Modelling Oil Price Volatility

by Diego Eduardo Santilli

Student # 4753243

Major Paper presented to the Department of Economics of the University of Ottawa in partial fulfillment of the requirements of the M.A. Degree

Supervisor

Professor Kathleen M. Day

Ottawa, Canada

ECO7997

December 2008

Table of Contents

Acknowledgements

I would like to thank the professors that assisted me during my graduate courses, especially to Leslie Shiell, who raised in me a keen interest for the economics of the environment, and suggested the topic for this paper, and to Kathleen Day for her priceless combination of professional assistance and human kindness. This gratefulness is also extended to my co-workers at HRSDC, especially to Erwin Gómez Gómez and Mohsen Bouhaisa, to whom I almost entirely owe the programming skills necessary to run the econometric work, and the understanding of the theory behind the applied econometric techniques.

It is to my family and friends that I dedicate this work. They have been my support and inspiration throughout this journey.

Abstract

This report measures the volatility of oil prices using univariate GARCH models. The resulting conditional variances (or conditional standard deviations) constitute a measure of uncertainty within the oil market.

Among the several prices that are taken into account by the literature for this sort of analysis, price per barrel has been selected to run the regressions. The sample consists of 753 monthly observations of WTI crude oil price over the period January 1946 to September 2008. The mean equations are ARMA(p,q) processes from which the residuals are taken to form the (conditional) variance equations. The model that best fits the complete sample is a GARCH(2,2), but it is only able to forecast in the short-run. A sub-sample of 271 observations for the period March 1986 to September 2008 is taken and the process is repeated. The resulting GARCH(1,1) process is the best-fitting model, able to forecast both the short-run and long-run.

The findings, which point to increasing uncertainty of oil prices, should raise questions about the effectiveness of policies that may not be compatible with highly volatile variables.

Keywords: oil prices, volatility, ARMA, GARCH.

1. Introduction

Recent economic history has been marked by events that provoked or were induced by shocks in a variety of markets. The price of crude oil is a paradigmatic example: after years of relative stability, the crisis of 1973 represented the beginning of an era characterized by an increasing uncertainty regarding its price. The linkages between a commodity as crucial as crude oil and other commodities and variables made price volatility a key issue in the economics field. As a result, many economists and econometricians began to develop models in order to measure and capture volatility. The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is nowadays one of the most preferred models to describe uncertainty. This paper applies it using up-to-date available data on nominal crude oil prices, in order to understand the years following the invasion of Iraq by American forces in 2003, and the sharp price downturn of 2008.

This paper is organized in the following way: Section 2 introduces the problem of volatile oil prices, the main historical events that produced the most significant shocks, and provides a brief review of research that has delved into the study of volatility in oil prices or of other variables with similar behaviour through the use of GARCH models. Section 3 describes the fundamentals of the model, namely stationarity, differenced time-series, unit root tests, ARMA processes and GARCH models, their assumptions, properties, and restrictions. Section 4 presents the data used in this report. Section 5 provides a step-by-step description of the procedures employed to derive the error series, and then to design the best GARCH models. Section 6 analyzes the results of the previous section and distils the main findings in terms of policy-making, particularly when a policy cannot properly respond to the rapid fluctuations in crude oil prices. Areas for further research are also suggested. Section 7 concludes.

2. Literature Review on Oil Price Volatility Modelling

The importance of oil price stability

Whether or not oil price volatility is a determining factor in real economic activity is a question frequently discussed and debated in the literature. Intuitively, one may establish a turning point in history stemming from the 1973 crisis, and pose the question of how precisely oil prices affected the economy before and after that pivotal year. Hamilton (1983) studied the period between the Second World War and 1973, finding that economic activity was strongly correlated with oil prices over the course of that period. He suggested, however, that though the relationship is not spurious, oil cannot be blamed as the sole factor for economic recessions in the U.S. Not surprisingly, but somehow putting into context the prediction of Hamilton about what would happen after 1973, this relationship was weakened after the crisis. Hooker (1996) discovered that the weakening between the price of oil and the US economy was especially in evidence during the 1986 decline in crude prices.

However, it is not the price of oil itself, but rather it is uncertainty that may affect economic activities. The theory of investment itself supports this argument, from the classic work by Keynes to more recent studies such as Bernanke (1983). More recently, Elder and Serletis (2006) also confirmed the theory; they utilized conditional standard deviations as a measure of volatility within a multivariate GARCH-in-Mean

VAR model to provide evidence that oil price volatility led to significant negative effects on production in the USA between 1980 and 2006.[1] Such negative effects are even more pronounced within the durable goods industries of the manufacturing sector.

There is an obvious transfer of money from oil-importing countries to oil-exporters and that flow is correlated with prices. That said, the uncertainty issue is not exclusively the concern of importing nations, but is also a characteristic of exporter economies. Uncertainty affects their liquidity and their industrial and financial activities, as was observed in the case of Venezuela (Claessnes and Varangis, 1994). Several models have been proposed as a tool to deal with uncertainty in oil prices for exporting countries. For example, Engel and Valdés (2000) focus on a fiscal strategy based on intergenerational oil distribution, precautionary savings, and adjustment costs.

Perspectives on oil – empirical volatility

Price volatility needs to be measured by formal methods; in the existing literature, researchers and statisticians overwhelmingly recommend doing so rather than only observing empirical data and making intuitive appreciations. It is relevant however to look back on the recent history and see how some events may have affected oil prices, and how these affected the entire economy.

Brook et al. (2004) and Fattouh (2005) propose a chronology that focuses on the last 35 years' socio-economic events and that is exceptionally useful. The chronology of events begins in 1973 when the price of oil ascended dramatically as a response to the oil embargo and cut in crude oil production that was part of Arab retaliation against

1 It will be explained later that two equations constitute a basic GARCH model: the *mean equation* and the *variance equation*. The special case mentioned above is called "GARCH-in-Mean", and consists of only one equation (the mean one) that has the variance equation included in it.

American support to Israel. After the dramatic 1973 shock, the cut in oil supply from the Middle East was partially offset by an increase in production and exports from Iran, but the 1978 Iranian Revolution decreased access to supplies yet again. In 1980, Iran and Iraq engaged in a war that would last many years and the price of oil continued climbing. In 1986, oil prices collapsed to about half of their value when crude oil production was boosted in the North Sea, which was not controlled by the Organization of Petroleum Exporting Countries (OPEC) and by Saudi Arabia (an OPEC member); in addition, OPEC adopted another method to set oil prices. On the demand side, that decade was characterized by an effort on the part of consumers to increase the efficiency of the use of oil, and the proliferation of other fuels such as coal.

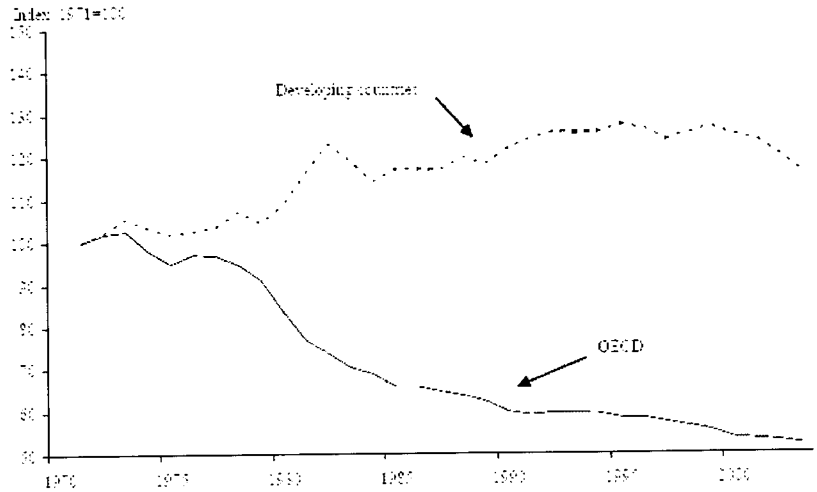
It is worth highlighting that, contrary to the conventional wisdom, Iraq's invasion of Kuwait and the dissolution of the Soviet Union did not dramatically affect the price of oil; there was certainly a supply problem, but both events were at least partially mitigated by increasing production in other OPEC and non-OPEC countries. The major problem at play in 1990 was not production levels, but rather the level of uncertainty. What would have happened if Iraqi troops had not withdrawn from Kuwait? The UN-authorized actions of coalition forces to expel the Iraqi military included ground action and swift air strikes that led prices to drop down to $20/barrel in 1991.

This period in the chronology was characterized by a move to substitute oil for the production matrix, leading to a fall in the "oil intensity of production", defined as "the total primary oil use per unit of output (GDP)".[2] Figure 1 below, which appears in

2 This definition was adopted by the Organization for Economic Cooperation and Development (OECD) and the International Energy Agency (IEA). See for instance Brook et al., (2004, p.9).

Brook et al. (2004, p.9), clearly demonstrates how oil intensity of production has fallen since the 1973 crisis.

Figure 1: Oil intensity of production, worldwide



Source: Brook et al. (2004, p.9).

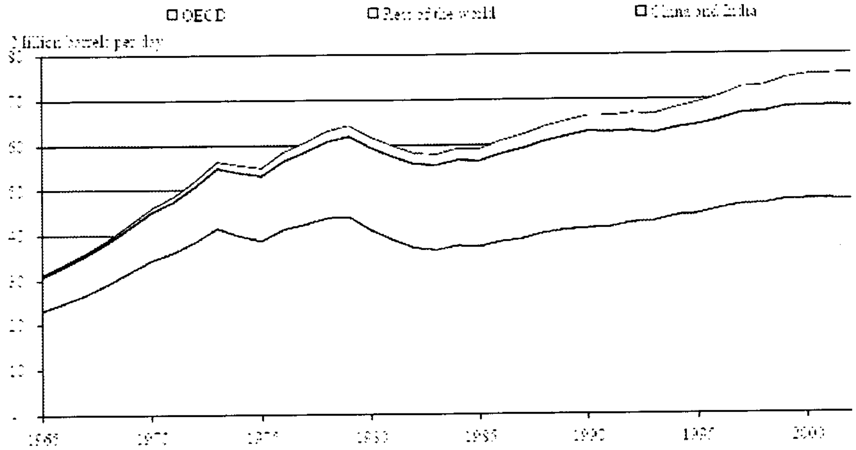
Oil prices continued to fluctuate, reaching a low of about $31/barrel in 1998, and a high of $37/barrel in 2000.[3] After 2000, oil prices again experienced a decrease that had among its causes the fear of an economic recession after US attacks; but the subsequent invasion of Iraq shocked the price again.

