the economics literature with regard to optimal congestion pricing. The author examined different models of congestion and factors that
might be expected to complicate the analysis (pricing across road lanes, driver heterogeneity, alternative routes, etc.).

Model

The model employed in this paper is an applied or computable general equilibrium model. This class of models allows for the extension of the analysis beyond basic theoretical considerations used to determine the existence and qualitative nature of an equilibrium solution while “abstract[ing] from many complications that might obscure the key effects.” Balistreri and Markusen (2009). An applied general equilibrium analysis allow the research to focus on quantitative considerations and to give more than just general insights into policy considerations. The extension of the basic theoretical model to incorporate multiple interacting economic agents in a theoretically consistent framework is one of the chief strengths of CGE modeling. Admittedly, however, these models suffer the drawback of abstraction. To keep the model sufficiently convex and estimable, some restrictions must be placed upon the dimensionality of the problem. For example, instead of modeling individual economic agents, the models are restricted to modeling different classes of agents. However, the structure of these models imposes theoretical consistency and derives baseline characteristics from known empirical results.

The framework of the model incorporates many aspects of the monocentric city model popularized by Alonzo (1964) and Muth (1968) among others. As noted above, Solow (1973) employed such a framework to examine congestion costs and land rents. Safirova (2002) used this framework to examine impacts from telecommuting. The model employs a general equilibrium framework with transportation costs that depend on road quality and quantity. Transportation costs occur as direct cash costs and increased travel times brought about by road congestion. This is a monocentric city model with central business district (CBD) where production occurs, and an outer residential district (RD) populated by households. A
single firm is located in the CBD. Three households are located and evenly spaced in the RD. The location of a particular household is denoted \( x_i \). The location of a general household or the firm is simply denoted by \( x \).

The city is located on a flat, featureless plain and the market for land is competitive. The land within the CBD is owned by landlords who rent to the city government at a rental rate of \( R_a \) per unit. The city government only rents land that it can lease to firms. Households are endowed with land in the RD. There are two distinct, fixed boundaries within our area of interest. One boundary is the boundary between the CBD and the RD, which lies at a distance \( c \) from the city center. The second boundary is the boundary at the edge of the city, which lies at a distance \( s \) from the city center. The baseline area of interest includes one two-lane road that runs from the city center at 0 to the city boundary at \( s \). Both \( c \) and \( s \) are exogenous in the model.

The city government functions to rent land from landowners, lease land to firms, and return any excess rents back to households. The city government is also responsible for building and maintaining roads. Roads must be maintained to delay their deterioration, which leads to higher transportation and travel costs. Increasing the quantity of roads works to ease congestion. The city pays for road construction and maintenance through taxes.

The Firm

The firm produces the composite consumption good, \( Z \), using labor, \( L \), (in hours), land, \( E \), (in units), and transportation, \( T \), (in trips). The firm operates in perfectly competitive input and output markets. The firm is located in the CBD a prescribed distance from the CBD/RD boundary. The firm transports its output to the CBD/RD border to sell to households. The transport distance for the firm is equal to \( \frac{c}{2} \).

The production function of the firm exhibits constant returns to scale technology and is
\( (10) \quad Z = AL^\alpha E^\beta T^\gamma, \text{ where } \alpha + \beta + \gamma = 1 \)

where \( L \) is the total amount of labor hours used, \( E \) is the total units of land rented, and \( T \) is the total trips made to move the firm’s output.

All market transactions are assumed to occur at the CBD/RD border. The firm transports its output and sells it at the border, and all households travel to and from the CBD/RD border to supply labor to the firm and purchase goods. This simplifies the analysis should one allow for multiple firms by avoiding trying to match specific firms to workers, or having to incorporate the notion of of effective wage adjusted for workers’ travel costs in the CBD. Each firm earns zero profits.

\( (11) \quad \pi = Z - wL - r_x E - tT = 0, \text{ where } r_x \) is the rental price of land at location \( x \) and \( w \) is the wage

\( (12) \quad t = \left[ (t_{sf} \left( 1 + \delta_f \left( \frac{CF_l}{N_l} \right)^{K_f} \right)) \right] \tau_T + \tau_f \right] (c - x) \)

where \( x \) is the location of the firm, \( c \) is the distance from the boundary to the city center, \( t_{sf} \) is the free-flow cost of travel, \( 1 + \delta_f \left( \frac{CF_l}{N_l} \right)^{K_f} \) is the cost of traffic congestion\(^{15} \), \( \tau_T \) is the firm’s cost per hour of transport, \( \tau_f \) is the cash cost per mile of transport, and \( N_l \) is the number of one-way road lanes. The congestion cost, \( 1 + \delta_f \left( \frac{CF_l}{N_l} \right)^{K_f} \), is measured by the ratio of number of vehicles to the number of vehicles that can be accommodated on the roadway at the freeflow rate of speed (capacity), \( \frac{1}{t_{sf}} \). It is assumed that an increase in the number of lanes, \( N_l \), will reduce congestion proportionally by increasing the maximum capacity at the freeflow rate of speed.

Households

Residents of the city are identical except for their household location. Households maximize utility functions

\( (13) \quad u_i = Z_i^\lambda E_i^\mu B_i^\nu \text{ where } \lambda + \mu + \nu = 1 \)

\(^{15}\text{This is the standard congestion function in this literature. See for example Brueckner (2007).}\)
$Z_i$ is consumption of the composite good, $E_i$ is consumption of land, and $B_i$ is hours of leisure. In this model, land serves as a proxy for housing. One member of the household earns a competitive wage $w$ working at the CBD/RD border. Each household splits time between leisure, $B_i$, labor, $L_i$, and transportation (commuting), $T_i$. Households are also endowed with land, $E_i$.

\[(14) \quad H = L_i + B_i + T_i\]

Note that for households $T_i$ denotes hours of travel, whereas for firms $T$ denotes number of trips. Houses split income on expenditures for the consumption good, and cash travel costs. Household income includes labor income, land rent plus rental income, $RI$, received from the city government. Thus the household’s budget constraint is

\[(15) \quad w(H - T_i - B_i) + r_x E_i + \frac{RI}{N} = \tau_h (x_i - c) + Z_i\]

where $RI$ is rental income collected by the government and refunded to the household, $\tau_h$ is the cash cost of travel, $x$ is the location of the household, and $r_x$ is the rental rate of land.

Transportation

It is assumed that transportation, both inside the CBD and inside the RD, is costly. Costs include cash costs of travel and transport in addition to time costs. Time costs are affected by traffic congestion. The CBD is inhabited by the firm that must move back and forth from their particular location to the CBD/RD boundary to sell their output. The firm incurs transport costs of the form

\[(16) \quad C = \left[ t_{sf} \left(1 + \delta_f (\frac{CF_i}{N_{lf}})^{K_f}ight) \right] \tau_t + \tau_f \left( c - x \right) \]

where $x$ is the location of the firm, $c$ is the distance from the boundary to the city center, $t_{sf}$ is the free-flow cost of travel, $\delta_f (\frac{CF_i}{N_{lf}})^{K_f}$ is the cost of traffic congestion, $\tau_t$ is the firm’s cost per hour of transport, $\tau_f$ is the cash cost per mile of transport, and $N_{lf}$ is the number of one-way road lanes in the CBD.
Households travel through the RD back and forth to work. Travel is costly to households who must pay cash costs of travel and incur travel time costs, which reduce household income. Cash costs are

\[ \tau_h (x_i - c) \] (17)

whereas travel time costs are

\[ t_{sh} \left( 1 + \delta_h \left( \frac{CF_h}{N_{lh}} \right)^{K_h} \right) (x_i - c) \] (18)

where the former is the cash cost as noted above, and the latter is the congestion cost. 

\( x_i \) is the location of the household, \( c \) is the distance from the boundary to the city center, \( t_{sh} \) is the free-flow cost of travel, \( \delta_h \left( \frac{CF_h}{N_{lh}} \right)^{K_h} \) is the cost of traffic congestion, \( \tau_h \) is the cash cost per mile of travel, and \( N_{lh} \) is the number of one-way road lanes in the RD.

As noted above households travel from their homes to the CBD/RD boundary where they work. They travel over roads that must be constructed using a fixed portion of their land endowment. Thus households must trade-off welfare gained from the land they possess with the ability to supply their labor in the competitive market for the purpose of earning income in order to purchase the composite good for consumption. In order to participate in the labor market, each household must commit leisure hours and roads. Households are given a modest ability to substitute leisure for roads using an elasticity of substitution equal to 0.2. Also noted above, household’s transportation costs are time costs associated with commuting and directly reduce their available leisure hours.

In the CBD the land is owned by “agricultural landowners” who lease some land to the government, which in turn leases land to the firm for use in production. The firm uses land directly in the production of the composite good, and to produce roads in a fixed proportion. The firm then uses the created supply of roads to generate transportation, which is also used in the production of the composite good with an associated cost per trip including the cash cost of travel and a congestion cost.

General Equilibrium Model
The model is solved using the General Algebraic Modeling System (GAMS)\textsuperscript{16}. The model consists of forty equations constituting fifteen zero-profit conditions, twenty market clearing conditions, and five income-balance conditions. For this initial model calibration prices and market quantities are all set equal to one.

After the initial equilibrium condition is calibrated, several counterfactual scenarios are explored. The first such scenario is to introduce a spatial structure to the model. This is accomplished by moving the firm away from the CBD/RD boundary toward the city center, and moving the three households away from the CDB/RD boundary to different locations. The size of the city is exogenous with the distance from the city center to the CBD/RD boundary equal to 2.5 miles, and the distance from the city center to the city boundary equal to 17.5 miles. The firm is moved to a distance of 1.25 miles from the CBD/RD boundary, and the three households are moved to be equidistant from one another - household 1 at 3.75 miles, household 2 at 7.5 miles and household 3 at 11.25 miles from the CBD/RD boundary. The increased distances imply larger transportation costs for both the firm and the households.

