2013

Crude oil prices: speculation versus fundamentals

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

CRUDE OIL PRICES: SPECULATION VERSUS FUNDAMENTALS

by

MAREK KRZYSZTOF KOLODZIEJ

B.Sc., University of Rhode Island, 2004
M.A., University of Illinois, 2006

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2013

Approved by

[redacted]

First Reader [redacted]
Robert K. Kaufmann, Ph.D.
Professor of Geography

[redacted]

Second Reader [redacted]
Ian Sue Wing, Ph.D.
Associate Professor of Geography

[redacted]

Third Reader [redacted]
Dana Bauer, Ph.D.
Assistant Professor of Geography

For my Parents.

ACKNOWLEDGEMENTS

First, I would like to thank Robert Kaufmann for his continual encouragement, econometric advice and friendship. The countless hours we spent discussing energy modeling, time series econometrics, and how to survive the stresses of a full-time job while completing the degree requirements will be fondly remembered. I will be forever indebted to him for his mentorship and wisdom.

The completion of this work would have been impossible without Ian Sue Wing's support. The two years that I spent working for Ian as a research assistant funded by a U.S. DOE grant marked a period of incredible intellectual growth for me. Ian opened my eyes to the potential of operations research and optimization-based modeling, for which I will be eternally grateful.

I would like to thank Nalin Kulatilaka for his financial theory insights, without which the econometric results found in this dissertation would have lacked a solid interpretation.

I would like to express my gratitude for Dana Bauer's and Cutler Cleveland's input into this work. I am also grateful for Dana's resource economics mentorship over the course of my Ph.D. program, as well as for her emotional support and faith in me in times of high stress.

I would like to extend my thanks to Houston Stokes at the University of Illinois, who over the course of my Master's program taught me most of the econometrics I know today, as well as to Doug Reynolds at the University of Alaska-Fairbanks, who taught me a great deal about institutional issues pertaining to the oil and gas industry, and with whom I published three peer-reviewed papers. I am also indebted to Craig Schmidt, who taught me a lot of neat modeling tricks that while not relevant to this dissertation, are very much relevant to my growth as a predictive analyst.

Navigating the BU bureaucracy would have been impossible without the continued support from Chris DeVits and Christian Cole. Chris and Christian managed to transform Kafkaesque situations into Swiss clockwork many times, and he always did it with tons of good humor.

I am deeply indebted to Swapnil Shah, Ken Kolkebeck and Badri Raghavan at FirstFuel Software for enabling me to put my skills to practical use. The ability to make a difference in the real world by deploying energy efficiency auditing technologies restored my enthusiasm for research and my motivation to complete the Ph.D.

The quintessential Zen career advice is to "choose a job you love, and you will never have to work a day in your life." I would like to thank Stephen Harrison, Prakash Teli, Siddharth Kundalkar, and the entire India Team (especially Neerav Kulshreshtha, Basavaraj Kalloli, Sameer Khare, Shreyas Pai, Amit Dixit and Ameeth Paatil) for helping me discover a passion for both software engineering and computer science.

I would like to thank my wonderful non-IT colleagues at FirstFuel Software, Tarik Borogovac, Indran Ratnathicam, John Tehrani, Domenic Armano, Dan Foley, Chris Tomasini, Cara Giudice, Janet Desaulniers and John MacPhee for making work both fun and fruitful.

Life would have been meaningless and unbearable without true friendships. I would like to thank Marek Mroziewicz, Ewa Kozinska, Jagoda Gandziarowska, Anna Senkara, Michal Jedrzejewski, Ric McIntyre, Jim Starkey, Sally Honda, Dan Craig, Mike Mann, Jamie Baldwin, Britt Briber and many others for giving my life a purpose.

This dissertation is dedicated to my parents, whose incredibly hard work, sacrifice, and unconditional love words simply cannot begin to describe. I hope that you feel as loved back as I feel loved by you.

CRUDE OIL PRICES: SPECULATION VERSUS FUNDAMENTALS

(Order No.)

MAREK KRZYSZTOF KOLODZIEJ

Boston University Graduate School of Arts and Sciences, 2013

Major Professor: Robert K. Kaufmann, Professor of Earth and Environment

ABSTRACT

Beginning in 2004, the price of crude oil fluctuates rapidly over a wide range. Large and rapid price increases have recessionary consequences and dampen long-term infrastructural investment. I investigate whether price changes are driven by market fundamentals or speculation.

With regard to market fundamentals, I revisit econometric evidence for the importance of demand shocks, as proxied by dry maritime cargo rates, on oil prices. When I eliminate transportation costs from both sides of the equation, disaggregate OPEC and non-OPEC production, and allow for more than one cointegrating relation, I find that previous specifications are inconsistent with arguments that demand shocks play an important role. Instead, results confirm the importance of OPEC supply shocks.

I investigate two channels by which speculation may affect oil prices; the direct effect of trader behavior and changes in oil from a commodity to a financial asset. With regard to trader behavior, I find evidence that trader positions are required to explain the spread between spot and futures prices of crude oil on the New York Mercantile Exchange. The inclusion of trader

positions clarifies the process of equilibrium error correction, such that there is bidirectional causality between prices and trader positions. This creates the possibility of speculative bubbles.

With regard to oil as a commodity and/or financial asset, I use a Kalman Filter model to estimate the time-varying partial correlation between returns to investments in equity and oil markets. This correlation changes from negative to positive at the onset of the 2008 financial crisis. The low interest rates used to rescue the economy depress convenience yields, which reduces the benefits of holding oil as a commodity. Instead, oil becomes a financial asset (on net) as the oil market changed from contango to backwardation.

Contradicting simple political narratives, my research suggests that both market fundamentals and speculation drive large oil prices. Chinese oil demand is not responsible for large increases in oil prices; nor are they caused by behavioral idiosyncrasies by oil traders. Finally, oil will be treated largely as a financial asset so long as interest rates are held near their all-time lows.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 2

CHAPTER 3

LIST OF FIGURES

CHAPTER 2

CHAPTER 4

LIST OF ABBREVIATIONS

CAPM	Capital Asset Pricing Model
CFTC	Commodity Futures Trading Commission
COP	ConocoPhillips
CPI	Consumer Price Index
CVAR	Cointegrated Vector Autoregression
EIA	Energy Information Administration
FOB	Free On Board
FRED	Federal Reserve Economic Data
GDP	Gross Domestic Product
NYMEX	New York Mercantile Exchange
OLS	Ordinary Least Squares
OPEC	Organization of Petroleum Exporting Countries
PADD	Petroleum Administration for Defense Districts
SBC	Schwarz Bayesian Criterion
S&P	Standard & Poor's

SPR	Strategic Petroleum Reserve
TRTH	Thomson Reuters Tick History
UNCTAD	United Nations Conference on Trade and Development
VAR	Vector Autoregression
WTI	West Texas Intermediate
XOM	ExxonMobil

1. INTRODUCTION

Beginning in 2004, the price of crude oil fluctuates rapidly over a wide range. From \$60 per barrel in 2004, the price for the West Texas Intermediate (WTI) crude oil peaks at \$148 in July 2008. Only six months later, the price drops under \$40. In January 2012, the price moves beyond \$100, and in July 2012, it drops below \$80. Large and rapid price increases may have negative macroeconomic consequences (Hamilton, 1983; Burbidge and Harrison, 1984; Mork *et al.*, 1994; Cuñado and Pérez de García, 2003; Jiménez-Rodríguez and Sánchez, 2005). Just as importantly, growing volatility itself introduces considerable uncertainty that complicates infrastructural planning and thereby dampens long-term investment and aggregate demand (Lilien, 1982; Bernanke, 1983; Hamilton, 1988; Ferderer, 1996).

Despite the considerable secular increase in both the level and volatility of the price of oil in the past decade, many fundamental questions regarding the source of these changes remain unanswered. Are market fundamentals (strong global aggregate demand) or speculation the driving force behind the secular increase in the price of oil between 2004 and 2008? Does the source of the fundamental-driven price changes (supply, aggregate demand, oil-market-specific precautionary demand) affect the macroeconomic outcomes and the timing of the dynamic adjustment to the back to equilibrium? Is speculation the cause of increased volatility? Do trader positions in the oil futures market affect oil prices, or is the speculative component associated with an optimal portfolio allocation strategy, in which money between from equity and commodity markets?

Kaufmann (2011) argues that speculation plays a significant role in the 2008/2009 price spike and collapse based on three changes: (1) a significant increase in private crude oil inventories, (2) a breakdown in the cointegrating relationship between spot and far month futures prices that is inconsistent with arbitrage opportunities, and (3) predictive failures by an econometric model of oil prices that is based on economic fundamentals. Masters (2008) finds that non-commercial traders now buy and sell oil as a financial asset, and Medlock and Jaffe (2009) note that non-commercial traders now hold about 50 percent of oil futures positions, and that this percentage grows steadily in recent times. If speculators act as fully informed participants (in the rational expectations/efficient market sense), then their role may be benign (Weiner, 2002). However, when they take an extrapolative strategy or if herding or contagion occurs, their presence in the market can amplify volatility. Cifarelli and Paladino (2010) find evidence for the latter scenario to be consistent with reality over the past several years.

However, many of these studies are merely suggestive. In order to quantify the role of speculators, I need to examine the role of trader positions on spot and futures oil prices, both in the long and the short run. One aspect of such an analysis would be to see whether trader positions Granger-cause oil prices, whether the causality goes in the opposite direction, or whether there is evidence for bidirectional causality. Studies that have been published to date that examine this issue indicate that trader positions do not Granger-cause oil prices, however they are limited by several difficulties. First, they ignore the presence of unit roots in the data, thereby allowing for a high probability of spurious regressions (Granger and Newbold, 1974). Second, even if these regressions are not spurious, using an econometric technique that does not account for cointegration and error correction (such as a vector autoregression) generates a misspecified model, per the Granger Representation Theorem (Granger and Lee, 1990). Therefore, the conclusions of these studies are likely biased. Another problem with the studies that had been done to date is omitted variable bias – a bivariate vector autoregression or a distributed lag model lacks many other explanatory variables that should be part of the *information* set. These omissions may bias the point estimates of the included variables if they are correlated with the missing variables. This biases F statistics that are used to test for Granger causality. Furthermore, models with omitted variables have correlated residuals, which again invalidates econometric inference used to determine Granger causality. Therefore, it seems that the existing literature is inconclusive, and could benefit from better econometric specifications to improve inference. In this dissertation, I examine the evidence for causality among oil prices and trader positions, and I find evidence for bidirectional causality, which is at odds with the existing literature which finds no causality from trader positions to oil prices, and therefore no evidence for the influence of speculation on the price discovery process. The separation of commercial and non-commercial positions, as well as the inclusion of oil inventories and of spot and futures prices provides a richer information set that generates cointegrating relationships, and is thereby sound from an econometric point of view. Furthermore, a cointegrated vector autoregression (CVAR) model allows me to separate the short- and long-run impacts, resulting in a more informative analysis.

Another issue that has not been explored in depth is the dynamic evolution of the potential influence of speculation on the price discovery process. The cointegrated vector autoregression finds cointegration for the entire period spanned by the dataset, however this may be misleading if there are brief (but potentially substantively significant) changes in the relationship between the variables. Rolling window correlations between U.S. stock market returns and the returns on commodities show a strong regime change in 2007, which coincides with the full automation of U.S. stock exchange trading in September 2007, and hence an increase in the influence of high-frequency and algorithmic trading. Even if one examines the E-mini S&P500 index (which was always fully electronic) as opposed to the standard S&P500 index, dramatic changes that start in 2007. For instance, the number of trades on the E-mini S&P500 increase from 8 million to 11 million between 2001 and 2004, and is *stable at about 11* million until 2006. However, the number of trades doubled to 22 million in 2007, more than doubled the following year to 49 million, and subsequently doubled again to 107 million in 2010. A parallel change occurs in oil markets, where the number of trades of WTI on NYMEX increases from 485 thousand in 2001 to 920 thousand in 2005, to 12 million in 2007 and to 42 million in 2011 (Bicchetti and Maystre, 2012, in press). Consistent with these parallel changes, Bicchetti and Maystre identify a large regime change that took place in 2007 in terms of significantly increased correlations between stock and commodity market returns, using hourly, 5-minute, 10-second and 1-second data. At such high frequencies, economic fundamentals become irrelevant because industrial production is reported monthly, GDP is reported quarterly, and even oil inventories are reported weekly. Therefore, if there exists a high correlation between stock market and commodity market (e.g. oil price) returns, one may conclude that commodities have become but one more asset in the market players' optimal portfolio allocation strategy. The U.S. Energy Information Administration reaches similar conclusions (EIA, 2012).

The issue with Bicchetti and Maystre's analysis is that correlations are likely biased due to omitted variables. If there are other factors that are correlated with one of the variables in the simple bivariate correlation, then the estimate of the correlation will be biased by the lack of the omitted variable(s). Estimating a multivariate regression gives partial correlations associated with the variables in question, given the conditioning set of the other variables in the model. If the information set is complete, these partial correlations will be unbiased. The second issue with Bicchetti and Maystre's analysis is that they look at arbitrary subperiods to determine the changing correlations. In my research, I use the Kalman filter to determine the statistically optimal state variable (regression coefficient) variances, and hence the optimal "memory" of the coefficients' histories. This will also allow me to estimate the time-varying standard errors for these coefficients, and hence the time-varying statistical significance of the variables in the model.

Besides speculation, the importance of macroeconomic fundamentals in the oil price discovery process merits further investigation. As I mentioned above, there exists a very rich literature that identifies and measures the negative macroeconomic effects of oil price increases. This literature starts in 1983 with Hamilton's seminal paper, and continues to grow through the present day (e.g. Jiménez-Rodríguez and Sánchez, 2005). However, more recent studies point to the fact that the measurement of the influence of oil prices on aggregate demand may be biased due to model misspecification, because the approach taken to date implicitly assumes that oil prices are exogenous. In reality, oil prices are determined not only by speculation and political events, but also, and hopefully by supply and demand conditions in the oil market. Treating oil price changes exclusively as supply shocks is not realistic, because both supply and demand determine the equilibrium price. Precluding the influence of demand shocks involves assuming a completely stationary demand curve. Ignoring the influence of aggregate demand on oil prices and attributing oil price changes to pure supply shocks does not seem sensible. In addition to oil supply and demand, there exists an unobservable precautionary demand component, which may influence oil prices in a different way and with a different dynamic. Several papers describe this topic (Barsky and Kilian, 2002; Hamilton, 2003; Barsky and Kilian, 2004; Rotemberg 2007). It was not until Kilian's 2009 paper that this hypothesis is analyzed in a formal econometric manner. Kilian (2009) estimates a monthly model of oil production, oil prices, as well as a dry-bulk maritime shipping index (as a proxy for global aggregate demand). Kilian's findings seem to undermine the findings of the supply-centered literature of the past three decades. Kilian concludes that supply shocks have a negative macroeconomic effect but that it is small and transitory, while the effects of aggregate demand are positive despite the higher oil prices (since they stem from high aggregate demand), and they are long lasting. Furthermore, increases in precautionary demand have a negative macroeconomic effect that is larger and more long-lasting than that generated by supply shocks.

Kilian's research sheds light on the importance of separating supply and demand factors when analyzing the sources and consequences of oil price shocks, however his model suffers from many shortcomings. First, the variables have unit roots, so estimating a vector autoregression (VAR) may result in spurious regressions and divergent impulse response functions at worst, or in very long lags and serious residual autocorrelation and biased inference at best. I will use a cointegrated vector autoregressive (CVAR) model to address these issues. Kilian also uses an aggregate measure of oil production, but different production-setting strategies make the relative contributions of OPEC and non-OPEC to the overall supply stream important (Kaufmann *et al.*, 2004; Smith, 2005). In particular, I find OPEC oil production to be exogenous to the model (though influencing the other variables), with non-OPEC production adjusting to changes in prices and other factors. This is consistent with the widely held view that non-OPEC producers try to maximize the net present value of their resources, while OPEC producers have political and other motives driving production changes. In a static model, the separation of the oil supply streams would be less important, however in a dynamic model such as a CVAR, this separation allows me to identify multiple cointegrating relations, short-run adjustment parameters, weak exogeneity, etc. Also, Kilian uses a potentially poor measure of oil prices – his measure, the refiner's acquisition cost, includes tariffs, wharfage, demurrage, and other costs that cause this price to diverge from the wellhead price. Moreover, refiner's acquisition cost includes the cost of maritime transportation, thereby resulting in the inclusion of transportation costs on both sides of some of Kilian's VAR equations. By using a the free-on-board (FOB) oil price that does not include these distortions, I will attempt to identify a more accurate relationship between the variables. Kilian also includes all variables except for oil production in levels in his VAR, with oil production specified in first differences. He does not explain why this mix of levels and first differences was chosen, but it would seem reasonable for all the variables to be either consistently in levels or first differences. Given that oil production has a unit root just like the other variables, including it in first differences precludes cointegration and creates a possibility of a spurious regression. I will estimate my model consistently in levels in the cointegrating vectors, and in first differences in the short-run parameters, which is how the CVAR methodology is supposed to work (Juselius, 2006). Finally, given the use of a VAR with variables containing unit roots, Kilian predictably uses long lags due to strong residual autocorrelation (which may even be indicative of a spurious regression). He chose the lag length of twenty four, i.e. of exactly two years' worth of lags. This seems to be a strange coincidence, given that lag length should be chosen based on a trade-off between fit, loss of degrees of freedom, and the statistical significance of the lagged variables. I will use a lag selection scheme such as the Schwartz Bayesian Criterion and Hannan-Quinn Criterion to choose an optimal lag length. All these considerations and modifications to Kilian's model will allow me to determine how robust his model is to an alternative specification that is consistent with econometric theory. As described in chapter 2, the improved specification supports the existing supply-shock literature more so than Kilian's own findings.

