Application of Demand Analysis Framework to Understand the Price and Volume Movements of Exchange Traded Funds (ETFs)

Chengcheng Fei

Department of Agricultural Economics

Texas A&M University

feiccheng@tamu.edu

Chen Gao

Department of Agricultural Economics

Texas A&M University

ChenFengV4@tamu.edu

Erin M. Hardin

Department of Agricultural Economics

Texas A&M University

ehardin@tamu.edu

Senarath Dharmasena

Department of Agricultural Economics

Agribusiness, Food and Consumer Economics Research Center

Texas A&M University

sdharmasena@tamu.edu

Selected paper prepared for presentation at the Southern Agricultural Economics Association's 2016 Annual Meeting, San Antonio, Texas. February 6-9, 2016.

Abstract

Exchange Traded Funds (ETFs), allow for the study of the demand for investment in a commodity or an industry surrounding a commodity. They are comparable to mutual funds where they separate ownership among shares and are a combination of several assets. ETFs are marketable securities traded on stock exchanges and are available in following industries: commodities, consumer, financial, health, natural resources, real estate, technology, and utilities. Objective of this paper is to assess the usefulness of using an ETF's volume as a proxy for demand and the price as the independent variable to determine if demand relationships exist between the two measures. Daily price and volume data for 21 oil and natural gas ETFs was collected from Bloomberg for five years (March 2010 - December 2014). The energy ETFs were categorized into oil, natural gas and general energy. A system of demand equations was developed for aforementioned energy ETFs using seemingly unrelated regression method. Preliminary results show that, own-price elasticity of demand for oil and natural gas is -0.69 and -0.27. Also, we find that natural gas and oil are substitutes, which is generally observed in energy markets. Consequently, the use of ETFs as proxies for their industries shows promise.

Application of Demand Analysis Framework to Understand the Price and Volume Movements of Exchange Traded Funds (ETFs)

Introduction

There exists a substantial amount of research on the correlation between the volume and price of a stock. Research on understanding the actual demand of a stock using this information is far less explored. The goal of this paper is to assess the usefulness of using a stock's volume as a proxy for its quantity demand and the price of a stock as the independent variable to determine if a demand relationship exists between the two measures.

In addition to modeling the demand, a relatively new marketable security will be used in place of the standard publicly traded company. Exchange Traded Funds, known in short as "ETFs," allow for the study of the demand for investment in a commodity or an entire industry surrounding a commodity. ETFs were first introduced in 1989. They are comparable to mutual funds in that they separate ownership among shares and are a combination of several assets tracking usually an index, a commodity, a bond or a basket of assets. In contrast to mutual funds, ETFs are marketable securities traded on stock exchanges similar to common stock.

The introduction of ETFs in the marketplace has made investment in more difficult assets such as futures contracts, simpler. Previously, investors not interested in managing a futures portfolio due to it higher degree of maintenance were left out of the market for many commodities or were restricted to investing in companies associated with the production of the commodity. With ETFs investors can now own futures contracts of commodities without the responsibility of constantly managing their position. Additionally, ETFs allow investors to own diversified portfolios without going through a broker.

Today, there are well over 1,400 ETFs available totaling $1.794 trillion in assets. ETFs span several industries with major concentrations being commodities, consumer, financial, health, natural resources, real estate, technology, utilities, and others. The largest groups in number of ETFs and asset value are commodities and natural resources. Considering the number of categories of ETFs and the size of the ETF market overall, the energy sector under the commodity concentration will be the focus of this paper. Energy ETFs have made investment into the energy market easier and more efficient leading to the belief that there exists substantial demand for these assets. Additionally, trading volume is relatively large which is necessary for a robust analysis.

While the market and popularity of these products continue to grow, we realized little financial and econometric analysis has been expanded to include them. By expanding on the present literature of price and volume and considering ETFs, this paper is expanding on two areas of needed research.

Previous research on the relationship between price and volume has focused primarily on establishing correlation. Research has also only focused on common stocks. The earliest work with Godfrey, Granger and Morgenstern (1964) showing that daily volume correlates positively with the difference between the daily high and daily low price. In addition to the consensus that there exists a relationship between price and volume, much of the previous research believes that the correlation is tied to information flow.

Smirlock and Starks (1984) investigate the information arrival process by modeling the price and volume relationship using Granger Causality. They test the possibility of two kinds of information processes. The first is a simultaneous process where investors receive the

information simultaneously and then revises their expectations and trade. The next process, sequential, allows for intermediate equilibria. Traders receive information sequentially and trading occurs after each reception leading to a series of price and volume outcomes. Results show strong support for the sequential information flow. Crouch (1970) found positive correlations between the absolute values of daily price changes and daily volumes for both market indices and individual stocks. The results of Crouch's (1970) work to market indices supports our expansion to a model using products, which are an index for a sector of a market. In our research, we see a simultaneity issue in our directed graphs. We believe a sequential process exists but data limitations prevent us from capturing it. Wood, McInish and Ord (1985) use transactions data. Transactions data captures the outcome of every trade made. Using this high of a frequency of data may allow their model to capture the sequential pattern which ours misses.

Smirlock and Starks (1984) also find that the causal relationship seen between absolute price changes and volume is stronger in periods surrounding earnings announcements. In future research, we could compare periods of time approaching contract expiration for the underlying futures contracts to further out months. Since price volatility tends to increase in periods approaching contract expiration, we may see interesting patterns in the price and volume changes.