In general, there have been shocks over the timeline that have affected oil price positively and negatively. However, it is also true that an ascending trend in nominal prices exists: the price for instance never returned to the 2001 values, let alone 1971 levels. As shown in figure 2 below from BP Statistical Review of World Energy (2004), that upward trend is expected to continue (Brook et al., 2004) as oil dependence shows little indication of being significantly reduced, with the arrival of new, energy-hungry actors –China and India– playing an increasingly important role in the share of demand.

---

[3] Figure 3 below summarizes the behaviour of nominal oil prices over time.

Figure 2: Oil consumption, worldwide



Source: Source: BP Statistical Review of World Energy, 2004.[4]

From a more formal point of view, if a mean had been established in oil nominal prices, the market never returned to it and the expected value of the error term is not zero. This suggests that one is dealing with a non-stationary process, which will be further explained in the next section and be formally tested in Section 5.

Oil scarcity as causal of price volatility: a myth?

Contrary to popular belief that oil reserves will be completely exhausted within the near future, some authors reject the contention and in turn argue that oil extraction will not stop as a result of scarcity of the non-renewable natural resource, but rather when price reaches a level that no one is willing to pay, as happened with coal in Europe (Adelman, 2004). Whether the price of oil is more closely related to the true or false alarms about a declining supply or the result of market speculation activities related to commercial/financial intermediation and consumption than it is to production (Slade,

[4] Data available at http://www.bp.com/liveassets/bp_internet/globalbp/globalbp_uk_english/reports_and_publications/statistical_energy_review_2008/STAGING/local_assets/downloads/spreadsheets/statistical_review_full_report_workbook_2008.xls

1991) is beyond the scope of this paper, which will focus on a simple univariate time-series analysis to model crude oil price volatility.

Measuring oil price volatility

Volatility is presented in the literature usually as one of the issues econometricians have to deal with in order to analyze more complex models. A simple analysis of recent oil price volatility is that of Kuper (2002), who measured volatility through the use of conditional standard deviations from GARCH models. These were applied to the daily and monthly prices of Brent crude barrel, using the general GARCH framework of

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \lambda_j \sigma_{t-j}^2 ,$$

where σ_t^2 is the conditional variance at period t, α_0 is a constant, and α_i and λ_j are the coefficients for the squared errors (ε_{t-i}^2) and the conditional variances (σ_{t-j}^2) of previous periods, respectively. Two different models were obtained; the first analyzed a time frame of 20 years (1982 to 2002). In that model, the chosen values of p and q were 1 and 3 respectively. The second model was applied for a longer time period (1970-2002). For this model the author chose $p = q = 1$. These values were obtained by observing the Schwarz Information Criterion. In many cases, the estimates for $\hat{\alpha}_1 + \hat{\lambda}_1$ in the GARCH (1,1) model are close to unity, which means that the model is not covariance stationary. As a result, the model can be used only to describe short-term volatility. The data used in that paper rejected the unit root hypothesis for the period 1982-2002 but could not reject the hypothesis for that of 1970-2002. Finally, the conditional standard deviations obtained were used to design a parameter of uncertainty. The monthly estimation from daily variances was calculated as the difference between the maximum and the minimum daily and monthly standard deviations identified within a year.

The main objective of the present report is to reproduce the methodology proposed by Kuper (2002) with some alterations, namely a difference in sample size and data frequency, and the type of oil considered (further explained in Section 4).

A similar study based on conditional variances (or conditional standard deviations) was carried out by Kuper and van Soest (2003), with the addition of an analysis of how oil price volatility may influence investment decisions. Through the application of the GARCH (1,3) and (1,1) models (for the periods 1982-2002 and 1970-2002 respectively) once more, and the use of price elasticity of energy use and the elasticity of substitution between labour and energy, they examined the effects of volatility in decision making. These effects became more pronounced when oil prices dropped.

Denni and Frewer (2006) applied alternative GARCH models to analyze the relationship between crude oil and its products' prices. They found that these two variables are co-integrated when studying the effects of Brent crude oil in European markets. They used a weekly time-series that includes observations from January 1990 to September 2005. [5]

Fong and See (2002) used daily data for the period 1992-1997 (around 1500 observations) and carried out a regular GARCH and a more advanced "regime switching GARCH" model in order to explain regime shifts in future returns, namely whether the

[5] Cointegration is a special case of a Vector Autoregression (VAR) that accepts nonstationary variables. See Enders (2004) for a detailed explanation.

data are affected by changes in the regimes. They concluded that by adding this effect to the model, the accuracy of short-term forecasting is enhanced.

Theoretical underpinnings in GARCH models

GARCH models are only one among many techniques used to study volatility, and their use is always taken cautiously by researchers. One of the limitations found in pure GARCH models is their inability to capture certain nonlinearities. Pagan and Schwert (1990) studied the effectiveness of the model with data on stock volatility in the period 1835-1925, and found that an exponential GARCH (EGARCH) model is more able to capture the asymmetries of the process. Another example is Nwogugu (2006a; 2006b), who critiques the fundamentals of all ARMA and GARCH models (and their derivatives) as these models do not provide good measures of risk, an issue especially prevalent in emerging markets. He proposes (and provides foundations for the construction of) new models able to capture elements that are absent in a regular GARCH model, by sketching "aggregations of preferences of market participants, and or optimization problems" that cannot be accurately defined in a rigid model like GARCH (Nwogugu, 2006b, 1749).

Acknowledging these caveats, the literature on GARCH models as a means to describe volatility continues developing. Recently Fan et al. (2008) ran several options to analyze both Brent and WTI crude oil prices, finding TGARCH(1,1,1) the best fitting model for WTI and GARCH(1,1) for Brent. They found significant volatility clustering (periods characterized by either relatively low or high volatility) in both types of oil. The WTI prices, however, appear to be more volatile historically than Brent prices (except during the Gulf War), which constitutes another motivation to analyze the WTI price in this report. They also find leverage effects on WTI, but not on Brent, suggesting that the asymmetric responses are significant in the latter case.

Costello et al. (2008) also put the effectiveness of GARCH under the microscope through a review of previous findings that showed an ARMA process with historical simulation (called in that paper "C&M Method") was able to provide better VaR (Value at Risk) forecasts than a GARCH model. For these tests they used daily data on Brent crude prices, starting from May 20, 1987 and ending on January 18, 2005. They concluded that a well designed semi-parametric GARCH model can perform even better than C&M, and that previous findings cannot be attributed to a structural weakness of the GARCH models.

Macro and microeconomic effects of oil price volatility

The relationship between price volatility and inflation has been thoroughly studied in the economic literature. As an example, Levy Yeyati (1996) started from the hypothesis that there exists a positive correlation between inflation and volatility, and a negative correlation between price volatility and investment, based on the assumption that the high volatility of prices translated into high levels of volatility for capital returns. The study proved this hypothesis to be consistent with the empirical data in short-term dynamics.

Castillo et al. (2007) expand the concept of uncertainty affecting the real economy developed by the New Keynesians, using a second order solution to analyze specifically oil prices as determinants of US inflation. Their model is consistent with the empirical data and shows a positive relationship between inflation and oil price volatility and a negative relationship between these two and the degree of oil substitutability.

The previously mentioned work of Hamilton (1983) demonstrated that there existed a negative correlation between US gross domestic product (GDP) and oil price changes between 1948 and 1980. However, his result could be considered to have considerable selection bias because in that period price increases were mainly observed. The natural question is whether these results would hold if a price decrease was observed. Then, the longer period 1948-1989, characterized by the important oil price drop in 1986, was studied by Mork (1989), who showed that the correlation between oil price and GDP growth is statistically significant when oil prices increase, but statistically insignificant when they decrease. One can argue that the oil price drop in 1986 marked the beginning of the oil price volatility era, because up to that point the obvious trend had been one of increases.

It is worth highlighting that the difference between these empirical findings is not caused by different methodologies or theoretical frameworks, but by a more extended number of observations in the later studies. The purpose of this paper is to further increase the number of observations, and their frequency. The expectation is that a higher level of volatility will be observed the greater the frequency of the data. Lee et al. (1995) found, leaving aside the asymmetric effects of prices changes (i.e. the impacts of suppressive oil prices shocks are greater when shocks are upwards than downwards), these shocks have a greater impact on an economy that is characterized as being stable before the blow than those already coping with a high degree of price volatility.

3. Theoretical Framework

A wide variety of studies in the economic literature have selected GARCH models as a means to analyze the volatility of variables such as commodity prices, GDP, the interest rate, inflation, etc. This section reviews how the model of this report is built and its basic elements, fundamentally following the scheme used by Enders (2004).

Stationarity

Let $\{lnP_t\}$ be the time-series of the natural logarithm of crude oil prices, and $\{lnp_t\}$ its realization. This process is said to be *stochastic* because the real value of the series is unknown (random) until the realization. $\{lnP_t\}$ is *weakly stationary* (or covariance stationary) if the following conditions are true:

$E(lnP_t) = E(lnP_{t+m}) = \mu$	the mean is constant, or *time independent*;
$\text{var}(lnP_t) = \text{var}(lnP_{t+m}) = \sigma^2$	the variance is also time independent;
$\text{cov}(lnP_t, lnP_{t+m}) = \sigma_m$	the covariance depends on the distance between the two points in times.

Only weak stationarity is needed for the purposes of this paper and will be referred to simply as "stationarity". A series that is stationary without any modification is called integrated of order zero [$I(0)$]. But if some test for stationarity shows this not to be the case, the series should be modified using different techniques. In this paper, *differencing* will be used. The purpose of this technique is to transform a non-stationary process into a *difference-stationary process* or a *stochastic trend.* As differencing is applied to the data as many times as necessary to obtain a stationary process, the resulting series will be called integrated of order *d*, where *d* is the number of times differencing is applied. Formally, $dlnP_t \sim I(d)$, where $dlnP_t$ represents the series lnP_t after differencing has been applied . The next section will show that only one difference was necessary in this paper to make the series stationary; consequently, $dlnP_t \sim I(1)$. Now it can be formally stated that

$$dlnP_t = lnP_t - lnP_{t-1}.$$

The Unit Root Tests

These tests are useful to determine if a process is trend or difference-stationary, by posing the null hypothesis that the series has a unit root. The Augmented Dickey-Fuller test and the Phillips-Perron test are widely used in econometric analysis and will be applied in this report.

Augmented Dickey-Fuller Test (ADF): Considering the process $lnP_t = \rho lnP_{t-1} + \varepsilon_t$ the null hypothesis is that $\rho = 1$, so that the process is a random walk. Dickey and Fuller (1979) were able to find a distribution for this "special case" t-test. If the null hypothesis is rejected, the alternative suggests that lnP_t is an AR(1) process with zero mean ($\mu = 0$). Dickey and Fuller adjusted the original test to be useful should the case arise that the series lnP_t is autocorrelated, by altering the basic equation to obtain

$dlnP_t = \gamma lnP_{t-1} + \sum_{i=1}^{p} \xi_i dlnP_{t-i} + \varepsilon_t$, where *p* is the number of lags necessary to have no autocorrelation in the residuals ε_t.