The remaining counterfactual scenarios retain the same spatial structure, but increase traffic congestion and allow for systematic road quality deterioration. In this simple model, congestion can only be alleviated directly through construction of additional road lanes, which are funded by taxes collected by the government. Road quality deterioration will occur at an empirically established rate in the absence of maintenance expenditures - also funded through government collected taxes. Road quality affects transportation costs by reducing the free-flow traffic speed. A simple rule is employed that relates maintenance expenditures to road quality (and thus average speed), which is that an X percent level of full maintenance expenditures results in a 100-X percent of road deterioration that would have occurred absent any maintenance (thus fully funded road maintenance results in no

\footnote{http://www.gams.com/}
road quality deterioration). New traffic lanes will be constructed whenever congestion as
defined above exceeds a certain threshold. Thirty six different scenarios were examined
for deviations from the baseline and initial spatial equilibria (this encompasses twenty-two
equilibrium solutions for each scenario). The values chosen for the scenarios are shown in
Table 1.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Congestion Threshold (%)</th>
<th>Maintenance Level (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>2</td>
<td>3.75%</td>
<td>25.00%</td>
</tr>
<tr>
<td>3</td>
<td>7.50%</td>
<td>65.00%</td>
</tr>
<tr>
<td>4</td>
<td>15.00%</td>
<td>85.00%</td>
</tr>
<tr>
<td>5</td>
<td>30.00%</td>
<td>95.00%</td>
</tr>
<tr>
<td>6</td>
<td>60.00%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Table 21. Threshold values for road congestion and maintenance. Congestion thresholds
denote congestion levels that trigger new road construction, and maintenance thresholds
denote proportions of full road maintenance to maintain current road quality.

Table 21 shows, for example, that for a congestion threshold of 0 percent (a new lane is built
immediately and each year) six different levels of road maintenance are examined.

Data

A variety of data are used to provide for values used in the initial equilibrium and
subsequent counterfactual scenarios. The coefficients used in the production function for
the composite good were derived from Safirova (2002), $\alpha = 0.855$ and $\beta = 0.095$, and
Friedlaender and Spady (1980), $\gamma = 0.05$. The coefficients used in the common household
utility function were obtained from Safirova (2002), $\lambda = 0.32$, $\mu = 0.03$ and $\nu = 0.65$.

Several parameters related to congestion and transportation cost are defined in the model.
Cash costs per hour of transport for both the firm and households were set equal to one in
the spirit of pricing in the applied general equilibrium framework used in GAMS ($\tau_h = 1$
and $\tau_t = 1$). The cash cost per mile for the firm is estimated from Barnes and Langworthy
(2003) to be about 80 cents per mile, $\tau_f = 0.8$. That figure used the average of the costs for
highway and non-highway transport estimated in 2003 and updated for general increases in the price level and an increase in the price of fuel by approximately 2.5\textsuperscript{17} times.

The congestion cost parameters, $K_h = 3$, $\delta_h = 0.1$, $K_f = 10$, and $\delta_f = 0.05$ were obtained from Small and Verhoef (2007) and are based upon parameters developed by the Federal Highway Administration. The congestion function for households is in line with that for a highway, whereas the congestion function for the firm is that used for an arterial.

The increase in the value of the congestion parameters, $CF_h$ and $CF_f$, was based upon an annual rate of growth of 2.57 percent. This number was derived from the population growth of the ten fastest growing urban areas in the United States between 1990 and 2010\textsuperscript{18}. This figure is slightly less than the average annual growth in vehicle registrations in the United States from 1960 to 2004,\textsuperscript{19} which was approximately 2.68 percent. This also assumes no attempt to improve congestion. Figure 1\textsuperscript{20} shows congestion in the United States from 1990 through 2009. Figure 1 shows the average level of congestion for various urban areas in the United States grouped according to population. We note that congestion increased steadily over a decade starting roughly in 1994, but started improving after 2005. These trends could be influenced by a variety of factors including economic conditions, government efforts to reduce congestion, and changes in fuel prices. The data show average annual growth of between 0.8 and 1.5 percent, which is less than the roughly 2.6 percent assumed in this model. The difference can be attributed to focusing on a very few high population growth areas and particular roadways with more severe congestion problems.

\textsuperscript{17}The average price of diesel fuel in 2003 was $1.51/gallon and is currently $3.75/gallon. Source U.S. Energy Information Agency.


\textsuperscript{19}Source http://www.infoplease.com/ipa/A0908125.html

\textsuperscript{20}Source http://www.bts.gov/publications/national_transportation_statistics/excel/table_01_71.xls
Figure 28:
Average road congestion by city size for the period 1990 to 2009. Values greater than indicate road use greater than road capacity.

The model also employs a function for road quality deterioration. Abu-Lebdeh et al. (2003) estimated equations (5.1) and (5.2). The authors used data on Michigan roads collected by the Michigan Department of Transportation to estimate road surface wear over the life of the road.

\[
DI_{th} = (1.02 \times DI_{t-1,h}) + (0.4 \times AGE_{t-1}) + 0.194
\]

\[
DI_{tf} = (1.407 \times DI_{t-1,f}) + (0.25 \times AGE_{t-1}) + 0.63
\]

Here \( DI_{t,h} \) is the distress index at time \( t \) for the household road, and \( AGE_{t-1} \) is the age of the road the previous period. Equation (21) is the relationship used for roads in the RD, while (5.2) is that for roads in the CBD. Both equations were estimated for new or fully reconditioned roads. The roads in the CBD are non-highway roads, whereas the roads in the RD are assumed to by highway roads.

As the road surface deteriorates, the average speed begins to decline. Karan et al. (1976) estimated the relationship between average vehicle speed and road quality. Equations (5.3) and (5.4) show these relationships for roads in the RD and CBD respectively.
\( VEL_h = 30.7368 + (1.037 \times (1 - \frac{DI_h}{100}) \times 10) - (11.2421 \times \frac{CF_h}{NI_h}) + \left(0.0062 \times \frac{1}{ts_{h}^{\text{max}}} \times 1.609344\right) \)

\( VEL_f = 30.7368 + (1.037 \times (1 - \frac{DI_f}{100}) \times 10) - (11.2421 \times \frac{CF_f}{NI_f}) + \left(0.0062 \times \frac{1}{ts_{f}^{\text{max}}} \times 1.609344\right) \)

\( VEL_i \) is the average road speed for roads in the CBD, \( i = f \), and roads in the RD, \( i = h \).

\( VEL_i \) is measured in kilometers per hour, so appropriate scaling is included. \( DI_i \), which are obtained from equations (5.1) and (5.2), are scaled by a factor of ten because Karan et al. (1976) used the Ride Comfort Index (RCI) as their measure of road quality, which is based on a ten-point scale. The values for \( ts_{i}^{\text{max}} \) used are the maximum hours per mile provided for in the baseline (50 mph for RD roads and 35 mph for CBD roads).

The final parameters used are values for tax rates set to collect revenues for road construction and maintenance. The rates used are a 0.50 percent tax on output for road construction and a 0.15 percent tax on output for annual road maintenance. The construction tax is levied each time a new road lane is constructed, while the road maintenance tax is levied each period at prescribed levels. For example, 80 percent of the tax may be levied, which would result in a 20 percent decline in road quality. As noted above, this serves to examine impacts from foregoing road maintenance in order to reduce taxes and government expenditures. The tax rates were calculated by comparing total state spending on road construction and maintenance\(^{21}\) as a proportion of gross domestic product\(^{22}\) from 2001 to 2008. These tax rates are expected be somewhat conservative in that national new road construction does not include a significant fraction of all existing roads, yet in this model adding the first new lane essentially doubles the quantity of roads, and anecdotal evidence suggests that road maintenance tends to be under-funded.\(^{23}\)

**Results**

---

\(^{21}\) Source http://www.fhwa.dot.gov/policy/ohim/hs0Y/sf2.htm where Y=1 to 8.

\(^{22}\) Source http://www.stlouisfed.org/fred2/

\(^{23}\) See for example http://knowledgecenter.csg.org/kc/content/condition-us-roads-bridges
In order to investigate the impacts from traffic congestion and deteriorating roads, the model was used to simulate a road over its expected 20 year lifespan. Over that 20 year period traffic volumes were allowed to grow and the road surface was allowed to deteriorate. During each year the government was able to alleviate congestion by constructing new roadways, which were modeled here as simple lane additions. The government was also able to maintain the road by paying for maintenance activities that could keep the road in like new condition over the life of the road, or the government could pursue a maintenance policy that allows the road quality to decline fully or partially\textsuperscript{24}.

Congestion Mitigation

One topic of this paper is to examine the economic effects of traffic congestion and congestion mitigation through roadway construction. Congestion levels for different road construction scenarios are shown in Figure 2. It is observed that congestion levels decline each time a new, additional road lane is constructed, which is an assumption built into the model. Under the 0\% policy, a new road is constructed each year, while under other policy levels, new roads are constructed when congestion reaches a particular specified threshold (3.75\%, 7.5\%, 15\%, 30\% and 60\%). Although constructing a new road (or road lane) is assumed to reduce congestion, there are associated costs. Taxes are levied by the government to pay for the road construction, and land is taken from households and other productive uses.

\textsuperscript{24}It is possible to use this model to compare outcomes if the government expends zero resources on road construction and maintenance versus a policy where the government fully funds road construction and maintenance with the parameters of the model. Over the life of the road, as road quality worsens and traffic congestion increases (no government maintenance or road construction) average speeds decrease almost 20 percent in the RD and over 25 percent in the CBD. Land prices are initially higher under a government policy of no maintenance and no construction, but eventually increasing transportation costs force land prices down. Population density decreases for residents located at household 2 by 13 percent and household 3 by 26 percent when transportation costs increase. Population density at the city center increases by approximately 20 percent. Wages are about two to three percent lower when transportation costs are allowed to increase as households located farther from the CBD/RD border reduce their labor supply. Taxes and land required to build and maintain roads reduce output by about two percent per year, but reduce consumer welfare by over three percent.
The 60% line, which is the upper most line in the figure, shows the evolution of traffic congestion with no mitigation efforts. If this is compared with the information from Figure 1, we note that the absolute most severe level of congestion was a value of 1.3,\textsuperscript{25} so a value of 1.6 might seem somewhat large for congestion values that have been observed. However, the values in Figure 1 are averages for different city-types, so it is quite likely that some roads within a particular city would be observed to have congestion values as high as 1.6.