Considerable review was done on the literature concerning the relationship between the different energy products primarily oil and natural gas. Bachmeier and Griffin's (2006) work gives a sufficient overview of the market integration between oil, coal and natural gas. They found weak integration between markets and did not find a primary energy market. Villar and Joutz (2006), on the other hand, find oil and natural gas are compliments in production and

substitutes in consumption. They also find that due to the relative sizes of the markets cause an asymmetric relationship between the two markets. For example, a 1 month temporary shock to WTI of 20% has a 5% contemporaneous impact on natural gas prices and dissipates to 2% in two months. A 20% permanent shock in WTI lead to a 16% increase in the Henry Hub price one year out, holding all else equal. Therefore, we see that oil prices may have an influence on the long run development of natural gas prices but the converse is not true. They also look at the relationship of the two commodities from the perspective of supply and demand. The net effect of an increase in the price of oil leads to an ambiguous change in the production or supple of natural gas. The demand for natural gas has a positive net effect for changes in price of oil.

Data & Methodology

Daily price and volume data for 21 oil and natural gas ETFs was collected from Bloomberg for five years from March 2010 to December 2014. Summary statistic of the ETF data are in Table 1. The price of each ETF is in dollars and volume is calculated as the number of shares of a stock exchanged in a day. ETF selection within the oil and gas category was restricted to those that were actively traded. Therefore young ETFs or those with a significant amount of missing data were dropped from the sample. For ETFs with a small amount of data missing, the moving average was calculated and used for the missing dates. Additionally, data for the construction of the instrumental variables for the prices of the ETFs were needed. Macro data for quarterly gross domestic product, monthly consumer price index, daily returns for 3-

month Treasury bill rates, daily S&P 500 prices and daily NASDAQ prices were collected from

Datastream for five years.

Table 1. Summary Statistics ETFs Price and Volume Data

	GAZ		RJN		OLO		JJE		OIL		UNL		UNG	
	Price	Volume	Price	Volume	Price	Volume	Price	Volume	Price	Volume	Price	Volume	Price	Volume
Mean	4.67	157966.40	6.43	77329.14	12.94	13788.13	18.87	7636.29	22.77	869161.05	23.61	26789.44	29.87	6612351.01
StDev	2.76	224880.02	0.85	116109.25	1.79	25193.15	3.40	20296.82	3.60	1230253.15	8.70	39920.89	14.30	5588102.91
95 % LCI	4.49	143822.46	6.37	70026.39	12.83	12156.00	18.65	6324.65	22.55	791783.63	23.07	24278.60	28.97	6260884.34
95 % UCI	4.84	172110.35	6.48	84631.89	13.06	15420.27	19.09	8947.94	23.00	946538.47	24.16	29300.29	30.77	6963817.68
Min	1.20	4560.00	3.40	104.00	6.50	1.00	8.80	1.00	9.44	41955.00	12.47	162.00	13.10	729204.00
Median	3.19	95524.00	6.61	41750.00	13.20	5668.00	18.53	2440.00	23.15	503546.00	19.25	15526.00	22.28	4811660.00
Max	12.18	2983197.00	7.97	1873955.00	16.81	313465.00	26.35	456780.00	29.98	11722165.00	45.88	611328.00	70.64	54614952.00
Skewness	0.89	5.27	-1.74	5.30	-1.64	5.49	-0.65	11.36	-1.79	4.09	0.91	5.93	0.96	2.82
Kurtosis	-0.62	41.51	3.37	53.21	3.67	44.94	1.05	210.01	4.10	21.21	-0.52	57.19	-0.38	13.15
	DBO		DBE		USO		UHN		USL		UGA			
	Price	Volume	Price	Volume	Price	Volume	Volume	Price	Volume	Price	Volume	Price		
Mean	26.42	376027.79	27.09	103239.04	34.81	9966778.56	31.32	4920.68	40.68	41539.35	50.37	55463.11		
StDev	3.76	463574.33	3.48	215064.99	5.11	8504046.61	4.02	6210.48	5.05	60770.03	9.44	71077.48		
95 % LCI	26.19	346871.04	26.87	89712.41	34.49	9431912.08	31.06	4525.53	40.37	37717.19	49.77	50992.65		
95 % UCI	26.66	405184.54	27.31	116765.66	35.13	10501645.04	31.57	5315.83	41.00	45361.52	50.96	59933.56		
Min	12.57	11421.00	14.96	2702.00	15.96	919762.00	18.91	11.00	22.94	318.00	29.45	2349.00		
Median	27.01	239490.00	28.14	48005.00	35.38	7620638.00	32.53	2927.00	41.48	21812.00	52.97	33798.00		
Max	34.26	7476688.00	33.45	3658090.00	45.15	76460976.00	38.67	58135.00	51.26	584788.00	65.71	899816.00		
Skewness	-1.77	5.09	-1.66	7.71	-1.76	2.62	-1.03	3.32	-1.48	4.15	-0.63	4.63		
Kurtosis	4.12	51.68	2.85	87.36	4.01	9.70	0.40	16.55	2.99	23.41	-0.85	33.29		

A correlation matrix for the price and quantity indexes for each ETF was calculated and used to determine the causal graphs of the ETFs. The GES Algorithm, a scoring metric algorithm which considers causal sufficiency and faithfulness to search over equivalence classes scoring each to find the best model is used to determine the relationship between categories and indices. A penalty value of 0.75 was used. Two results were surmised from the graphs. Firstly, many "bi-directed" graphs were seen in the result. One conclusion that can be made is that there may exist a simultaneity problem. The simultaneity problem may arise from the issue of even the daily data not being in small enough increments of time to catch or understand the underlying relationships that are occurring. Considering the rapid movement of data with the use of algorithmic and high frequency trading this is highly possible. The movement of information is likely faster than a daily average can capture.