Phillips-Perron Test (PP): The test developed by Phillips and Perron (1988) is also useful where autocorrelation is present, and represents a non-parametric modification of the original DF test.

Stationary Autoregressive Moving-Average models – ARMA(p,q)

A general ARMA(p,q) process is given by the following equation:

$y_t = \alpha_0 + \sum_{i=0}^{p} \alpha_i y_{t-i} + \sum_{i=0}^{q} \beta_i \varepsilon_{t-i}$, where the characteristic roots are assumed to be within the unit circle. As usual, the units are normalized so that $\beta_0 = 1$ and therefore ε_t is the disturbance at the current time t.

Examining the right-hand side of the equation, the first term is the intercept; the second is the autoregressive part (difference equation); and the third part is the moving-average component. It could be noted that if $q = 0$, the model will be a pure AR(p), whereas if $p = 0$ it would be the case of a pure MA(q). The process is stationary if the mean and variance are constant in the long-run. Both the covariances and the autocorrelation coefficients must be time-independent. Note that under a pure stationary ARMA process, the assumption of a constant variance implies homoskedasticity.

Autoregressive and Generalized Autoregressive Conditional Heteroskedasticity models – ARCH(q) and GARCH(p,q)

The variance of a series is considered a key parameter of volatility. If a series presents periods of higher volatility than others, it means that the variance is not constant across the series; consequently, the assumption of homoskedasticity is violated and the modelled homoskedastic ARMA processes are no longer sufficient to describe the series.

Originally proposed by Engle (1982), an ARCH model is able to estimate simultaneously the mean *and* the variance of a process. The general equation is:

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \ldots + \alpha_q \varepsilon_{t-q}^2 + v_t \text{, or simply}$$

$$\varepsilon_t^2 = \delta_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + v_t,$$

where ε_t^2 is the square of the estimated residuals, ε_{t-i}^2 is the past realization of the squared error term, and v_t is a white-noise process, or the conditional variance. Now the square of the residuals is no longer assumed constant, so that the expected value (forecast) becomes:

$$E_t(\varepsilon_{t+1})^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t+1-i}^2 .$$

It can be clearly seen that the magnitude of the forecasted residuals is related to their past realizations, as shown in the following equation:

The residuals can come from different sources; in this report they are obtained from an ARMA process that fits best for the variable $dlnP_t$. In fact, it can be observed that the ARCH model is an AR(q) process on the squared residuals. The ARCH model is applicable if the best model for the series is an AR process. The main characteristics are:

$E(\varepsilon_t) = 0$	(null unconditional expectation of ε)
$E(\varepsilon_t \varepsilon_{t-i}) = 0$	(errors with zero mean and uncorrelated)
$E(\varepsilon_t^2 \mid \varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \ldots) = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t+1-i}^2$	(conditional variance is related to errors from previous periods).

It can be observed that the unconditional variance is unaffected, but the size of the conditional variance is directly related to the size of the realized error of the previous period.

If instead, the best model to describe the series is an ARMA process, a generalized ARCH(p,q) model (GARCH), designed by Bollerslev (1986), should be employed to analyze the volatility. The two main equations that underlie a GARCH process are:

$$\varepsilon_t = v_t h_t^{1/2} \qquad \text{(assumes that the variance of } v_t \text{ is equal to unity),}$$

and

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$$

where in both equations h_t is the conditional variance. It is straightforward to define the right-hand side of the equation; in the latter, the first term is the intercept, the second is the MA (ARCH) component and the third is the AR (GARCH) component, hence a GARCH process is a special case of an ARMA process to model the conditional variance.

Mean and variance equations

As was mentioned before, the error utilized for the GARCH process design can come from a variety of sources and it is very common to obtain them from another ARMA process. If that is the case, it should be noted that despite the general notation p,q to define the orders of the models, the order of the ARMA model (mean equation) need not be equal to the order of the corresponding GARCH model.

The properties of any GARCH model are:

$$E_{t-1}\varepsilon_t^2 = h_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j h_{t-j}$$

$Ev_t = 0$, so it follows that $E\varepsilon_t = E(v_t\sqrt{h_t}) = 0$

$$E\varepsilon_t\varepsilon_{t-j} = E(v_t\sqrt{h_t}v_{t-j}\sqrt{h_{t-j}}) = 0 \ \forall j \neq 0.$$

4. The Data

Selection of crude oil blend:

The majority of the literature has focused on Brent Crude and West Texas Intermediate Crude (WTI) prices to perform econometric analyses in general and analyses of volatility in particular, as they serve as a reference for the principal worldwide oil markets. Brent is traded on the London International Petroleum Exchange (IPE) and is the most widely used price in European markets, whereas WTI is traded on the New York Mercantile Exchange (NYMEX) and influences markets in North and South America as a price benchmark.

This report uses a WTI oil prices database, as opposed to Kuper (2002), who employed Brent oil prices data. The main purpose of this report is to expand on that work and other studies by applying more recent data and visualizing any changes in volatility that might have been caused by the most recent economic shocks. It can be observed and intuitively assumed that there exists a high correlation between the two prices (Brent and WTI), and therefore the level of volatility found in this report about WTI and its predictive power would have some usefulness in the Brent environment, provided similar modelling criteria are considered.

Selection of the frequency

i) Limitations of daily data

GARCH models are very sensitive to the data and frequency (Kuper and van Soest, 2003). This makes the highest frequency available a tempting option for the test run in this report. If following Kuper (2002), daily observations should be used, although major limitations exist: for instance, the farthest available data are from 1986. This is good to capture the price drop that took place that year, but obviously not enough to capture the events of the previous decade, including the 1973 oil embargo. Alternatively, monthly data are available for a broader timeframe, namely 1946 to the present, which represents a time-series of 753 *nominal* observations.[6] The timeframe is therefore much larger than that of Kuper (2002), but the frequency is lower. Consequently, when considering the information available on WTI crude prices, one should choose between working with daily or monthly time-series data, which implies a trade-off between number of observations and time span (a monthly series entails one observation a month, whereas a daily series means around 20 or 22 observations per month).

The biggest problem arising from a daily time-series on oil prices is that the number of observations per year changes, not only because of the different number of weekdays and weekends each year may have, but also the holidays on which markets do not operate.

The data available at the U.S. Energy Information Administration (EIA) on daily prices of WTI-Cushing, Oklahoma and Brent-Europe are for the periods 1986-2008

[6] The term "nominal" is used because when dealing with autoregressive / moving average processes, the number of *effective* observations is reduced by a number equal to the order of such process. Moreover, if the equation needs to be differenced, the number of effective observations will also be reduced. This will be put in practice in Section 5.

and 1987-2008 respectively. These observations include blanks for weekdays and holidays. To deal with these blanks, economists can chose among several options, such as establishing an average between last and next quotation, using the last available price (notice that even though there is no market in a holiday the price is indeed that of the day before), adding a dummy variable to the missing observations (whose estimator is one if it was a missing observation or zero otherwise), or simply ignoring such observations (Greene, 2008). Nonetheless, any of these or other more complex techniques implies certain distortion in the data and is beyond the scope of this paper.

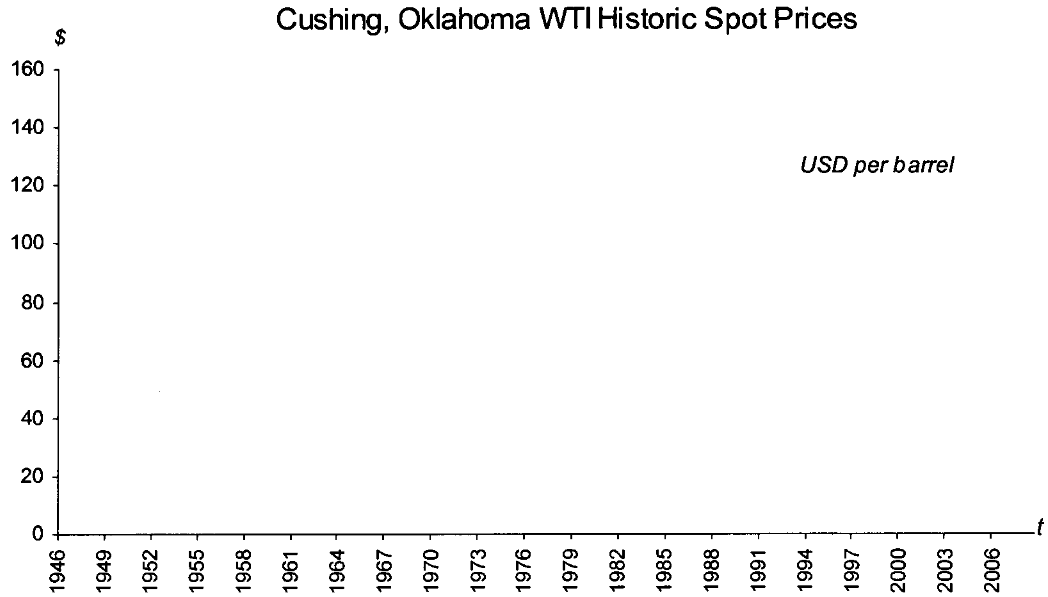
ii) Trade-off between daily and monthly data

Another time-series on WTI oil spot prices has been obtained from the Research Division of the Federal Reserve Bank of Saint Louis (FRB), in October, 2008[7]. The sample includes monthly prices for the period January 1946 to September 2008, constituting a total of 753 observations. These observations are depicted in figure 3 where the variable price is named P_t, while the logarithmic version is shown in figure 4 and the variable is referred to as lnP_t. The natural logarithmic form will be used in the development of this report as it is more relevant in terms of economic analysis.

This time-series has a number of observations that is obviously larger than that offered by the EIA, hence the risk of a time span that does not include enough changes in volatility is minimized. Moreover, one interesting fact in favour of monthly data is that unit root tests perform better with a larger time span, even if that means fewer observations (Perron, 1989). For all these reasons the monthly time-series on WTI crude oil spot prices provided by the FRB is employed in this report.