What is the practical importance of this research? Specifically, are there potential policy implications of these findings? In order to answer this question, one has to distinguish between oil price shocks driven by economic fundamentals from those caused by speculation. In the case of macroeconomic fundamentals, one needs to make a further distinction between supply and demand shocks. If high prices are driven by positive demand shocks, there would not be a need for intervention because any recessionary effect of high oil prices would bring the economy back towards a secular trend since it was already performing above the trend (booming). Negative demand shocks cause prices to fall, so there would not be a negative macroeconomic consequence of a negative oil demand shock. On the supply side, a negative supply shock, e.g. caused by military conflicts or OPEC production quota changes, could have unfavorable economic consequences, and there may be policy implications of the finding that such shocks matter. For instance, frequent and large supply shocks might call for higher inventory levels in consuming nations, beyond what profit-maximizing oil distributors would be willing to store. A case in point is the Strategic Petroleum Reserve (SPR) maintained by the U.S. Department of Energy. Focusing on speculation, one needs to distinguish benign from harmful speculation, and oil-market-specific speculation from cross-market capital flows due to portfolio optimization. As I mentioned above, oil-market-specific speculation consistent with the rational expectations hypothesis is harmless because prices do not consistently diverge from the valuation dictated by market fundamentals in such a situation. This kind of harmless speculation is actually beneficial, because it provides liquidity needed by the hedgers, who use the futures markets to guarantee prices at some future point in time. On the other hand, speculation that causes divergence of the price of oil away from market fundamentals, e.g. due to market bubbles, would be considered harmful and reducing market efficiency, and may call for government intervention if the expected costs of such intervention are lower than the expected benefits. Finally, speculative activity associated with cross-market capital flows, e.g. between the equity and commodity markets, can be considered harmless or even beneficial if it leads to risk reduction due to portfolio optimization consistent with the capital asset pricing model (CAPM). If cross-market capital flows cause price divergence away from the balanced portfolio predictions of the CAPM, then this kind of speculative activity can be seen as inefficient and calling for intervention provided that the cost-benefit analysis indicates a net benefit of the proposed intervention. All in all, the identification of the source of the large increase in the level and volatility of the price of oil has clear policy implications.

2. OIL DEMAND SHOCKS RECONSIDERED: A COINTEGRATED VECTOR AUTOREGRESSION

2.1 Introduction

The hypothesis that large, abrupt increases in real oil prices contribute to recessions can be traced back to a seminal paper by Hamilton (1983). Since then, many studies corroborate the adverse macroeconomic consequences of oil price shocks (e.g. Burbidge and Harrison, 1984; Mork *et al.*, 1994; Cuñado and Pérez de García, 2003; Jiménez-Rodríguez and Sánchez, 2005). Much of this literature implicitly assumes that oil price shocks originate from exogenous supply-side disruptions. More recent studies explore the possibility that these shocks stem not only from shifts in supply, but also shifts in demand. Although this strand of the literature is fairly new, there are theoretical arguments and descriptive historical accounts that support this view (e.g. Barsky and Kilian, 2002; Hamilton, 2003; Barsky and Kilian, 2004; Rotemberg, 2007).

To support arguments for the importance of demand-side oil price shocks, Kilian (2009) estimates a vector autoregressive (VAR) model that specifies the first difference of global oil production, real oil prices, and an index for dry bulk maritime freight costs (used to proxy macroeconomic activity/non-precautionary oil demand). A structural decomposition of the VAR disaggregates oil price shocks into oil supply, global aggregate demand, and oil market-specific, precautionary demand shocks, which are related to changing expectations about future supply. Impulse response functions indicate that the source of an oil price shock has important implications for both its macroeconomic impact and the dynamic response. During the 1973-2007 sample period, aggregate and oil market-specific demand shocks cause persistent and large oil price increases, while oil supply shocks generate small and transitory price increases. Moreover, while oil market-specific demand and supply shocks have a negative impact on economic activity, the evidence for aggregate demand shocks is mixed, and depends on the time horizon. Together, Kilian's findings imply that oil price shocks, which result from robust global aggregate demand, are not a cause for concern because the high price of oil dampens the global economy when it is already performing well above the secular trend, while oil supply shocks are not an issue due to their transitory and minor impact. Therefore, the real concern is oil market-specific demand shocks, which are associated with changes in expectations about future oil supply.

This paper recognizes the advantages of disaggregating oil shocks into supply, oil market-specific demand, and aggregate demand shocks, but evaluates the degree to which Kilian's findings about this disaggregation are robust. To do so, I estimate cointegrating vector autoregression (CVAR) models that explore four aspects of Kilian's specification. First, Kilian (2009) uses dry bulk maritime freight costs as a proxy for global economic activity and/or non-precautionary oil demand. There is little evidence for such a wide-ranging interpretation of dry bulk maritime freight costs. For example, Kaufmann (2011) finds that there is no statistically measurable relation between dry bulk maritime freight costs and oil consumption during the 1968-2008 the sample period. The lack of a relation begs the question, do dry cargo bulk maritime freight rates represent anything beyond transportation costs?

The interpretation of dry bulk maritime freight costs is critical because the measure of oil prices used by Kilian (2009), the refiner acquisition cost of imported crude oil, includes transportation costs (and wharfage, tariffs, and other charges). Put simply, measuring oil prices with the refiner acquisition cost puts transportation costs on both sides of the regression equations and therefore creates *a priori* a positive relationship between the refiner acquisition cost and dry bulk maritime freight costs. The second aspect of my analysis investigates this effect (and effects of including the costs of wharfage, tariffs, and other charges that may blur the connection between oil prices and economic activity), by estimating a model that measures oil prices with a free-on-board (FOB) price, which excludes transportation costs, as well as wharfage, demurrage, and tariffs.

Third, Kilian (2009) specifies oil supply as the first difference of a single aggregate of OPEC and non-OPEC production. But this aggregation may be inappropriate because OPEC and nonOPEC nations use different criteria to set production levels. Specifically, OPEC nations use strategic behavior (in part) to set production levels (e.g. Smith, 2005) while oil output in most non-OPEC countries is based primarily on economic considerations (e.g. Kaufmann *et al.*, 2004). Given these fundamental differences, aggregating global oil supply into OPEC and non-OPEC sources may obfuscate the importance of supply-side shocks. To investigate this effect, I estimate a model that specifies OPEC and non-OPEC production separately.

Fourth, Kilian's VAR specification includes a log of his detrended shipping index, a log of refiner acquisition cost deflated using the U.S. consumer price index (CPI), and the *first difference* of the log of global oil production. Kilian (2009) does not explain why oil production is specified using its (logged) *first difference*, as opposed to its *level*. Because univariate and multivariate tests suggest that the crude oil price and shipping cost time series are I(1) (see below) and the first difference of oil production is I(0), specifying the first difference of global oil production eliminates any long-run relation between oil production and oil prices and/or the shipping index *a priori*. As such, specifying the first difference of production reduces the possible impacts of supply side shocks. To investigate this effect, I estimate a model that specifies the level of OPEC and non-OPEC production (separately).

To evaluate the effect of these four issues on the results generated by Kilian (2009) I estimate a four variable system using a cointegrated vector autoregression (CVAR) model. I use this methodology because it is designed for non-stationary time series, it explicitly represents cointegration and error correction, it can identify more than one cointergating relation, and it can be used to explicitly test whether time series are endogenous/exogenous. As such, it represents a viable alternative to the vector autoregression used by Kilian (2009).

The results of the four variable CVAR model are largely inconsistent with those generated by Kilan (2009). In summary, the results suggest that OPEC and non-OPEC nations use different criteria to set production levels and that changes in OPEC production have significant price effects that are commonly associated with price shocks. There is a relation between oil prices and dry bulk maritime freight costs, but the loadings of the cointegrating relations and simulations of the system's dynamics suggest that this relation simply represent the effect of higher oil prices on transportation costs. Sensitivity analyses suggest that the differences in the results reported here and by Kilian (2009) are caused by including transportation costs in the measure of oil prices, aggregating OPEC and non-OPEC production, and using a very long lag length to estimate the VAR.

These results and the methods used to obtain them are described in five sections. In section 2, I describe the data and methodology used to estimate a CVAR model of the long- and short run-relations among oil prices, the index for dry bulk maritime shipping costs, OPEC oil production, and non-OPEC oil production. The classification of the variables as endogenous or exogenous, the number of cointegrating relations, and their make-up is described in section 3. Section 4 describes the cointegrating relations and interprets their meaning based on simulations that explore the dynamics of perturbations to oil prices, oil production, and dry bulk maritime freight costs. Section 5 concludes by arguing that the empirical evidence currently available is consistent with the large literature that describes the importance of supply shocks on oil prices.

2.2. Methodology

2.2.1 Data

Kilian's index for dry cargo bulk freight rates (*SHIP*) is obtained from the data set submitted by Kilian along with his paper to the *American Economic Review*[1]. (Logged) values for global monthly oil production by OPEC (*OPEC)* and non-OPEC nations (*NOPEC)* are obtained from the U.S. Energy Information Administration (EIA). The U.S. monthly, seasonally-adjusted consumer price index for all urban consumers and all items, published by the U.S. Bureau of Labor Statistics, is obtained from the Federal Reserve Economic Data database[2]. Monthly observations for the free-on-board (FOB) price of crude oil imported into the United States are obtained from the US EIA. These values are deflated using the CPI, logged, and used as the measure for real oil prices (*PRICE*). All data series are available for the period over which Kilian estimates his original model, i.e. January 1973 through November 2007. The CPI series is used as-is, i.e. with the index centered on 100 for the years 1982-1984[3]. Each time series is normalized by subtracting its mean and dividing by its standard deviation to ease the interpretation of regression coefficients and to improve the numeric precision of floating-point calculations.

The time series properties of the four time series is determined using the Augmented Dickey Fuller statistic (Dickey and Fuller, 1979; Said and Dickey, 1984). The lag length is chosen using

[1] Data found in the ZIP file located at http://www.aeaweb.org/aer/data/june09/20070211_data.zip. See Kilian (2009) for a detailed discussion of the contruction of hi index.
[2] http:/research.stlouisfed.org/fred2/series/CPIAUCSL?cid=9.
[3] http:/stats.bls.gov/cpi/cpifaq.htm

the Akaike Information Criterion (Akaike, 1974) to evaluate up to eight lags, which is based on the usual metric $T^{1/3}$, where T represents the number of time series observations. Results indicate that the null hypothesis of a unit root is not rejected for any time series at any of the conventional significance levels (1-10%). These results are not sensitive to the lag length used. These results also are consistent with multivariate tests of stationarity (with two cointegrating relations—see below), which strongly reject ($p < .000$) the null hypothesis of stationarity, except for non-OPEC production ($p < 0.13$). Together, these results strongly suggest that the time series are I(1).

2.2.2 Statistical Methodology

To address the presence of stochastic trends in the data (and hence cointegration and error correction), to determine which variables are exogenous and endogenous, and to identify more than one cointegrating relation (if present), I use a cointegrated vector autoregressive (CVAR) model (instead of a VAR) to analyze the relation among dry cargo bulk maritime freight rates (*SHIP*), non-OPEC production (*NOPEC*), OPEC production (*OPEC*), and the real price of crude oil (*PRICE*). The cointegrated vector autoregressive (CVAR) time-series model is discussed in detail by Juselius (2006). Unlike the traditional vector error correction model (VECM), the CVAR allows for the presence of more than one linear combination of non-stationary variables that is stationary, i.e. multiple cointegrating relations. Each cointegrating relation has its own set of variable coefficients (cointegrating vector) that generates a stationary linear combination of variables present in the particular cointegrating relation. A CVAR with two lags (1 lagged first difference per variable) can be concisely specified as follows:

$$\Delta x_t = A_0 \Delta w_t + A_i \Delta w_{t-1} + \Gamma_1 \Delta x_{t-1} + \Pi(x'_{t-1} w'_{t-1}, 1)' + \varepsilon_t \qquad (2.1)$$

in which x_t is a vector of variables whose behavior is being modeled endogenously, w_t is a vector of exogenous variables, Γ_1, Π, A_0, A_1 are matrices of regression coefficients, Δ is the first difference operator ($\Delta x_t = x_t - x_{t-1}$) and $\varepsilon_t : Niid(0, \Omega)$.

When the time series are nonstationary, the long-run matrix can be formulated as:

$$\Pi = \alpha\beta' \tag{2.2}$$

in which α is a $p \times r$ matrix of adjustment coefficients (also known as *loadings*) and β' is an $r \times p$ matrix of cointegrating vector coefficients. The term Πx_{t-1} represents the error correction mechanism (ECM). The number of cointegrating relations present is given by the rank (r) of Π.

2.3 Results

I select the lag length using two statistical criteria, the multivariate Schwarz Bayesian (SBC) and the Hannan-Quinn criteria (Dennis *et al.*, 2005; Juselius, 2006). Starting with a maximum lag length of eight ($T^{1/3}$), both criteria indicate that two lags is optimal. A value of two lags implies that there will be one lagged first difference per variable per equation (in addition to the error correction mechanism) because the CVAR model is based on first differences. The optimal lag length of two stands in sharp contrast with the twenty-four lags specified in the VAR estimated by Kilian (2009). The effect of specifying twenty-four lags is explored in section 4.

Because there can be more than one cointegrating relation, the number of cointegrating relations (rank of Π) is determined empirically by calculating the λ_{trace} statistic for every possible cointegrating rank (see Juselius (2006) for more information). The λ_{trace} statistics indicate that the CVAR model contains two cointegrating relations and two common stochastic trends.

After choosing the rank, additional testing is done to determine whether variables are endogenous or weakly exogenous. By weakly exogenous, I mean that a variable does not respond to disequilibria in the two cointegrating relations. To evaluate whether a variable is weakly exogenous and therefore can be excluded from the x vector in equation 2.1, I test restrictions that make all elements of α associated with an individual equation equal to zero. I reject restrictions that make *PRICE* ($\chi^2(2) =$ 6.7, p < 0.037), *NOPEC* ($\chi^2(2) =$ 27.7, p < 0.000), or *SHIP* ($\chi^2(2) =$ 23.7, p < 0.000) weakly exogenous. Conversely, I fail to reject ($\chi^2(2) =$ 2.81, p > 0.24) a restriction that makes *OPEC* weakly exogenous. Furthermore, the relevant elements of the A matrices indicate that OPEC oil production does not respond to changes in the first differences of the other variables in a statistically significant manner, which suggests that *OPEC* is strongly exogenous. Based on these results, *OPEC* is transferred from the x to the w vector[4]. In addition to explicitly testing for exogeneity, reducing the number of endogenous variables improves the efficiency of identifying the long-run structure of a CVAR (Greenslade *et al*., 2002).

Next I identify the cointegrating relations by imposing overidentifying restrictions on the cointegrating vectors. I fail to reject ($\chi^2(1) = 0.73$, p > 0.40) restrictions that eliminate *NOPEC* in the first cointegrating relation and eliminate *SHIP* and *OPEC* in the second cointegrating relation. Any additional restriction is rejected at p < .05.

---

[4] Keeping OPEC in x does not significantly change the results reported below

2.4 Discussion

2.4.1 Base Case

The first cointegrating relation indicates a positive long-run relation between real oil prices and oil production by non-OPEC nations (Table 2.1). The positive relation suggests a supply relationship. This suggestion is consistent with the relevant elements of α. If the first cointegrating relation is normalized such that β_{11} = 1.0, the value of α_{11} = -0.010. This indicates that non-OPEC production adjusts to real oil prices, albeit very slowly. Conversely, the value of α_{13} is not statistically different from zero, which indicates that price does not adjust to disequilibrium in the first cointegrating relation. Furthermore, Γ_{31} is not statistically different from zero (Table 2.2), which indicates that prices do not adjust to changes in the first difference of non-OPEC production. Together, these results indicate that non-OPEC oil production does not have a direct short- or long-run effect on oil prices, which is consistent with the widely accepted hypothesis that non-OPEC producers are price takers.

The second cointegrating relation includes dry bulk maritime freight costs (*SHIP*), the real price of oil (*PRICE*), OPEC production (*OPEC*), and a constant. The meaning of the second cointegrating relation and what *SHIP* represents therein is open to competing interpretations. The positive long-run relation between dry bulk maritime freight costs (*SHIP*) and the real price of oil (*PRICE*) is consistent with the hypothesis that *SHIP* represents oil demand and that demand shocks raise oil prices. Conversely, the positive relation is inconsistent with the hypothesis that *SHIP* represents oil demand and/or macroeconomic activity because higher oil prices should not raise oil demand and/or economic activity. The negative relation between OPEC production and prices is consistent with the notion of supply shocks. The positive relation between OPEC production and the shipping index suggests that higher rates of OPEC production increase the cost of shipping oil in particular and transportation costs (i.e. the dry freight cost-of-shipping index) in general.

I evaluate these competing interpretations by examining the way in which the second cointegrating relation loads (elements of α) into the equations for $\Delta SHIP$ and $\Delta PRICE$ (Table 2.1). If I normalize the second cointegrating relation by *SHIP*, α_{22} is -0.062. The negative sign indicates that disequilibrium in the first cointegrating relation moves *SHIP* towards the equilibrium value implied by that cointegrating relation. This equilibrium value can be calculated by solving the second cointegrating relation for *SHIP* as follows:

$$SHIP_{t-1} = 0.842PRICE_{t-1} + 0.422OPEC_{t-1} + 0.393 \qquad (2.3)$$

The positive coefficient associated with *PRICE* indicates that higher oil prices increase the cost of shipping non-energy goods, such as "grain, oilseeds, coal, iron ore, fertilizer, and scrap metal" (Kilian 2009, p. 1056), which Kilian (2009) uses to construct the shipping index. As such, this loading is consistent with the hypothesis that SHIP simply represents transportation costs. This interpretation is boolstered by the positive value (0.151) of Γ_{23} which represents the short-term effects of oil prices on shipping costs. For example, the USDA (2005) finds a very strong correlation between bunker adjustment surcharges for agricultural containers and lagged values of world oil prices. Together, these results are consistent with the hypothesis that *SHIP* is nothing more than a measure of transportation costs and that its positive relation with oil prices simply represents the effect of oil prices on the cost of moving freight.

Equation 2.3 also indicates that higher rates of OPEC production increase the cost of shipping non-energy goods. This effect is amplified by the positive coefficient (0.106) of A_1 that is associated with the lagged first difference of OPEC production in the SHIP equation (Table 2.3). Together, these positive relations may represent the effect of increasing OPEC oil shipments in particular on shipping costs in general, perhaps by increasing demand for bunker fuels.