Secondly, the results of the directed graph showed highly endogenous prices and highly exogenous volume. There was also a great deal of relation among the prices and the quantities with little interaction between the two groups. The direct graph for the ETFs can be seen in Graph 1. Similar results were found in Smirlock and Stark's paper, (1984). Their results from their Granger causality tests show there exists a relationship where price causes volume and volume causes price. As a result, instrumental variables were created to reduce the effects of the endogeneity problem for all ETF prices. Several macro variables were considered for the IVs: GDP, CPI, S&P 500 and NASDAQ. Due to the high correlation of the macro factors (the lowest correlation coefficient being .9133) and to remain consistent with the frequency of the ETF data, the S&P 500 was chosen. In only one case, GDP performed better than S&P 500 and was therefore used in the estimation. The 3-month T-Bill was also included as it is included in many of the ETFs portfolios. Additionally, the underlying energy futures contracts were included.

Graph 1. Directed Graph of ETFs with Penalty Value of 1 and No Knowledge Imposed.
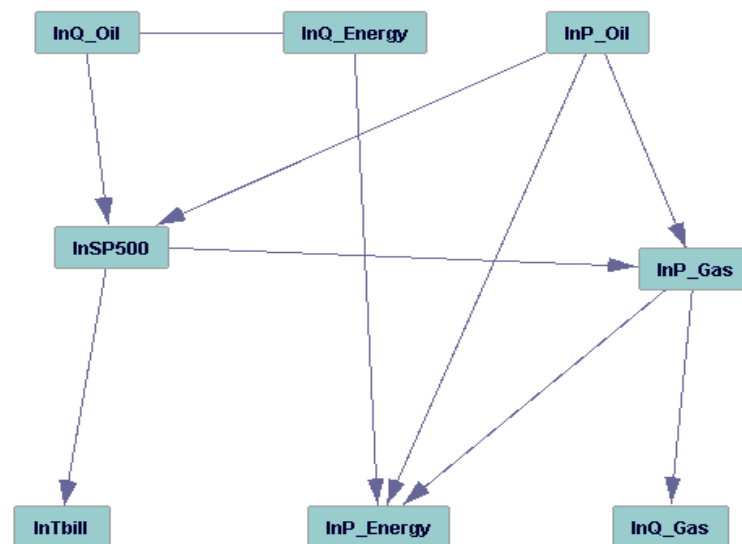
The NYMEX Crude Oil Futures Continuous, the NYMEX Henry Hub Natural Gas Futures Continuous and the NYM E-Mini Heating Oil Continuous futures contracts were used specifically for each underlying commodity to represent the markets as a whole. The natural logarithm was taken of the dependent and all independent variables. There also existed issues of heteroskedasticity. A generalized autoregressive conditional heteroskedasticity (GARCH) model was used to address and minimize this problem. GARCH takes the following form:

$$\ln P = f(P_{i,t-1}, S\&P\ 500, T\text{-Bill}_{3mo}, \text{Futures Prices}) + \varepsilon_t$$

$$\text{where } \varepsilon t = \sigma_2 z_t$$

$$\text{and } \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Due to the differences in investments of each ETF, a uniform function could not be created for all the prices. The resulting IV equations are given as follows:

(i) $\ln(\hat{P}_{GAZ}) = f\left(\ln(P_{GAZ,t-1}), \ln(S\&P\ 500), \ln(1 + Tbill_{3mo}), \ln(P_{Nat\ Gas})\right) +$ $GARCH\ (1,1)$

(ii) $\ln(\hat{P}_{RJN}) = f\left(\ln(P_{RJN,t-1}), \ln(S\&P\ 500), \ln(1 + Tbill_{3mo}), \ln(P_{Crude\ Oil}),\right.$ $\left. \ln(P_{Nat\ Gas})\right) + GARCH\ (1,2)$

(iii) $\ln(\hat{P}_{OLO}) = f\left(\ln(P_{olo,t-1}), \ln(S\&P\ 500), \ln(P_{Crude\ Oil})\right) + GARCH\ (1,1)$

(iv) $\ln(\hat{P}_{JJE}) = f\left(\ln(P_{JJE,t-1}), \ln(S\&P\ 500), \ln(P_{Crude\ Oil}), \ln(Nat\ Gas)\right) +$ $GARCH\ (1,1)$

(v) $\ln(\hat{P}_{OIL}) = f\left(\ln(P_{OIL,t-1}), \ln(S\&P\ 500), \ln(P_{Crude\ Oil}), \ln(P_{Nat\ Gas})\right) +$ $GARCH\ (1,1)$

(vi) $\ln(\hat{P}_{UNL}) = f\big(\ln(P_{UNL,t-1}), \ln(S\&P\ 500), \ln(1 + Tbill_{3mo}), \ln(P_{Nat\ Gas})\big) +$

$GARCH\ (1,2)$

(vii) $\ln(\hat{P}_{UNG}) = f\big(\ln(P_{UNG,t-1}), \ln(S\&P\ 500), \ln(1 + Tbill_{3mo}), \ln(P_{Nat\ Gas})\big) +$