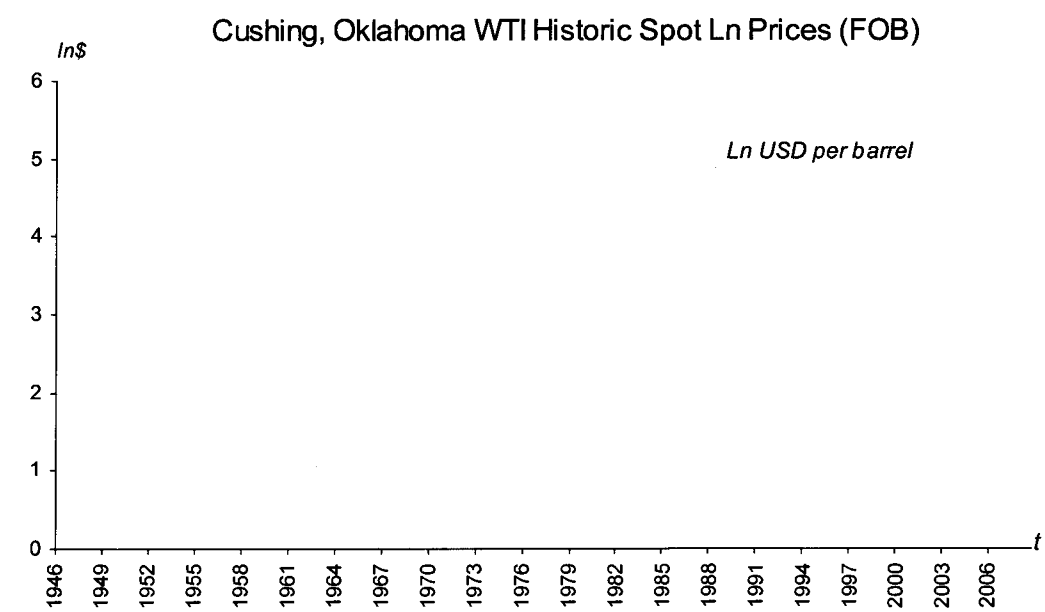
[7] Data available online at http://research.stlouisfed.org/fred2/data/OILPRICE.txt

Figure 3



Primary data source: Federal Reserve Bank of Saint Louis, 2008

Figure 4



Primary data source: Federal Reserve Bank of Saint Louis, 2008

5. Modeling Oil Price Volatility[8]

Unit root tests

At the first sight of figure 4 above, it can be seen that the data are not stationary and some periods are more volatile than others. However, as advised by Enders (2004) the visual examination of the curves should not replace the formal testing.

The first assignment is to test whether or not a unit root exists in the series. There are several options to consider. The modified Dickey-Fuller Test (DF-GLS) of Elliott, Rothenberg and Stock (1996), the Augmented Dicky-Fuller Test (ADF) and the Phillips-Perron Test (PP) were applied to analyze the null hypothesis that the series has a unit root. The results of the three tests are summarized in table 1 below:

Table 1: Unit root test on the ln prices of crude WTI

Tests on lnP_1			Tests with no trend		Tests with a trend	
			t-Statistic	Prob.	t-Statistic	Prob.
DF(GLS)	Lag Length: 1 (Automatic based on Modified SIC, MAXLAG=19)					
	Elliott-Rothenberg-Stock DF-GLS test statistic		1.47195		-2.11376	
	Test critical values:	1% level	-2.56805		-3.48000	
		5% level	-1.94125		-2.89000	
		10% level	-1.61642		-2.57000	
ADF	Lag Length: 1 (Automatic based on Modified SIC, MAXLAG=19)					
	Augmented Dickey-Fuller test statistic		-0.63736	0.8595	-2.13587	0.5243
	Test critical values:	1% level	-3.43883		-3.97032	
		5% level	-2.86517		-3.41581	
		10% level	-2.56876		-3.13017	
PP	Bandwidth: 6 (Newey-West using Bartlett kernel)		Adj. t-Stat			
	Phillips-Perron test statistic		-0.51571	0.8854	-1.95477	0.6244
	Test critical values:	1% level	-3.43882		-3.97030	
		5% level	-2.86517		-3.41580	
		10% level	-2.56876		-3.13016	

[8] The descriptive statistics, regressions and other econometric analysis will be carried out using the popular econometric and forecasting software package EViews © 5.0, made available by Quantitative Micro Software (QMS).

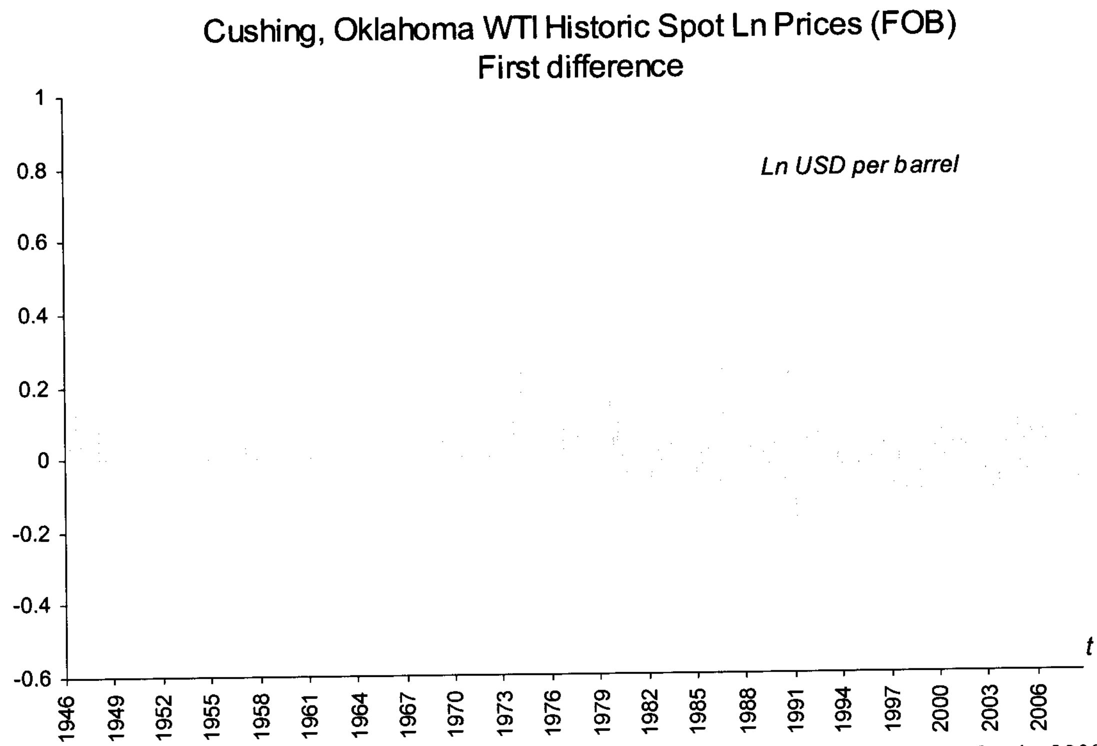
The results show that all the observed critical values / p-values for tests that include an intercept and no trend are not able to reject the null hypothesis at 1%, 5% or 10% significance levels. The series has therefore been proven to be non-stationary. Similar test were run with the inclusion of a trend, and similar results were obtained, meaning that the series is not stationary around a trend either.

The following step is taken to modify the data in order to make it stationary. It is proposed then that the series is *difference-stationary*, which means that the trend is not deterministic but stochastic. The process is called differencing and is applied as many times as necessary to make of lnP_t a stationary series $d_u lnP_t$, where u is the number of times differencing should be applied to obtain a stationary series. Formally it can be said that for a first difference ($u = 1$)

$$d_1 lnP_t = lnP_t - lnP_{t-1}.$$

The resulting line graph is shown in figure 5 below. It can be seen now how the mean tends to be constant and around zero.

Figure 5



Primary data source: Federal Reserve Bank of Saint Louis, 2008

To formally test this, the unit root tests are once again performed, obtaining the results shown in table 2 below:

Table 2: Unit root test on the first difference of ln prices of crude WTI

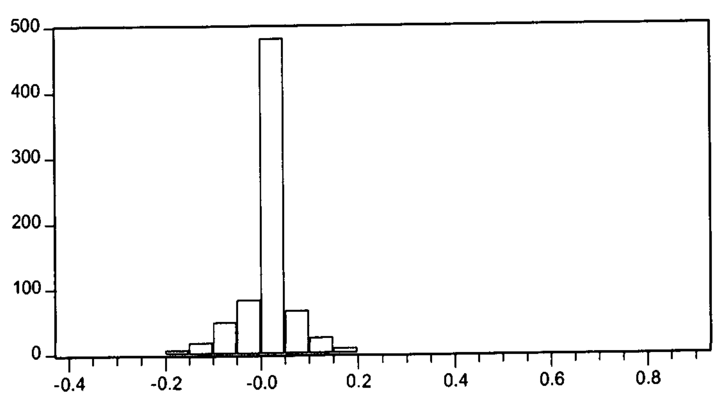
Tests on $d_1 lnP_t$			Tests with no trend		Tests with a trend	
			t-Statistic	Prob.	t-Statistic	Prob.
DF(GLS)	Lag Length: 1 (Automatic based on Modified SIC, MAXLAG=19)					
	Elliott-Rothenberg-Stock DF-GLS test statistic		-6.95571		-4.53003	
	Test critical values:	1% level	-2.56809		-3.48000	
		5% level	-1.94125		-2.89000	
		10% level	-1.61641		-2.57000	
ADF	Lag Length: 1 (Automatic based on Modified SIC, MAXLAG=19)					
	Augmented Dickey-Fuller test statistic		-7.24499	0.0000	-7.25974	0.0000
	Test critical values:	1% level	-3.43895		-3.97049	
		5% level	-2.86523		-3.41589	
		10% level	-2.56879		-3.13021	
PP	Bandwidth: 6 (Newey-West using Bartlett kernel)		Adj. t-Stat			
	Phillips-Perron test statistic		-22.5428	0.0000	-22.5263	0.0000
	Test critical values:	1% level	-3.43883		-3.97032	
		5% level	-2.86517		-3.41581	
		10% level	-2.56876		-3.13017	

With these p-values we are able to reject the null hypothesis of a unit root. The new series is stationary and as no more differencing is necessary, it can be called *integrated of order 1*. Formally, $d_1 lnP_t \sim I(1)$. The monthly time-series on WTI crude oil prices is "*non-stationary on its level*", but "*stationary on its first difference*".

The possibility of working with an $I(1)$ series has some advantages. For instance, recalling the fact that the differential of a variable's logarithm is the rate of change of that variable, then the vertical axis of figure 5 times 100 can be interpreted as the monthly percentage variations of oil prices. We see small shocks up to 1973, the year of the embargo. During the 1973 crisis, an increase of about 90% is observed, representing the beginning of a period of much higher levels of volatility. As explained before, 1986 is the year in which oil prices fall to about half their value, and the figure shows one of the highest shocks in terms of price increases and decreases of about 25% and 40% respectively (always keeping in mind that monthly variations are being taken into account only).

The $d_1 lnP_t$ series (from now on simply called $DlnP_t$ by assuming $D = d_1$) provides the descriptive statistics shown in figure 6 below.

