ERASMUS SCHOOL OF
ECONOMICS

Erasmus
University
Rotterdam

Master Thesis

On the predictability of WTI crude oil futures volatility

A two-state switching-volatility model based on fundamentals and technics.

Johannes Matthijs Begeer
415373

A Master Thesis in the Social Sciences
Discipline of Economics

Submitted to Erasmus University School of Economics
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE IN ECONOMICS AND BUSINESS ECONOMICS

MAY 2016

Abstract

This research investigates whether a certain volatility state on the WTI market for crude oil, either the current one or one in the near-future, can be explained using a switching-regression model. Focusing on weekly data over 2000-2015, four fundamental factors - the S&P 500 index, the rolling-month NYMEX gold future, the Baltic Dry Index, and the US Dollar Index – and the rolling-year WTI spread, the more technical factor, are used as explanatory parameters. While separate analyses of the impact of fundamental or technical factors on the crude oil price respectively have been performed before, this research is, to the author's best knowledge, the first in combining both types of factors into one model related to crude oil volatility. Based on the results of this research, the following can be concluded: simultaneous changes in simultaneous deviations in the rolling-year WTI futures spread, S&P 500 and US Dollar Index seem to increase the probability for a certain WTI volatility state, both the current one as well as the one for next week. For the current state, a simultaneous absolute deviation in these three parameters larger than 2.74 percent significantly increases the chance of a volatility switch in the current period. Furthermore, a simultaneous return larger than 3.10 percent now will significantly increase the chance of a volatility switch in the next period. These results could significantly improve risk management on the crude oil market. The outcome from the lagged model is particularly useful as it provides market participants with the probability of risks affecting their investments in the near-future based on the current developments on connected markets.

Table of content

Figures and tables

Figures

Tables

Introduction

By the time this research started, around January 2016, WTI crude oil prices were settling at decade-low prices for front-month delivery; the result of a steady price decline, which started from a post-financial crisis high in 2014. Since then, in just a few months' time, prices have increased by no less than 80 percent – with analysts' expectations still remaining bullish.

During the 21st century, the market for crude oil has experienced multiple of these price swings.[1] These developments have a significant impact within all levels of the economy, which makes the topic of this research, identifying certain factors which increase or decrease the probability of a particular volatility state on the crude oil market, highly relevant. Fundamental factors such as the supply-demand relationship, or market factors like the gold price, the state of the world economy, and the stock market could be responsible. On the other hand, the financialization of the crude oil market through futures contracts has introduced speculative behaviour to the market, driving prices away from their fundamentals. Analysing these seemingly irrational pricing patterns could also provide useful information.

While separate analyses of the impact of fundamental or technical factors on the crude oil price respectively have been performed before, this research is, to the author's best knowledge, the first in combining both types of factors into one model related to crude oil volatility. For the period between 2000 and 2015, this research is designed to answer the question as to what drives oil prices movements in the short-term. Specifically:

Is a combination of fundamental and technical factors able to predict, or at least explain, a certain state of volatility on the WTI crude oil market, either the current or one in a future period?

Four fundamental factors - the S&P 500 index, the rolling-month NYMEX gold future, the Baltic Dry Index, and the US Dollar Index – and the rolling-year WTI spread, the more technical factor, will be used as explanatory parameters. Here, the combined impact of the spread, the S&P 500 and the US Dollar Index on the WTI volatility state is expected to be the most significant. As a proxy for this volatility, the 5-day forthcoming standard deviation in returns of the rolling-month WTI futures contract will be used.

As mentioned earlier, the covered period is one with significant high and lows, both for the input parameters as well as in terms of the price and volatility of the WTI crude oil contracts. Hence, one could expect structural breaks in the relations between the aforementioned factors and crude oil. In order to account for this, a two-state volatility switching regression model will be used. The two states describe

---

[1] For instance, between the start of 2004 and mid-2008, prices more than quadrupled from $32 to $145, while December of that same year settled at $37. This was followed by an average price of around $100 over 2011-2014, before settling in 2015 close to $35.

a high- and a low-volatility state, where probabilities are attached to the impact of deviations in the earlier-mentioned factors on the current and future state of WTI volatility. Essentially two models are estimated, one for the current state of the volatility based on current deviations in the assessed input factors and another the current impact on WTI's second moment resulting from previous period's changes in fundamentals and the rolling-year spread. The latter result could be particularly useful as it provides market participants with risks affecting their investments in the near-future.

Based on the results the following can be concluded:

- Simultaneous changes in BDI or the rolling-month NYMEX gold future, as part of a combined model with the rolling-year WTI spread, the US Dollar Index and the S&P 500, do not seem to have a statistically significant impact on the volatility state of the WTI rolling-month future, neither the current one nor the one for next period.