Disequilibrium in the second cointegrating relation also loads into the equation for oil prices such that the price of oil adjusts to the long-run value implied by the second cointegrating relation, which can be calculated by normalizing the second cointegrating relation by the element associated with *PRICE* and solving for *PRICE* as follows:

$$PRICE_{t-1} = 1.34 SHIP_{t-1} - 0.567 * OPEC_{t-1} + 0.527 \tag{2.4}$$

The positive coefficient associated with *SHIP* is consistent with the hypothesis that dry bulk maritime freight costs represent oil demand and that demand shocks have a positive long-run effect on real oil prices. Lastly, the negative coefficient associated with *OPEC* is consistent with the notion that OPEC supply shocks have a negative effect on prices—higher rates of OPEC production reduce prices as they did in 1986 and lower rates of OPEC production increase prices as they did in the 1970's and 1980's.

Because the elements of α, β, Γ_1, A_0, and A_1 do not generate unambiguous results regarding the interpretation of *SHIP* and its positive relation with oil prices, I simulate the impulse response functions that are generated by a structural MA version of the identified CVAR model. The structural MA model is specified to match the ordering imposed by Kilian (2009) on the VAR. Specifically, the ordering is such that oil prices respond to all shocks while non-OPEC production does not respond to the permanent shock.

The CVAR contains two cointegrating relations and one weakly exogenous variable (*OPEC*), therefore, model contains two endogenous stochastic trends, which correspond to transitory shocks, and one exogenous stochastic trend, which corresponds to a single permanent shock. The first transitory shock has the greatest immediate effect on non-OPEC production (blue line), increasing output by 0.06 standard deviations, while reducing oil prices (*PRICE*, green line) by 0.016 and transportation costs (*SHIP*, red line) by 0.009 (Figure 2.1). Given the magnitude and sign of these effects and the way in which the effects of the transitory shock fade, I interpret the first transitory shock as a non-price induced non-OPEC supply shock[5]. This shock has the immediate effect of increasing non-OPEC production, which generates smaller reductions in the well-head price of oil, and these reductions have an even smaller effect on transportation costs. Both of these effects fade over time as the shock to non-OPEC production declines.

The second transitory shock has the greatest immediate effect on the shipping index, increasing transportation costs by .15 standard deviations (red line, Figure 2.2). At the same time, the second transitory shock reduces oil prices (*PRICE*, green line) by 0.05 standard deviations. There is no immediate effect on non-OPEC production. Given the magnitude and sign of these effects, I interpret the second transitory shock as a shock to transportation costs. Consistent with this interpretation, the negative relation between shipping costs and the FOB price of oil represents the degree to which a change in transportation costs is passed on refiners, who buy crude oil. *Ceteris paribus,* an increase in transportation costs raises the price of crude oil to refiners. This is the positive relation between the transportation costs and the refiners acquisition cost that Kilian (2009) interprets as the price increasing effect of a demand shock.

But I measure crude oil prices with the FOB price of crude oil. Using this measure, I expect a negative relation between transportation costs and FOB prices if changes in transportation costs

[5] Non-OPEC production changes for reasons not related to price. For example, Russian output declines due to the collapse of the Former Soviet Union and starts to rise again once property relations are reestablished.

are not fully passed to refiners. Under these circumstances, oil producers will bear some of the increase in transportation costs, and this will reduce their net price, as measured by the FOB price.

The degree to which increases in transportation costs are borne by producers is analogous to the differential incidence of energy taxes on energy prices. Empirical analyses indicate that the portion of gasoline or diesel taxes appear in the consumer price depends in part on capacity (Marion and Muehlegger, 2011). Similarly, oil refiners may bear only a portion of unanticipated increase in transportation costs. The rest is borne by oil producers, which has the effect of reducing the net (FOB) price that is received by energy producers.

The permanent shock has its greatest immediate effect on oil prices 0.11 (green line, Figure 2.3), with a lesser positive effect on shipping costs 0.09 (red line) and non-OPEC production 0.007 (blue line). Given the magnitude and sign of these effects, I interpret the permanent shock as a shock to oil prices (a so-called price shock). According to this interpretation, unanticipated increases in oil prices raise transportation costs via transportation fuel charges. This interpretation is consistent with the pattern by which prices and shipping costs decline after their respective peaks. Transportation costs start their decline sooner and the decline in transportation costs is faster than oil prices. The increase in oil prices will slow economic activity, and this will slow shipping activity. The reduction in shipping activity will slowly reduce the transportation costs. Nonetheless, transportation costs remain high because oil prices remain high. That is, a permanent increase in oil prices generates a permanent increase in transportation costs because oil prices are a large component of shipping costs. As described by Büyükşahin *et al.* (2008): "at typical 2005 (*mid 2008)* charter rates, the bunker fuel used to propel ships accounted for approximately one third (*one half*) of dry-cargo shipping costs."

The permanent increase in transportation costs mitigates against interpreting the permanent shock as the demand shock described by Kilian (2009). Consistent with arguments made by Kilian (2009), the positive relation between *SHIP* and *PRICE* could represent a demand shock that raises oil prices. But a significant increase in oil prices should (eventually) slow economic activity, as represented by *SHIP*. But the negative effect of a permanent increase in oil prices is not represented in Figure 2.3. That is, *SHIP* continues to rise after the initial increase in oil price. Eventually, *SHIP* does decline, but only slightly, from a peak of 0.159 five months after the shock to 0.139 twelve years after the shock. If one interprets the permanent shock as an unanticipated increase in oil demand due an unanticipated increase in economic activity, one would then have to argue that the resultant large increase in oil prices has little if any effect on economic activity.

The statistical results of the CVAR model and impulse response functions are inconsistent with the results described in Kilian (2009) in two important ways:

- Reductions in OPEC production raise oil prices, which is consistent with a large literature on the effects of supply shocks.
- Demand shocks, as represented by an increase in the transportation costs, reduce the FOB price of oil, which is inconsistent with the positive effect of demand shocks (as proxied by an increase in *SHIP* that is described by Kilian (2009).
- Increases in oil prices increase the shipping index compiled by Kilian (2009), which is only consistent with interpreting the positive relation between *SHIP* and *PRICE* as the effect of higher oil prices on the cost of moving goods.

These results beg the question, why do the results of the CVAR differ from those generated by the VAR? In these following sections, I attempt to identify the causes for the differences by comparing results of CVAR's that retain individual aspects of the specification used by Kilian

(2009).

2.4.2 The Measure of Oil Price

Kilian (2009) uses the refiner's acquisition cost of imported oil to measure oil prices. This measure creates two difficulties; (1) transportation costs are on both sides of the equations, individually as *SHIP* or as a part of refiner's acquisition cost, (2) it blurs the connection between supply and price because price includes transportation costs, tariffs, wharfage and demurrage, none of which accrue to producers. To investigate these possible effects, I re-estimate the CVAR with the inflation corrected refiner's acquisition cost of imported oil instead of the FOB price.

Repeating the procedure described in section 2, I find that the optimal lag length is two, there are two cointegrating relations and one exogenous stochastic trend and OPEC oil production is (weakly and strongly) exogenous. Changes in α, β, Γ_1, A_0, and A_1 matrices alter the long- and short-run relations in a way that highlights the confusion introduced by including transportation costs, wharfage, and demurrage, in the measure of oil prices. In the first cointegrating relation, the element associated with oil prices no longer is statistically different from zero ($t = 1.37$, $p > 0.17$). This implies that the relatively straightforward relation between non-OPEC oil production and oil prices is disrupted by including transportation costs, wharfage, and demurrage in the measure of oil prices. In other words, the costs transportation, wharfage, and demurrage obfuscate the stochastic trend in FOB oil prices that cointegrates with non-OPEC oil supply. For the second cointegrating relation, α_{22} is still significant, but α_{32} no longer is statistically different from zero. Furthermore, Γ_{32} is not statistically different from zero (Γ_{32}= -0.035, $t = 1.12$, $p > 0.18$), which indicates that real oil prices do not respond to short-run changes in *SHIP*. Together, these results are inconsistent with the hypothesis that dry bulk maritime freight costs can be used to represent oil demand and/macroeconomic activity or that changes in *SHIP* can directly generate changes in *PRICE*.

2.4.3 Disaggregating Oil Supply

Kilian (2009) aggregates OPEC and non-OPEC oil production into a single variable and specifies the first difference of this aggregate[6]. The CVAR estimated in the base case suggests that OPEC and non-OPEC production behave differently. OPEC production is weakly exogenous (i.e. *OPEC* does not respond to changes in *SHIP*, *PRICE*, or *NOPEC*), whereas non-OPEC production is endogenous and is positively related to price.

Consistent with the different behaviors, aggregating oil production undermines one of the most basic results of the base-case—that the system contains two cointegrating relations. The rank test statistics reject the null hypothesis of zero cointegrating relations ($\lambda_{trace} = 48.15, p < 0.002$) but fails to reject the null hypothesis of only one cointegrating relation ($\lambda_{trace} = 10.2, p > 0.62$). The one cointegrating relation that remains corresponds to the second cointegrating relation in the base model–the supply relation associated with the first cointegrating relation disappears. This disappearance is consistent with the notion that non-OPEC producers are price takers that set production based on price, OPEC producers use some other criteria not included in the model, and that aggregating supply decisions, which use different criteria, obfuscates the relation between non-OPEC supply and price that is captured by

[6] I do not evaluate the effects of taking the first difference of aggregate supply. By definition, an I(0) time series (the percent change in total oil supply) can not have a long-run relation with I(1) variables (*SHIP, PRICE*).

the first cointegrating relation in the base model.

2.4.4 Lag Length

In contrast to the two lags specified in the CVAR model described above, Kilian (2009) estimates a VAR with twenty-four lags, which seems to correspond to a two-year period[7]. To assess this difference, I relax the "rule of thumb" and use to the SBC and Hannan-Quinn criteria to consider up to twenty-five lags (twenty-four lagged first differences). The values of both criteria grow monotonically with lag length longer than two (one lagged first difference). Thus, using twenty-four lags in the CVAR is not consistent with the statistical criteria I use to choose the lag length.

To examine the effect of choosing Kilian's lag length, I estimate the CVAR with twenty-four lagged first differences. This longer lag length changes one of the most basic results—that the system contains two cointegrating relations. The rank test statistics barely reject the null hypothesis of zero cointegrating relations ($\lambda_{trace} = 43.3, p < 0.05$) but fails to reject the null hypothesis of only one cointegrating relation ($\lambda_{trace} = 20.8, p > 0.17$). This change is consistent with a reduction in model efficiency at long lag lengths (Juselius, 2006).

2.5 Conclusion

Taken together, the results of this analysis suggest that the importance of demand shocks and the unimportance of supply shocks described by Kilian (2009) are not robust to alternative

[7] Kilian (2009) does not offer any objective basis for the choice of twenty-four lags.

specifications that are consistent with some widely held assumptions about the working of the world oil market. Although clever, the index for dry cargo bulk freight costs does not seem to proxy monthly macroeconomic activity and/or non-precautionary oil demand—the index for dry cargo bulk freight costs simply measures transportation costs. As such the positive relation with oil prices simply represents the effect of higher oil prices on transportation costs. It needs to be noted that this finding in and of itself does not diminish the importance of demand shocks – my results suggest that Kilian's finding does not support the demand shock hypothesis, because Kilian's shipping index is not a measure of aggregate demand. Therefore, more research is needed to determine the relative importance of aggregate demand shocks on the price of oil, provided that better proxies for aggregate demand become available at the global level.

Furthermore, the results described here beg the interpretation of an oil demand shock. While exogenous political or weather events can generate significant reductions in oil demand (e.g. the earthquake and tsunami that hit Japan in March 2011), it is more difficult to tell a compelling a story for a positive oil demand shock. Oil is not consumed in isolation, it is consumed by capital equipment. Given that total capital stock changes little from one year to the next, it is difficult to describe a scenario in which a growing economy generates an oil shock—rarely does oil consumption exceed expectations by a large percentage. Indeed, despite claims that a demand shock played a significant role in the 2007-2008 spike in oil prices, global consumption for oil grew less rapidly than forecast by the US Energy Information Administration in 2006.

Of course, oil consumption, which is directly observable, is not equivalent to oil demand, which is not directly observable. So, it is possible that oil demand grew sharply during the 2007-2008 oil price spike, but oil consumption grew only at trend. This mismatch is only possible if the own price elasticity for oil demand is greater than the income elasticity. That is, oil demand grew sharply, but consumption was held back by the price effect. Such an effect is inconsistent with a large body of empirical results which indicate that the absolute value of the income elasticity for oil is much greater than the own price elasticity for oil.

Instead, the results of this analysis are roughly consistent with the importance of oil supply shocks. As described by existing literature, price-taking non-OPEC nations use criteria different from those used by OPEC nations. And those criteria are not included in the CVAR estimated here, as indicated by the finding that OPEC production is exogenous. Furthermore, OPEC decisions have a long-run relation with oil prices, which is consistent with previous efforts to identify the effect of OPEC production on price (e.g. Gately *et al.*, 1977; Kaufmann *et al.*, 2004; Kaufmann *et al.*, 2008; Chevillon and Rifflart, 2009). Together, these effects are consistent with the importance of both the positive price effects of OPEC supply shocks in the 1970's and early 1980's and negative price effects of the OPEC supply shocks in the mid 1980's.

Lastly, I emphasize that the results presented here are simply meant to evaluate the evidence for the role of demand shocks as presented by Kilian (2009). Because the results suggest that dry cargo bulk freight cargo rates do not proxy economic activity or oil demand, the CVAR estimated here is missing important aspects of the demand side of the oil market. Furthermore, the CVAR fails to represent the depletion of non-OPEC production and so cannot simulate the lack of increased production since 2004, despite significant price increases. As such, the model results should not be interpreted as well-behaved model of the relation among oil supply, oil demand, and oil prices.[8]

[8] These short-comings are one reason I do not run further experiments with the CVAR, such as impulse response functions.

Table 2.1 – Long-run relations as given by the elements of the α *and* β matrices

	Cointegrating Relation #1	**Cointegrating Relation #2**
Cointegrating Vector (β)		
NOPEC	1.000	--
SHIP	--	1.000
Price	-0.395^{+}	-0.844**
OPEC	--	-0.422**
Constant	-0.887	-0.394*
Loadings (α)		
NOPEC equation	-0.010**	-0.008*
SHIP equation	0.032**	-0.062**
Price equation	-0.008	0.024*

Table 2.2 – Short-run (Gamma) matrix of endogenous variables

	$\Delta NOPEC_{t-1}$	$\Delta SHIP_{t-1}$	$\Delta PRICE_{t-1}$
NOPEC equation	-0.287 (-6.163)	0.053 (3.441)	-0.026 (-1.201)
SHIP equation	-0.201 (-1.444)	0.273 (5.917)	0.151 (2.352)
PRICE equation	-0.076 (-0.805)	0.013 (0.431)	0.459 (10.602)

Table 2.3 - Short-run effects of exogenous variable (OPEC)

	$\Delta OPEC_t$	$\Delta OPEC_{t-1}$
NOPEC equation	0.045 (2.691)	0.020 (1.184)
SHIP equation	0.006 (0.119)	0.106 (2.143)
PRICE equation	-0.032 (-0.942)	0.024 (0.706)