$GARCH\ (1,2)$

(viii) $\ln(\hat{P}_{DBO}) = f\big(\ln(P_{DBO,t-1}), \ln(S\&P\ 500), \ln(P_{Crude\ Oil})\big) + GARCH\ (1,1)$

(ix) $\ln(\hat{P}_{DBE}) = f\big(\ln(P_{DBE,t-1}), \ln(1 + Tbill_{3mo}), \ln(P_{Crude\ Oil})\big) + GARCH\ (1,2)$

(x) $\ln(\hat{P}_{USO}) = f\big(\ln(P_{USO,t-1}), \ln(S\&P\ 500), \ln(P_{Crude\ Oil}), \ln(P_{Nat\ Gas})\big) +$

$GARCH\ (1,1)$

(xi) $\ln(\hat{P}_{UHN}) = f\big(\ln(P_{UHN,t-1}), \ln(1 + Tbill_{3mo}), \ln(P_{Crude\ Oil}), \ln(P_{Nat\ Gas})\big) +$

$GARCH\ (1,1)$

(xii) $\ln(\hat{P}_{USL}) = f\big(\ln(P_{USL,t-1}), \ln(GDP), \ln(P_{Crude\ Oil})\big) + GARCH\ (1,1)$

(xiii) $\ln(\hat{P}_{UGA}) = f\big(\ln(P_{UGA,t-1}), \ln(S\&P\ 500), \ln(1 + Tbill_{3mo}), \ln(P_{Crude\ Oil}),$

$\ln(P_{Nat\ Gas})\big) + GARCH\ (1,1)$

The Bayesian Information Criterion (BIC) was used to select the most optimal estimator. Once the IV equations were chosen, a Durbin-Wu-Hausman test was used to determine the consistency of the estimators. In four cases, the test reported IVs that were not statistically significant. More efficient IV could not be estimated and therefore the estimated IVs were used. Although weak, the IVs still performed better in reducing the endogeneity issues compared to using the original values.

The ETFs were categorized into three groups: oil, natural gas and general energy ETFs. Categories were chosen based on the similarities of the holdings in their respective portfolios and the associated correlation coefficients of the forecasted prices. Correlation coefficients of the

ETF price are available in table 3. A list of the grouped ETFs and their description is given in

Table 2. Price and volume indexes were calculated for each category using a weighted average

based on the market share of the individual ETFs. Summary statistics for the indexes is given in

Table 4. The market share is calculated as follows:

$$w_i = \frac{P_i Q_i}{M} \qquad \text{where } M = \sum_i P_i Q_i$$

Table 2. ETF Descriptions and Groups

Natural Gas		
Ticker Symbol	**Name**	**Description**
GAZ	iPath Dow Jones-UBS Natural Gas Subindex Total Return ETN	Relates to a single commodity, natural gas (currently the Henry Hub Natural Gas futures contract traded on the NYMEX).
UNL	United States 12 Month Natural Gas Fund	The underlying assets of the fund consist of natural gas futures contracts.
JJE	iPath Exchange Traded Notes Dow Jones - AIG Energy Total Return Sub-Index ETN Series A	Composed of four energy-related commodities contracts (crude oil, heating oil, natural gas and unleaded gasoline) traded on U.S. exchanges.
UNG	United States Natural Gas Fund, LP	The underlying assets of the fund consist of natural gas futures contracts.
Oil Category		
Ticker Symbol	**Name**	**Description**
OLO	PowerShares DB Crude Oil Long ETN	Reflects the performance of certain crude oil futures contracts plus the returns from investing in 3 month United States Treasury Bills.
OIL	iPath Exchange Traded Notes S&P GSCI Crude Oil Total Return Index Medium-Term Notes Series A	Reflects the returns that are available through an unleveraged investment in the West Texas Intermediate (WTI) crude oil futures contract plus earning from the Treasury Bill rate of interest on funds invested in such underlying contracts.
DBO	PowerShares DB Oil Fund	A rules-based index composed of futures contracts on Light Sweet Crude Oil (WTI) and is intended to reflect the performance of crude oil.
USO	United States Oil Fund, LP	Tracks changes in the price of light, sweet crude oil, as measured by the changes in price of the futures contract on light, sweet crude oil traded on the NYME.
USL	United States 12 Month Oil Fund, LP	Reflects the changes in percentage terms of the price of light, sweet crude oil, as measured by the changes in the average of the prices of 12 futures contracts on crude oil traded on the NYME.
Energy Category		
Ticker Symbol	**Name**	**Description**
RJN	ELEMENTS Exchange Traded Notes Rogers International Commodity Index - Energy Total Return	Represents the value of a basket of 6 energy commodity futures contracts and is a sub-index of the Rogers International Commodity Index.
DBE	PowerShares DB Energy Fund	Composed of futures contracts on some of the most heavily traded energy commodities in the world: Light Sweet Crude Oil (WTI); Heating Oil; Brent Crude Oil; RBOB Gasoline; and Natural Gas.
UHN	United States Diesel-Heating Oil Fund	Reflects the changes in percentage terms of the price of heating oil as measured by the futures contract on heating oil traded on the NYME that is the near month contract to expire.
UGA	United States Gasoline Fund, LP	Reflects the changes in percentage terms of the price of gasoline as measured by the futures contract on unleaded gasoline traded on the NYME.