Figure 29:
Road congestion levels occurring under different new road construction congestion thresholds (The 0% threshold implies a new road lane is constructed annually while the 60% threshold implies a new road lane is constructed when the congestion measure reaches 0.6).

Figure 30 shows the actual observed\textsuperscript{26} government expenditures on road construction from 2001 through 2008 in the United States. We observe that government spending on new road construction increased starting in 2005 and remained at the higher levels in 2007-2008. The implication is that governments continue to try and alleviate traffic congestion through road construction.

\textsuperscript{25}It should be noted that one must be added to the congestion values shown in Figure 2 to compare with the values shown in Figure 1.

\textsuperscript{26}Source http://www.fhwa.dot.gov/policy/ohim/hs0Y/sf2.htm where Y=1 to 8.
Figure 30:

The effect of road congestion on welfare is shown in Figure 4. We observe that increased congestion is associated with a decrease in consumer welfare. Figure 4 also highlights the relative importance of road maintenance. A combination of high traffic congestion levels and no road maintenance results in welfare nearly 0.5 percent lower. Figure 4 also shows how road quality deteriorates more at higher levels of traffic congestion.
Figure 31:
The impact on consumer welfare at different levels of congestion when roads are fully maintained (100%) and when road maintenance is neglected (0%) .

The impact from congestion on production is shown in Figure 5. Over the range of traffic congestion examined, output has a range of from one to over two and one-half percent, which shows larger impacts than welfare. One must keep in mind that road quality is not held constant in Figure 5. Thus for the case where there is no road maintenance, the rapid decline in output at higher levels of traffic congestion is reflecting increasingly poor road quality, and reinforces the theme that the costs of road maintenance outweigh the benefits until a certain threshold of road quality deterioration is reached. Additionally, the impacts on production output are more severe than impacts on household welfare.
Overall it is clear that traffic congestion is costly, both in terms of household welfare and production output. What is also clear is that wear and tear on road surfaces exacerbates the costs from traffic congestion over time (as much as 1.6 percent of production output).

Road Maintenance
Figure 33: Road quality in the RD for different levels of annual road maintenance (0% indicates no road maintenance while 95% indicates road maintenance at the 95 percent of the level of road maintenance required to maintain the road at current level or quality.

The other major topic of this paper is to examine economic impacts of road quality deterioration. This section shows specific impacts that occur from different levels of road maintenance. Figure 6 shows the state of the road measured by the distress index (DI) in the RD given different annual levels of maintenance. Recall from the discussion in Section 3 that the distress index function is such that average vehicle speeds cannot approach zero, which implies that even after 20 years of no road maintenance, the distress index barely reaches 100. It is likely that this aspect of the model is unrealistic, but that is likely due to the fact that governments rarely pursue a policy of zero road maintenance over the entire life of the road.
Figure 34: Consumer welfare for different road conditions absent significant increases in traffic congestion. Distress Index of 0 is associated with a new road surface.

Figure 7 shows the linear trend of impact on household welfare from road quality deterioration given annual construction of new roads to alleviate traffic congestion. Welfare declines about 0.05 percent for every increase of 10 units in the distress index.

Figure 8 provides a comparison by looking at output differences observed from a policy of road maintenance at a level of 0, 25, 65, 85, 95, and 100 percent of the amount needed to keep the roadway in like new condition. For output the annual production and consumption costs of increased taxation outweigh the benefits obtained from reduced transportation costs for nine years, but after a certain level as road quality continues to degrade, the costs on gross domestic product tend to be much greater than the savings from foregoing annual maintenance.
Figure 35:
Production output for different levels of road quality. Output declines as road quality deteriorates, but the costs of road maintenance outweigh the benefits of lower transportation costs for approximately nine years.

I also examine production output for different levels of road quality. Figure 9 shows the relationship between road quality and output. The dashed, red line is the trend for production output associated with road quality when there is annual road construction to mitigate traffic congestion. Every 10 unit increase of the road distress index is associated with approximately a 0.2 percent drop in output.
Figure 36: The impact on production output levels as road quality deteriorates. These results assume that a new road is constructed annually to alleviate traffic congestion. A 10 percent decrease in road quality results in about a 0.2 percent decline in production output.

The analysis indicates that governments have leeway in how they approach their policies of road maintenance. This model suggests that governments have seven to nine years before the indirect costs associated with neglecting road maintenance will outweigh the economic costs imposed by funding required for maintenance. Figure 10 shows average annual road conditions for different types of road in the U.S. from 1990 to 2008. The trends observed would seem to indicate that governments are engaging in a policy of less than full maintenance since the observed average road quality is not particularly good and has improved only for interstate and intrastate highways.
This is probably not surprising given the relatively small role transportation plays in production and welfare. However, it may well be that neglecting road maintenance will result in significantly higher repair costs in the future.

Figure 11 shows different wage levels that arise as a result of different levels of road maintenance. Recall that as congestion increases and road quality decreases, households tend to supply less labor in the labor market, so wages increase. Thus in Figure 11 we observe that wages are higher when the government invests in less road maintenance. The general increase observed for all maintenance levels occurs as a result of increasing transportation costs from increasing traffic congestion.

Implications
Figure 38:
A comparison of wage changes given different levels of annual road maintenance. Percentage values indicate that road maintenance is funded at the percentage of full maintenance required to maintain the road in its current condition. Wages increase over time because of increasing traffic congestion.
The current state of roads and projections through 2019 for road quality are shown in Figure 12 for both roads in the residential district and roads in the central business district. Figure 13 shows congestion levels over time. Projections are based upon the continuation of observed road maintenance policies. Currently it appears that road quality is somewhat worse in the central business district than in the residential district. However, over the next eight years, because of roadway usage patterns the road quality in the CBD becomes much worse over time than the road quality in the RD with similar levels of maintenance. Congestion levels vary between 1.10 and 1.15, which impacts travel costs, but is not too severe. Note that congestion levels drop when additional roads are constructed.

Figure 39:
Shows projected road conditions given current trends for the residential district and the central business district.
Figure 40:
Shows projections for traffic congestion from 2011 to 2019. Congestion levels decrease when a new road is constructed, but traffic volumes continue to increase erasing any gains from congestion mitigation.

The differences in road quality deterioration in the CBD and RD do not necessarily translate into average speed reductions. Figure 14 shows current and projected average speeds. Speeds in the CBD are measured on the right axis and RD speeds are measured on the left axis. Despite a nearly four-fold decrease in road quality in the CBD, average speeds only drop about 10 percent over the same period.
Figure 41:
Shows current and projected average vehicle speeds from 2011 through 2019 in the CBD (right axis/solid line) and RD (left axis/dashed line). Average speeds decline approximately 1.5 mph in the RD and less than 4.0 mph in the CBD.

Figure 15 shows that, given current policies, the city is expected to become slightly less spread out. The increase in transportation costs from road quality deteriorating causes some households to move toward the city center. This is one consequence of alleviating congestion through the use of congestion road tolls - any time transportation costs increase, cities contract toward the city center.
Figure 42:
Changes in city density as a result of projected traffic congestion and road quality changes. City center density will increase about two percent.

Figure 43 and Figure 44 show the projected impacts of increased transportation costs on land prices. Figure 16 shows land prices faced by the firm in the CBD. Because road construction requires land, land prices rise in years where the government constructs new roads to alleviate growing traffic congestion. The solid trend line indicates that land prices increase at an increasing rate whenever more land is used for roads. Figure 44 also shows that land prices for the firm in the CBD exhibit a decreasing trend, which is being pulled up as more land is being used for roads.
Figure 43:
Change in land prices faced by the firm in the central business district (CBD) over time. New road construction causes higher land prices as land is taken for new roads.

Household land prices in the RD do not appear to vary much over time except when land is taken for road construction. Road construction somewhat reduces commuting costs for households, which reduces their demand for land more than taking land for roads reduces the supply. It is also apparent that road quality has less of an impact on household land prices than land prices for the firm in the CBD. This occurs because CBD road quality deteriorates much more rapidly than RD road quality.
Figure 44: Change in land prices faced by households in the residential district (RD) over time. New road construction causes temporarily lower land prices as transportation costs decline resulting in lower demand for land by households.

Road quality impacts on production output are shown in Figure 45. As road quality deteriorates, average speeds drop and output declines as the cost of production increases. The decline is more severe when government engages in road construction to alleviate traffic congestion with an annual loss of approximately 1.5 percent of output. The average annual decrease in output from road quality deterioration is approximately 0.05 percent, which would be approximately $7 billion for the United States. The annual cost to production from road construction would be approximately $200 billion.
Figure 45:
Projected impacts on production output. As road quality deteriorates, the production output trend is declining with or without new road construction.

Output appears to decline as road quality deteriorates and Figure 46 shows that employment (as measured by labor supplied) also declines even though wages remain almost constant. We observe in Figure 46 labor supplied by household along with linear trends. All three trends have negative slope, which indicates that labor supply is decreasing over time as road quality worsens. Wages are also shown in Figure 46, but apart from movements caused by road construction, wages are flat.
Projected impacts on labor supply and market wage. As road quality worsens, labor supplied declines as transportation costs rise, even in the face of fairly constant wages.

Conclusion

This paper has presented a general equilibrium model of a monocentric city with deteriorating road quality in the face of traffic congestion. Results have show that there are different impacts from the application of different government policies aimed at addressing increasing transportation costs that arise from increased traffic congestion and the decline in the quality of roadways. More specifically, the model shows that although both traffic congestion and deteriorating road quality impose economic costs on individuals and society, those costs are often much less than the opportunity costs imposed on the economy when the government pursues aggressive road construction and maintenance policies.