Figure 2.1 – Shock to non-OPEC production

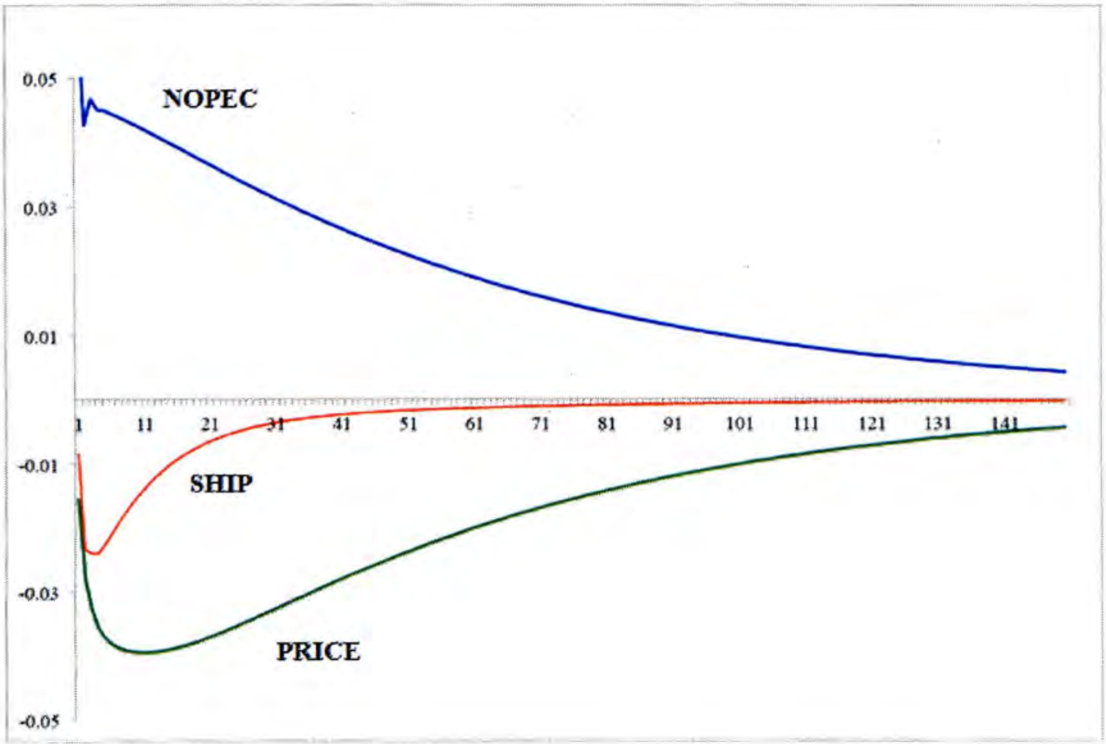


Figure 2.2 – Shock to transportation costs

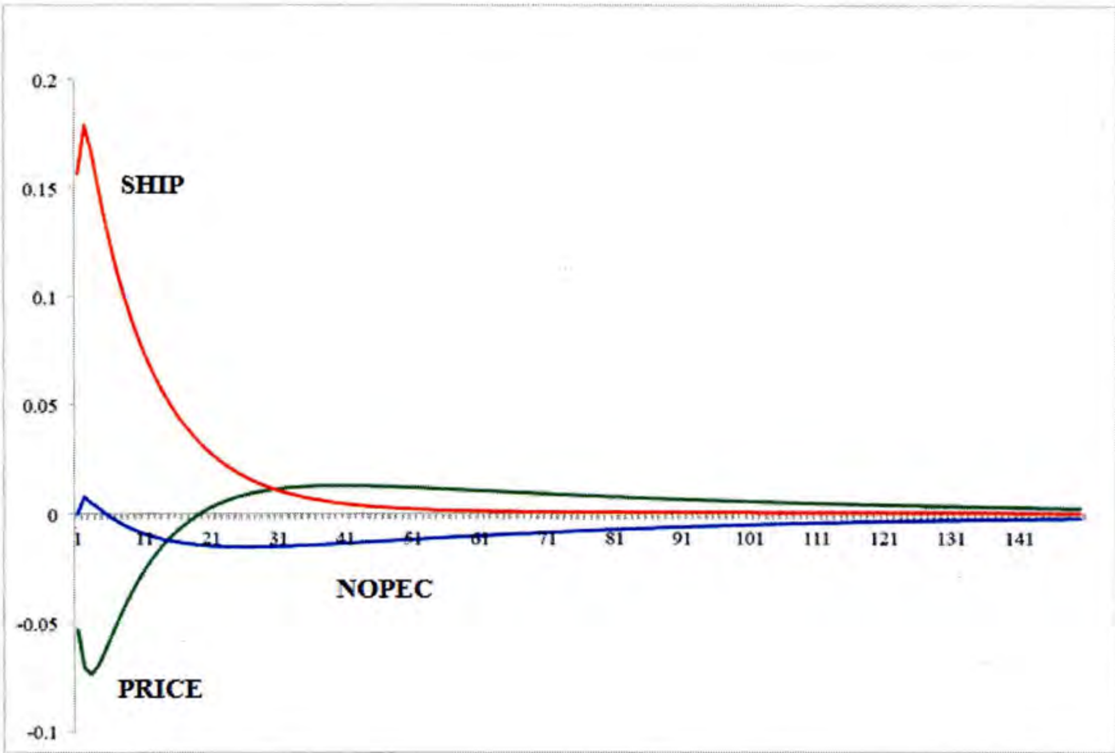
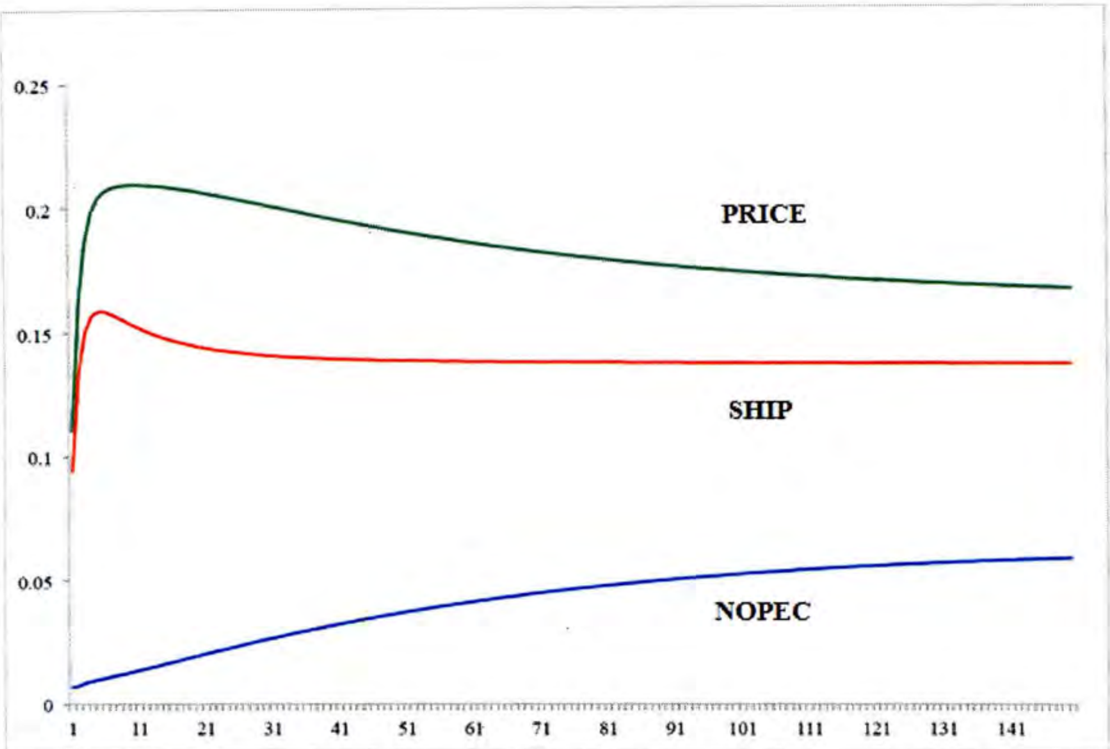


Figure 2.3 – Shock to FOB prices



3. THE RELATION AMONG TRADER POSITIONS AND OIL PRICES: BEYOND CAUSAL ORDER AND MARKET EFFICIENCY

3.1 Introduction

Beginning in 2004, the price of crude oil fluctuates rapidly over a wide range. From $60 per barrel in 2004, the price for West Texas Intermediate (WTI) peaks at $148 in July 2008. Only six months later, the price drops under $40. In January 2012, the price moves beyond $100. And in July 2012, the price drops below $80.

These large price swings probably are not caused solely by volatility—oil prices are not especially volatile relative to other commodities (Regnier, 2007). Instead, these large price swings are attributed to two general causes; speculative expectations and market forces. Kaufmann (2011) argues that speculation plays a significant role in the 2008/2009 price spike and collapse based on three changes; (1) a significant increase in private crude oil inventories that reverses more than twenty years of steady decline, (2) a break-down in the cointegrating relationship between spot and far month future prices that is inconsistent with arbitrage opportunities, and (3) statistical and predictive failures by an econometric model of oil prices that is based on market fundamentals.

But such changes are only suggestive. More direct analyses of speculation's role examine the relation between trading behaviors and oil price. Masters (2008) argues that noncommercial traders now buy and sell oil as a financial asset. And the importance of these traders is increasing. Medlock and Jaffe (2009) note that about 50 percent of the outstanding positions in US futures markets is held by noncommercial traders. Much of this increase may be due to institutional investors (e.g. sovereign wealth funds), who use commodity markets to diversify their portfolios. Commodities in general, and crude oil in particular, are a good hedge against inflation because commodity returns are positively correlated with inflation and a negative correlation with returns on stocks and long-term bonds make commodities a good hedge against investments in equity markets (Gorton and Rouwenhorst, 2006).

If crude oil now is an option for wealth allocation, the effect of trader activity on oil prices depends on the information that is used to buy and sell oil (Weiner, 2002). If traders make decisions independently, their role is benign (Weiner, 2002). Conversely, if traders act as poorly informed participants, either by extrapolating past price changes (i.e. technical analysis) or by following the actions of other traders (i.e. herding or stampeding behavior), traders can amplify market volatility. Consistent with the latter, Medlock and Jaffe (2009) note that the run-up in oil prices is highly correlated with the increasing importance of noncommercial traders.

Efforts to quantify a possible relation between trading behavior and oil prices follow two general approaches. One approach analyzes how traders make decisions. For example, Kraepels (1999) argues that "of the hundreds of fund managers and commodity traders, the vast majority are "systems traders," relying upon the analysis of price trends for their trading decision and paying little if any attention to the fundamentals of the markets in which they are trading." This characterization is supported by econometric results that provide "convincing evidence of positive feedback trading in the oil market" (Cifarelli and Paladino, 2010). Both of these results are consistent with the hypothesis that speculative expectations play a significant role in the large price swings.

But this conclusion is inconsistent with results that are generated by another approach, which examines the causal relation between crude oil prices and trader positions. To date, most of this literature seeks to answer the question, do trader positions 'Granger cause' oil prices and/or do oil prices 'Granger cause' trader positions? If trader positions 'Granger cause' oil prices, trader positions could generate speculative movements that drive oil prices far from the level suggested by market fundamentals. Conversely, the role of traders is seemingly diminished by findings that oil prices 'Granger cause' trader positions.

To date, econometric analyses generally indicate that oil prices 'Granger cause' trader positions while there is little evidence that trader positions 'Granger cause' oil prices. This statistical ordering is first reported by Sanders *et al.* (2004), who use a vector autoregression (VAR) to analyze the relation between commitments of traders (as compiled by the Commodity Futures Trading Commission) and market returns. They find that the net long position held by commercial hedgers declines following price increases. However, traders' net positions do not lead market returns. These general results are replicated by several studies, including the ITF (2008) and Büyükşahin and Harris (2011).

Econometric results that indicate a causal relation from oil prices to trader positions do not invalidate the hypothesis that trader positions have an important effect on oil prices. First, the technique used to analyze causal order, a vector autoregression (VAR), may not be ideal for time series of oil prices and trader positions that contain a stochastic trend (see below). When non-stationary time series are present, VAR modeling techniques cannot represent cointegration and error correction (Engle and Granger, 1987; Stock and Watson, 1988; Sims, Stock and Watson, 1990).

Furthermore, statistical results regarding causal order depend in large part on the information set of conditioning variables. The results described above are generated from VARs that include oil prices and trader positions only. Even those who argue for the importance of speculation

3.2.4 Ten Models

Equation 3.4 is used to generate ten CVAR models, each of which includes the five measures of trader positions, inventories in PADD 2, and the spot price, but differ according to the the maturity of the futures contract and the interest rate. For all models, one of the two prices is the WTI spot price. The other is the price for a futures contract, with either a one-, two-, three-, six- or twelve-month maturity. Each of these five combinations of a spot and futures pries is estimated twice, once with the interest rate on the three-month Treasury Bill and once with the interest on the six-month Treasury Bill. Both interest rates are tested because there is no *a priori* reason to prefer one. Thus, the two possibilities for Treasury bill interest rates, combined with five possibilities for futures contract maturities, generate ten CVAR models. Each of these ten models is 'named' based on the maturity date of the futures contract the Treasury Bill.

3.3 Results

For all models, the smallest values of both the Hannan-Quinn and Schwarz Bayesian criteria indicate a lag length of two. A value of two lags implies that there is one lagged first difference per variable per equation ($\Gamma_1(\Delta x_{t-1}, \Delta w_{t-1})$ in addition to the error correction mechanism because the CVAR model is based on first differences.

For all models, the λ_{trace} statistic strongly ($p < .001$) rejects the null hypothesis that there are zero cointegrating relations. This implies that there is one or more long-run relations among oil prices, trader positions, oil inventories, and interest rates. For all ten CVARs, the λ_{trace} statistic indicates that there are four cointegrating relations. In all cases, the variables that make up the four cointegrating relations are identified by imposing a set of restrictions on elements of the cointegrating vectors that is far from the $p = 0.05$ critical value (Table 3.1).

For each CVAR model, one or more of the time series are weakly exogenous (Table 3.1). The variable(s) that is weakly exogenous varies among models, and includes spot oil prices (five models), futures prices (five models), and crude oil inventories (two models).

Finally, differences in the interest rate on the three- and six-month Treasury bill generally have little effect on the results. The most notable difference concerns the exogenous price variable in the models for the three-month futures contract. Specifically, the futures prices is weakly exogenous in the model that uses the interest rate from the three month Treasury bill whereas the spot price is exogenous in the model that uses the interest rate from the six month Treasury bill. This difference should not be over-emphasized. As indicated in Table 3.1, tests of weak exogeneity are far from the critical value in all cases—the exogenous price variable is decided by a secondary criterion that the price with the smallest value for the test of weak exogeneity is made weakly exogenous.

3.4 Discussion

All ten CVAR models contain at least one cointegrating relation that can be interpreted as a cointegrating relation for the endogenous oil price (Tables 3.2-3.3). These cointegrating relations include the spot price, a futures price, oil inventories, and a measure of trader positions, and load into the equation for the endogenous oil price such that the endogenous oil price error corrects towards the equilibrium value that is implied by this cointegrating relation. In the model for the near-month contract (and three-month Treasury Bill) the first cointegrating relation includes both the spot and futures price, oil inventories, and the short position by non-commercial traders (Table 3.2), and disequilibrium in this cointegrating relation loads ($\alpha = -1.06$) into the equation for the first difference of the spot price in a way that moves the spot price towards the long-run value implied by the first cointegrating relation (Table 3.2). As such, this cointegrating relation can be interpreted as a cointegrating relation for the spot price for WTI.

All of the cointegrating relations for price contain one measure of trader positions. In addition to rejecting ($p < .05$) the null hypothesis that the coefficient associated with the trader position is zero, the inclusion of trader positions is evaluated by testing a restriction that eliminates the measure of trader position from the cointegrating relation for price in the over-identified model. In all cases, the test statistic rejects the null hypothesis that eliminating trader positions does not change the likelihood function given the rank condition (Table 3.1). This rejection suggests that spot price, future prices, and oil inventories do not cointegrate. As such, the traditional cash and carry model does not describe completely, the relation between spot and futures prices for WTI. Specifically, trader positions have information about the long-run difference between spot and futures prices that is not contained in inventories and/or interest rates. This result clearly suggests that trader positions play a role in price discovery.

The effect of trader positions on price is as expected. The sign on the coefficient associated with a long position indicates a positive relation with the endogenous oil price whereas the coefficient associated with short position indicates a negative relation with the endogenous oil price. For example, solving the first cointegrating relation in the model of the near-month future price (three-month Treasury Bill) for the spot price as follows:

$$S_t = -0.002 * I_t + 1.004 * F_{1|t} - 0.004 * Non-Commercial\_Short_t$$

indicates that short positions by noncommercial traders (along with oil inventories *I*) have a negative long-run relation with spot prices. And because this cointegrating relation loads into the equation for the first difference of spot prices, an increase in short positions by noncommercial traders will reduce spot prices.

Positions by commercial traders (as opposed to noncommercial traders) are present in eight of the ten models. At first glance, the importance of commercial positions seems to contradict the importance of speculation by non-commercial traders. Nonetheless, activity by commercial traders may create a bubble in the oil market (Tokic, 2011). According to this hypothesis, the advantages of portfolio-diversifying investments in commodities (Gorton and Rouwenhorst, 2006) push institutional investors to invest in crude oil even after prices move beyond the upper value of a price range that is consistent with market fundamentals. As this happens, commercial hedgers lose their informational advantage, which pushes them to remove their short hedges. This reduces their ability to arbitrage inefficiencies in the traditional cash and carry model. And by doing so, oil producers become positive feedback traders, buying the crude oil futures as the price of crude oil rises. This explanation is consistent with the importance of commercial long positions in six of the ten models.

The loadings of the cointegrating relation for price have information about the causal order among oil prices, trader positions, and oil inventories. In all cases (other than the models for the two-month futures contract), the cointegrating relation for price loads into the equation for the first difference of the trader position in a way that moves the trader position towards the long-run value that is implied by the cointegrating relation for price. This result, coupled with results that the cointergation relation for price loads in the equation for the endogenous oil price, indicate that adjustments between trader positions and oil prices run in both directions, as opposed the unidirectional causal relation from oil prices to trader positions that is suggested by previous analyses (Sanders *et al.*, 2004; ITF, 2008; Büyükşahin and Harris, 2011). This difference may be generated by the inclusion of variables other than prices and the use of a CVAR rather than a VAR.

The cointegrating relations for price also illustrate the relation between prices and oil inventories. The cointegrating relations for price show a negative long-run relation between inventories and the spot price when the spot price is endogenous. Consistent with market fundamentals, an exogenous increase in oil inventories depresses spot prices. This relation often is the basis of media reports that relate movements in spot or future prices to government announcement regarding inventories (e.g. New York Times, 1993).

Conversely, the cointegrating relations for price in the models for futures contracts with maturity dates of three, six, or twelve months load into the equations for both the endogenous futures price and oil inventories in a way that moves both variables toward the equilibrium value that is implied by the cointegrating relation. In these cointegrating relations, there is a negative relation between oil inventories and the exogenous spot price. This relation is consistent with the notion that an unexpected short-run increase in demand (e.g. an unusual period of cold weather) will reduce oil inventories (i.e. a draw on stocks) and cause spot prices to rise (Pindyck, 2001).

The result that adding a trader positions (and oil inventories) to the traditional cash and carry model generates cointegration between the spot and future price may explain previous results, which suggest that spot and futures prices do not cointegrate. For example, Moos and Al-Loughani (1994) find that spot and futures prices for WTI do not cointegrate and they argue that this lack of cointegration is due to a time varying risk premium. Gulen (1998) and Maslyuk and Smyth (2009) try to fix this omission by allowing for a structural break in the relation between spot and futures prices. This approach generates results that indicate cointegration, but allowing for a structural break may simply be a statistical 'fix' for problem that is caused by omitted variables. Our results clearly indicate that spot and futures prices cointegrate when their long (and short) run relations are estimated as part of a system that includes trader positions, oil inventories, and interest rates.

The inclusion of variables other than spot and futures prices and the use of a CVAR rather than a VAR allow us to revisit an important aspect of the efficient market hypothesis, the prediction hypothesis. The prediction hypothesis, which postulates that price discovery occurs largely in the futures markets, can be evaluated based on the classification of spot and futures prices as exogenous/endogenous and the loadings of the cointegrating relation for price. For models of the near- and two-month futures contract, the futures price is weakly exogenous and the cointegrating relation for price loads into the equation for the first difference of the spot prices. These results suggest that price discovery occurs in the futures market and that spot prices 'error correct' to changes in futures prices. As such, these results are consistent with the prediction hypothesis. In addition to the standard arguments made by the prediction hypothesis, the near month and two month futures contracts are traded heavily and so may have information about supply and demand in world oil markets that goes beyond the physical supply and demand for a single crude that may be losing its ability to serve as a benchmark (Fattouh, 2007). Furthermore, the futures price is more transparent than the spot price because there is no requirement for public disclosure of spot contracts. Instead, spot prices are generated by price reporting agencies that use different methodologies (Fattouh, 2011). This result is consistent with those generated by Schwarz and Szakmary (1994) who find that futures prices dominate price discovery for light sweet crude oils. But a leading role for futures prices contradicts results that spot prices lead the near month and three month futures prices with no relation with six and nine month futures prices (Quan, 1992).