Figure 6: Descriptive statistics of the time-series *DlnP*$_t$



Series: DLOGP
Sample 1946M08 2008M09
Observations 746

Mean	0.005904
Median	0.000000
Maximum	0.852587
Minimum	-0.396009
Std. Dev.	0.064160
Skewness	3.154300
Kurtosis	47.44744
Jarque-Bera	62644.52
Probability	0.000000

The mean is close to zero and the maximum and minimum values are those described above: the increase in 1973 and the decrease in 1986.

The skewness is close to 3.15; this positive value indicates that the sample is asymmetric and that the right tail is longer than the left one. This can be clearly seen in figure 6 above, and constitutes a major difference with the skewness found by Kuper (2002) of about -1.25. A negative number means that the tail is longer on the left side of the mean. The reason for this difference is the choice of sample period. It should be emphasized that the present sample includes the crisis of 1973, as opposed to that of Kuper (2002); this shock produced the value of 0.8 in the Gauss curve, "dragging" it to the right and hence producing a positive skewness.

The kurtosis is higher than 3, which is the value of a normal distribution. Hence, this distribution is peaked or thicker (leptokurtic) relative to the normal, as seen in the figure.

The following is the analysis of the autocorrelation and partial autocorrelation functions (ACF and PACF). Looking at the correlogram in figure 7 it can be seen that the ACF (also called correlogram) tends to have a symmetric decay the greater the lag is. This could be a sign of a low-order autoregressive process (AR). The fact that the values do not drop quickly to zero eliminates the chances of a low order moving average process (MA), but it could also be a hint of an absolute AR process (no MA element). The possibility of an absolute MA process (no AR element) is unlikely because the ACF does not asymptotically tend to zero. In addition, the Q-statistics show that the null hypothesis of no autocorrelation is rejected for all orders of lags. The errors seem to behave as a white-noise process.

Figure 7: Series Correlogram

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.177	0.177	23.485	0.000
		2	-0.007	-0.040	23.526	0.000
		3	-0.007	0.002	23.559	0.000
		4	-0.040	-0.041	24.760	0.000
		5	0.024	0.039	25.177	0.000
		6	-0.047	-0.063	26.846	0.000
		7	-0.001	0.021	26.848	0.000
		8	0.026	0.019	27.363	0.001
		9	0.064	0.062	30.510	0.000
		10	0.045	0.017	32.023	0.000
		11	0.049	0.047	33.813	0.000
		12	-0.027	-0.047	34.380	0.001
		13	-0.074	-0.056	38.529	0.000
		14	-0.014	0.007	38.687	0.000
		15	-0.090	-0.088	44.877	0.000
		16	-0.089	-0.065	50.936	0.000
		17	0.008	0.032	50.984	0.000
		18	0.010	-0.005	51.055	0.000
		19	0.021	0.003	51.403	0.000
		20	0.081	0.079	56.432	0.000

Mean equation: Selection of an ARMA(p,q) model

At this point it is suspected that the process is both autoregressive and moving average and the order of both is not high. This conclusion comes after visualizing the ACF and PACF above. It can be clearly seen that the ACF decays notably after lag 1, and same appreciation holds for the PACF, which respectively indicates that the q and p orders of an ARMA process is around 1. Nevertheless, the general model

$$DlnP_t = c + \sum_{i=1}^{p} a_i DlnP_{t-i} + \sum_{j=1}^{q} \beta_i \varepsilon_{t-j}$$

is tried for $p = 0,2,\ldots,6$ and $q = 0,2,\ldots,6$, giving a total of 48 possible equations (excluding the trivial case such that $p = q = 0$). The most important results in terms of the selection are considered here. These are the Schwarz Information Criterion (SIC, also known by Bayesian Information Criterion BIC), the Akaike Information Criterion (AIC), the F-statistic and its p-value, and the t-values of the coefficients.

It is important to highlight that even though the first observation of the sample is that of January 1946, all the regressions performed in this report start with the observation corresponding to August, 1946. The reason is that in order to have good selection criteria, the number of usable observations should be the same across all models, and the more lags they have the more observations they lose.

The preliminary information is shown in table 3. The F-statistics and their probabilities have been omitted as all of them are highly significant even at the 1% level of significance (all p-values are less than 0.01), thus all the models are *overall* significant; i.e., have explanatory power.

Table 3: Information criteria for possible ARMA processes

ARMA(*p*,*q*)	*SIC*	*AIC*	ARMA(*p*,*q*)	SIC	AIC
0,1	-2.675269	-2.687641	3,4	-2.64102	-2.69051
0,2	-2.666521	-2.685078	3,5	-2.63237	-2.68805
0,3	-2.657759	-2.682503	3,6	-2.62421	-2.68607
0,4	-2.652991	-2.683921	4,0	-2.649503	-2.68043
0,5	-2.646067	-2.683183	4,1	-2.65425	-2.69137
0,6	-2.639948	-2.683249	4,2	-2.64661	-2.68992
1,0	-2.672436	-2.684808	4,3	-2.63702	-2.68651
1,1	-2.67	-2.68856	4,4	-2.63211	-2.68778
1,2	-2.66926	-2.694	4,5	-2.62821	-2.69007
1,3	-2.66122	-2.69215	4,6	-2.62461	-2.69266
1,4	-2.65523	-2.69234	5,0	-2.641572	-2.678688
1,5	-2.64646	-2.68976	5,1	-2.64548	-2.68878
1,6	-2.63759	-2.68708	5,2	-2.63776	-2.68725
2,0	-2.665523	-2.684081	5,3	-2.63057	-2.68624
2,1	-2.66919	-2.69394	5,4	-2.62911	-2.69097
2,2	-2.66055	-2.69148	5,5	-2.61942	-2.68746
2,3	-2.65343	-2.69054	5,6	-2.62215	-2.69638
2,4	-2.64699	-2.69029	6,0	-2.636576	-2.679878
2,5	-2.63818	-2.68766	6,1	-2.63685	-2.68633
2,6	-2.62933	-2.685	6,2	-2.629	-2.68468
3,0	-2.656677	-2.681421	6,3	-2.62479	-2.68665
3,1	-2.66033	-2.69126	6,4	-2.62505	-2.69309
3,2	-2.6523	-2.68942	6,5	-2.61828	-2.69251
3,3	-2.64935	-2.69265	6,6	-2.61018	-2.6906

The three best models according to each criterion have been coloured. Ideally both SIC and AIC would select the same models, although in the present analysis, this is not the case and big discrepancies exist. The strengths and weaknesses of each criterion, together with the p-values of the selected models' coefficients will help determine which one fits the data best. Looking at table 4, it can be seen that there is at least one process where the coefficients' p-values are greater than 0.05. As these coefficients do not contribute to the model, they will not be given further consideration. As such, the process ARMA(5,6) is discarded.

There is another reason not to consider ARMA(5,6): the process was selected only by the AIC, and this criterion is known to be biased towards selecting over-parameterized models.

After considering that the SIC values are very close to each other, the fact that the ARMA(1,1) process has the lowest p-values, and the facts that the SIC works better with larger samples and it is the best criteria to find the most parsimonious models (Enders, 2004), the ARMA(1,1) model is selected to constitute the mean equation from which the errors will be taken to form the variance equation. This decision corresponds to the fact that the GARCH(1,1) process is the most commonly used to study variables where the volatility is persistent over time (Enders, 2004). [9]

[9] Recall that a GARCH process is essentially an ARMA process applied to the conditional variance.

Table 4: Information on pre-selected ARMA processes

ARMA(*p*,*q*)	SIC	AIC	Variable	Coefficient	St. error	P-value
0,1	-2.675269	-2.687641	C MA(1)	0.005914 0.197928	0.002764 0.035935	0.0327 0.0000
1,0	-2.672436	-2.684808	C AR(1)	0.005869 0.183182	0.002829 0.036131	0.0384 0.0000
1,1	-2.67	-2.68856	C AR(1) MA(1)	0.005887 -0.517946 0.679732	0.002550 0.115981 0.099456	0.0212 0.0000 0.0000
1,2	-2.66926	-2.694	C AR(1) MA(2) MA(1)	0.005824 -0.749217 0.949533 0.112324	0.002708 0.148931 0.154795 0.056407	0.0318 0.0000 0.0000 0.0468
2,1	-2.66919	-2.69394	C AR(1) AR(1) MA(2)	0.005847 -0.617869 0.102042 0.821004	0.002760 0.119277 0.049419 0.112611	0.0345 0.0000 0.0393 0.0000
5,6	-2.62215	-2.69638	C AR(1) AR(2) AR(3) AR(4) AR(5) MA(1) MA(2) MA(3) MA(4) MA(5) MA(6)	0.005155 -0.981425 -0.163805 -0.046601 0.844785 0.820570 1.188141 0.340324 0.067380 -0.887307 -1.017516 -0.131860	0.002524 0.102194 0.176570 0.180386 0.171733 0.089172 0.107936 0.194945 0.205651 0.198438 0.118507 0.042958	0.0415 0.0000 0.3539 0.7962 0.0000 0.0000 0.0000 0.0813 0.7433 0.0000 0.0000 0.0022

Variance equation: GARCH(p,q) modelling for the complete sample

The squared errors of the ARMA(1,1) model are used to build the GARCH equation

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$$

Orders up to 3 are considered for both the ARCH and GARCH components, giving a total of nine possibilities. The first selection criterion is for those models with significant coefficients (p-value less than 0.05). After observing all the p-values, models 1,2, 1,3, 2,3, and 3,3 are discarded. The outstanding models and their main features are shown in the table 5 below:

Table 5: Information on pre-selected GARCH processes

GARCH(p,q)	SIC	AIC	Variable	Coefficient	St. error	P-value
1,1	-3.393630	-3.418245	C RESID(-1)^2 GARCH(-1)	7.15E-05 0.436138 0.732528	6.01E-06 0.030856 0.012435	0.0000 0.0000 0.0000
2,1	-3.416136	-3.446905	C RESID(-1)^2 GARCH(-1) GARCH(-2)	0.000108 0.274696 1.203109 -0.405107	6.30E-06 0.012870 0.009220 0.004464	0.0000 0.0000 0.0000 0.0000
2,2	-3.477801	-3.514723	C RESID(-1)^2 RESID(-2)^2 GARCH(-1) GARCH(-2)	3.99E-05 0.142257 0.152094 1.195607 -0.396279	7.62E-07 0.025563 0.019368 0.001939 0.001290	0.0000 0.0000 0.0000 0.0000 0.0000
3,1	-3.205802	-3.242724	C RESID(-1)^2 GARCH(-1) GARCH(-2) GARCH(-3)	0.000491 0.505157 0.564304 -0.348265 0.366174	6.03E-05 0.076642 0.066420 0.066542 0.043626	0.0000 0.0000 0.0000 0.0000 0.0000
3,2	-3.333160	-3.376236	C RESID(-1)^2 RESID(-2)^2 GARCH(-1) GARCH(-2) GARCH(-3)	0.000549 0.169538 0.596043 -0.067708 0.093496 0.291789	4.50E-05 0.034448 0.074992 0.008874 0.011343 0.038277	0.0000 0.0000 0.0000 0.0000 0.0000 0.0000