I examined the economic costs of road quality deterioration and traffic congestion in three contexts. First I looked at impacts that result as a consequence of traffic congestion
and congestion mitigation through the construction of new roads for a given level of road maintenance. If traffic congestion is left unabated, household welfare declines between 0.3 and 0.5 percent depending on road maintenance. Similarly, production output drops between 1.2 and 2.9 percent, which shows that taking land away from productive use has much higher opportunity costs than taking land from households.

When I focus on the impact of road maintenance, I find that as the road distress index (DI) increases by ten units, welfare drops by about 0.05 percent, and output declines by about 0.2 percent. However, if I use the model to approximate increasing transportation costs over time, the economic costs of maintenance outweigh the increasing cost of transportation for approximately nine years. Thus, governments can neglect road maintenance for about nine years before the economic harm from poor road quality outweighs the economic costs of paying for road maintenance.

The model can also be used to examine economic impacts if currently observed policies with regard to road maintenance and traffic congestion continue from 2011 through 2019. In the CBD the road distress index will increase from 20 to 80, while in the RD the distress index will increase from 5 to about 30. Congestion will be between 1.10 and 1.15 over the period. These trends will decrease average speeds over 10 percent in the CBD and only about 0.3 percent in the RD. The population density in the city center will increase about two percent. Land prices in the RD will remain largely unchanged, but output will be lower by about 0.04 percent annually. Labor supply will decrease as transportation costs increase, but wages will remain unchanged.

Clearly the model in its current framework is more suggestive of an intertemporal interpretation. No attempt has been made to adjust for time preferences or discounting. Still it is clear that for an existing, productive, developed economy transportation plays a relatively small role in production and overall welfare, which allows for policies that allow transportation costs to increase over time.
Chapter 3

Introduction

At the time of this writing petroleum prices are again rapidly increasing and many researchers are engaged in an effort to explain and predict these prices. Additionally, the relationship between oil and gasoline prices is still uncertain. For example a recent article in the Wall Street Journal\textsuperscript{27} reported on the possible impacts of speculation on oil prices and their ultimate impacts on retail gasoline prices. This paper seeks to examine retail gasoline pricing in the context of a seven equation vector error correction model using daily, weekly and monthly data.

The questions to be addressed include whether macroeconomic performance is useful in forecasting retail gasoline prices, what are the impacts on retail gasoline prices from oil and gasoline price volatilities as well as prices, particularly looking at oil futures prices, and how does temporal aggregation affect the conclusions with respect to and forecasts of retail gasoline prices in the context of a vector error correction model. As discussed by Tsay (2010), often financial data are associated with conditional heteroskedasticity. Because of this another topic examined here is the impact of the presence of conditional heteroskedasticity on a vector error correction model framework. Additionally, most econometric analysis assumes the existence of some underlying data generating process and given the possibility that many of these processes are higher frequency processes, it is somewhat important to gauge the feasibility of using models involving daily data.

Results indicate that information about macroeconomic performance does impact retail gasoline prices, but that the impacts are more pronounced in daily and weekly models. It is also the case that oil futures prices have significant effects on retail gasoline prices. This

is especially important given the increased volatility in oil futures markets. Price volatilities appear to have only limited effects on retail gasoline prices. Oil futures price volatility has significant negative impacts on retail gasoline prices using daily, weekly and monthly data, but gasoline price volatility only appears to have significant impacts using daily data. Analysis seems consistent across temporally aggregated data, but conclusions can differ depending on which type of data is used.

The remainder of this paper: Section 2 presents relevant literature, Section 3 discusses the data used, Section 4 presents the model, Section 5 gives the model results, Section 6 discusses the impacts of conditional heteroskedasticity, and Section 7 concludes.

Literature

Oil and Gasoline Prices

There have been a number of recent papers which examine oil and gasoline pricing. Alquist, Vigfusson and Kilian (2011) is an encyclopedic effort to examine forecasts of oil prices, and raises some important issues. Those issues include the distinction between looking at real or nominal prices, how prices are defined, the analytical methods employed, the sample period and the forecast evaluation process. These authors examine stability notions for different sample periods, and reject the efficacy of using oil futures prices, at least, to forecast oil spot prices.

---

28 The differences occur as a result of impacts that change over time. For example, gasoline price volatility has significant positive impacts on retail gasoline prices for a few days, but over intermediate time frames, the impacts are insignificant and negative, but over longer (monthly) time frames the insignificant impacts are positive. Also, a particular model might have very good very short-run daily forecasting ability, but a very bad forecasting ability for longer time periods. Thus using temporally aggregated data might cause one to reject an otherwise viable model.

29 For example, the price of oil can be spot or futures, West Texas Intermediate or Brent, or even the acquisition cost for refineries.
Those authors also question the usefulness of U.S. macroeconomic aggregate information in forecasting the nominal price of oil. They cite Kilian and Vega (2011) for the proposition that macroeconomic variables are not useful in forecasting the nominal spot price of oil. Chatrath et. al. (2011) offer contrary evidence that macroeconomic variables can be useful for predicting oil prices when accounting for the effects of oil inventories. Those authors examine oil prices in the context of a stock-flow model, which allows for macroeconomic information to impact the desirable level of oil inventories.

A few interesting papers have examined speculation in the context of energy markets. Parsons (2010) examines the notion of a price bubble in oil markets from 2003 to 2008, and criticizes the view that a lack of identifiable oil storage implies that the 2003-2008 oil price increases were not a bubble. The author points to the increasing importance of oil as a financial asset and changing oil market fundamentals as the reasons for the large oil price increases from 2003 through 2008. Buyuksahin and Harris (2011) used disaggregated data on trader positions and a calculation of a Working’s Speculative Index to conclude that while speculation increased over the sample period, it did not decrease afterward, and changes in trading positions of groups of traders do not influence prices. Ederington et. al. (2011) provides a survey of recent literature on the nature of oil prices from 2003 through 2008. They conclude that the general consensus is that market fundamentals were likely associated with the price changes over that period, but that evidence suggests that the role of speculation was uncertain.

Radchenko (2005) looked at the impact of oil price volatility on gasoline prices. He employed a bivariate vector autoregression with weekly data using measures of oil price volatility and gasoline price asymmetries in response to changing oil price volatilities. The author uses three measures of oil price volatility: two rolling standard deviations (one of 13 weeks and one of 26 weeks), and a GARCH\textsuperscript{30} process (using estimated standard errors as volatilities).

\textsuperscript{30}Results indicate that the rolling standard deviations and GARCH processes produce similar outcomes.
Kilian (2008) employs a structural model to examine gasoline prices. The author employs a five equation recursive vector autoregression on “the percent change of world production of crude oil, the measure of global real economic activity proposed in Kilian (2008a), the real price of imported crude oil, the real price of gasoline and other motor fuels in the U.S., and the percent growth rate of the quantity of gasoline and motor fuels consumed in the U.S.” Kilian (2008, p. 5). The results indicate that variation in gasoline prices are largely a result of demand shocks.

Kilian (2010) takes a similar approach to look at retail gasoline prices using the monthly data over the period 1974 through 2005. Results suggest that short-run gasoline price changes are influenced significantly by refining capacity, but that in the long-run gasoline prices are most heavily influenced by world petroleum demand.

Some authors have examined disaggregate data. For example, Hosken et. al. (2008) uses a panel of individual retail gasoline sellers to examine pricing patterns over time. The authors note a number of interesting features with regard to individual station pricing, but provide no explanation for the sources of price changes. Hastings (2004) used a panel of retail gas stations in southern California to show that third-party competition in retail gasoline markets leads to lower gasoline prices. Lewis (2009) looked at daily wholesale and retail gasoline price data for 90 cities for 2004 and 2005. The authors focus was to measure impacts on gasoline prices associated with hurricanes Rita and Katrina. His analysis concludes that price responses are asymmetric, rising rapidly and falling slowly, and supports the findings of Hastings (2004) that price competition leads for faster price responses.

As current research shows, energy pricing remains an engaging area of research given the recent increase in the volatility of oil and gasoline prices, the fairly recent emergence of the importance of energy commodities as financial assets, and the still uncertain links between oil and gasoline pricing.

Temporal Aggregation
If we are given a particular data generating process, but our frequency of observation is of a lower frequency, how does that impact the inferences we make? This is a very important question in the context of petroleum pricing models as discussed above.

Temporal aggregation refers to the process of taking high frequency data and using it to generate lower frequency data series. In the present analysis the data is available as daily, weekly, or monthly series. Granger (1988) provided an early survey of this problem. Granger noted instances of univariate time series properties of high frequency data being systematically sampled, or aggregated into lower frequencies. Granger and Siklos (1995) examined temporal aggregation and cointegration in the context of seasonality. Marcellino (1999) examined some consequences of temporal aggregation in the context of multivariate models. That analysis showed that a variety of issues can arise in the context of temporal aggregation. For example, the author determined that although unit root and cointegration properties survived the temporal aggregation process, most empirical properties vary with temporal aggregation and forecast errors increase.

Given these analyses, one must inquire as to how temporal aggregation affects the inferences that can be gleaned from the class of vector autoregression models used to examine petroleum pricing. These class of models have become relatively important for describing relationships in the energy market. For example Kilian (2011) is an attempt to categorize the properties of these structural models. However, he does not address issues in vector autoregression models associated with temporal aggregation.

Data

31 These include exogeneity, causality, seasonal unit roots, impulse response functions and trend-cycles decompositions.
The data used in this paper include daily, weekly, and monthly data for retail gasoline prices, the first contract for the futures price of oil, the three-month treasury bill rate, the 10- and 20-year treasury bond rates, and Moody’s BAA and AAA corporate bond rates. Futures price data and weekly/monthly retail gasoline price data were obtained from the U.S. Energy Information Agency\(^{32}\). Daily retail gasoline price data were obtained from the Oil Price Information Service (OPIS)\(^{33}\). All interest rate data were obtained from the Federal Reserve bank of St. Louis’ FRED\(^{34}\).