Conversely, the prediction hypothesis is inconsistent with results generated by models of futures contracts with six or twelve-month maturities. In these models, the spot price is weakly exogenous and the cointegrating relation for price loads into the equation for the first difference of futures prices. These results suggest that price discovery occurs in the spot market and that far-month futures prices 'error correct' to changes in spot prices. This reversal may be explained by the expiration dates on the futures contracts. Trading in these far month markets is thin and therefore their price evolution may follow prices on more heavily traded spot markets (either directly, or as transmitted to spot prices via near- and two-month futures contracts). This result contracts those reported by Silvapulle and Moosa (1994) who find that futures prices (near month, three month, and six month maturities) lead spot prices, with some evidence or a smaller nonlinear causal relationship from spot prices to futures prices.

Beyond the direction of causality, the rate at which the difference between spot and futures prices reaches the equilibrium varies systematically with the maturity of the futures contract (Table 3.4). Disequilibrium in the price difference between the spot and near-month contract is eliminated very quickly, as indicated by the -1.06 value for the alpha that loads the cointegrating relation for price into the equation for the first difference of spot prices. This is consistent with arbitrage opportunities between spot and near month contracts. The opportunity for arbitrage declines as the expiry date for the futures contract extends further into the future and this is reflected by a slowing in the rate of adjustment as the maturity of the future contract extends beyond two months (Table 3.4).

An addition to adjustments via the cointegrating relation for price, futures prices adjust to spot prices via short-run effects that are quantified by the elements of Γ. The elements of Γ

that are associated with the first difference in spot prices are positive for the equation for the first difference of futures contracts with two, three, six, and twelve month maturities. The values imply that about 20-30 percent of a change in the spot prices appears the next period in the futures price. This effect is absent for the near month contract, in which the loading of the cointegrating relation ensures rapid adjustment. Spot prices do not have a statistically significant short-run effect on futures prices at any maturity.

3.5 Conclusion

The results generated by the ten CVAR models suggest that the long-run cointegrating relation between spot and futures prices for WTI is possible only if the models include oil inventories and trader positions. The inclusion of the latter implies that trader positions may play an important role in price discovery. This result provides additional evidence that speculation plays a role in price discovery. A role for speculation is re-inforced by results that suggest a bi-directional adjustment between prices and trader positions, as opposed to the one way causal relation from oil prices to trader positions that is suggested by previous analyses.

While these results are consistent with a role of speculation in the recent spike and collapse in oil prices, the results reported here do not answer several important questions, such as how important is the effect of trader position on spot or futures prices and do their effects change over time. To answer these effects, future research will use impulse response functions to quantify the effects of trader positions and recursive estimation techniques to examine the degree to which the price effects of trader positions, oil inventories, and/or interest rates change over time.

Table 3.1 - Diagnostic statistics for the CVAR models

	WTI(1)	WTI(2)	WTI(3)	WTI(6)	WTI(12)
3-month T-bill rate models					
WTI (spot)	27.78^{**}	4.52	1.32	0.78	3.89
Commercial long	11.17^{*}	17.05^{**}	19.96^{**}	21.54^{**}	21.22^{**}
Commercial short	10.63^{*}	11.93^{*}	13.64^{**}	16.39^{**}	18.26^{**}
Noncomm. Long	15.98^{**}	14.70^{**}	14.12^{**}	13.26^{**}	13.10^{*}
Noncomm. short	18.36^{**}	16.75^{**}	15.26^{**}	13.32^{**}	13.75^{**}
Noncom. Spread	13.15^{*}	8.43^{+}	8.36^{+}	8.17^{+}	7.94^{+}
3-month T-bill rate	11.65^{*}	12.63^{*}	12.85^{*}	11.56^{*}	10.32^{*}
Inventories	5.45	17.57^{**}	20.07^{**}	21.45^{**}	22.61^{**}
WTI (futures)	2.98	0.58	1.23	4.67	12.36^{*}
Coint. relations	4	4	4	4	4
Ident. restrictions	$\chi^2(23) = 25.90$	$\chi^2(19) = 17.71$	$\chi^2(19) = 18.63$	$\chi^2(19) = 16.69$	$\chi^2(19) = 18.34$
Exclude trader position	$\chi^2(1) = 4.11^{*}$	$\chi^2(1) = 6.40^{*}$	$\chi^2(1) = 5.14^{*}$	$\chi^2(1) = 6.86^{**}$	$\chi^2(1) = 7.10^{**}$
Unbiasedness	$\chi^2(1) = 4.40^{*}$	$\chi^2(1) = 7.25^{**}$	$\chi^2(1) = 5.58^{**}$	$\chi^2(1) = 8.53^{**}$	$\chi^2(1) = 9.78^{**}$
6-month T-bill rate models					
WTI (spot)	27.97^{**}	4.52	1.45	1.16	4.50
Commercial long	10.88^{*}	17.06^{**}	20.03^{**}	21.87^{**}	21.77^{**}
Commercial short	11.73^{*}	13.16^{*}	14.96^{**}	17.76^{**}	18.98^{**}
Noncomm. Long	14.73^{*}	13.36^{*}	12.80^{*}	12.25^{*}	12.41^{*}
Noncomm. short	18.26^{**}	16.76^{**}	15.31^{**}	13.69^{**}	14.35^{**}
Noncom. Spread	16.43^{**}	12.09^{*}	12.03^{*}	11.82^{*}	11.55^{*}
6-month T-bill rate	12.32^{*}	13.36^{**}	13.62^{**}	12.57^{*}	11.73^{*}
Inventories	6.01	18.46^{**}	20.97^{**}	22.44^{**}	23.39^{**}
WTI (futures)	3.29	0.97	1.71	5.19	12.72^{*}
Coint. relations	4	4	4	4	4
Ident. restrictions	$\chi^2(23) = 27.10$	$\chi^2(19) = 18.65$	$\chi^2(19) = 19.20$	$\chi^2(19) = 17.75$	$\chi^2(19) = 20.08$
Exclude trader position	$\chi^2(1) = 4.04^{*}$	$\chi^2(1) = 6.98^{**}$	$\chi^2(1) = 8.97^{**}$	$\chi^2(1) = 7.42^{*}$	$\chi^2(1) = 7.50^{**}$
Unbiasedness	$\chi^2(1) = 4.39^{*}$	$\chi^2(1) = 8.01^{**}$	$\chi^2(1) = 10.29^{**}$	$\chi^2(1) = 9.40^{**}$	$\chi^2(1) = 10.26^{**}$

Test statistics are statistically significantly different from zero at the: **1%, *5%, +10% level.

Table 3.2 - Results of CVARs with 3-month Treasury bill rate

	WTI(0)	C_Long	C_Short	N_long	N_Short	Spread	TB3N	Inventory	WTI(#)
WTI(1)									
CR#1†	1.000†	--	--	--	0.004*†	--	--	0.002*	-1.004**
CR#2	--	--	--	1.000†	-0.997**	--	--	--	--
CR#3	--	-0.102**	--	--	-0.897**	1.000	--	--	--
CR#4	--	--	-0.119**	--	-0.876**	1.000	--	--	-0.005*
WTI(2)									
CR#1†	1.000†	--	0.017*	--	--	--	--	0.019**†	-1.016**
CR#2	--	--	--	1.000†	-0.993**	--	--	--	--
CR#3	--	0.038**	--	1.000	-1.044**	--	--	--	--
CR#4	--	--	0.045**	--	1.000	-1.053**	--	--	0.002*
WTI(3)									
CR#1†	-0.971**†	-0.029*†	--	--	--	--	--	-0.034*†	1.000**
CR#2	--	--	--	1.000†	-0.993**	--	--	--	--
CR#3	--	-0.046**	--	--	-0.948**	1.000	--	--	--
CR#4	--	--	-0.054**	--	-0.939**	1.000	--	--	-0.003*
WTI(6)									
CR#1†	-0.952**	-0.045*†	--	--	--	--	--	-0.067*†	1.000†
CR#2	--	--	--	1.000†	-0.993**	--	--	--	--
CR#3	--	-0.053**	--	--	-0.944**	1.000	--	--	--
CR#4	-0.003*	--	-0.062**	--	-0.933**	1.000	--	--	--
WTI(12)									
CR#1	-0.922*†	-0.066*†	--	--	--	--	--	-0.110**†	1.000†
CR#2	--	--	--	-0.991†	1.00**	--	--	--	--
CR#3	--	-0.062**	--	--	-0.945**	1.000	--	--	--
CR#4	0.042*	--	1.000**	--	12.924**	-13.863	--	--	--

Test statistics are statistically significantly different from zero at the: **1%, *5%, +10% level.

Variables in red are exogenous to the CVAR

-- Variable excluded from the cointegrating relation based on overidentifying restrictions.

† Cointegrating relation loads into the equation for that variable in a way that moves the variables towards the equilibrium value implied by the cointegrating relation (p< 0.10).

Table 3.3 - Results of CVARs with 6-month Treasury bill rate

	WTI(0)	C_Long	C_Short	N_long	N_Short	Spread	TB6N	Inventory	WTI(#)
WTI(1)									
CR#1'	1.000†	--	--	--	0.004*†	--	--	0.002*	-1.004**
CR#2		--	--	1.000†	-0.998**	--	--	--	--
CR#3	--	-0.114**	--	--	-0.886**	1.000	--	--	--
CR#4	--	--	-0.132**	--	-0.863**	1.000	--	--	-0.006*
WTI(2)									
CR#1'	1.000†	--	0.017*	--	--	--	--	0.019*†	-1.017**
CR#2	--	--	--	1.000†	-0.996**	--	--	--	--
CR#3	--	0.064**	--	--	1.000	-1.068**	--	--	--
CR#4	--	--	0.076**	--	1.000	-1.082**	--	--	0.004*
WTI(3)									
CR#1'	-0.969**	-0.030*†	--	--	--	--	--	-0.033*†	1.000†
CR#2	--	--	--	1.000†	-0.998**	--	--	--	--
CR#3	--	-0.072**	--	--	-0.927	1.000	--	--	--
CR#4	--	--	-0.083**	--	-0.912**	1.000	--	--	-0.004*
WTI(6)									
CR#1'	-0.950**	-0.046*†	--	--	--	--	--	-0.067*†	1.000†
CR#2	--	--	--	1.000†	-0.997**	--	--	--	--
CR#3	--	-0.074**	--	--	-0.926**	1.000	--	--	--
CR#4	-0.004*	--	-0.086**	--	-0.911**	1.000	--	--	--
WTI(12)									
CR#1	-0.920**	-0.068*†	--	--	--	--	--	-0.110*†	1.000†
CR#2	--	--	--	-1.000†	1.00**	--	--	--	--
CR#3	--	-0.078**	--	--	-0.924**	1.000	--	--	--
CR#4	-0.039*	--	1.000**	--	9.884**	-10.887	--	--	--

Test statistics are statistically significantly different from zero at the: **1%, *5%, +10% level.

Variables in red are exogenous to the CVAR

-- Variable excluded from the cointegrating relation based on overidentifying restrictions.

† Cointegrating relation loads into the equation for that variable in a way that moves the variables towards the equilibrium value implied by the cointegrating relation (p< 0.10).

Table 3.4 - Error correction and short-run effects

	WTI(1)	WTI(2)	WTI(3)	WTI(6)	WTI(12)
3-month Treasury bill					
α_{spot}	-1.06**	-0.29*	0.17+	--	--
α_{Future}	--	--	--	-0.09+	-0.07*
Γ_{spot}	-0.12	-0.01	-0.04	-0.03	-0.05
Γ_{Future}	0.13	0.30*	0.30*	0.20*	0.16*
6-month Treasury bill					
α_{spot}	-1.06	-0.29*	--	--	--
α_{Future}	--	--	-0.15+	-0.09+	-0.07*
Γ_{spot}	-0.12	-0.01	0.03	-0.03	-0.05
Γ_{Future}	0.13	0.30*	0.22*	0.20*	0.16*

Test statistics are statistically significantly different from zero at the: **1%, *5%, +10% level.

4. Crude oil: Commodity and/or financial asset?

4.1 Introduction

Beginning in 2004, the price of crude oil fluctuates rapidly over a wide range. From \$60 per barrel in 2004, the price of West Texas Intermediate (WTI) crude oil peaks at \$148 in July 2008, falls to \$40 six month later, and rises to \$107 in March 2012. In July, 2012, the price of oil drops below \$80.

This price volatility imposes costs on the macroeconomy. These costs are partly associated with an asymmetric macroeconomic response to oil price changes, whereby oil price increases slow economic growth or cause recessions, while oil price decreases do not increase the growth rate of real GDP (Mork, 1989). This phenomenon is observed both in the United States and other OECD countries (Mork *et al.*, 1994; Lee *et al.,* 1995). This asymmetric response is generated by a variety of mechanisms (Darby, 1982; Hamilton, 1988; Bohi, 1991).

Beyond this asymmetry, oil price volatility causes sectoral shocks that affect the macroeconomy via investment decisions that affect long-lasting infrastructure, the economy's potential output, and the elasticity of demand and substitution for oil and other forms of energy. Bernanke (1983) shows that oil price volatility induces firms to postpone costly, irreversible investments. The causes for this delay are straight-forward—increasing uncertainty about oil prices makes it rational to postpone decisions because waiting increases the probability of obtaining additional information that would lead to better decisions.

The potentially large macroeconomic impact of increased price volatility begs the question of its cause. The literature posits three possible causes: market fundamentals, speculative expectations, or correlations among markets due to capital flows (Kaufmann, 2011). If market fundamentals are responsible for large changes in prices (e.g. Kaufmann *et al.*, 2008; Kilian, 2009), policy intervention is justified only if the macroeconomic benefits of reducing volatility exceed the costs of suppressing information about changes in market fundamentals. The threshold for intervention is considerably lower if price volatility reflects changes in speculative expectations of future prices or changes in capital flows.

There is evidence that speculation is responsible for a portion of the volatility in oil prices. Masters (2008) finds that non-commercial traders buy and sell oil as a financial asset and that non-commercial traders hold about 50 percent of oil futures positions at the time of the study's publication, and that this percentage grows steadily over the past decade (Medlock and Jaffe, 2009). The ratio of non-commercial to commercial traders in the futures market is important because the former do not hold oil for physical delivery, rather non-commercial traders buy and sell to profit from price movements – in other words, they hold these assets for speculative reasons. By contrast, commercial traders are involved in producing, processing, selling and buying oil, and so they reduce exposure to price changes via hedging. Given these different goals, the substantial increase in the ratio of non-commercial to commercial traders over the past decade indicates a greater presence of speculators in the oil market.

But an increase in the presence of speculators does not automatically increase oil price volatility. If speculators act as predicted by the rational expectations and the efficient market hypotheses, their actions are benign (Weiner, 2002). Conversely, price volatility is enhanced if the behavior by non-commercial traders can be described by herding, technical analysis, contagion, and/or extrapolation. Consistent with these possible behaviors, Kraepels (1999) argues that "of the hundreds of fund managers and commodity traders, the vast majority are "systems traders," relying upon the analysis of price trends for their trading decision and paying little if any attention to the fundamentals of the markets in which they are trading." This characterization is supported by econometric results that provide "convincing evidence of positive feedback trading in the oil market" (Cifarelli and Paladino, 2010). Similarly, Kolodziej and Kaufmann (in review) find that trader positions (and oil inventories) are needed to explain the cash-and-carry relationship between spot and futures prices.

Alternatively, the volatility of oil price could be enhanced by capital flows among markets that are designed to diversify portfolios (e.g. between stock and commodity markets) if non-commodity markets are inefficient compared to predictions made by the rational expectations/efficient market hypotheses. Specifically, large capital flows between equity and commodity markets could alter commodity prices beyond levels indicated by market fundamentals. These flows could be motivated by empirical observations that commodities are a good hedge against investments in equity markets because commodity prices in general, and crude oil prices in particular, are positively correlated with inflation and negatively correlated with returns on stocks and long-term bonds (Gorton and Rouwenhorst, 2006).

Several studies present evidence for correlations between prices in equity and commodity markets. UNCTAD (2011) shows that starting in 2004, correlations between oil prices and financial assets increase, as measured by correlations between daily data in 30-day rolling-windows. Analyzing correlations estimated from one-year rolling windows for daily data, Tang and Xiong (2011) find that correlations between the prices for crude oil and various commodities strengthen over time. Büyükşahin and Robe (2011) also identify a sharp increase in equity-commodity market cross-correlations that start in the fall of 2008, but find no evidence for a longer-term secular trend. Analyzing daily data, Bicchetti and Maystre (2012) find that although daily data suggest a positive correlation between S&P500 and WTI that starts in 2005, these correlations do not appear in high-frequency data before 2008. In that year, a structural break occurs and correlations rise from below 0.1 to 0.5 and remain at approximately that level thereafter. The U.S. Energy Information Administration reaches similar conclusions using lower-frequency data (EIA, 2012). Bicchetti and Maystre (2012) find correlations using high-frequency data (1-hour, 5-minute, 10-second and 1-second frequencies) for WTI crude oil and several other commodities (corn, soybeans, wheat, sugar, live cattle) and the E-mini S&P500 stock market index. Furthermore, their results suggest a structural break in 2008 that marks a substantial increase in commodity and equity market cross-correlations. These studies suggest either on-going changes or structural breaks in the cross-market linkages and/or speculative influences, regardless of whether the correlations are caused by a general financialization of the oil market (Tang and Xiong, 2011) or due to increasing participation by specific types of traders, such as hedge funds (Büyükşahin and Robe, 2011).

Here I examine the relation among daily returns to oil prices, equity prices, and commodity markets by modifying previous efforts in two important ways; expanding the number of variables, which ameliorates omitted variable bias, and estimating the expanded model using the Kalman Filter, which reduces uncertainty associated with using OLS to estimate equations from rolling windows. Specifically, the model is expanded to include equity prices for an oil-producing company, ConocoPhillips. This expansion allows the model to separate changes in oil prices due to market fundamentals from other sources of change. In theory, the equity price for oil producing firms represents the present value of rents due to the extraction of crude oil from proved reserves. As such, a correlation between returns to direct investments in crude oil and oil-producing firms should represent long-run changes in prices due to market fundamentals. Furthermore, the residual of this correlation should be 'white noise.' But if cross market capital flows (or speculation) play a significant role in oil price discovery, these oil price movements will not be included fully in equity prices of oil producing firms. Price movements not included in the equity prices of oil producing firms will impart a pattern to the regression error from the correlation between returns to direct investments in crude oil and indirect investments in oil-producing firms. And this pattern may be related to equity indices if cross-market capital flows and/or speculative expectations affect oil prices.