In this case it can be seen that both the SIC and AIC provide similar results about the fit of these models. Their values are lowest for GARCH(2,2), which makes this model the best one to describe volatility of the sample being studied. Some more testing is still required to prove so. A more detailed evaluation of the regression is shown in table 6 below:

Table 6: GARCH(2,2) output (for complete sample)

Dependent Variable: RESID_11	Included observations: 751 after adjustments
Method: ML - ARCH	Convergence achieved after 31 iterations
Date: 11/15/08 Time: 16:58	Variance backcast: ON
Sample (adjusted): 1946M03 2008M09	

$$h_t = \alpha_0 + \sum_{i=1}^{2} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{2} \beta_j h_{t-j}$$

Variance Equation				
Variable	Coefficient	Std. error	z-statistic	p-value
α_0	3.99E-05	7.62E-07	52.29485	0.0000
α_1	0.142257	0.025563	5.565078	0.0000
α_2	0.152094	0.019368	7.852998	0.0000
β_1	1.195607	0.001939	616.5081	0.0000
β_2	-0.396279	0.001290	-307.1789	0.0000

R-squared[10]	-0.004420	Mean dependent var	6.87E-06
Adjusted R-squared	-0.011161	S.D. dependent var	0.062874
S.E. of regression	0.063224	Akaike info criterion	-3.514723
Sum squared resid	2.977940	Schwarz criterion	-3.477801
Log likelihood	1325.778	Durbin-Watson stat	1.987486

Diagnostic tests	Statistic	Prob.
Correlogram – Q-statistic (4 lags)	0.8138	0.937
Correlogram – Q-statistic (4 lags) on squared residuals	0.4406	0.979
ARCH LM test (4 lags)	0.11157	0.97847
Wald test: $H_0 : \alpha_1 + \alpha_2 + \beta_1 + \beta_2 = 1$ (F-statistic)	115.4871	0.0000
Wald test: $H_0 : \alpha_1 + \alpha_2 + \beta_1 + \beta_2 = 1$ (Chi-square)	115.4871	0.0000

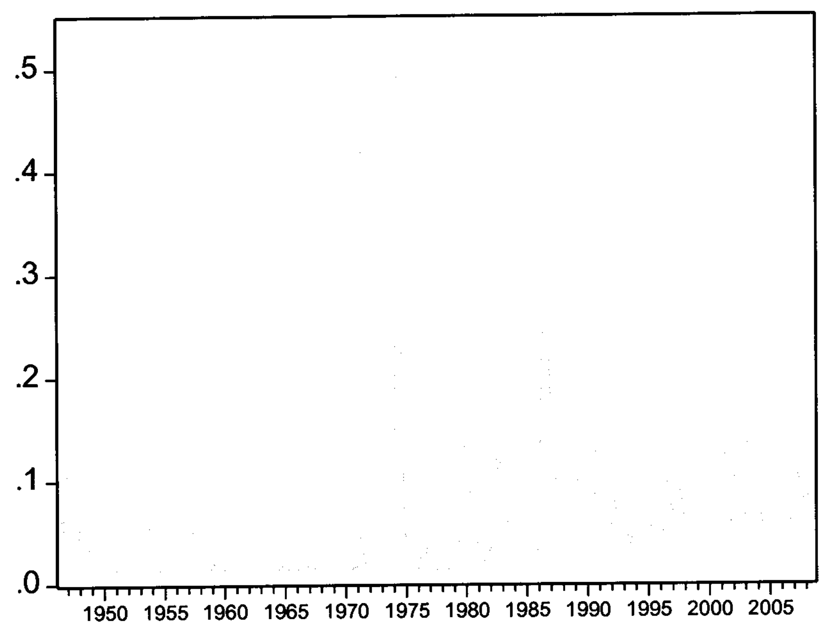
The three first diagnostic tests are on the residuals. None of them is able to reject the null hypothesis of the residuals being a white-noise process. In other words, the errors are not autocorrelated up to 4 lags, which indicates that the GARCH(2,2) model is adequate for the WTI crude oil price volatility in the period January 1946 – September 2008, using monthly frequency. Figure 8 below illustrates the result.

However, when analyzing the coefficients estimates, it can be observed that the sum of them is in fact greater than unity (around 1.09). In addition, examining the results of the Wald test, it can be seen that the null hypothesis that the sum of restricted coefficients (excluding the intercept) is equal to unity is rejected at the 1% level of

[10] R^2 can be positive in ARMA regressions. See Gray (1996)

significance. Consequently, forecasting for the long-run will make the standard deviation (the measure of the volatility) tend to infinity. This conclusion is made simply because a GARCH model is basically an ARMA process on a conditional variance, and as such the model should be stationary, otherwise the model would be explosive in the long run. As a result, the GARCH(2,2) model applied to the entire sample can only perform as a predictor in the short-run.

Figure 8: GARCH(2,2), conditional standard deviation (complete sample)



Variance equation: GARCH(p,q) modelling for a sub-sample

Many econometric studies considered structural breaks in 1980 because in the first years of that decade deregulation allowed oil prices to fluctuate freely (Karrenbrock, 1991). The Chow test was applied to test the model stability; the null hypothesis of coefficients being constant in different sub-samples was rejected at the 1%, 5%, and 10% levels of significance. This was intuitively predicted since one can, at a first glance at figure 4, discern several slopes and intercepts for different periods of tranquility / turbulence.

The purpose then is to sketch a GARCH model capable of forecasting volatility in the long term. A possibility could be to choose a sub-sample such that a) it discards the period of relatively low volatility that took place before 1973, and b) considers the oil embargo of 1973 and the price drop in 1986 (the two most important shocks to the crude oil market) outliers. The lowest price during the drop in 1986 occurred in March of that year. This sample is tentatively chosen as the first of the new sub-samples, which includes then 271 observations.

The procedure to get an optimal GARCH process is the same as the one explained in the previous subsection (and for this reason it will not be explained in great detail). The best fitting mean equation (suggested by the ACF, PACF, the information criteria, and the low p-values of the coefficients) is an AR(1) process; its errors are extracted to form the GARCH variance equation. The best fitting model for the variance equation is a GARCH(1,1) process, and the regression output is shown in table 7 below.[11]

[11] The most relevant steps carried out to get this result can be seen in the appendix of this paper.

Table 7: GARCH(1,1) output (for 1986M03-2008-M09 sub-sample)

Dependent Variable: RESID_10	Included observations: 271
Date: 11/16/08 Time: 22:13	Convergence achieved after 15 iterations
Sample: 1986M03 2008M09	Variance backcast: ON
Method: ML - ARCH (Marquardt) - Normal distribution	
$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$	

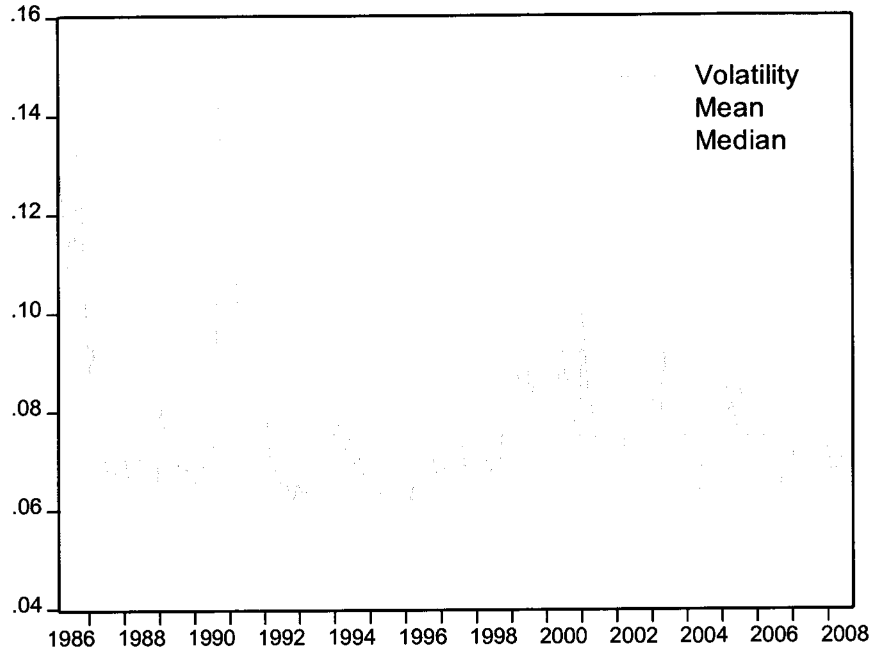
Variance Equation				
Variable	Coefficient	Std. error	z-statistic	p-value
α_0	0.001075	0.000593	1.812010	0.0700
α_1	0.130682	0.059927	2.180689	0.0292
β_1	0.682567	0.129802	5.258514	0.0000

R-squared	-0.000945	Mean dependent var	1.34E-14
Adjusted R-squared	-0.019689	S.D. dependent var	0.080808
S.E. of regression	0.081600	Akaike info criterion	-2.303056
Sum squared resid	1.777826	Schwarz criterion	-2.223727
Log likelihood	320.3672	Durbin-Watson stat	1.949915

Diagnostic tests	Statistic	Prob.
Correlogram – Q-statistic (4 lags)	4.5296	0.339
Correlogram – Q-statistic (4 lags) on squared residuals	1.4439	0.837
ARCH LM test (4 lags)	0.5562	0.8631
Wald test: $H_0 : \alpha_1 + \beta_1 = 1$ (F-statistic)	2.9347	0.0879
Wald test: $H_0 : \alpha_1 + \beta_1 = 1$ (Chi-square)	2.9347	0.0867

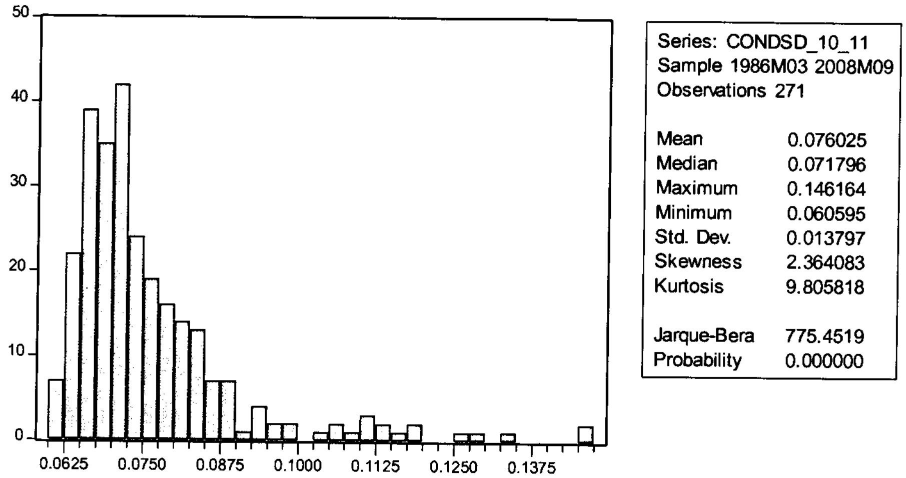
The conclusions are similar to those from the GARCH(2,2) model obtained in the previous subsection: the Q-tests and the ARCH-LM test reject the null hypothesis of errors being correlated. Therefore there is a strong indication that errors of the GARCH model are a white-noise process. The Wald test also rejects the null hypothesis of restricted coefficients such that the sum of them is equal to unity at a 10% level of significance but cannot reject it at lower levels. However, a big difference is noticed when comparing the coefficients of the previously sketched GARCH(2,2) with the present GARCH(1,1). The sum of the slope coefficients of the former was greater than one, whereas the sum of those of the latter is equal to 0.1897 < 1. The findings described in this paragraph permit the conclusion that the model is stationary, and adequate to work with good predictive power in both short and long-run. The conditional standard deviation is shown in figure 9 below.