Table 3. Correlation Coefficients of the ETF groups

Nat Gas	GAZ	UNL	UNG	JJE
GAZ	1.00			
UNL	0.96	1.00		
UNG	0.94	0.99	1.00	
JJE	0.86	0.87	0.85	1.00

Energy	UGA	RJN	UHN	DBE
UGA	1.00			
RJN	0.73	1.00		
UHN	0.81	0.86	1.00	
DBE	0.74	0.98	0.92	1.00

Oil	OLO	DBO	USO	USL	OIL
OLO	1.00				
DBO	0.99	1.00			
USO	0.95	0.95	1.00		
USL	0.98	0.99	0.92	1.00	
OIL	0.9503	0.9484	0.9995	0.9171	1

Table 4. Summary Statistic of Indices

Variable	N	Mean	Std	Dev	Sum	Minimum
ln(Q- Oil)	1211	15.553	0.668	18835	13.514	17.679
ln(Q- Natural Gas)	1211	15.366	0.708	18609	13.440	17.798
ln(Q-Energy Group)	1211	10.885	0.938	13182	8.569	15.070
ln(P- Oil)	1210	3.542	0.103	4285	2.970	3.787
ln(P- Natural Gas)	1210	3.316	0.418	4012	2.644	4.231
ln(P- Energy Group)	1210	3.463	0.264	4190	2.158	3.995
ln(SP500)	1211	7.286	0.190	8823	6.930	7.645

Three equation incomplete demand system was entered using seemingly unrelated regression procedure (Zellner 1962). The outcome of this estimation showed somewhat erroneous estimates. As a result, the energy category was dropped from the model due to volatile price data. The non-stable prices are believed to be due to the inclusion of price data from three different energy groups- oil, natural gas and heating oil. A graph of the data can be seen in Graph 2.

Graph 2. Price Index of Energy Group ETFs
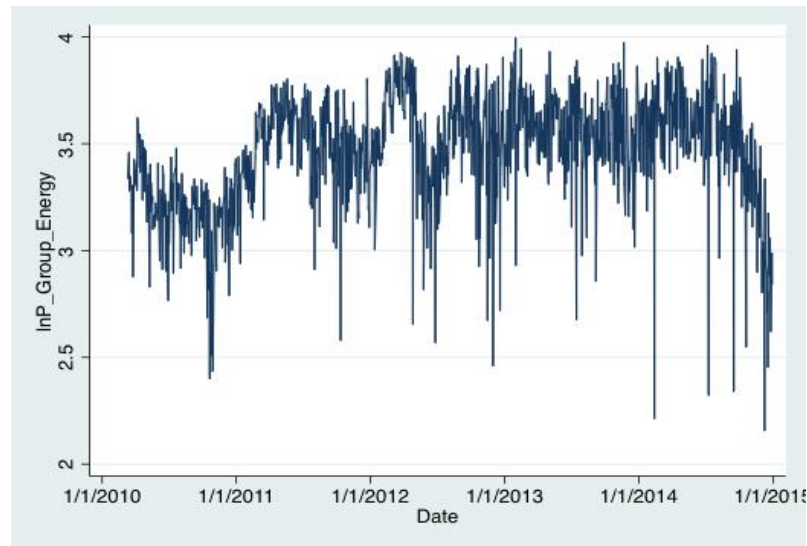
In 2012, a shale gas boom occurred causing a significant decline in the price of natural gas where prices have since remained. Due to potential structural breaks in natural gas the estimation process was broken into three partitions in time- March 2010- February 2011, March 2011- December 2011 and January 2012- December 2014. The graph of the natural gas price data and breaks is shown in Graph 3. The Bai Perron Structural Break test was used to designate the breaks in the data. The dates of the breaks determined by the test were consistent with the historical occurrence of the new technology. Price of oil over the five year time period remained relatively stable.

Graph 3. Structural Break Points in Natural Gas Data

The natural logarithm of each of the three prices indices; oil, natural gas and energy; and the natural logarithm of S&P 500 were regressed on the natural logarithm of the volume index. The S&P 500 is included as a proxy for expenditures. Not only the double logarithm system provides the best fit but also has the advantage of producing the elasticities directly from the estimated coefficients. A two-equation SUR model was estimated for the two equations. ARMA functions were included to address issues of autocorrelation of the error term and dependent variable. Using a graph of the residuals as well as the statistical significance of the lags, the appropriate number of lags was determined for each equation. Graphs of the residuals are available in the appendix.

Results

Results for the four periods and the entire five-year period will be discussed in this section. Similar to Villar and Joutz (2006), we found non-significant relationships for the price of natural gas affecting oil. Not until the third and fourth period (December 2013 to December

2014) did this relationship become significant. The possible cause of this shift in significance could be due to the overall increase of production and demand of natural gas. Villar and Joutz(2006) reason that because natural gas is a more regionally priced system and oil, on the other hand, is priced globally, we would not see natural gas affecting the price of oil. Today, natural gas is still largely only shipped within countries and therefore we do not see global pricing like we do in oil. Yet as the production and demand for natural gas becomes greater, we expect to see more of a bidirectional relationship. The two goods also show signs of being substitutes in demand. As prices of oil increase many people will simply substitute natural gas for their production processes. As a result the price of natural gas becomes more of a factor when oil is priced.

Substitutability between the two goods is consistent with our own expectations and those of previous work in studying the relationship between the two goods. Villar and Joutz's (2006) also find positive relationships between changes in the price of oil and the demand for natural gas. In all period except period four (December 2013 to December 2914) the model showed positive significant coefficients for oil prices and natural gas volume. In period three, the cross price elasticity of natural gas with respect to oil was negative but not statistically significant.