Daily data for retail unleaded gasoline prices covers the period from February 5, 1996 through January 31, 2011 whereas all other data (daily futures prices, interest rates and weekly and monthly data) are available prior to February 1996. Thus the retail unleaded gasoline data provides the bound for the sample period used in this paper, February 5, 1996 through January 31, 2011. Although the weekly and monthly data series are complete for all variables for the sample period, over the 5,475 days included, all of the daily data series have missing observations. The daily retail unleaded gasoline price series included 582 missing observations, the daily interest rate series have 1,565 missing observations, the daily oil futures price series has 1,721 missing observations. For the interest series, the missing days are weekend and holiday days, which is similarly true for the oil futures price series. The missing daily gasoline price series are mostly weekends early in the series (1996 through 1999), but there are instances where several days in a row are missing (the maximum run of missing days is 10 days).

Pavia-Miralles (2010) notes various methods that can be employed to fill in missing time series observations. Most of that analysis was concerned with missing a relatively significant proportion of the series. The methods chosen for this analysis are as follows. For interest rate series (BAA_AAA, TERM, and TBILL) and oil futures prices, the missing observations are completed with the last prior value (so Saturday and Sunday are completed using the Friday

\(^{32}\)http://www.eia.doe.gov/petroleum/data.cfm
\(^{33}\)http://www.opisnet.com
\(^{34}\)http://research.stlouisfed.org/fred2/
rate value). For the retail unleaded gasoline prices, missing observations were computed using interpolation accounting for observed average daily seasonality patterns. For example, if a five-day run of observations are missing starting with Tuesday and ending Saturday, the formula \( \left( \frac{\text{Sun}}{\text{Mon}} \left( \frac{1}{T+1} \right) \right)^T \) \( \cdot \) \( \text{Day} \) \( \cdot \) \( (1 + a) \), where \( \text{Sun} \) is the next available observation, \( \text{Mon} \) is the first preceding available observation, \( T \) is the length of the run of missing observations, \( \text{Day} \) is the missing observation, \( t = \{1, \cdots, 5\} \), and \( a \) is the average percentage change for the day of week associated with the missing observation.

Model

The general model is a vector error correction model for seven variables: the difference between the corporate BAA and AAA bond rates (called the default spread), denoted BAA_AAA, the difference between the three-month treasury bill rate and the average of the 10- and 20- year treasury bond rates (called the term rate), denoted TERM, the three-month treasury bill rate, denoted TBILL, the price of the first futures contract for oil, denoted OIL_FUT, the price volatility of the first futures contract of oil, denoted VOIL_FUT, the retail unleaded gasoline price, denoted GAS_RET, and the retail unleaded gasoline price volatility, denoted VGAS_RET.

Use of the interest rate variable, BAA_AAA, TERM, and TBILL is motivated by a long history in the finance literature and is intended to capture macroeconomic effects on oil and gas prices. The three month treasury bill rate has been shown to work as a proxy for future economic activity, the difference between Moody’s BAA and AAA corporate bond rates have been shown to track long-term business cycle activities\(^{35}\), and the difference between short- and long-run treasury rates (TERM) are associated with short-term business cycle activity\(^{36}\).

\(^{35}\)See Fama and Schwert (1977) and Fama (1981)
\(^{36}\)See Fama and French (1989). For more recent examples, see Chen, Roll and Ross (1986), Bailey and Chan (1993), and Chordia and Shivakumar (2002).
Another motivation for employing these variables (BAA_AAA, TERM, and TBILL) in this analysis is the relative scarcity of daily data for macroeconomic variables.

The price volatilities for oil futures and retail gasoline prices were calculated as rolling variances. For daily data a rolling 28-day period was used to calculate daily variances, for weekly data a rolling four-week period was used, and for monthly data a rolling four-month period was used. All models were estimated over the same data range, roughly March 1996 through December 2010 with the remaining available data held back to examine forecast errors. The oil futures prices and retail unleaded gasoline prices series were transformed using natural logarithms. All other series were left in levels.

The vector error correction models were specified and estimated using methods outlined in Breitung et. al. (2004) and Lütkepohl (2006). The series were investigated for the presence of unit roots using the Augmented Dickey-Fuller test. Next a vector autoregression was estimated to choose the appropriate lag length for each model. Four criteria were used: the Akaike Information Criteria, the Forecast Prediction Error, the Schwartz Criterion, and the Hannan-Quinn criterion. For each suggested lag, several diagnostic tests were examined to help choose the most appropriate lag length for each model.

To test for heteroskedasticity the Lagrange multiplier, multivariate ARCH-LM test was employed (see Lütkepohl (2006)) for a more detailed discussion. The null hypothesis of no ARCH effects is tested for each candidate model. The Jarque-Bera test is used to examine whether or not the residuals from each model are normally distributed. Serial correlation is tested using the Breusch-Godfrey Lagrange multiplier test. Finally, the possibility of structural changes was tested using the cumulative sums of residuals from OLS regressions (see e.g., Zeileis et. al. (2001)). Once a suitable lag length was chosen, the Johansen cointegration test was performed.

After determining the appropriate number cointegration relationships, separate vector error correction models were estimated for each level of temporal aggregation. A recursive\textsuperscript{37}

\textsuperscript{37}The initial goal was to specify and estimate separate structural vector error correction models for each of the temporal levels (daily, weekly, and monthly). However, the model with daily data failed to converge,
vector error correction model was employed using the sequence: three-month treasury bill rate -> BAA minus AAA corporate bond rates -> three-month treasury bill rate minus average of 10- and 20-year treasury bond rates -> volatility of oil futures prices -> volatility of gasoline futures prices -> oil futures prices -> retail unleaded gasoline prices.

This sequence is largely gathered from Tischer (2011) which examined oil and gasoline price responses to hurricane events in the Gulf of Mexico. Impulse response analysis from that paper suggested that oil futures price volatilities were largely responsible for the observed increases in gasoline prices that occur when a hurricane enters the Gulf of Mexico. The final step in the analysis is to examine impulse response functions and the forecast error variance decomposition with respect to retail unleaded gasoline prices.

Results

The model results are presented for daily, weekly and monthly data, which are then compared for consistency.

Daily

Figure 47 provides plots of the seven data series.

\underline{which is not surprising given seven variables requiring 21 restrictions, over 5000 observations, and 38 lags.}\n\underline{The weekly and monthly models had no such computational problems, but in order to remain consistent across temporal levels, use of structural vector error correction models were rejected.}
Figure 47:
Plots of daily data for each data series.

Most of the series are skewed right, the volatility measures are highly leptokurtic while the prices are platykurtic (flat-topped).

Table 22. Descriptive statistics for daily data series.

Table 22 gives the descriptive statistics for the daily data series.

The Augmented Dickey-Fuller test was used to examine the stability of the series and test for the possible presence of unit roots.
Table 23. Augmented Dickey-Fuller tests for unit roots. (***) indicate significant at 1% level.

Table 23 shows the ADF tests indicate that nearly all of the series have unit roots. The series for gasoline price volatility appears to be \( I(0) \), and the series for oil futures price volatility, and retail gasoline prices are \( I(1) \), significant at the 5% level, but not the 1% level.

A vector autoregressive process was specified to estimate the appropriate lag structure for the model. The Akaike Information Criteria (AIC) and Forecast Prediction Error (FPE) criteria suggest a lag length of 38 days. The Hannan-Quinn (HQ) and Schwarz Criterion suggested a lag length of 8 days. These two possibilities were further investigated in light of the goal to specify a model to forecast retail unleaded gasoline prices.

Table 24 gives the results of standard diagnostic tests.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Deterministic Terms</th>
<th>Lags</th>
<th>Test Value</th>
<th>Critical Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBAA_AAA</td>
<td>constant, trend</td>
<td>38</td>
<td>-3.11</td>
<td>-3.96 -2.86</td>
</tr>
<tr>
<td>( \Delta )DBAA_AAA</td>
<td>constant, trend</td>
<td>37</td>
<td>-2.98**</td>
<td>-2.98 -2.86</td>
</tr>
<tr>
<td>OTERM</td>
<td>constant, trend</td>
<td>55</td>
<td>-1.61</td>
<td>-1.95 -1.62</td>
</tr>
<tr>
<td>( \Delta )TERM</td>
<td>constant, trend</td>
<td>55</td>
<td>-2.73**</td>
<td>-1.95 -1.62</td>
</tr>
<tr>
<td>OTERM</td>
<td>constant, trend</td>
<td>55</td>
<td>-2.73**</td>
<td>-1.95 -1.62</td>
</tr>
<tr>
<td>OTERM</td>
<td>constant, trend</td>
<td>55</td>
<td>-2.73**</td>
<td>-1.95 -1.62</td>
</tr>
<tr>
<td>DCRU</td>
<td>constant, trend</td>
<td>55</td>
<td>-2.52**</td>
<td>-1.95 -1.62</td>
</tr>
<tr>
<td>DCRU</td>
<td>constant, trend</td>
<td>55</td>
<td>-2.52**</td>
<td>-1.95 -1.62</td>
</tr>
<tr>
<td>DCRU</td>
<td>constant, trend</td>
<td>55</td>
<td>-2.52**</td>
<td>-1.95 -1.62</td>
</tr>
<tr>
<td>DCRU</td>
<td>constant, trend</td>
<td>55</td>
<td>-2.52**</td>
<td>-1.95 -1.62</td>
</tr>
<tr>
<td>DCRU</td>
<td>constant, trend</td>
<td>55</td>
<td>-2.52**</td>
<td>-1.95 -1.62</td>
</tr>
</tbody>
</table>

Table 24. Diagnostic tests for VAR(K) for daily data series.

The results of the diagnostic tests given in Table 24 show that the model is not particularly well specified. The absence of serial correlation and multivariate autoregressive conditional heteroskedasticity can be rejected\(^{38}\) from both models. Additionally, as is often the case with high-frequency data, the normality of the residuals from the model can be rejected.