Figure 9: GARCH(1,1), conditional standard deviation (1986-2008 sub-sample)



Extracting the dependent (estimated) variables (i.e. the estimated conditional variances) and taking the square root permits the performance of some descriptive analysis on the conditional standard deviation. The average volatility registered in the period March 1986 to September 2008 is 7.60%. The positive skew means that the mean is greater than the median, and the median is greater than the mode. The longer right tails can clearly be seen in figure 10 below, as well as the longer than normal tails evidenced by a kurtosis of 9.805818 > 3. This result is not surprising; it is known that the mean tends to be dragged by extreme observations, and figure 10 clearly shows several observations on the extreme right, with the two on the extreme likely representing the shocks observed in 1986 and 1991.

Figure 10: descriptive statistics of the volatility (for the sub-sample)



After performing these econometric regressions, it is reasonable to conclude that a sample as large as the one initially proposed for this report will produce results with limited relevance for policy-making. As explained in previous sections, the last fifty years were characterized by periods of relative stability, and others very volatile, consequence of the oil embargo in 1973, the oil price drop in 1986, conflicts in Middle East, terrorist attacks, speculation, etc. A shorter sample able to capture the manner in which oil markets currently work will produce more effective results, namely the ability to produce forecasts with greater accuracy.

6. Further discussion - Oil price volatility influencing policy-making

This report exclusively used monthly data, while several studies mentioned in previous sections have estimated GARCH models using different frequencies. Though the outcomes may differ, they remain consistent with the theory in that the resulting measures of volatility are a function of the sample frequency (Engle and Patton, 2001).

The findings of this report, especially those obtained using the narrower sample 1986-2006, suggest that in the long-run the crude oil market will be volatile. It is expected that prices will increase and not return to the expected mean because of strong positive drifts, as found by Krichene (2006); his findings using the sub-sample 1986-2008 are able to predict prices returning to the expected mean, although it does not prevent the process from being jumpy. The same conclusion about a future of high volatility was made by Askari and Krichene (2008), who carried out a similar analysis but used a daily time-series of 1130 observations of *future* oil prices from January 02, 2002 to July 07, 2006. They do not however see prices able to return to the mean, but rather higher rises than falls. The effectiveness of the GARCH(1,1) model is limited; future research should always consider the most recent available data and the economic, financial, and social effects that affect data in order to obtain the GARCH model with best fit, able to produce more accurate forecasts for future volatility.

Having found a univariate model able to show crude oil price volatility does not prevent volatility from representing an obstacle to policy-making. The sensitivity of the oil markets to shocks (obviously caused by variables not considered in this model), the potential asymmetry between response to "good news" and "bad news", and the permanence of the volatility after a shock are problematic issues. The latter can be visualized by observing the thickness of the "waves" of figure 9, as presented above; some cases, such as the invasion of Iraq, produced a shock in the markets that increased the conditional volatility from about 7% (very close to its median value) to almost 15%; it took about one year for the volatility to go back to the levels before the invasion.

Imagine a policy-maker considering, for instance, setting up a tax rate on a given good whose price is greatly affected by shocks in crude oil markets. An interesting case could be an "environmental" tax, such as the Carbon Tax that has been instituted in British Columbia. [12] This is a revenue-neutral tax, in that the objective is not to collect money from the tax payers, but to make pollutant-producing goods less attractive, by making them more expensive. Both the price and tax work in the same direction, meaning that an increase on either would reduce the desire to purchase a polluting good. Assuming a Pigouvian tax rate that produces a situation of equilibrium in the emission levels is attained, that equilibrium would be lost if prices change. The government should then alter the rate to achieve equilibrium again. The immediate question is: how fast can the government react in the real world? And thinking of the goods that produce emissions, the other question is: how sensitive are they to shocks in the crude oil market?

The next two subsections will analyze these matters by reviewing the recent literature about a) the relationship between the prices of crude oil and gasoline, and b) inflationary effects of oil price shocks.

Relationship between oil price and gasoline price

Although one might think that oil prices and gasoline prices are highly correlated (for instance oil price changes have an instantaneous and proportional effect on gasoline prices), some literature has studied this relationship in detail, arriving at interesting results. Karrenbrock (1991) established three measures of asymmetry between the prices of oil and gasoline: the speed at which one price affects the other, the proportion of the oil price change that is passed to the final consumer, and a combination

[12] Information on British Columbia Carbon Tax can be found at the official Provincial website http://www.sbr.gov.bc.ca/individuals/Consumer_Taxes/Carbon_Tax/carbon_tax.htm

of both. The report was then able to demonstrate that gasoline prices respond more quickly to oil price increases than to oil price decreases. It was also noted that the gasoline consumer bears the weight of oil prices in a greater degree when the oil price goes up than when it goes down. The response of gasoline to oil in term of prices is therefore asymmetric.

Godby et al. (2000) however, were not able to find such an asymmetric response in the Canadian market. Although their sample period is shorter than that of Karrenbrock (1991) and the observations are for Canadian provinces, a threshold autoregressive model (TAR) tested the null hypothesis of a symmetric response, allowing the possibility that such a threshold is positive rather than zero, as previously assumed. The hypothesis was not rejected, and therefore they find no concrete reasons to call responses from gasoline prices to oil prices asymmetric.

In a more recent report, Radchenko (2005) put forward the concept of the *oligopolistic coordination theory*, and explains that according to it, "an increase in the price volatility leads to a faster response of gasoline prices to an oil price decrease and a reduction in the degree of asymmetry in the gasoline price response". The finding is new in the sense that the degree to which the gasoline price responds to oil price variations is based on the current volatility, and not necessarily based on the sign of such variations. Asymmetry of prices and volatility hold a robust negative relationship and the oligopolistic coordination theory is a plausible explanation for it. It is of interest to highlight that the models chosen to measure the response of gasoline prices are a vector autoregressive model (VAR), and partial adjustment model (PAM); as for proxies for oil price volatility, rolling standard deviations and a GARCH (1,1) model are preferred.

Relationship between oil price and other goods - Inflation

Regnier (2007) compared the volatility of crude oil with products manufactured in the U.S. in general, and crude commodities in particular. The generally accepted idea is that the crude oil shocks have indeed affected other commodities, especially energy-related commodities. The evidence suggests that the volatility suffered by these products after the 1973 shock for instance, was not as high as that experienced by crude oil. By the end of that decade, the volatility of U.S. products' prices had returned to its 1970s level, as opposed to that of crude oil, which remained high and increasing. A similar observation applies to the events of the 1980s. She noted, however, that crude oil was less volatile than other energy commodities until 1986, contradicting the general (consumer) perception.

An interesting approach to this matter is to see how prices in general reacted to shocks in oil markets; in other words, one can ask whether inflationary processes (generally observed by analyzing the CPI) were accelerated, indifferent or retarded by oil shocks.

The general belief is that oil price shocks increase inflation, and this is probably based on the inflationary spiral that took place after the events of 1973 drastically increased oil prices. LeBlanc and Chinn (2004) argued that the acceleration of the inflationary process was caused essentially by high oil intensity at that time. As the oil intensity was reduced over the years with the introduction of new technologies, the relationship between oil price shocks and inflation was reduced. Empirical examination of data from the U.S., Europe, and Japan supports their argument. In fact, they explain that what happened after 1973 is seen by many economists as a rearrangement of relative prices, rather than inflation. They conclude that later oil price shocks had modest incidence on inflation in the analyzed countries. The idea is shared by Trehan (2005), who highlighted the importance of the *expected* inflation after 1973, and the consequent monetary policy. Applying empirical data for January 01 to October 17 (2005), he finds that markets did not expect inflationary shocks in that period, despite the fact that WTI crude oil increased its price about 50% of the value at the beginning of that year. Recently, Cavallo (2008) studied core inflation by estimating a Phillips curve model, and also found that the impacts of crude price shocks were low in Europe, and even less significant in the U.S., the U.K., and Canada, attributing this to successful policy-making to procure a low inflation environment.

Caveats for policy-makers

The last two subsections and the empirical findings of this report represent a caveat for policy-makers. The general question is: do crude oil prices ever return to their mean? Considering only recent years (as this report did by using data from 1986), the findings suggest that eventually oil prices would have the ability to return to their *trend*, but that does not prevent the process from being highly volatile.

Volatility should be considered as a separate issue from price increases. They have been shown to be able to exist independently; observing figure 3 and figure 9, one can visualize for instance the events of 1986 that ended in a crude oil price drop with high volatility, whereas the events of 2001 increased the price, yet with volatility measures near their mean.

Revisiting the example of a zero-revenue environmental tax, a policy-maker should not only consider price increases in crude oil (they could eventually follow a trend); she will also –and perhaps more importantly– consider how volatile prices will be before they go back to equilibrium, if that in fact happens. A tax implies setting up a rate, and that rate should be able to follow these adjustments in oil markets. If long periods of volatility and uncertainty are forecast, it would be a good idea to consider alternative environmental policies, such as emission permits. This would be worthwhile as the price of these permits would immediately be adjusted to the permit's market, so that any shock on oil prices would be absorbed by such capacity.

Further research on volatility

This paper has used GARCH techniques to obtain models able to forecast the oil price volatility. The selected method was a univariate GARCH process, which may have arguable effectiveness due to the lack of many other factors that can affect the oil price volatility (recall that ARMA models consider only past realizations and errors as explanatory, plus a white-noise disturbance term). A more complex model could improve the forecasting power by incorporating other explanatory variables like inflation indexes or CPI, parameters of economic activity levels, terms of trade, etc., and using dummy variables to better represent the slope and intercept of the mean equation, and stress the influence of presidential changes, significant historical events, terrorist attacks, wars, revolutions, changes in monetary/fiscal policies, etc. that can work as structural breaks in the model. It should be reiterated also that the estimates of any regression are a consequence of the observed data, its length, and its frequency. In periods of high volatility like the one crude oil is presently experiencing, constant monitoring of new data is essential to continue enriching the model and its forecasting power.