In periods one through three, the own price elasticities (OPE) for oil were not statistically significant. In the fourth period, the OPE coefficient was -1.959 indicating oil is a relatively elastic good. Similar to previous research done, this would make sense as the production and use of natural gas has grown significantly, we see many people moving between the use of oil or natural gas easily. Therefore, substituting natural gas for oil when prices rise is not unlikely or relatively difficult.

Returns for the S&P 500 contract was used as a proxy for wealth level of individuals. In seven of the eight regressions, negative income elasticities were generated. Considering the use and demand for energy products such as oil and natural gas, this outcome does not seem plausible. Three of these coefficients were not statistically significant.

In general R^2 and adjusted R^2 increase as time progresses meaning the relationships become stronger. One reason for this may be related to the relative increase in size of the natural gas that has occurred. As we have indicated before, interaction between the two sectors has possibly become greater and therefore their interactions have likely become more apparent.

Conclusion

This paper not only assesses the relationship between price and volume of certain stocks but also looks more closely at the relationship between ETFs under the energy sector category. ETFs allow us to assess relationships outside simply looking at companies and provides insight into interaction across different energy production types.

Substantial consideration was given to properly modeling the structure of the industry across time and across energy groups. Following the correction of error issues and adjusting the model for structural changes, the outcomes still show some illogical outcomes for the elasticity calculations. Considering the outcomes of the model, the largest contributor likely lies in an endogeneity problem. This issue is apparent in the directed graphs shown previously.
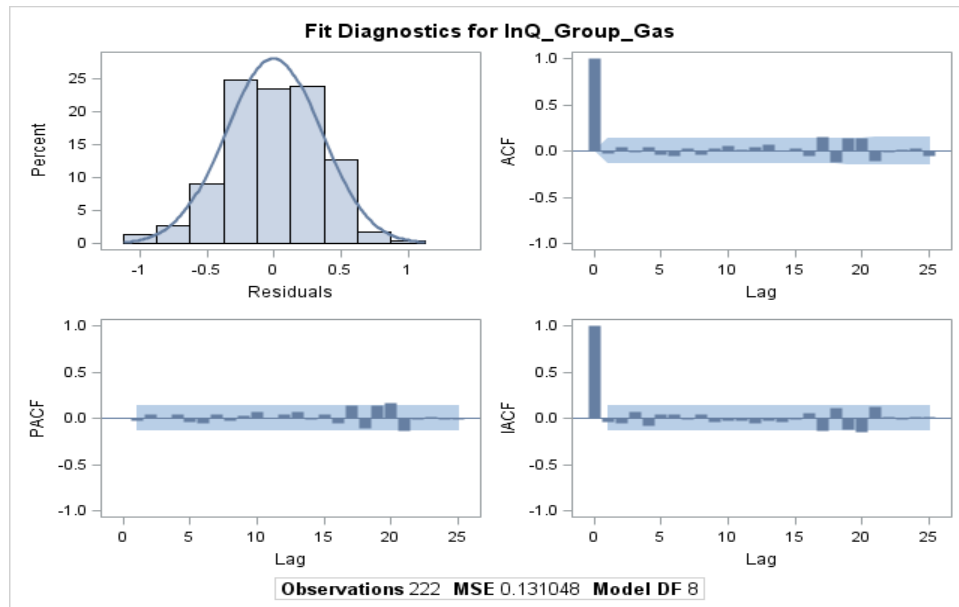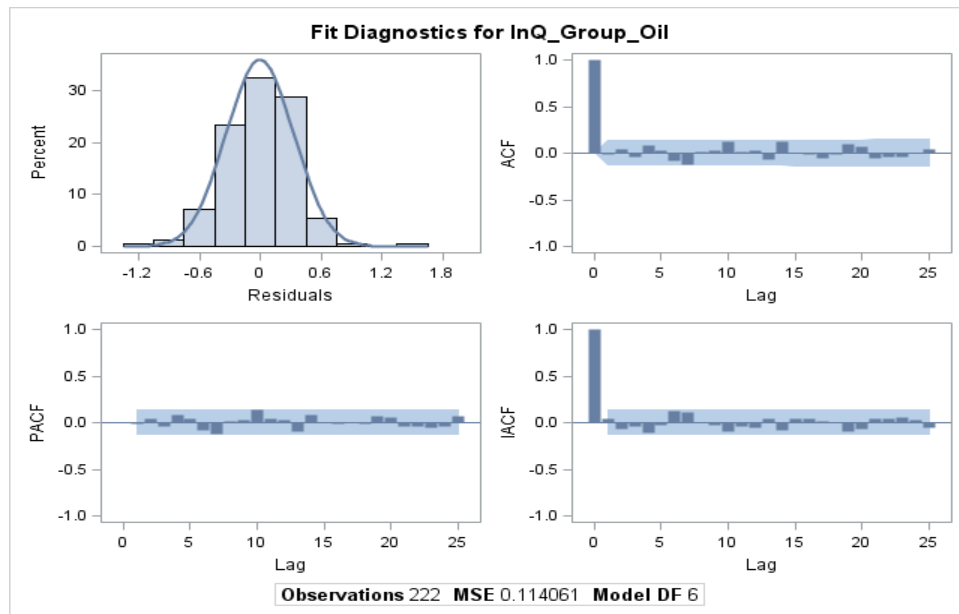
Future research will largely focus on this problem and assess the potential of omitted variables not captured here. Additional problems may be attributed to the inability to capture the true relationship between ETFs possibly due to a simultaneity problem. Despite these issues, the

use of ETFs as proxies for their specific industries or a commodity shows promise for expansion