\(^{38}\)The residuals were tested for autoregressive components. The presence of significant AR terms was rejected for each series using the AIC, and no estimated AR terms were greater than 0.005 in absolute value.
A comparison of the two models suggests that the lag length of 38 days is slightly more preferable.

Another reasonable test is to look for evidence of possible structural changes over the time period of this analysis. Figure 48 shows the OLS-CUSUM plots for the VAR(38). See, e.g., Zeileis et. al. (2002). Absent a structural change, the data should trend along the 0-axis. Despite insignificant deviations from time-to-time, this is generally the case with all seven series.

Given the relatively better choice of the model with 38 lags, the next step was to test for cointegration relationships in the series. The Johansen test was used and the results are provided in Table 25.
The results given in Table 25 suggest the presence of three cointegration relationships in the data series and the analysis proceeds with the assumption of three cointegration relationships for the daily data.

A vector error correction model with 38 lags (VECM(38)) was then estimated using the daily data. Diagnostic tests reveal potential misspecifications of the model. Such evidence is somewhat expected given the nature of the data series and the likelihood that the series have errors that follow a GARCH pattern.

Figure 49 shows the squared residuals for the gasoline price series (DGAS_RET), which shows typical volatility clustering patterns associated with GARCH\(^39\) processes.

\[^{39}\text{The implication is that the true model is of the form } \Delta y_t = \alpha \beta y_{t-1} + \sum_{i=1}^{K} \Delta y_{t-i} + \epsilon_t \text{ where } \epsilon_t = H_t^{1/2} \eta_t, \eta_t \sim N(0, I) \text{ and the elements } h_{ij,t} \text{ of } H_t^{1/2} \text{ are such that they follow a multivariate GARCH process.}\]
The larger question for this analysis is how well the model works to predict or forecast retail unleaded gasoline prices, and the dynamic effects within the model on gasoline prices. For the latter, impulse response functions and forecast error variance decompositions are employed. Figure 50 and Figure 51 show the orthogonal impulse response functions for gasoline prices (DGAS_RET). That information given shows several significant impacts on retail gasoline prices from other variables in the model.
Figure 50:
Responses of retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks.
Figure 51:
Responses of retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks.

Figure 52:
Forecast error variance decomposition for daily retail gasoline prices (DGAS_RET).
Figure 52 shows the forecast error variance decomposition. The forecast error variance decomposition shows that initially almost all of the fluctuation in gasoline prices occurs as a result of shocks in gasoline prices and price volatilities, but that after a couple of weeks oil futures prices account for about 30 percent of that variation.

Figure 53 shows the last 100 values for retail gasoline prices (DGAS RET) and the fitted values obtained from the VECM.

![Daily VECM(38) - Fitted vs. Actual](image)

Figure 53:
Actual versus fitted for retail gasoline prices for the last 100 observations.

The model appears to fit the data fairly well and seems to pick up the underlying data trend as well as some of the larger swings. The last 31 observations for the month of January 2011 were not used in the model estimation so that the forecast produced by the model could be evaluated using out-of-sample values. Figure 54 shows the plot of this actual versus forecast data.
Figure 54: Retail gasoline prices for January 2011 versus forecasted prices.

Although the model appears to do a very good job of forecasting the out-of-sample data (the mean squared prediction error is $0.0075$ per day), one should note that this month did not experience particularly wide swings in gasoline prices that are observed in later months of 2011 and 2012.

Weekly

The analysis techniques employed on the daily data is repeated using weekly data. In the earlier discussion of temporal aggregation the two methods grouped under the concept of temporal aggregation are lower frequency sampling and true temporal aggregation (for example adding monthly sales to obtain quarterly sales). The data used in this analysis is period averaged. Thus, weekly data is the average of one week of daily observations.
Figure 55:
Plots of weekly data for each data series.

Figure 55 shows the plots of the weekly series.

There are 775 weekly observations for the seven data series. Table 26 shows the descriptive statistics.

Table 26. Descriptive statistics for weekly data series.

<table>
<thead>
<tr>
<th></th>
<th>WBAA_ARA</th>
<th>WTERM</th>
<th>WTBILL</th>
<th>WCYL_FUT</th>
<th>WNGAS_RET</th>
<th>WVGAS_RET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.07</td>
<td>3.0654</td>
<td>1.9474</td>
<td>-0.1130</td>
<td>0.54736</td>
<td>0.00202</td>
</tr>
<tr>
<td>Variance</td>
<td>0.24</td>
<td>3.9448</td>
<td>1.7520</td>
<td>0.3671</td>
<td>0.14504</td>
<td>0.00005</td>
</tr>
<tr>
<td>Skew</td>
<td>2.8487</td>
<td>0.2140</td>
<td>0.0466</td>
<td>0.1126</td>
<td>-0.29</td>
<td>10.07</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>9.24</td>
<td>-1.50</td>
<td>-1.31</td>
<td>-1.02</td>
<td>-1.11</td>
<td>140.99</td>
</tr>
<tr>
<td>Max</td>
<td>3.47</td>
<td>6.22</td>
<td>4.17</td>
<td>1.22</td>
<td>1.3997</td>
<td>0.1255</td>
</tr>
<tr>
<td>Min</td>
<td>0.52</td>
<td>0.02</td>
<td>-0.4200</td>
<td>-1.396</td>
<td>-0.1221676</td>
<td>0.00000036</td>
</tr>
</tbody>
</table>
The weekly data show very similar patterns to the daily data. As with the daily data, the Augmented Dickey-Fuller tests were used to test for possible unit roots in the seven series. Table 27 gives the results of the ADF tests.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Deterministic Terms</th>
<th>Lags</th>
<th>Test Value</th>
<th>Critical Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>WBAA_AAA</td>
<td>constant, trend</td>
<td>5</td>
<td>-3.3</td>
<td>-3.96 -3.81 -3.12</td>
</tr>
<tr>
<td>GYWBAA_AAA</td>
<td></td>
<td>4</td>
<td>-8.74***</td>
<td>-2.59 -1.95 -1.62</td>
</tr>
<tr>
<td>WTERM</td>
<td>constant, trend</td>
<td>12</td>
<td>-1.78</td>
<td>-3.96 -3.41 -3.12</td>
</tr>
<tr>
<td>GYTERM</td>
<td></td>
<td>11</td>
<td>-7.17***</td>
<td>-2.06 -1.95 -1.62</td>
</tr>
<tr>
<td>WTBILL</td>
<td>constant, trend</td>
<td>36</td>
<td>-3.2</td>
<td>-3.96 -3.41 -3.12</td>
</tr>
<tr>
<td>GYWBILL</td>
<td></td>
<td>30</td>
<td>3***</td>
<td>-2.88 -1.95 -1.62</td>
</tr>
<tr>
<td>WOIL_FUT</td>
<td>constant, trend</td>
<td>27</td>
<td>-3.27</td>
<td>-3.96 -3.41 -3.12</td>
</tr>
<tr>
<td>GYWOIL_FUT</td>
<td></td>
<td>26</td>
<td>-5.40***</td>
<td>-2.59 -1.95 -1.62</td>
</tr>
<tr>
<td>WVGAS_RET</td>
<td>constant, trend</td>
<td>5</td>
<td>-3.74</td>
<td>-3.96 -3.41 -3.12</td>
</tr>
<tr>
<td>GYWVGAS_RET</td>
<td></td>
<td>4</td>
<td>-10.07***</td>
<td>-3.43 -2.86 -2.57</td>
</tr>
<tr>
<td>WVOIL_FUT</td>
<td>constant, trend</td>
<td>37</td>
<td>-2.53</td>
<td>-3.43 -2.86 -2.57</td>
</tr>
<tr>
<td>GYWVOIL_FUT</td>
<td></td>
<td>36</td>
<td>-5.54***</td>
<td>-2.59 -1.95 -1.62</td>
</tr>
<tr>
<td>WVGAS_RET</td>
<td>constant, trend</td>
<td>31</td>
<td>-5.54***</td>
<td>-3.43 -2.86 -2.57</td>
</tr>
<tr>
<td>GYWVGAS_RET</td>
<td></td>
<td>30</td>
<td>-7.57***</td>
<td>-2.59 -1.95 -1.62</td>
</tr>
</tbody>
</table>

Table 27. Augmented Dickey-Fuller tests for unit roots. (*** indicates significant at 1% level.

As expected the weekly data displays the same unit root characteristics as the daily data. The VAR lag selection criteria suggested three possible lag lengths. The AIC and FPE selected 7 lags, the HQ selected 5 lags, and the SC selected 2 lags. As with the analysis of the daily data, each suggested lag length was evaluated.

Table 28 shows the results of the diagnostic tests (serial correlation, normality, and ARCH effects).

<table>
<thead>
<tr>
<th>Model</th>
<th>SS</th>
<th>prob</th>
<th>JB</th>
<th>prob</th>
<th>MARCH</th>
<th>prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>k=2</td>
<td>1602.2</td>
<td>0.005</td>
<td>843.244</td>
<td>0.000</td>
<td>16,827.30</td>
<td>0.000</td>
</tr>
<tr>
<td>k=5</td>
<td>1395.4</td>
<td>0.000</td>
<td>674.050</td>
<td>0.000</td>
<td>19,581.6</td>
<td>0.000</td>
</tr>
<tr>
<td>k=7</td>
<td>1332.7</td>
<td>0.000</td>
<td>637.951</td>
<td>0.000</td>
<td>18,404.70</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 28. Diagnostic tests for VAR(K) for weekly data series.

As with the daily series, the models for the weekly series are not particularly well specified showing signs of serial correlation, non-normality and ARCH effects. Again, the model with a longer lag length appears somewhat preferable.

The results of the OLS-CUSUM are shown in Figure 56, and, as with the daily data, there does not appear to have been any structural changes in the data over the sample period.
The results of the Johansen test for cointegration relationships are shown in Table 29. Again the result correspond to those for the daily data, and three cointegration relationships are suggested. Thus a vector error correction model with seven lags (six lags with the first differencing) and three cointegration relationships was estimated. Figure 57 shows the same volatility clustering that was observed with the daily data.