7. Conclusion

This report has carried out different GARCH models to analyze the volatility of crude oil prices. A GARCH process is chosen under the assumption that the past volatility affects the present *conditional* variance (taken as a measure of volatility). The best fitting model for the complete monthly series (January 1946 to September 2008) on WTI crude oil spot prices is a GARCH(2,2) model, which can forecast efficiently in the short-run. For a sub-sample March 1986 – September 2008 the preferred model is GARCH(1,1), capable of predicting in both the short and long-run. These results correspond with the intuition that the complete sample (which includes about 27 years of stability, followed by 35 years of high volatility) will constitute an explosive model, meaning that a long-run forecast variance will tend to infinity.

The model obtained for the "volatile" sub-sample gives a monthly average volatility of about 7.6%. This unquestionably high level of volatility raises questions about how effectively a policy (for instance an environmental tax based on the price of gasoline and other pollutant goods related to crude oil), can follow such an unstable process.

8. References

Adelman, M. A., 2004. "The Real Oil Problem", Massachusetts Institute of Technology.

Askari, Hossein, and Noureddine Krichene, 2008. "Oil price dynamics (2002-2006)", *Energy Economics*, 30(5), 2134-2153.

Bernanke, Ben S., 1983. "Irreversibility, Uncertainty, and Cyclical Investment", *The Quarterly Journal of Economics*, 97(1), 85-106.

Bollerslev, Tim, 1986. "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 31(3), 307-327.

Brook, Anne-Marie, Robert Price, Douglas Sutherland, Niels Westerlund, and Christophe André, 2004. "Oil Price Developments: Drivers, Economic Consequences and Policy Responses", *Economics Department Working Papers*, No. 412, Organisation for Economic Co-operation and Development (OECD). Available on-line at http://lysander.sourceoecd.org/vl=3657251/cl=13/nw=1/rpsv/workingpapers/18151973/wp_5lgsjhvj75hk.htm. Date of access: October, 2008

Castillo, Paul, Carlos Montoro, and Vicente Tuesta, 2007. "Inflation Premium and Oil Price Volatility", *Centre for Economic Performance*, No 782.

Cavallo, Michele, 2008. "Oil Prices and Inflation", *FRBSF Economic Letter*, No. 2008-31.

Claessnes, Stijn, and Panos Varangis, 1994. "Oil Price Instability, Hedging, and an Oil Stabilization Fund", *Policy Research* Working Paper No. 1290, The World Bank – International Economics Department.

Costello, Alexandra, Ebenezer Asem, and Eldon Gardner, 2008. "Comparison of Historically Simulated VaR: Evidence from oil Prices", *Energy Economics*, 30(5), 2154-2166.

Denni, Mario, and Geoffrey Frewer, 2006. "New Evidence on the Relationship between Crude Oil and Petroleum Product Prices", Working Paper No. 61, Università degli Studi Roma Tre - Dipartimento di Economia.

Dickey, D. A. and W. A. Fuller, 1979. "Distribution of the Estimators for Autoregressive Time-series with a Unit Root", *Journal of the American Statistical Association*, 74, 427-431.

Elder, John, and Apostolos Serletis, 2006. "Oil Price Uncertainty", College of Business, North Dakota State University: Fargo, ND.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock, 1996. "Efficient Tests for an Autoregressive Unit Root", *Econometrica*, 64(4), 813-836.

Enders, Walter, 2004. *Applied Econometric Time-series*, Second edition, Hoboken, NJ: John Wiley & Sons, Inc.

Engel, Eduardo, and Rodrigo Valdés, 2000. "Optimal Fiscal Strategy for Oil Exporting Countries", *International Monetary Fund – Middle East Department*, IMF Working Paper No 118.

Engle, Robert F., 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation," *Econometrica*, 50(4), 987-1008.

Engle, Robert F., and Andrew J. Patton, 2001. "What Good is a Volatility Model?", Department of Finance, NYU Stern School of Business, and Department of Economics, University of California: La Jolla, CA.

Fan, Ying, Yue-Jun Zhang, Hsien-Tang Tsai, and Yi-Ming Wei, , 2008. "Estimating 'Value at Risk' of Crude Oil Price and its Spillover Effect using the GED-GARCH Approach", *Energy Economics*, 30(6), 3156–3171.

Fattouh, Bassam, 2005. "The Causes of Oil Price Volatility", *Middle East Economic Survey*, 48:13.

Fong, Wai Mun, and Kim Hock See, 2002. "A Markov Switching Model of the Conditional Volatility of Crude Oil Futures Prices", *Energy Economics*, 24(1), 71-95.

Godby, Rob, Anastasia M. Lintner, Thanasis Stengos, and Bo Wandschneider, 2000. "Testing for asymmetric pricing in the Canadian retail gasoline market", *Energy Economics*, 22(3), 349-368.

Gray, Stephen F., 1996. "Modeling the conditional distribution of interest rates as a regime-switching process", *Journal of Financial Economics*, 42(1), 27-62.

Greene, William H., 2008. *Econometric Analysis*, sixth edition, Upper Saddle River, NJ: Pearson Education, Inc.

Hamilton, James D., 1983. "Oil and the Macroeconomy since World War II", *Journal of Political Economy*, 91(2), 228-248.

Hooker, Mark A., 1996. "What happened to the oil price-macroeconomy relationship?", *Journal of Monetary Economics*, 38(2), 195-213.

Karrenbrock, Jeffrey D., July, 1991. "The Behavior of Retail Gasoline Prices: Symmetric or Not?", *Journal of the Federal Reserve Bank of St. Louis*, 19-29.

Krichene, Noureddine, 2006. "Recent Dynamics of Crude Oil Prices", *International Monetary Fund – African Department*, IMF Working Paper No 299.

Kuper, Gerard H., 2002. "Measuring oil price volatility", University of Groningen, Research Institute *SOM* Research Report No. 02C43.

Kuper, Gerard H., and Daan P. van Soest, 2003. "Does oil price uncertainty affect energy use?", University of Groningen, *CCSO Centre for Economic Research*, CCSO Working Paper No. 200306

LeBlanc, Michael, and Menzie D. Chinn, 2004. "Do High Oil Prices Presage Inflation? The Evidence from G-5 Countries", *Santa Cruz Center for International Economics (University of California, Santa Cruz)*, paper 04'04.

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

Levy Yeyati, Eduardo, 1996. "Investment in Inflationary Economies", *International Monetary Fund – Western Hemisphere Department*, IMF Working Paper No 105.

Mork, Knut Anton, 1989. "Oil and the Macroeconomy When Prices Go Up and Down", *Journal of Political Economy*, 97(3), 740-744.

Nwogugu, Michael, 2006a. "Further Critique of GARCH/ARMA/VAR/EVT Stochastic-Volatility Models and Related Approaches", *Applied Mathematics and Computation*, 182(2), 1735-1748.

Nwogugu, Michael, 2006b. "Volatility, Risk Modeling and Utility", *Applied Mathematics and Computation*, 182(2), 1749-1754

Pagan, Adrian R., and G. William Schwert, 1990. "Alternative models for conditional stock volatility", *Journal of Econometrics*, 45(1-2), 267-290.

Perron, Pierre, 1989. "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis", *Econometrica*, 57(6), 1361-1401.

Phillips, P. C. B., and P. Perron, 1988. "Testing for a Unit Root in Time-series Regressions", *Biometrika*, 75(2), 335-346.

Radchenko, Stanislav, 2005. "Oil price volatility and the asymmetric response of gasoline prices to oil price increases and decreases", *Energy Economics*, 27(5), 708-730.

Regnier, Eva, 2007. "Oil and energy price volatility", *Energy Economics*, 29(3), 405-427.

Slade, Margaret E., 1991. "Market Structure, Marketing Method, and Price Instability", *Quarterly Journal of Economics*, 106(4), 1309-1340.

Trehan, Bharat, 2005. "Oil Price Shocks and Inflation", *FRBSF Economic Letter*, No. 2005-28.

Appendix

Table 8: Regression output for the AR(1) process (reduced sample)

Dependent Variable: DLOGP
Method: Least Squares
Date: 12/10/08 Time: 15:24
Sample: 1986M03 2008M09
Included observations: 271
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.007322	0.006036	1.213122	0.2261
AR(1)	0.217605	0.056986	3.818555	0.0002

R-squared	0.051419	Mean dependent var	0.007034
Adjusted R-squared	0.047892	S.D. dependent var	0.079664
S.E. of regression	0.077733	Akaike info criterion	-2.263715
Sum squared resid	1.625421	Schwarz criterion	-2.237131
Log likelihood	308.7334	F-statistic	14.58136
Durbin-Watson stat	1.984583	Prob(F-statistic)	0.000167
Inverted AR Roots	.22		

Figure 11: Correlogram for the AR(1) process (reduced sample)

Date: 12/10/08 Time: 15:34
Sample: 1986M03 2008M09
Included observations: 271
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.007	-0.007	0.0147	
		2	-0.120	-0.120	3.9907	0.046
		3	0.042	0.040	4.4698	0.107
		4	-0.078	-0.093	6.1546	0.104
		5	-0.057	-0.049	7.0733	0.132
		6	-0.051	-0.076	7.7882	0.168

Table 9: Regression output for the GARCH(1,1) process (reduced sample)

Dependent Variable: RESID_10
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 12/10/08 Time: 15:24
Sample: 1986M03 2008M09
Included observations: 271
Convergence achieved after 15 iterations
Variance backcast: ON
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.001134	0.004469	-0.253714	0.7997
	Variance Equation			
C	0.001075	0.000593	1.812010	0.0700
RESID(-1)^2	0.130682	0.059927	2.180689	0.0292
GARCH(-1)	0.682567	0.129802	5.258514	0.0000

R-squared	-0.000214	Mean dependent var	2.29E-16
Adjusted R-squared	-0.011453	S.D. dependent var	0.077589
S.E. of regression	0.078032	Akaike info criterion	-2.306989
Sum squared resid	1.625769	Schwarz criterion	-2.253822
Log likelihood	316.5971	Durbin-Watson stat	1.984157

Note: The information in this table is similar to that of table 7. In this case, the GARCH(1,1) model is preferred not only because both AIC and SIC define it as the one with best fit, but also because all the other models' outputs returned coefficients with high p-values; i.e., individuals t-tests cannot reject the null hypotheses of slope coefficients being equal to zero at 5% level of signifficance. Those models have been discarded.