- On the other hand, simultaneous deviations in the rolling-year WTI futures spread, S&P 500 and US Dollar Index seem to increase the probability for a certain WTI volatility state, both the current one as well as the one for next week.

- For the current state, a simultaneous absolute deviation in these three parameters larger than 2.74 percent significantly increases the chance of a volatility switch in the current period. Furthermore, a simultaneous return larger than 3.10 percent now will significantly increase the chance of a volatility switch in the next period. These measures are at least significant at the ten-percent level, most even at the one-percent level.

This research has the following structure: the following section provides a comprehensive overview of the existing literature on the inefficiencies in the crude oil market which form the basis for this research, the theoretical relations between the aforementioned fundamental and technical factors and the crude oil market, and the appropriate methods to forecast and analyse volatility. The second section describes the data, section 3 the methodology and in section 4 the main results will be analysed.

Literature review

As mentioned earlier, the focus of this research is contributing towards a better understanding of the price dynamics on the crude oil market, developments in crude oil's volatility in particular. Specifically, if and how historical and current data can be used in order to effectively predict future prices movements. Bettman et al. (2009), among others, emphasize how finance professionals, both academic and commercial, roughly rely on two approaches towards tackling this type of research problem: an analysis based on underlying fundamentals which influence the instrument's dynamics: *the fundamental approach,* or, in apparent contrast, one which assumes the explanatory and predictive power of the asset's characteristics[2] itself: *the technical approach.*

The limitations of Nobel laureate Fama (1965)'s well-known efficient market hypothesis are the point where fundamentalists and technicians meet each other, as will become apparent below. The proposition of the efficient market hypothesis in its strongest form (Fama, 1991) has significant associations with the random walk theory; the idea that the price of an asset, at any point in time, is as likely to decrease as it is to increase (Samuelson, 1965). Combined, their statement is as follows: if markets are efficient, the flow of information unhampered and new information immediately reflected in asset prices, then tomorrow's price change is a mere reflection of tomorrow's news and thus unaffected by today's variations. Given the randomness and unpredictability of news, price changes therefore must be a purely random and unpredictable process as well (Malkiel, 2003). As a consequence, any research or analysis, on finding unbiased predictors for futures prices, including this one, is pointless, since this is based on the absence of total randomness and hence the very notion of market inefficiencies and some predictability (Harrington, 2003). Or, as Malkiel (1973) puts it: "a blindfolded chimpanzee throwing darts at the Wall Street Journal could select a portfolio that would do as well as the experts."

While the strong-form efficient market hypothesis as described above may hold for certain markets, the proposition that markets, and their participants, are efficient and unpredictable by nature has been contested heavily (e.g. Summers, 1986; Lo and MacKinlay, 1999; Zunino et al., 2010). Fama seemingly (1991) acknowledges this by roughly subdividing the aforementioned efficient market hypothesis in three categories: strong – the one discussed above -, medium, and weak. Under the weak form, all past price movements have already been incorporated in current prices, making technical analysis redundant. The same applies for fundamental analysis under medium-form efficiency where all publicly available information is priced-in as well. The strong form even prevents profitable insider trading, since under this presumed level of market efficiency even private information is reflected in the market prices next to historical and public information. Hence, for this research it is of great relevance to get a better understanding of the relative efficiency and predictability of the oil market, since this determines the most effective approach towards tackling the aforementioned research problem.

---

[2] For instance, price patterns, whether a forward premium exists, or the market is in contango or backwardation.

First of all, what are these so-called inefficiencies and how do they arise? Peters (2003) distinguishes between rather *simple* and more *complex* inefficiencies. The former type is mainly based on a market participant having information which no one else has. The value of this information should be readily apparent to the participant and allow him to retrieve a risk-free return; in the sense that it creates a short-term momentum which lasts until the information gets public and arbitrage removes the opportunity. Kahneman and Tversky (1973) indeed show that people tend to react uniformly to an unambiguous situation like the one described above, hence the medium-strong efficient market hypothesis is expected to hold for simple inefficiencies: in the long run such an inefficiency should be eliminated. Consequently, this is the type of inefficiency which is used as proof by proponents of the efficient market hypothesis (e.g. Malkiel, 2003). On the contrary, one can observe inefficiencies which are more complex in nature. The basis for this type of inefficiencies is the subjective interpretation of publicly available information, which is mainly caused by heterogeneity, and hence disagreements, among market actors (Peters, 2003). This leads to situations where actors take ambiguous factors as facts, or objective circumstances as ambiguous; inefficient behaviour. In terms of characteristics these inefficiencies are the total opposite of simple ones: they are unable to produce risk-free returns, cannot be arbitraged away and therefore, in general, are long-term in nature (Dorner, 1996). Ergo, this is the type of inefficiency which provides arguments to opponents of the efficient market hypothesis.