Table 29. Johansen cointegration tests for weekly data with 7 lags.

<table>
<thead>
<tr>
<th></th>
<th>test</th>
<th>10pct</th>
<th>5pct</th>
<th>1pct</th>
</tr>
</thead>
<tbody>
<tr>
<td>m0</td>
<td>290.34</td>
<td>141.01</td>
<td>146.76</td>
<td>156.49</td>
</tr>
<tr>
<td>r=1</td>
<td>174.97</td>
<td>110.42</td>
<td>114.90</td>
<td>124.75</td>
</tr>
<tr>
<td>r=2</td>
<td>199.32</td>
<td>83.21</td>
<td>87.31</td>
<td>96.59</td>
</tr>
<tr>
<td>r=3</td>
<td>27.7</td>
<td>69.14</td>
<td>82.90</td>
<td>90.05</td>
</tr>
<tr>
<td>r=4</td>
<td>30.44</td>
<td>39.06</td>
<td>42.44</td>
<td>48.45</td>
</tr>
<tr>
<td>r=5</td>
<td>11.73</td>
<td>22.76</td>
<td>25.82</td>
<td>30.45</td>
</tr>
<tr>
<td>r=6</td>
<td>2.98</td>
<td>10.49</td>
<td>12.25</td>
<td>15.26</td>
</tr>
</tbody>
</table>
The orthogonal impulse response functions for weekly retail gas prices (WGAS_RET) are shown in Figure 58 and Figure 59. The impulse responses of weekly retail gasoline prices are quite similar to their daily price counterparts (over the first four weeks). There are some differences in the significance of some responses, but overall the response trends for the 30-day daily price responses are mirrored by the 4-week weekly price responses.
Figure 58:
Responses of weekly retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks.
Figure 59:
Responses of weekly retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks.

Figure 60:
Forecast error variance decomposition for weekly retail gasoline prices (WGAS_RET).
Figure 60 shows a graph of the four week forecast error variance decomposition. The forecast error variance decompositions for weekly gasoline prices are also similar to the daily versions. The response from gas price shocks accounts for the largest share of error variance, but that declines over the four week period as oil futures price shocks increase. Figure 61 shows the actual versus fitted values for the weekly retail gasoline data.

To illustrate the performance of the models suggested by the lag selection criteria, Figure 61 shows the fitted versus actual data for the last six weeks of 2010. This graphical perspective shows that the seven-week lag seems slightly superior to the other two.

Figure 62 shows the forecast versus actual performance over the first four weeks of 2011. The forecasts is where we begin to observe the effects of temporal aggregation. The weekly
price forecast appears much worse than the forecast of the daily model (the weekly MSPE is $0.045, six times larger than the daily MSPE).

Figure 62: Weekly retail gasoline prices for January 2011 versus forecasted prices.

The forecast performance of the weekly model seems rather poor, but as we see below, the forecast performance over 26 weeks is much worse (the MSPE for the 26-week forecast is $0.462).

Monthly

The analysis of the monthly series proceeds in the same manner as for the weekly and daily series. The plots of the individual data series are shown in Figure 63. The 178 observations for monthly series all resemble the general trends observed in the weekly and monthly counterparts. The monthly data is obviously “smoother” than the weekly or daily series, which is to be expected as the aggregated data are period averages of the daily data.
Figure 63: Plots of monthly data for each data series.

Table 30 shows the descriptive statistics for the monthly data. Again they are similar to those for the daily data.

<table>
<thead>
<tr>
<th></th>
<th>WBAI_AAA</th>
<th>WTERM</th>
<th>WTBILL</th>
<th>WOIL_FUT</th>
<th>WGS_RET</th>
<th>WOIL_FUT</th>
<th>WVGAS_RET</th>
<th>WVGAS_RET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.01</td>
<td>1.9459</td>
<td>3.0631</td>
<td>-0.1124</td>
<td>0.04812</td>
<td>0.01525</td>
<td>0.02598</td>
<td>0.02598</td>
</tr>
<tr>
<td>Variance</td>
<td>0.24</td>
<td>1.7941</td>
<td>3.9543</td>
<td>0.3870</td>
<td>0.14983</td>
<td>0.00265</td>
<td>0.02718</td>
<td>0.02718</td>
</tr>
<tr>
<td>Skew</td>
<td>2.8152</td>
<td>0.0439</td>
<td>-0.2127</td>
<td>0.4125</td>
<td>0.29</td>
<td>8.35</td>
<td>7.16</td>
<td>7.16</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>8.96</td>
<td>-1.32</td>
<td>-1.50</td>
<td>-1.03</td>
<td>-1.11</td>
<td>41.24</td>
<td>59.17</td>
<td>59.17</td>
</tr>
<tr>
<td>Max</td>
<td>3.38</td>
<td>4.07</td>
<td>6.17</td>
<td>1.16</td>
<td>1.3883</td>
<td>0.4027</td>
<td>0.0361</td>
<td>0.0361</td>
</tr>
<tr>
<td>Min</td>
<td>0.55</td>
<td>-0.33</td>
<td>0.0300</td>
<td>-1.3120</td>
<td>-0.053656</td>
<td>0.000062</td>
<td>0.0100087</td>
<td>0.0100087</td>
</tr>
</tbody>
</table>

Table 30. Descriptive statistics for monthly data series.

The monthly data was examined for the presence of unit roots using the Augmented Dickey-Fuller tests. These test results are shown in Table 31.
Table 31. Augmented Dickey-Fuller tests for unit roots. (***) indicate significant at 1% level.

The unit root analysis mirrors those of the daily and weekly data with unit roots suggested for six of the seven series (all except the gasoline price volatility series).

The VAR lag selection process suggested possible lags of two months (HQ and SC) and six months (AIC and FPE). Thus the two suggested models were evaluated. Table 32 shows the results of the diagnostic tests.

Table 32. Diagnostic tests for VAR(K) for monthly data series.

Table 32. Diagnostic tests for VAR(K) for monthly data series.

Once again the series show evidence of serial correlation and nonnormality, but unlike for the daily and weekly series, the null hypothesis of ARCH effects can be rejected with a p-value of one.

The tests for cointegrating relationships suggested three cointegrating relationships for a lag length of two, but only two cointegrating relationships for a lag length of six. The results for both candidate models are shown in Table 33.
Table 33. Johansen cointegration tests for monthly data with 2 and 6 lags.

In light of the results presented in Table 33, use of a lag length of two was chosen. This was done to reflect the general theoretical conclusion that cointegration relationships survive temporal aggregation and both the daily and weekly models exhibited three cointegrating relationships. An examination of the model fit, shown in Figure 64, supports this choice.

![Monthly VECMs - Fitted vs. Actual](image)

Figure 64:
Actual versus fitted for retail gasoline prices for the last 12 months for proposed lags of 2, and 6 months.
The fit of both models is shown for the data covering 2010. Both lag length candidates supply similar fits of the actual data.

Figure 65 and Figure 66 show the impulse responses for monthly retail gasoline prices for a 12-month period. The impulse response functions for the monthly model specification are relatively close to the weekly counterparts. However, the impulse responses for monthly retail gasoline price volatilities show some differences.

**Figure 65:**
Responses of monthly retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks.
Figure 66:
Responses of monthly retail gasoline prices to economic, gasoline price volatility and oil price and price volatility shocks.

Figure 67:
Forecast error variance decomposition for monthly retail gasoline prices (MGAS_RET).
The forecast error variance decomposition is shown in Figure 67. There are some slight differences in the forecast error variance decomposition between the monthly series and the weekly series. The monthly model shows less of the error variance attributable to the retail gasoline prices and relatively more to the oil futures price. Also, some of the economic series account for more of the variation than the weekly model would suggest (taking the place of the oil futures price volatility). The implications for the model forecasts are shown in Figure 68.

Figure 68:
Monthly retail gasoline prices for January and February 2011 versus forecasted prices.

The forecasts for the two months after the sample deviate significantly from the actual observed retail gasoline prices (the FPE for January is $0.035, which is better than the weekly model forecasts, but much worse, again, than the daily model forecasts). The forecast performance does not improve. The model forecast for the entire 12 months of 2011 has an MSPE of $0.501 because this model lacks the ability to track the rapid price increase experienced during 2011.
The sections above support much of the existing literature with respect to the effects of temporal aggregation. All data periods show the presence of similar unit roots, and estimated cointegrating relationships are largely consistent\textsuperscript{40} across different data aggregations.

The models’ forecast performance varied considerably with the daily forecast being far superior. This is not unexpected as prior research indicates that forecast properties do not tend to survive temporal aggregation. This suggests that a model might be rejected based on forecast performance simply because the data was temporally aggregated to a lower frequency than the true data generating process. Figure 20 shows the 180-day ahead forecast for retail gasoline prices along with 95% the confidence interval and the corresponding monthly prices gleaned from the monthly data series.

![Figure 20: 180-day forecast for retail gasoline prices.](image)

**Figure 20:** 180-day forecast for retail gasoline prices.

Figure 69 also shows that the daily point forecasts vary significantly from the actual prices that occurred during the first half of 2011, although all of the monthly snapshot retail gasoline prices are within the 95% confidence interval.

**Impacts from Conditional Heteroskedasticity**

\textsuperscript{40}The single exception was the monthly data for the model incorporating 6 lags when there were two estimated cointegration relationships instead of the expected three.
What is impact on the estimation of a VECM process when error terms follow a GARCH process? In order to provide some insight into this question, a simulation was conducted using a four equation vector error correction data generating process with known coefficients and a known cointegration relationship (r=2 in this instance). This model was specified with two different error processes. In one the error terms follow an unknown GARCH process, and in the other, the error terms are distributed as four univariate normal random variables with mean zero and standard error of 0.05.