Second, I estimate this expanded model using the Kalman filter. Previous studies estimate the relation between equity and oil prices from rolling windows, but the optimal width of the rolling window is unknown, may be time-varying, and therefore, may influence the results generated by these estimates. Sensitivity analysis can be used to assess the effect of window widths on the point estimates, however there are no objective criteria to choose the results generated by one set of rolling windows. To reduce this source of uncertainty, I estimate models using both rolling windows and a Kalman filter model that estimates the optimal state variable (regression coefficient) variances as well as the overall equation (measurement) variance.

Results indicate that there is a positive correlation between returns to the spot price of WTI and ConocoPhillips throughout the sample period and this correlation allows us to quantify a change in the correlation between returns to WTI and the S&P 500, which is negative between 2003 and 2008 and positive from 2009 forward. This change is explained by a large reduction in interest rates in the fourth quarter of 2008, which reduces the convenience yield earned on physical holdings of crude oil. This reduction in convenience yields is associated with a change from backwardation to contango in crude oil markets. These two changes alter the source of returns such that crude oil is converted from a commodity to a financial asset, which is termed the financialization of oil.

These results, and the methods used to obtain them, are described in four sections. The second section describes the data sources, as well as the statistical models used to perform the analysis. The third section discusses the econometric findings. The fourth section concludes.

4.2. Methodology

4.2.1. Data

Observations for the price of WTI on the New York Mercantile Exchange (*WTI*), the price of ConocoPhillips stock on the New York Stock Exchange (*COP*), and the Standard and Poor's 500 index (*SP500*) are obtained from the Thomson-Reuters Tick History (TRTH) database. TRTH provides data for 45 million different financial instruments across over 400 exchanges, for both quotes and trades at the millisecond and microsecond frequencies. Because the number of trades at these high frequencies is extremely small (in fact, trades do not occur every millisecond but only once in a while), these data are down-sampled (using sample means) to daily frequencies. Daily data for the spot price of gold are obtained from Globex.

Data are available for October 7, 1996 through August 21, 2011. The sample period is defined by the longest period for which observations are available for all time series. All time series are transformed from levels into log-differences. This transformation is used because correlations among market returns are more relevant than correlations among levels. Furthermore, the logarithmic transformation of the variables compresses the variable ranges, which improves convergence of the Kalman filter maximum likelihood algorithm.

The daily frequency is chosen to avoid the effects of structural changes that are associated with the implementation of around-the-clock electronic trading. As described by Bicchetti and Maystre (2012), round-the-clock electronic trading starts in September 2007. Before then, transaction volumes are very low. Therefore, performing a very high-frequency analysis could potentially involve estimating pre- and post-fully-electronic regimes, in which case one might have in-sample data availability and regime change issues.

4.2.2 Models

To analyze the relation among oil and equity prices, I estimate models using OLS applied to rolling windows and using the Kalman filter. The latter is chosen because it allows me to determine the optimal degree of coefficient persistence and smoothing (as in separating the noise from the signal). The Kalman filter model also allows us to measure both partial correlations (regression coefficients) and the standard errors, which are used to determine the statistical significance of the partial correlations. Quantifying confidence intervals for these time-varying estimates is imperative. Confidence intervals are estimated from rolling-window regressions as well, but the standard errors could be affected by the window size. Specifically, if the window is too wide, too much persistence will be included. To assess these potential effects, I compare results generated by the Kalman filter model to regressions estimated from one-year rolling windows.

The rolling-window regression model can be written as

$$Y_t = X_t\beta_t + \varepsilon_t \text{ estimated for } t\epsilon[t_0, t_0 + w],\ t_0\epsilon[1, T - w] \tag{4.1}$$

in which T represents the number of observations and w is the width of the rolling window. That is, I take a subsample of w observations, and estimate the model with a subset that starts at $[1, w + 1]$ and ends at $[T - w, T]$.

The Kalman filter with time-varying coefficients is specified as:

$$s_{t+1} = d_t + T_t s_t + R_t \eta_t \tag{4.2}$$

$$y_t = c_t + Z_t s_t + e_t \tag{4.3}$$

in which $s_t \in R^m$, $y_t \in R^k$ (with k being 1 in this case), $c_t \in R^k$, $T_t \in R^{m\times m}$, $Z_t \in R^{k\times m}$, $R_t \in R^{m\times n}$, $\eta_t \in R^n$, $e_t \in R^k$, $\eta_t \sim N(0, Q_t)$, $e_t \sim N(0, H_t)$, i and Q_t and H_t are positive-definite. In the analysis below, it is assumed that e_t and η_t are independent.

Equation 4.2 is called the state or transition equation and equation 4.3 is called the measurement equation. The meaning of the state vector s_t depends on the application (e.g. a noisy signal to be filtered in an electrical engineering application). In the statistical application of linear regression, the state vector represents the regression coefficients. For my model:

$$\Delta ln(WTI_t) = \beta_{0,t} + \beta_{1,t}\Delta ln(GOLD_t) + \beta_{2,t}\Delta ln(SP500_t)$$
$$+ \beta_{3,t}\Delta ln(COP_t) + e_t, \qquad e_t \sim N(0, \sigma_e^2) \tag{4.4}$$

the state vector will be

$$s_t = \begin{bmatrix} \beta_{0,t} \\ \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{bmatrix} \tag{4.5}$$

and the states (regression coefficients) evolve according to a random walk

$$\beta_{i,t+1} = \beta_{i,t} + \eta_{i,t}, \qquad \eta_{i,t} \sim N\left(0, \sigma_{\eta_i}^2\right), \quad i \in [0,3] \tag{4.6}$$

I can rewrite the model as follows:

$$\Delta ln(WTI_t) = [1 \quad \Delta ln(GOLD_t) \quad \Delta ln(SP500_t) \quad \Delta ln(COP_t)] \begin{bmatrix} \beta_{0,t} \\ \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{bmatrix} + e_t \tag{4.7}$$

$$\begin{bmatrix} \beta_{0,t+1} \\ \beta_{1,t+1} \\ \beta_{2,t+1} \\ \beta_{3,t+1} \\ \beta_{4,t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_{0,t} \\ \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \\ \beta_{4,t} \end{bmatrix}$$
$$+ \begin{bmatrix} \eta_{0,t} \\ \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \\ \eta_{4,t} \end{bmatrix} \tag{4.8}$$

This means that $T_t = R_t = I_4, \ c_t = 0, d_t = 0,$

$Z_t = [1 \quad \Delta ln(GOLD_t) \quad \Delta ln(SP500_t) \quad \Delta ln(COP_t)],$

$H_t = \sigma_e^2$ and $Q_t = diag\{\sigma_{\eta_0}^2, \sigma_{\eta_1}^2, \sigma_{\eta_2}^2, \sigma_{\eta_3}^2\}$.

Allowing for diffuse initialization (Tsay 2005, p. 503), I still need to determine $\sigma_{\eta_i}^2$, the state variances for the time-varying coefficients and σ_e^2, the measurement equation variance. These are estimated by maximizing the likelihood function

$$p(y_1, \ldots, y_T | \sigma_e, \sigma_\eta) = p(y_1 | \sigma_e, \sigma_\eta) \prod_{t=2}^{T} (y_t | F_{t-1}, \sigma_e, \sigma_\eta)$$

$$= p(y_1 | \sigma_e, \sigma_\eta) \prod_{t=2}^{T} (v_t | F_{t-1}, \sigma_e, \sigma_\eta) \qquad (4.9)$$

in which $y_1 \sim N(s_{1|0}, V_1)$, $v_t = (y_t - s_{t|t-1}) \sim N(0, V_t)$ and $F_{t-1} = y_1, \ldots, y_{t-1}$

Diffuse initialization implies $\Sigma_{1|0} = \infty$, which implies $s_1 \sim N(y_1, \sigma_e^2)$. I can then estimate the vector σ_η^2 and the scalar σ_e^2 using maximum likelihood procedures. One approach is to set $\sigma_e^2 = 1$ and have the σ_η^2 vector adjust to maximize the likelihood function – that will concentrate the measurement error variance. Another option is to let both σ_e^2 and σ_η^2 vary. I use the latter to estimate the variances. To prevent convergence to a local optimum, I use the gradient-free simplex solver. To make sure that the simplex solver identifies a global maximum, I also use a genetic algorithm to maximize the likelihood function. The results indicate that they are the same to within the specified tolerance limit. This suggests that the two methods find the same (global) maximum and that the results reported in the next section are not sensitive to the method used to maximize the likelihood function.

4.3. Results

The maximum likelihood-based variance estimates are as follows:

$$\sigma^2_{\eta_{gold}} = 2.2890E - 5$$

$$\sigma^2_{\eta_{S\&P500}} = 7.2156E - 4$$

$$\sigma^2_{\eta_{COP}} = 1.7304E - 4$$

$$\sigma^2_{e} = 5.4109E - 4$$

Because $\sigma^2_e \neq 1$, these are unconcentrated measurement equation variance estimates. These variance estimates are hard to interpret on their own without visualizing their impact on both the degree of filtering of the stochastically evolving model coefficients, as well as on their standard errors.

As indicated in Figure 4.1, the dependent variable (*WTI*) is positively correlated with the price of the ConocoPhillips stock (*COP*) over the sample period. Furthermore, this positive relation is statistically different from zero for the entire sample period. Conversely, the significance and sign of the relation between *WTI* and *SP500* changes over the sample period. Early in the sample period, 1997- 2002, there is no relation between *WTI and SP500*. Between 2003 and late 2008, the relation generally is negative. From 2009 through the end of the sample period in early 2012, the relation between *WTI and SP500* is positive. Lastly, the relation between WTI and gold goes from slightly negative to slightly positive during the sample period, although it is not statistically different from zero for much of the mid-sample period. These results are similar to those generated by estimating equation 4.3 from samples of one year rolling windows (Figure 4.2).

4.4. Discussion

Results generated by both the Kalman Filter and the regression model estimated from rolling windows indicate that there is a positive relation between returns to direct investments in crude oil (*WTI*) and indirect investments in crude oil via the oil-producing firm ConocoPhillips (*COP*) valuation throughout the sample period (Figure 4.1). The positive relation is consistent with the hypothesis that long-run expectations about the present value of rents generated by oil production are positively related to oil prices.

To test the hypothesis that the positive correlation between *WTI* and *COP* represents the effect of oil prices on the present value of rents generate by oil production, I estimate a model in which *COP* is replaced with the price of ExxonMobil stock (*XOM*) as traded on the New York Stock Exchange. Compared to ConocoPhillips, ExxonMobil is less reliant on exploration and production (E&P)–it is a more vertically integrated company, which provides refining services, wholesale and retail distribution of finished products, etc. Also, ExxonMobil has a large presence in the natural gas market, which blurs the correlation of its valuation with the price of crude oil[9]. Both of these differences suggest that the correlation between *WTI* and *XOM* should be weaker than the correlation between *WTI* and *COP*. Consistent with this hypothesis, the measured correlation between returns to *XOM* and *WTI* measured by the Kalman filter model is statistically insignificant throughout most of the sample period. The stronger correlation with ConocoPhillips implies that its correlation with *WTI* will eliminate the effect of market fundamentals from any correlation among returns to direct investments in oil, equity markets, and gold.

After accounting for the effect of market fundamentals on oil prices, correlations remain between returns to *WTI* and returns to direct investments in a basket of equities, *SP500*, and gold, (Figure 4.1). As described below, the changing correlation between returns to equities and crude

[9] Since 2006, prices of crude oil and natural gas in the US decouple due to large increases in the production of natural gas.

oil are consistent with the financialization of the crude oil market, which I define as a change in the dominant source of returns, from convenience yields, which are derived from holding crude oil as a commodity to the expected appreciation of crude oil as a financial asset.

The hypothesis that the main source of returns to crude oil changes from convenience yields to the expected appreciation of the asset is based on an interpretation of Figure 4.1 relative to the Capital Asset Pricing Model equation (equation 4.10), which represents the expected rate of return on investments. The relation between returns to investments in crude oil and equity markets is given by the following:

$$r_{oil,t} = r_{F,t} + \beta r_{e,t} \tag{4.10}$$

in which $r_{oil,t}$ is the rate of return on crude oil at time t, $r_{F,t}$ is the rate of return on risk-free investments, $r_{e,t}$ is the rate of return in equity markets, such as the S&P 500, and β is the correlation between returns to crude oil and equities. Ad defined by the capital asset pricing model, β is akin to the correlation coefficients generated by the Kalman Filter model and the OLS regression coefficients estimated from rolling windows.

Figure 4.1 indicates that there is a negative correlation between returns to investments in crude oil and the S&P 500 between 2003 and 2007. Without knowing the causal order between these variables, there are at least two possible explanations for the negative correlation. From an economic perspective, rising oil prices tend to slow economic activity (e.g. Hamilton 1983). As such, a period of steadily increasing oil prices could dampen the outlook for economic activity and hence the earnings of publicly traded firms. Although feasible, this explanation is not consistent with the changes in the levels of these two variables. Between 2003 and 2008, both oil prices and the S&P 500 rise rapidly (Figure 4.3). If the negative effect of rising oil prices dampens the outlook for economic activity, this effect would have to be very small because the S&P 500 rises strongly between 2003 and 2008.

Alternatively, the negative correlation between returns to *WTI* and *S&P500* between 2003 and 2008 may be generated by investments in crude oil as a hedge for investments in S&P 500 (Gorton and Rouwenhorst, 2006). If either a macroeconomic expansion or a speculative bubble causes equity values to appreciate faster than oil prices, investors may hedge some of these gains by transferring funds from equities to commodities. Conversely, if oil prices appreciate faster than returns to the S&P 500, noncommercial traders may transfer some of their funds from commodities to equities.

Beyond the negative correlation between returns to crude oil and the S&P 500, the absolute level of returns to investments in crude oil can be approximated by solving equation 4.10. During the 2003-2007 period, the value of β estimated by the Kalman filter model varies between -0.5 and -1.5. If I use an average value of -1.0, and assume that the risk free rate of return is 3 percent (based on values of short-run interest rates) and long-run returns to the S&P500 are about 6 percent, investments in crude oil have a return of about -3 percent. As such, this return lies along the security market line to the left of the origin in Figure 4.4. That is, investors recognize the value of diversification offered by WTI, and are willing to hold WTI even when the expected return to direct investments in crude oil is less than the risk free rate of interest.

This negative rate of return is consistent with the yield curve for futures contracts for crude oil. For most of the 2003-2007 period, the difference between far month and near month contracts for crude oil generally is negative (Figure 4.5). In other words, backwardation in the crude oil market generates a negative return to holding physical quantities of crude oil.

Backwardation in commodity markets is possible because convenience yields can create positive returns. For those who buy and sell crude oil as a commodity, the convenience yield is defined as an adjustment to the cost of carry in the non-arbitrage pricing formula for forward prices in markets with trading constraints. According to this definition, convenience yield can be calculated from the following formula:

$$C_t = r_{et} - \frac{\ln\left[\frac{F_{tT}}{S_t}\right]}{T-t} \tag{4.11}$$

in which C is convenience yield, F is the price for the futures contract for maturity T at time t, S is the spot price at time t, and T-t is the time till maturity for maturity t. Using this formula, convenience yields generally are positive during the 2003-2008 period. These positive values, along with the positive risk free rate of investment, may generate a small positive rate of return to those who hold crude oil as a commodity relative to the negative rate of return (given by the security market line in Figure 4.4) for those who hold crude oil as a financial asset.

The negative and positive sources of returns to investments in crude oil are altered by the 2008 financial crisis, which reduces the convenience yield, alters the yield curve for crude oil, and ultimately converts crude oil from a commodity to a financial asset. This change is signaled by a sharp rise in the correlation relation between returns to crude oil and equity markets, from strongly negative values in the third quarter of 2008 to strongly positive values that start in the first quarter of 2009 and remain positive thereafter (Figures 4.1 & 4.2).

The sharp change from a negative to positive correlation in the fourth quarter of 2008 coincides with the Chapter 11 bankruptcy protection filing by Lehman Brothers (September 15) and the subsequent deepening of the financial crisis. The fourth quarter of 2008 is characterized by significant efforts by the U.S. Treasury to inject capital into financial markets, purchase 'toxic assets,' and force buyouts of large financial institutions (e.g. Merrill Lynch buyout by Bank of America and Wachovia's purchase by Citigroup). Although this chain of events is preceded by bailouts (e.g. Bear Stearns in March 2008 or the rescue of Fannie Mae and Freddie Mac in July 2008), the crisis deepens dramatically after Lehman Brothers files for Chapter 11 bankruptcy protection. Many economists argue that the downfall of Lehman Brothers causes a market panic that generates a contagion in the entire financial system. In order to avoid such a panic, the US Federal Reserve pursues a lax monetary policy by lowering short-term interest rates nearly to zero.

Lowering short term interest rates (r_e) shifts the security market line down and the concurrent decline in the demand for crude oil reduces convenience yields. Both of these changes reduce the financial attractiveness of crude oil as a commodity and emphasize crude oil as a financial asset. Specifically, the reduction in interest rates and increase in oil inventories (Figure 4.6) sharply reduce the convenience yield of crude oil. By the start of 2009, convenience yields for holding crude oil are nearly zero. This eliminates the advantages of holding crude oil as a commodity.

With very small benefits of holding crude oil (due to low convenience yields) and the high cost of marginal physical storage (due in part to high inventory levels, see Figure 4.6), the spot price of crude oil drops relative to the futures price. Consistent with these changes, the crude oil market goes into contango during the first months of 2009 and remains there for the rest of the sample period (Figure 4.5). This long period of contango is relatively unprecedented in crude oil markets.