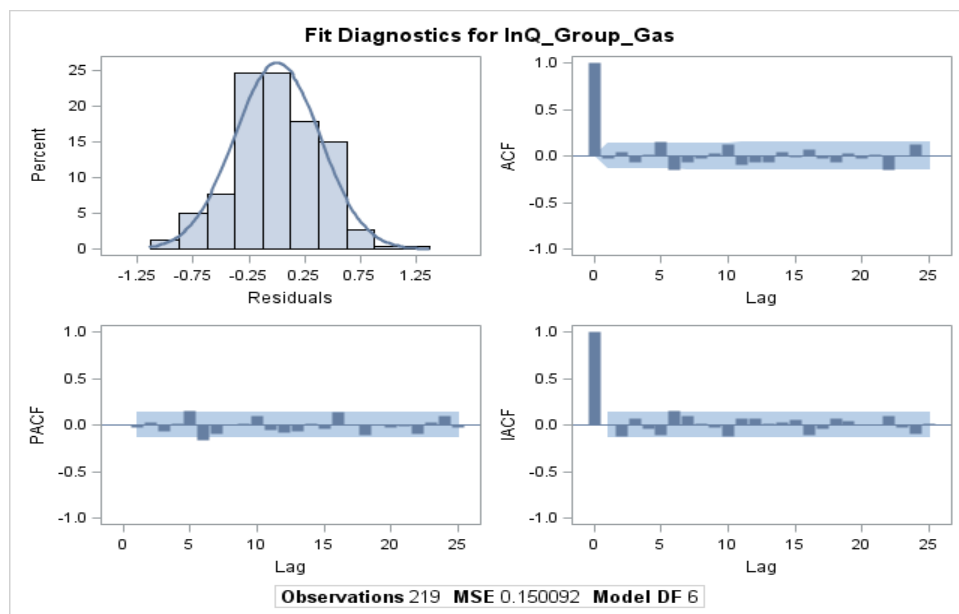to many topics in understanding relationships across industries and across commodity types.

Appendix

Graph of Residuals for the Oil and Natural Gas Models:

First Period

Fit Diagnostics for lnQ_Group_Oil

Fit Diagnostics for lnQ_Group_Gas

Second Period

Fit Diagnostics for lnQ_Group_Oil

Observations 219   MSE 0.128215   Model DF 7

Fit Diagnostics for lnQ_Group_Gas

Observations 219   MSE 0.150092   Model DF 6

Third Period

Fit Diagnostics for lnQ_Group_Oil

Observations 499   MSE 0.175168   Model DF 6

Fit Diagnostics for lnQ_Group_Gas

Observations 499   MSE 0.198405   Model DF 7

Fourth Period

Fit Diagnostics for lnQ_Group_Oil

Observations 267   MSE 0.213652   Model DF 6

Fit Diagnostics for lnQ_Group_Gas

Observations 267   MSE 0.254315   Model DF 7

Whole Data Set Regression Results

Parameter	Estimate	Approx std Err	t value	P
Constant_Q_Oil	55.130	6.735	8.190	<.0001
Oil_P	-0.687	0.314	-2.190	0.0289
Nat Gas_P	0.945	0.451	2.090	0.0365
S&P 500	-5.185	0.910	-5.700	<.0001
Oil_Q_LagAR1	1.249	0.037	33.930	<.0001
Oil_Q_LagAR2	-0.249	0.037	-6.760	<.0001
Oil_QM1	0.851	0.021	41.330	<.0001
Constant_Q_Nat Gas	2.835	4.974	0.570	0.5688
Oil_P	0.607	0.300	2.020	0.0434
Nat Gas_P	-0.275	0.375	-0.730	0.4638
S&P 500	1.594	0.607	2.630	0.0087
Oil_Q_LagAR1	0.983	0.050	19.800	<.0001
Oil_Q_LagAR5	0.183	0.036	5.090	<.0001
Oil_Q_LagAR6	-0.173	0.039	-4.460	<.0001
Nat Gas_Q_M1	0.821	0.060	13.750	<.0001

First Period

Parameter	Estimate	Approx std Err	t value	P
Constant_Q_Oil	42.022	10.074	4.170	<.0001
Oil_P	-0.244	0.587	-0.420	0.6784
Nat Gas_P	1.716	1.003	1.710	0.0886
S&P 500	-4.438	1.474	-3.010	0.0029
Oil_Q_LagAR1	0.917	0.042	22.120	<.0001
Oil_QM1	0.590	0.082	7.180	<.0001
Constant_Q_Nat Gas	-0.895	7.646	-0.120	0.9069
Oil_P	0.705	0.426	1.660	0.0994
Nat Gas_P	0.173	0.822	0.210	0.8338
S&P 500	1.743	1.149	1.520	0.1307
Oil_Q_LagAR1	0.641	0.244	2.630	0.0092
Oil_Q_LagAR5	0.330	0.068	4.820	<.0001
Oil_Q_LagAR6	-0.242	0.101	-2.400	0.0173
Nat Gas_Q_M1	0.474	0.274	1.730	0.0856