The data generating process was used to simulate 3,652 observations for four data series. the underlying data generating process was of the form $\Delta y_t = \nu + \alpha \beta y_{t-1} + \gamma_1 \Delta y_{t-1} + \gamma_2 \Delta y_{t-2} + \text{error}_t$ where the error term is either a GARCH process, or is normally distributed. The error terms were simulated 100 times and the mean values of the estimated coefficients were calculated. The GARCH error process was obtained by extracting the error terms from TBILL, TERM, DVOIL_FUT, and DGAS_RET from the final vector error correction model. These error terms were them randomly sampled in blocks of 100 observations to create the error terms to be appended to the known VECM process.

<table>
<thead>
<tr>
<th></th>
<th>Y1</th>
<th>Y2</th>
<th>Y3</th>
<th>Y4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>-0.379</td>
<td>0.081</td>
<td>0.243</td>
<td>-1.409</td>
</tr>
<tr>
<td></td>
<td>-0.174</td>
<td>-0.507</td>
<td>0.222</td>
<td>-1.106</td>
</tr>
<tr>
<td></td>
<td>-0.012</td>
<td>-0.002</td>
<td>0.371</td>
<td>-1.031</td>
</tr>
<tr>
<td>GARCH</td>
<td>-0.069</td>
<td>0.256</td>
<td>0.236</td>
<td>-1.144</td>
</tr>
<tr>
<td></td>
<td>0.154</td>
<td>0.508</td>
<td>0.301</td>
<td>-1.063</td>
</tr>
<tr>
<td></td>
<td>-0.363</td>
<td>0.001</td>
<td>0.371</td>
<td>-1.031</td>
</tr>
<tr>
<td></td>
<td>-0.011</td>
<td>0.003</td>
<td>0.010</td>
<td>-0.071</td>
</tr>
<tr>
<td>Actual</td>
<td>-0.362</td>
<td>0.082</td>
<td>0.243</td>
<td>-1.409</td>
</tr>
<tr>
<td></td>
<td>-0.174</td>
<td>0.256</td>
<td>0.236</td>
<td>-1.106</td>
</tr>
<tr>
<td></td>
<td>-0.363</td>
<td>0.001</td>
<td>0.371</td>
<td>-1.030</td>
</tr>
<tr>
<td></td>
<td>-0.012</td>
<td>0.003</td>
<td>0.010</td>
<td>-0.091</td>
</tr>
</tbody>
</table>

Table 34. Actual versus estimated coefficients for $\gamma_1$.

Table 34 compares the two sets of estimated coefficient for $\gamma_1$ for the model with normal errors and the model with GARCH errors. In 37.5 percent of the cases, the estimates with
GARCH errors have smaller deviations from the actual coefficients. Table 35 shows the same comparisons for $\gamma_2$.

<table>
<thead>
<tr>
<th></th>
<th>Y1</th>
<th>Y2</th>
<th>Y3</th>
<th>Y4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>0.000</td>
<td>0.121</td>
<td>0.042</td>
<td>-1.014</td>
</tr>
<tr>
<td></td>
<td>0.191</td>
<td>-0.196</td>
<td>0.019</td>
<td>-0.350</td>
</tr>
<tr>
<td></td>
<td>0.062</td>
<td>-0.244</td>
<td>0.106</td>
<td>-0.679</td>
</tr>
<tr>
<td></td>
<td>-0.057</td>
<td>0.012</td>
<td>0.009</td>
<td>0.105</td>
</tr>
<tr>
<td>GARCH</td>
<td>0.046</td>
<td>0.122</td>
<td>0.052</td>
<td>-1.001</td>
</tr>
<tr>
<td></td>
<td>0.199</td>
<td>-0.197</td>
<td>-0.001</td>
<td>-0.342</td>
</tr>
<tr>
<td></td>
<td>0.063</td>
<td>-0.246</td>
<td>0.103</td>
<td>-0.677</td>
</tr>
<tr>
<td></td>
<td>-0.056</td>
<td>0.012</td>
<td>0.007</td>
<td>0.115</td>
</tr>
<tr>
<td>Actual</td>
<td>0.051</td>
<td>0.122</td>
<td>0.043</td>
<td>-1.014</td>
</tr>
<tr>
<td></td>
<td>0.190</td>
<td>-0.196</td>
<td>0.012</td>
<td>-0.352</td>
</tr>
<tr>
<td></td>
<td>0.063</td>
<td>-0.246</td>
<td>0.103</td>
<td>-0.677</td>
</tr>
<tr>
<td></td>
<td>-0.057</td>
<td>0.011</td>
<td>0.011</td>
<td>0.106</td>
</tr>
</tbody>
</table>

Table 35. Actual versus estimated coefficients for $\gamma_2$.

Again, 37.5$^{41}$ percent of the estimates under GARCH errors are closer to the actual coefficient values.

Model diagnostics were also generated in order to compare the two models. Table 36 shows the tests for serial correlation, normality, and ARCH effect.

<table>
<thead>
<tr>
<th></th>
<th>B0</th>
<th>p-value</th>
<th>B1</th>
<th>p-value</th>
<th>BARCH1</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>115.2</td>
<td>0.784</td>
<td>6.16</td>
<td>0.430</td>
<td>0.001</td>
<td>0.114</td>
</tr>
<tr>
<td>GARCH</td>
<td>304.25</td>
<td>0.000</td>
<td>196.493</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 36. Diagnostic tests for model with normal errors versus model with GARCH errors.

Table 36 shows that the model with normal errors shows no serial correlation, is multivariate normally distributed, and has no ARCH effects. The model with GARCH errors cannot reject the nulls of serial correlation, and nonnormality, and does reject the null of no ARCH effects, which is suspected as a major source of the misspecification in the models presented in the main part of this paper.

How does the presence of GARCH errors impact the usefulness of the model? Figure 70 and Figure 71 show comparisons of impulse response functions.

$^{41}$To check the sensitivity of these results, a simulation was conducted with 500 repetitions and 50 percent of the estimates with GARCH errors were closer to the known parameter values.
Figure 70:
Impulse responses for 14 periods for both types of error terms. Impacts are shown on Y4 from impulses in Y1, and Y2 for the model with normal errors (left) and the model with GARCH errors (right).
Figure 71:
Impulse responses for 14 periods for both types of error terms. Impacts are shown on Y4 from impulses in Y3, and Y4 for the model with normal errors (left) and the model with GARCH errors (right).

The impulse response functions are fairly similar except for the responses to impulses from variable Y3, which are much larger under normal error terms.

Table 37. Forecast error variance decompositions for normal and GARCH errors.

Table 37 shows the forecast error variance decomposition. The FEVD results are much different. The model under normal errors has nearly all of the impact coming from changes
in Y4 for the entire period. The model under GARCH errors is split between Y1, Y2, and Y4. Another possible evaluation tool is the associated forecasts for each model.

Both models appear to provide reasonable forecasts. Mean squared prediction errors are given in Table 38. The MSPE is lower for the model with GARCH error processes. This was based on a 52-step ahead forecast after estimating the models withholding the last 52 observations.

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th>GARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>0.1105</td>
<td>0.0092</td>
</tr>
<tr>
<td>Y2</td>
<td>0.1654</td>
<td>0.0033</td>
</tr>
<tr>
<td>Y3</td>
<td>0.0962</td>
<td>0.0177</td>
</tr>
<tr>
<td>Y4</td>
<td>0.0311</td>
<td>0.0095</td>
</tr>
</tbody>
</table>

Table 38. Mean squared prediction errors for both error processes.

The presence of errors that follow a GARCH process is one of several possible misspecifications that can affect a vector error correction model analysis. Coefficient estimates and impulse response functions appear fairly stable between the two error schemes, but forecast error variance decompositions can be significantly different, while the GARCH error process forecasts are not inferior to those of the normal error process.

Conclusion

This paper has presented a simple recursive vector error correction model of retail unleaded gasoline prices using lagged values of interest rate variables, which have been shown in the finance literature to correspond with overall macroeconomic activity, the prices and price volatilities of the first contract for oil futures, and prices and price volatilities for the U.S. average retail unleaded gasoline. Three such models were estimated, one using daily data, one using weekly data, and one using monthly data.

There are several overall goals of this analysis. One goal is to examine whether macroeconomic performance is useful in forecasting retail gasoline prices via a mechanism that also
allows their impacts to flow through changes in oil futures price volatilities, retail gasoline price volatilities, oil futures prices, and retail gasoline prices. Another goal is to continue the examination of the impacts on retail gasoline prices from oil and gasoline price volatilities as well as prices, particularly looking at oil futures prices. A third goal is to examine the impacts from temporal aggregation in the context of a vector error correction model used for the purpose of providing a forecast of retail gasoline prices.

These models suggest that macroeconomic performance does impact retail gasoline prices. Significant responses in gasoline prices were found for all three macroeconomic variables, particularly for weekly and monthly data. The difference in corporate bond rates (BAA_AAA) is expected to be negatively correlated with economic activity, and we generally observe increases in this variable associated with falling gasoline prices. Expectations for the 90-day treasury bill rate and the difference between the average long-term treasury bond rate and the 90-day treasury bill rate are that these will be positively related to economic activity, just the model concludes.

Price volatilities in this model are not as significant at driving retail gasoline prices. Gasoline price volatility has a short duration positive impact on gasoline prices, whereas oil futures price volatility has an overall mixed but insignificant impact in the short-run and a largely negative impact in the longer term. It is clear, however, that retail gasoline prices are significantly impacted by oil futures prices.

Finally, the analysis supports prior research that temporal aggregation maintains certain aspects of the underlying data generating process (units roots and cointegrating relationships), but loses other aspects (forecasts). Looking at the forecast accuracy, the model appears to perform well for daily, short-run forecasts, but does quite poorly with weekly or monthly data and longer-run forecasts. One implication of this is that a researcher using temporally aggregated data might be tempted to reject a particular model formulation based on apparent poor performance that arises as a result of missing information.
Overall it is clear that care must be taken when viewing these results. The presence of errors that appear to follow a GARCH process can possibly influence the reliability of any results obtained. At the same time it is quite likely that this model is not a complete description of the energy markets.


Alonzo, W., 1964, Location and land use: toward a general theory of land rent, Harvard University Press.