The combination of contango and the reduction in short-term interest rates have the simultaneous effect of increasing the return to investments in crude oil as a financial asset. With the market in contango, positive returns are defined by the risk-free rate of interest and the positive slope of the yield curve for future contracts for crude oil. At the same time, the reduction in short term interest rates shifts the security markets line down to SML' (Figure 4.4). Both of these changes have the net effect of moving the returns to oil to the right side of the origin—it has positive returns net of convenience yields. Under these conditions, crude oil is a financial asset that has a positive relation with other financial assets, such as the equity markets. This positive relationship is represented by the positive value for β for periods after the fourth quarter of 2008 that are estimated by the Kalman Filter model and the OLS models estimated from rolling windows (Figure 4.1 & 4.2).

4.5. Conclusion

Using the time-varying Kalman filter, I find support for the finding of a regime change in the equity-oil market cross-correlations described in the existing literature (Büyükşahin and Robe, 2011; Tang and Xiong, 2011; UNCTAD, 2011; Bicchetti and Maystre, 2012; EIA, 2012). The Kalman filter results are roughly equivalent to 1-year rolling window regressions, though the former methodology provides an optimal degree of coefficient persistence that does not rely on the arbitrary choice of window width. I find that the past decade is characterized by two distinct regimes – one predominantly driven by market fundamentals (2004-2008:Q3) and one dominated by speculation and portfolio rebalancing (2008:Q4-2011). These results make sense from a financial theory point of view, as evidenced by the convenience yield and Capital Asset Pricing Model (CAPM) analysis. The positive equity-oil market cross correlations in 2008:Q4-2011 are consistent with low interest rates and high oil inventories, and the resulting oil futures market contango. Considering the timing of this regime change, the switch from a market fundamentals-driven regime (in which oil could be considered a commodity) to a financialized regime (in which oil is, on the net, an asset) need not imply any malevolent speculation, but rather portfolio rebalancing following the collapse of the real estate bubble and the ensuing recession. The unprecedented, near-zero interest rate levels, combined with recession-induced crude oil inventory build-up, depressed convenience yields, i.e. the gains from holding oil as a commodity. The net result of this change was the relative increase in value of holding oil as a financial asset.

Figure 4.1. Kalman filter point estimates

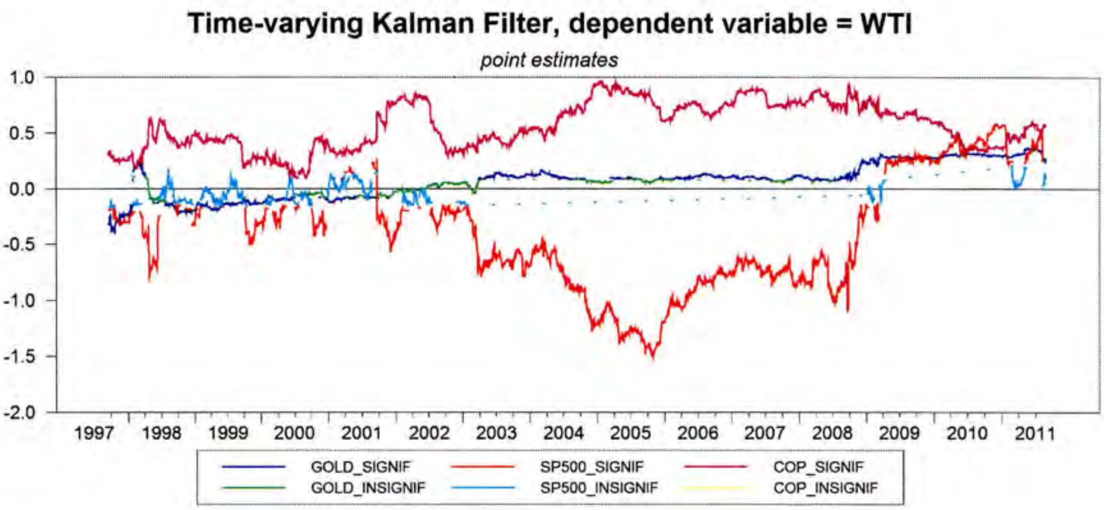


Figure 4.2. Rolling window regression estimates

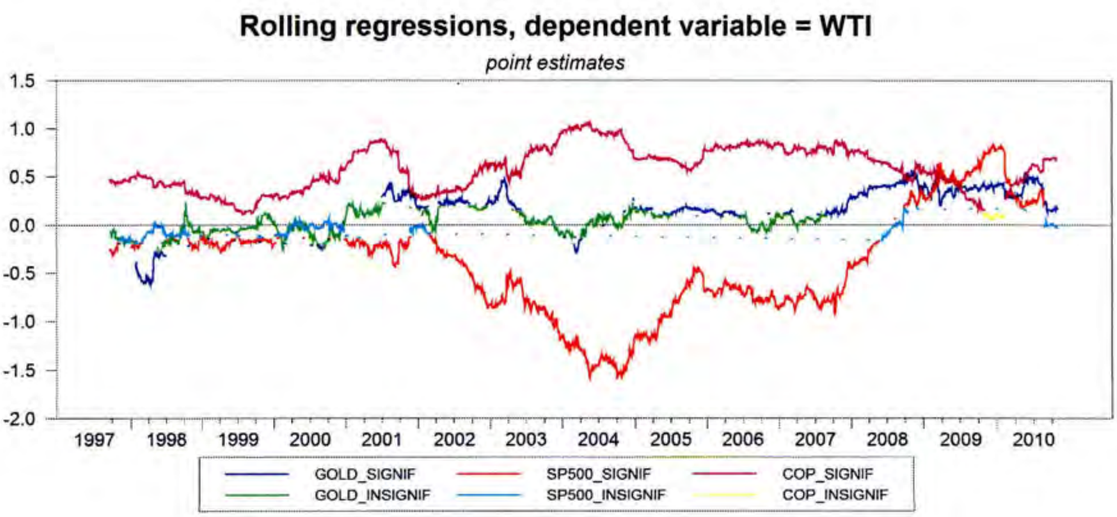


Figure 4.3. Normalized variable levels

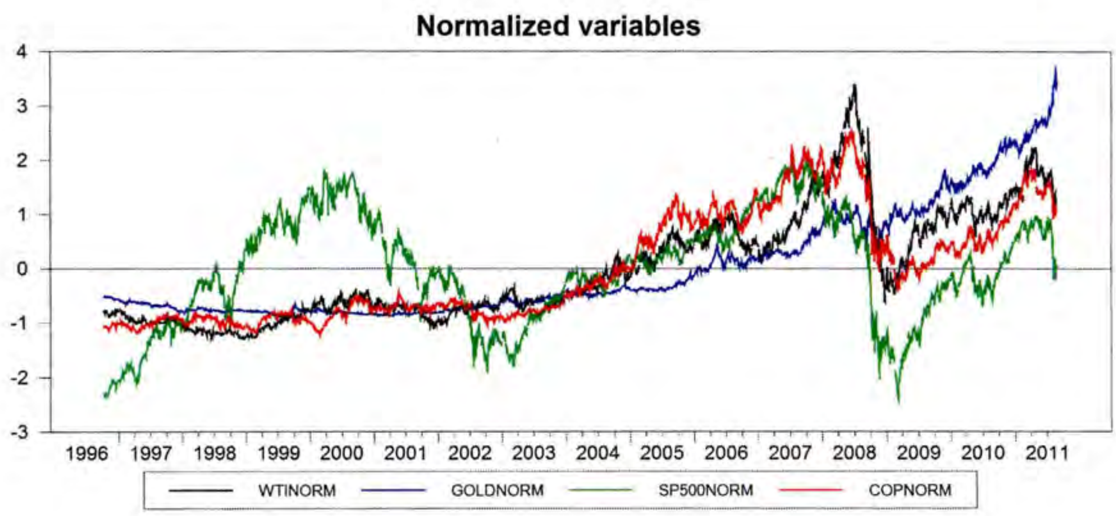
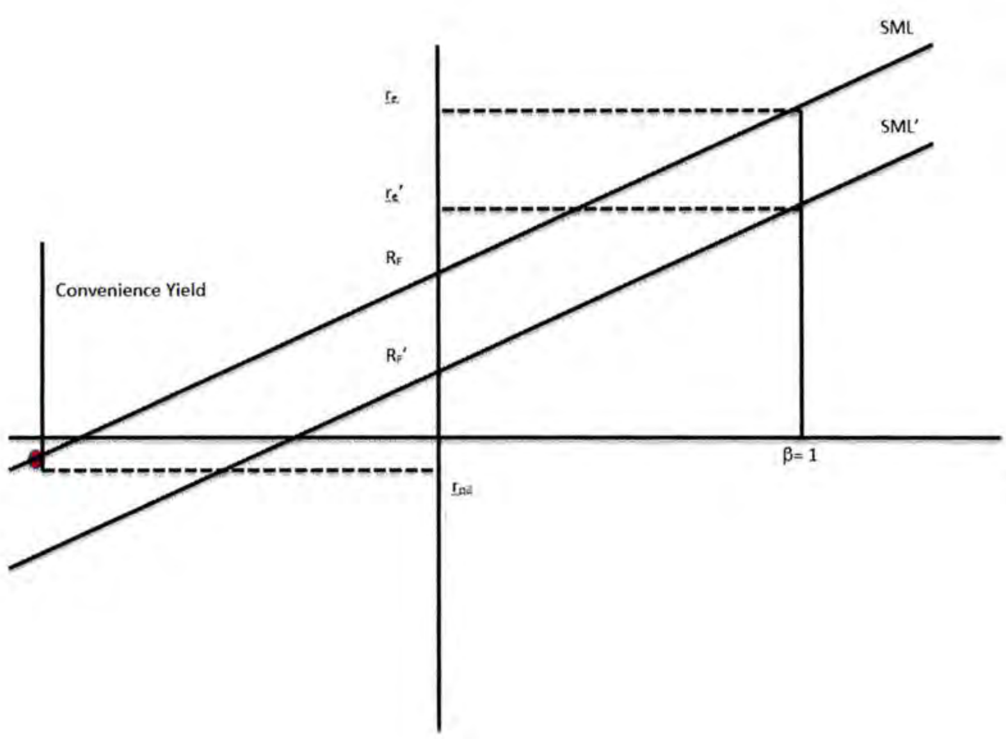
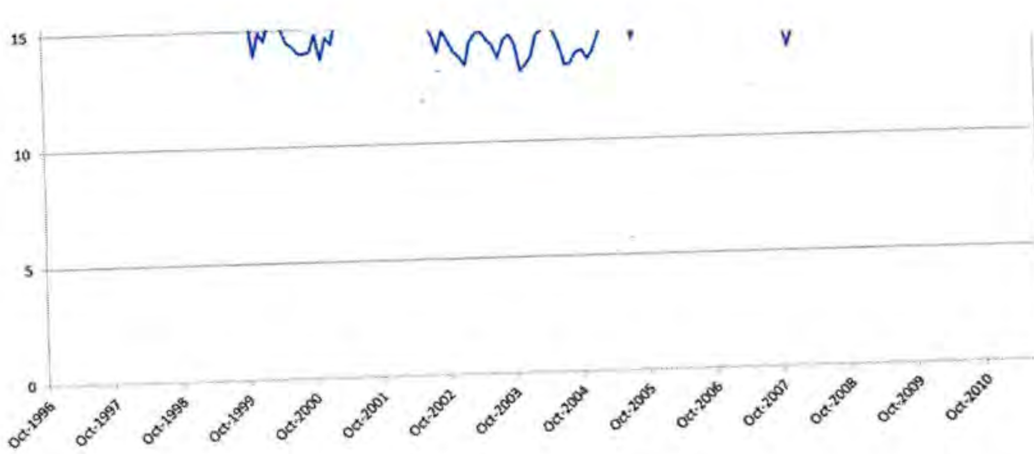


Figure 4.4. Capital Asset Pricing Model





5. CONCLUSION

This dissertation examines the quantitative and qualitative importance of speculation and fundamentals in the oil price discovery process. With regard to economic fundamentals, my research corroborates the long-standing literature on the relative importance of oil supply shocks (e.g. Hamilton, 1983; Burbidge and Harrison, 1984; Mork *et al.*, 1994; Cuñado and Pérez de García, 2003; Jiménez-Rodríguez and Sánchez, 2005) and undermines evidence for the relative importance of demand shocks. Specifically, I find that alternative, more econometrically appropriate specifications of Kilian's (2009) model, along with an alternative choice of variables (e.g. FOB price of oil instead of refiner's acquisition cost, and disaggregating global supply into OPEC and non-OPEC supply streams) lead to results that are inconsistent with Kilian's findings. In other words, Kilian's econometric support for the theoretical arguments found in the recent oil demand literature (Barsky and Kilian, 2002, 2004; Hamilton, 2003; Rotemberg, 2007) are not robust to alternative, arguably superior model specifications.

That said, I do provide additional support for the prevalent, rich oil supply shock literature. This is not to deny the originality of the present work. To the contrary, by examining Kilian's dataset, I explicitly model aggregate and oil-market-specific precautionary demand, thereby doing away with the assumption implicit in the earlier literature that all oil price changes are due to supply shocks, rather than to an interplay of supply and demand. Kilian's argument that one needs to simultaneously model supply and demand is well-taken, however his empirical findings are inconsistent with a proper econometric specification and variable selection. The use of the cointegrated vector autoregression (CVAR) instead of a standard VAR is noteworthy – in the presence of unit roots and cointegration, the former model can capture the long-run equilibrium, short-run adjustments towards the equilibrium, and can verify cointegration, thereby guarding against spurious regressions (Juselius, 2006).

At the same time, I would like to underscore that this research is by no means conclusive – Kilian's set of conditioning variables is limited, and a richer information set could lead to different conclusions. Also, I find that Kilian's dry cargo shipping index is a poor proxy for aggregate demand, and is relevant in Kilian's model mostly due to the oil price mark-up associated with shipping costs. Therefore, my invalidation of Kilian's findings via the use of a shipping-cost-free oil price measure (FOB price) simply weakens his oil demand argument by pointing out the issue of a poor demand proxy, but it does not resolve the problem by using a better conditioning variable, because global monthly aggregate demand data are simply not available. Although clever, Kilian's index for dry cargo bulk freight costs is is not a good correlate of global aggregate demand, and my model re-specification confirms this line of reasoning. It would be a mistake to conclude that oil demand is much less important relative to supply in terms of driving the oil price formation process and the macroeconomic outcomes associated with oil price changes. My findings simply point out the need for further research given the availability of better correlates of the market drivers in question.

In addition to providing empirical findings, my economic fundamentals-related research raises several important questions. First, even if I had a good proxy for aggregate demand (e.g. a monthly measure of PPP-adjusted gross world product), the question remains as to what a demand shock would actually entail. Oil is not consumed in isolation, it is used by capital equipment. Because the total capital stock changes little from year to year, and an increase the in capital utilization rate simply returns the economy back towards a secular trend following a recession, it would be hard to make the case for an abrupt, positive demand shock – unless the economy grows at a rate wildly exceeding expectations via extraordinarily high rates of investment. Also, considering the discussions surrounding the 2004-2008 increase in the price of oil found in both the scientific literature and in the popular press, one should note that global economic growth in general, and oil demand growth in particular, were considerably below the U.S. Energy Information Administration's forecasts. Of course, observed oil consumption is not equivalent to unobserved oil demand, and demand could grow sharply even as consumption changes little. If that were the case, it would imply that the income elasticity of oil demand is considerably less than the own-price elasticity of oil demand, which is inconsistent with the evidence.

The results of my oil market fundamentals-related research are roughly consistent with the importance of oil supply shocks. As described by the existing literature, price-taking non-OPEC nations use criteria different from those used by OPEC nations. And the criteria used by OPEC are not included in the CVAR, as indicated by the finding that OPEC production is exogenous. Furthermore, OPEC decisions have a long-run relation with oil prices, which is consistent with previous efforts to identify the effect of OPEC production on price (e.g. Gately *et al.*, 1977; Kaufmann *et al.*, 2004; Kaufmann *et al.*, 2008; Chevillon and Rifflart, 2009). Together, these effects are consistent with the importance of both the positive price effects of OPEC supply shocks in the 1970s and early 1980s and negative price effects of the OPEC supply shocks in the mid-1980s.

With respect to the speculative (non-commercial traders) or fundamentals-driven (commercial traders) influence of trader positions on oil prices, my CVAR models identify a long-run (cointegrating) relation between trader positions and oil prices. Furthermore, oil prices error-correct to the disequilibrium. This means that trader positions Granger-cause oil prices, which stands in contrast with simple, bivariate Granger causality based on VARs (Sanders *et al.*, 2004; ITF, 2008; Büyükşahin and Harris 2011). The cointegration and error-correction-permitting nature of the CVAR, combined with a richer set of conditioning variables (disaggregated commercial and non-commercial short and long trader positions, oil inventories, and spot and futures prices) probably reduce or eliminate the omitted variable bias present in bivariate Granger causality models. The CVAR is specifically designed to handle non-stationary time series that are cointegrated, so this multiple-cointegrating-vector framework is better suited than a simple VAR. Compared to a VECM, the CVAR allows many cointegrating vectors, each with their individual loadings and speeds of adjustment, while a VECM only allows a single cointegrating relation in each equation, along with its speed-of-adjustment parameter. The restrictive nature of a VECM could have generates different results, and in any case would not have been as informative as a model that allows for several long-run relations that can be restricted in a way that they are both statistically significant and interpretable from an economic and financial theory point of view.

In addition to finding a long-run equilibrium relation between trader positions and oil prices to which oil prices error correct, I find that this disequilibrium has a short-run effect on oil prices, with the direction of causality going from trader positions to oil prices. Again, this conclusion differs from that generated by previous, bivariate Granger causality research based on VARs. My research also finds that there exist long-run equilibria among trader positions within and between categories of trader positions, and that these additional long-run equilibria are limited to trader positions. For instance, it is pretty common to find a long-run relation between non-commercial short and long positions – this makes sense from a theoretical standpoint, since in the long run, the market cannot be long or short on the net. The presence of these long-run relations not only confirms basic financial principles, but it also separates this effect from other long-run relations of interest, which would not be possible in a simple VECM, in which the single cointegrating relation is a linear combination of separate CVAR cointegrating relations (the latter lends itself to an economic interpretation, while the former often does not).