Second Period Regression Results

Parameter	Estimate	Approx std Err	t value	P
Constant_Q_Oil	62.689	9.335	6.720	<.0001
Oil_P	-1.497	0.843	-1.780	0.0773
Nat Gas_P	-0.015	0.927	-0.020	0.987
S&P 500	-5.716	1.387	-4.120	<.0001
Oil_Q_LagAR1	1.132	0.110	10.250	<.0001
Oil_Q_LagAR2	-0.146	0.103	-1.420	0.1578
Oil_QM1	0.744	0.082	9.130	<.0001
Constant_Q_Nat Gas	9.222	6.382	1.440	0.1499
Oil_P	0.725	0.346	2.100	0.0373
Nat Gas_P	1.546	0.638	2.430	0.0161
S&P 500	-0.374	1.165	-0.320	0.7484
Oil_Q_LagAR1	0.712	0.204	3.490	0.0006
Nat Gas_Q_M1	0.561	0.235	2.390	0.0177

Third Period Regression Results

Parameter	Estimate	Approx std Err	t value	P
Constant_Q_Oil	19.436	3.295	5.900	<.0001
Oil_P	0.368	0.287	1.280	0.2007
Nat Gas_P	1.769	0.443	3.990	<.0001
S&P 500	-1.553	0.349	-4.450	<.0001
Oil_Q_LagAR1	0.393	0.045	8.700	<.0001
Oil_Q_LagAR2	0.069	0.045	1.540	0.1241
Constant_Q_Nat Gas	32.283	5.665	5.700	<.0001
Oil_P	-0.965	0.407	-2.370	0.018
Nat Gas_P	-0.564	0.665	-0.850	0.3968
S&P 500	-1.619	0.626	-2.580	0.01
Oil_Q_LagAR1	0.227	0.043	5.330	<.0001
Oil_Q_LagAR2	0.208	0.043	4.830	<.0001
Oil_Q_LagAR5	0.268	0.042	6.370	<.0001

Fourth Period Regression Results

Parameter	Estimate	Approx std Err	t value	P
Constant_Q_Oil	38.770	11.116	3.490	0.0006
Oil_P	-1.959	0.583	-3.360	0.0009
Nat Gas_P	-2.671	0.481	-5.560	<.0001
S&P 500	-1.098	1.345	-0.820	0.4149
Oil_Q_LagAR1	0.356	0.062	5.720	<.0001
Oil_Q_LagAR2	0.068	0.062	1.090	0.2747
Constant_Q_Nat Gas	31.558	17.169	1.840	0.0672
Oil_P	1.207	0.797	1.510	0.1312
Nat Gas_P	-1.535	0.760	-2.020	0.0445
S&P 500	-1.869	2.107	-0.890	0.3759
Oil_Q_LagAR1	0.168	0.056	3.010	0.0029
Oil_Q_LagAR2	0.271	0.058	4.700	<.0001
Oil_Q_LagAR5	0.296	0.057	5.190	<.0001

References

Bachmeier, Lance J. and Griffin, James M. 2006. Testing for Market Integration Crude Oil, Coal and Natural Gas. *The Energy Journal* 27, No.2, 55-71

Bai, Jushan and Perron, Pierre. 1998. Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica.* Vol 66, No. 1, 47-78

Crouch, R. L.1970. A Nonlinear Test of the Random-Walk Hypothesis. *American Economic Review* 60, 199-202.

Crouch, R. L.1970. The Volume of Transactions and Price Changes on The New York Stock Exchange. *Financial Analysts Journal*. 104-109

DeFusco, Richard A., Stoyu I. Ivanov, and Gordon V. Karels. 2011. The exchange traded funds' pricing deviation: analysis and forecasts. *Journal of Economics and Finance* 35, No. 2, 181-197.

Epps, Thomas W. 1976. Security Price Changes and Transaction Volumes: Theory and Evidence. *Econometrica: Journal f the Econometric Society.* Vol 65, No 4, 586-597

Godfrey, M. D., C. W. J. Granger, and O. Morgenstern. 1964. The Random Walk Hypothesis of Stock Market Behavior. *Kyklos*, Vol 17, Issue 1, 1-30

Hiemstra, Craig and Jones, Jonathan D. 1994. Testing for Linear and Nonlinear Granger Casusality in the Stock Price-Volume Relation. *The Journal of Finance.* Vol 49, 1693-1664

Jain, Prem C. and Joh, Gun-Ho. 1988. The Dependence between Hourly Prices and Trading Volume. *Journal of Financial and Quantitative Analysis.* Vol 23, No 3, 269-283.

Karpoff, Jonathan M. 1987. The Relation between Price Changes and Trading Volume: A Survey. *Journal of Financial and Quantitative Analysis.* Vol 22, no. 1

Smirlock, M., and L. Starks. 1984. A Transactions Approach to Testing Information Arrival

      Models. Unpublished Manuscript, Washington University

Smirlock, M., and L. Starks. 1985. A further Examination of Stock Price Changes and

      Transaction Volume. *The Journal of Financial Research*. Vol 8, No 3

Smirlock, M. and Starks, Laura. 1988. An Empirical Analysis of the Stock Price-Volume

      Relationship. *Journal of Banking and Finance.* 12, 31-41

Villar, Jose A., and Joutz, Frederick L. 2006. The relationship between crude oil and natural gas

      prices. *Energy Information Administration, Office of Oil and Gas*, 1-43.

Wood, R. A., T. H. Mclnish, and J. K. Ord. 1985. An Investigation of Transactions Data for

      NYSE Stocks. *Journal of Finance*, 60, 723-739.