Finally, my research finds that oil prices also have a short-run effect on trader positions. Both commercial and non-commercial traders react to oil price changes in a way that is consistent with positive feedback trading. Together with the classification of oil prices as endogenous or exogenous (spot or futures prices are respectively endogenous or exogenous, depending on the model), these findings suggest a mechanism by which changes in trader positions can have a significant effect on oil prices.

Lastly, my research identifies cross-commodity and equity-to-commodity market financial flows, which can be thought of as a portfolio rebalancing adjustment mechanism. The relationship between spot oil prices, commodity prices (as proxied by gold), equity valuation (S&P500 index) and oil market fundamentals (proxied by the price of the ConocoPhillips stock) changes significantly over time. Due to the arbitrary nature of rolling-window size selection in the case of rolling regressions, I use the Kalman filter which chooses an optimal stochastic evolution of the regression coefficients by choosing state and measurement equation variances to maximize the likelihood function. Interestingly, these results look strikingly similar to a 250-trading day rolling-window regression, though the Kalman filter estimates are preferred on theoretical econometric grounds. My research finds that the fourth quarter of 2008 marks a regime change in the relationships between the variables – the oil-equity relationship changes from one in which oil is valued primarily as a commodity (2004-2008:Q3) to one in which it is valued on the net as a financial asset (2008:Q4-present). Also, the partial correlation between oil and gold prices becomes large, statistically significant and positive after 2008:Q3, which probably suggests an increased general interest in commodity investments that follow the declines in the equity market and a decline in convenience yields. The presence of strong statistically significant correlations between oil prices, equity prices, the prices of other commodities (proxied by gold), and an oil company stock price (proxy for oil market fundamentals) suggests that both speculation (or at least pure financial flows) and fundamentals play a role in the price formation process between 1997 and 2011, although these correlations only became statistically significant in 2004 and they did not experience a regime change until the fall of 2008, which marks the intensification of the mortgage-related financial crisis and the associated global recession.

What are the policy implications of this research? With regard to economic fundamentals, my dissertation shows that the importance of oil supply shocks is consistent with the previous literature. To the extent that these disruptions are caused by wars (e.g. the Iraq War), natural disasters (e.g. hurricane Katrina), OPEC production quota decisions (e.g. the 1973-74 embargo) or other exogenous factors, private, profit-maximizing decisions by oil distributors may not be sufficient to maximize the social benefit due to the macroeconomic consequences of such disruptions. In this sense, an emergency stockpile such as the Strategic Petroleum Reserve (SPR) operated by the U.S. Department of Energy makes sense, although the size of the reserve that would be optimal would be hard to estimate, because the econometrically determined market response to a disruption of a given size only constitutes part of the picture. The macroeconomic response depends on the size of the shock, however the shock of a given size has an unknown probability of occurrence. Looking at the past volatility (in the GARCH sense) is not informative when one is dealing with unpredictable events such as wars. Nevertheless, even though the optimal size of the stockpile cannot be easily determined, the SPR does make sense from the supply shock point of view.

My revisiting of Kilian's (2009) research did not undermine the argument for the importance of aggregate demand shocks – I only pointed out that Kilian does not use a good proxy for aggregate demand. Thus, more research is needed once better aggregate demand proxy data become available. However, at a thought experiment level, I can still reach some policy conclusions. First, one need not worry about negative demand shocks, because they would cause a decrease in price. Such a decrease would stimulate the economy, although the asymmetry of the economic response to price decreases relative to price increases indicates that this response would be small in any case (Darby, 1982; Hamilton, 1988; Mork, 1989; Bohi, 1991, Mork *et al.*, 1994; Lee *et al.*, 1995). As far as positive demand shocks are concerned, it needs to be noted that they can only stem from a movement of the economy towards full potential, since investment only induces gradual changes in the energy-using capital stock. Oil producers should be able to accommodate the movement of the global economy towards its potential output, so positive demand shocks should not in principle cause large price changes. Even if the economy moves towards potential output, i.e. above its secular trend which is characterized by a less than maximum capital utilization, such a temporary boom would be tempered by higher oil prices in a self-correcting way. Therefore, even at a theoretical level, positive aggregate demand shocks should not have considerable recessionary effects.

On the speculation side, one needs to distinguish between the oil futures market-specific speculation and the equity-oil market capital flows. My research indicates that futures market-specific speculation shows potential for speculative bubbles, which indicates that prices could move away from market fundamentals, and such disequilibria could have social costs that could outweigh the benefits reaped by the market participants. This could call for government intervention if these bubbles generate a net social cost after accounting for the benefit to speculators, in other words, if the bubbles are not Pareto optimal. Such a determination would involve a detailed analysis that is beyond the scope of this dissertation. As far as equity-oil market capital flows are concerned, my research indicates that these flows are consistent with the predictions of the capital asset pricing model, and are therefore efficient. What my findings suggest is that the expansionary monetary policy pursued by the Federal Reserve to save the financial system following the financial crisis in the fall of 2008 had an unintended consequence of driving convenience yields to zero, and thereby causing a regime change of the returns from oil changing from commodity-driven returns (convenience yields) to purely financial returns (futures market contango). Therefore, in future recession-fighting efforts, the Fed will need to weigh the benefits of an expansionary monetary policy against the costs of financializing nonperishable commodity markets in general, and the oil market in particular. Considering that an expansionary fiscal policy raises interest rates while raising aggregate demand while an expansionary monetary policy lowers interest rates while increasing demand, an expansionary fiscal policy would not have the side effect of financializing the oil market that an expansionary monetary policy had in the 2008 financial crisis mitigation efforts.

REFERENCES

Akaike, Hirotugu (1974). "A new look at the statistical model identification." *IEEE Transactions on Automatic Control* 19 (6): 716–723.

Barsky, R., and L. Kilian (2002). "Do We Really Know that Oil Caused the Great Stagflation? A Monetary Alternative." In B. Bernanke and K. Rogoff, eds., *NBER Macroeconomics Annual*. Cambridge, MA: MIT Press.

Barsky, R., and L. Kilian (2004). "Oil and the Macroeconomy Since the 1970s." *Journal of Economic Perspectives* 18 (4): 115-134.

Bernanke, B. (1983). "Irreversibility, Uncertainty, and Cyclical Investment." *The Quarterly Journal of Economics* 98 (1): 85-106.

Bicchetti, D., and N. Maystre (2012). "The Synchronized and Long-lasting Structural Change on Commodity Markets: Evidence from High Frequency Data." Retrieved July 8, 2012 from http://mpra.ub.uni-muenchen.de/37486

Brenner, R., and K. Kroner (1995). "Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets." *Journal of Financial and Quantitative Analysis* 30 (1): 23-42.

Burbidge, J., and A. Harrison (1984). "Testing for the Effects of Oil-Price Prises Using Vector Autoregressions." *International Economic Review* 25 (2): 459-484.

Büyüksahin, B., Haigh, M., Harris, J., Overdahl, J., and M. Robe (2008). "Fundamentals, Trader Activity and Derivative Pricing." Retrieved Aug. 16, 2012 from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=966692

Büyüksahin, B., and J. Harris (2011). "Do Speculators Drive Crude Oil Futures Prices?" *The Energy Journal* 32 (2): 167-202.

Büyüksahin, B., and M. Robe (2011). "Speculators, Commodities and Cross-Market Linkages." Mimeo. Accessed July 8, 2012: http://www.ou.edu/content/dam/price/Finance/Oklahoma_conference/2011/Michel%20Robe%20paper.pdf

Bohi, D. (1991). "On the Macroeconomic Effects of Energy Price Shocks." *Resources and Energy* 13 (2): 145-162.

Carmona, R., and M. Ludkovski (2004). "Spot Convenience Yield Models for the Energy Markets." In G. Yin and Y. Zhang (eds.), *Contemporary Mathematics* 351: 65-80.

Chevillon, G., and C. Rifflart (2009). "Physical Market Determinants of the Price of Crude Oil and the Market Premium." *Energy Economics* 31 (4): 537-549.

Cifarelli, G., and G. Paladino (2010). "Oil Price Dynamics and Speculation: A Multivariate Financial Approach." *Energy Economics* 32 (2): 363-372.

Cuñado, J., and F. Pérez de García (2003). "Do Oil Price Shocks Matter? Evidence for Some European Countries." *Energy Economics* 25 (2): 137-154.

Dennis, J., H. Hansen, S. Johansen, and K. Juselius (2005). "CATS in RATS, Version 2." Evanston, IL: Estima.

Dickey, D., and W. Fuller (1979). "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association* 74 (366): 427–431.

Darby, M. (1982). "The Price of Oil and World Inflation and Recession." *American Economic Review* 72 (4): 738-751.

EIA (2012). "What Drives Crude Oil Prices?" Retrieved July 8, 2012 from http://www.eia.gov/finance/markets/financial_markets.cfm

Engle, R., and C. Granger (1987). "Co-integration and Error Correction: Representation, Estimation and Testing." *Econometrica* 55 (2): 251-276.

Fattouh, B. (2007). "WTI Benchmark Temporarily Breaks Down: Is It Really a Big Deal?" *Middle East Economic Survey* 49 (20).

Fattouh, B. (2011). "An Anatomy of the Crude Oil Pricing System." Oxford, United Kingtom: The Oxford Institute for Energy Studies.

Ferderer, J. (1996). "Oil Price Volatility and the Macroeconomy." *Journal of Macroeconomics* 18 (1): 1-26.

Garbade, K., and W. Wilber (1983). "Price Movements and Price Discovery in Futures and Cash Markets." *The Review of Economics and Statistics* 65 (2): 289-297.

Gately, D., J. Kyle, and D. Fischer (1977). "Strategies for OPEC's Pricing Decisions." *European Economic Review* 10 (2): 209-230.

Gorton, G., and K. Rouwenhorst (2006). "Facts and Fantasies about Commodity Markets." *Financial Analysis Journal* 62 (2): 47-68.

Granger, C., and T. Lee (1990). In G. Rhodes, Jr. and T. Fomby (eds.): "Advances in Econometrics: Cointegration, Spurious Regressions and Unit Roots." Greenwich, CT: JAI Press.

Granger, C., and P. Newbold (1974). "Spurious regressions in Econometrics." *Journal of Econometrics* 2: 111–120.

Greenslade, J., S. Hall, and S. Henry (2002). "On the Identification of Cointegrated Systems in Small Samples: a Modelling Strategy with an Application to UK Wages and Prices." *Journal of Economic Dynamics and C*ontrol 26 (9-10): 1517-1537.

Gulen, S. (1998). "Efficiency in the Crude Oil Futures Market." *Journal of Energy Finance and Development 3* (1): 13-21.

Hamilton, J. (1983). "Oil and the Macroeconomy since World War II." *Journal of Political Economy* 91 (2): 228-48.

Hamilton, J. (1988). "A Neoclassical Model of Unemployment and the Business Cycle." *Journal of Political Economy* 96 (3): 593-617.

Hamilton, J. (2003). "What Is an Oil Shock?" *Journal of Econometrics* 113: 363-398.

ITF (2008). "Interagency Task Force on Commodity Markets (2008) Interim Report on Crude oil". Retrieved on August 15, 2012 from http://www.cftc.gov/ucm/groups/public/@newsroom/documents/file/ itfinterimreportoncrudeoil0708.pdf

Jiménez-Rodríguez, R., and M. Sánchez (2005). "Oil Price Shocks and Real GDP Growth: Empirical Evidence for Some OECD Countries." *Applied Economics* 37 (2): 201-228.

Johansen, S. (1991). "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica* 59 (6): 1551-1580.

Juselius, K. (2006). "The Cointegrated VAR Model: Methodology and Applications." Oxford, United Kingdom: Oxford University Press.

Kaufmann, R. (2011). "The Role of Market Fundamentals and Speculation in Recent Price Changes for Crude Oil." *Energy Policy* 39 (1): 105-115.

Kaufmann, R., S. Dees, A. Gasteuil, and M. Mann (2008). "Oil Prices, the Role of Refinery Utilization, Futures Markets, and Non-linearities." *Energy Economics* 30 (5): 2609-2622.

Kaufmann, R., S. Dees, P. Karadeloglou, and M. Sanchez (2004). "Does OPEC matter? An Econometric Analysis of Oil Prices." *The Energy Journal* 25 (4): 67-90.

Kilian, L. (2009). "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." *American Economic Review* 99 (3): 1053-69.

Kolodziej, M., and R. Kaufmann (in review). "The Relation Among Trader Positions and Oil Prices: Going Beyond Causal Order." *Energy Economics*.

Kraepels, E. (2001). "Re-examining the Metallgesellschaft Affair and Its Implication for Oil Traders." *Oil and Gas Journal*, March 26.

Lee, K., S. Ni, and R. Ratti (1995). "Oil Shocks and the Macroeconomy: The Role of Price Variability." *Energy Journal* 16 (4): 39-56.

Lilien, D. (1982). "Sectoral Shifts and Cyclical Unemployment." *Journal of Political Economy* 90 (4): 777-793.

Marion, J., and E. Muehlegger (2011). "Fuel Tax Incidence and Supply Conditions." *Journal of Public Economics* 95 (9-10): 1202-1212.

Maslyuk, S., and R. Smyth (2009). "Cointegration Between Oil Spot and Futures Prices of the Same and Different Grades in the Presence of Structural Change." *Energy Policy* 37 (5): 1687-1693.

Masters, M. (2008). Testimony before Committee on Homeland Security and Governmental Affairs, United States Senate, May 20.

Medlock, K., and A. Jaffe (2009). "Who is in the Oil Futures Market and How Has It Changed?" James A. Baker III Institute for Public Policy, Rice University. Retrieved on July 8, 2012: http://www.bakerinstitute.org/publications/EF-pub-MedlockJaffeOilFuturesMarket-082609.pdf

Morck, R., E. Schwartz, and D. Strangeland (1989). "The Valuation of Forestry Resources Under Stochastic Prices and Inventories." *The Journal of Financial and Quantitative Analysis* 24 (4): 473-487.

Mork, K. (1989). "Oil and the Macroeconomy When Prices Go Up and Down: An Extension of Hamilton's Results." *Journal of Political Economy* 97 (3): 740-744.

Mork, K., Ø. Olsen, and H. Mysen (1994). "Macroeconomic Responses to Oil Price Increases and Decreases in Seven OECD Countries." *Energy Journal* 15 (4): 19-35.

New York Times (1993). "Rising domestic inventories lead to drop in oil prices." January 21.

Peck (1985). "The Economic Role of Traditional Commodity Futures Markets." In A. Peck (ed.), Futures Markets: Their Economic Roles. Washington, DC: American Enterprise Institute for Public Policy Research.

Pindyck, R. (2001). "The dynamics of commodity spot and futures markets: A Primer." *The Energy Journal* 22 (3): 1-29.

Purcell, W., and M. Hudson (1985). "The Economic Roles and Implications of Trade in Livestock Futures." In A. Peck (ed.), Futures Markets: Their Economic Roles. Washington, DC: American Enterprise Institute for Public Policy Research.

Quan, J. (1992). "Two-step Testing Procedure for Price Discovery Role of Futures Prices." *The Journal of Futures Markets* 12 (2): 139-149.

Regnier, E. (2007). "Oil and Energy Price Volatility." *Energy Economics* 29 (3): 405-427.

Rotemberg (2007). "Comment on Blanchard-Galí: The Macroeconomic Effects of Oil Price Shocks: Why are the 2000s so different from the 1970s." Cambridge, MA: National Bureau of Economic Research.

Said, E. and D. Dickey (1984). "Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order." *Biometrika* 71 (3): 599–607.

Sanders, D., K. Boris, and M. Manfredo (2004). "Hedgers, Funds, and Small Speculators in the Energy Futures Market: An Analysis of the CFTC's Commitment of Traders Reports." *Energy Economics* 26 (3): 425-445.

Schwarz, T., and A. Szakmary (1994). "Price Discovery in Petroleum Markets: Arbitrage, Cointegration, and the Time Interval of Analysis." *Journal of Futures Markets* 14 (2):147-167.

Silvapulle, P., and I. Moosa (1999). "The Relationship Between Spot and Futures Prices: Evidence From the Crude Oil Market." *Journal of Futures Markets* 19 (2): 175-193.

Sims, C., J. Stock, and M. Watson (1990). "Inference in Linear Time Series Models with Some Unit Roots." *Econometrica* 58 (1): 113-44.

Smith, J. (2005). "Inscrutable OPEC? Behavioral Tests of the Cartel Hypothesis." *The Energy Journal* 26 (1): 51-82.

Stock, J., and M. Watson (1988). "Testing for Common Trends." *Journal of the American Statistical Association* 83 (404): 1097–1107.

Tang, K., and W. Xiong (2011). "Index Investment and Financialization of Commodities." March, Princeton University, Mimeo.

Toda, H., and P. Phillips (1993). "The Spurious Effects of Unit Roots on Exogeneity Tests in Vector Autoregressions: An Analytic Study." *Journal of Econometrics* 59 (3): 229–255.

Tokic, D. (2011). "Rational Destabilizing Speculation, Positive Feedback Trading, and the Oil Bubble of 2008." *Energy Policy* 39 (4): 2051-2061.

Tsay (2005). "Analysis of Financial Time Series." Hoboken, NJ: Wiley.

UNCTAD (2011). "Price Formation in Financialized Commodity Markets: The Role of Information." New York and Geneva: United Nations, June.

USDA (2005). "Agricultural Container Indicators, Transportation Services Branch, Transportation and Marketing programs." Washington, DC: United States Department Of Agriculture.

Weiner, R. (2002). "Sheep in Wolves' Clothing? Speculators and Price volatility in petroleum futures." *The Quarterly Review of Economics and Finance* 42 (2): 391-400.

Yang, J., D. Bessler, and D. Leatham (2001). "Asset Storability and Price Discovery in Commodity Futures Markets: a New Look." *Journal of Futures Markets* 21 (3): 279-300.

CURRICULUM VITAE

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

- [redacted]

[redacted] [redacted]
[redacted]

- [redacted]
- [redacted]

[redacted]

[redacted] [redacted]
[redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted] [redacted]
[redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted] [redacted]
[redacted]

- [redacted]
- [redacted]

[redacted] [redacted]
[redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted] [redacted]
[redacted]

- [redacted]

[redacted] [redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]