Using the opposite of the aforementioned theory of Samuelson (1965), the inefficiencies described above could give rise to a certain level of predictability (Eom et al., 2008). A significant amount of research has been dedicated to other common predictable asset features. Mandelbrot (1963), for instance, was the first to show that small moves and large moves (of either sign) in financial asset returns tend to form *clusters*. Hence, a volatility move today will largely influence - expectations of - volatility for multiple future time periods. The bandwagon effect (Shiller, 2000; Kaizoji, 2000; Brown & Cliff, 2004) amplifies this proposition. Its argument, in short, is that asset returns predict future sentiment. As a consequence, increased bullishness (bearishness) can be expected after a market rise (fall). Verma and Verma (2007) reify this statement with their main findings that this effect is assumed to be greater (smaller) when the market is in contango (backwardation), and that the effect can be found for individual investors but not for institutional ones. Their second finding may be explained by the fact that institutional investors have superior knowledge over individual investors. In other words: not all information is public information and individuals seem to follow institutional investor sentiment and not vice versa (Verma & Verma, 2007). The clustering of an asset's second moments also implies that a certain period of high (low) volatility will be followed by a respective rise (fall) in volatility. According to Engle & Patton (2001), and in line with the preceding statement, there is a trend of mean reversion in volatility, meaning there is a normal level of volatility to which the long-run volatility level should always return.

The inefficiencies and predictable patterns described above are mainly related to the behaviour of the financial assets or the market actors rather than to the markets itself. Hence, these observed

patterns should be found on most financial asset markets where these actors are active, and variations in levels of efficiency between markets should be mainly due to differences between these markets' characteristics. For instance, it has been shown that emerging markets operate at lower efficiency levels compared to more developed ones (Bęben & Orlowski, 2001; Di Matteo et al., 2003). Knowing this, what does empirical evidence say about the predictability of crude oil markets, the WTI market in particular?

A significant amount of research has been dedicated to assessing the short-term efficiency of crude oil markets. Alvarez-Ramirez et al. (2010), for instance, using detrended fluctuation analysis (DFA) to account for non-stationarity, find too many deviations from random walk behaviour on the WTI spot market, thus emphasizing its short-term inefficiency. Their results are supported by Wang et al. (2011), following the same statistical method, who argue that WTI spot prices are not efficient when the considered time horizon is less than a month; Alvarez-Ramirez et al. (2002) even scale this down to an order of days to weeks. Also, a series of papers has been focusing on long-term crude oil market efficiency. Alvarez-Ramirez et al. (2008), although finding evidence of time-varying autocorrelation for one-month horizons rejecting the random walk hypothesis, find results consistent with the efficient market hypothesis for long time horizons. This last result is supported by Serletis and Andreadis (2004) for daily observations of WTI spot prices over 1990-2001. Lean et al. (2010), using a stochastic dominance approach, robustly find no arbitrage opportunities between WTI spot and futures prices. Tabak and Cajueiro (2007), who find weak-form efficiency for Brent oil but no efficiency for WTI and attribute this to different levels of regulation on both markets; less regulation being associated with higher levels of efficiency, with a visible trend of deregulation. Their results, however, following (Lo, 1991) are prone to biasedness since the R/S method used is relatively trend-sensitive to short-term dependencies. Wang and Liu (2010), though, accounting for non-stationarity through DFA, also conclude that crude oil markets have become more efficient over time. Several authors, in turn, have linked the crude oil market to other markets. Narayan et al. (2010), for instance, emphasize the relationship between the markets for gold and crude oil. Since investors hedge their investments through the gold market, and the oil market presumably predicts the gold market and vice versa, this should lead to inefficiencies on both markets. Kristoufek and Vosvrda (2014) find that commodity futures in general are predictable in the short-term but return over the long-run to their fundamental value, which differs from earlier-found results on stocks indices (Kristoufek and Vosvrda, 2013).

Perhaps a proposition in Graham and Dodd (1965) is the best summary of the above: market prices may be a voting system in the short run, but a weighing system in the long run; true value will win out in the end. A preliminary conclusion at this point should be that the crude oil market seems to be relatively predictable, but empirical evidence also reveals time-varying efficiency. Given this, could fundamental or technical analyses, or maybe a combination, then be able to forecast this "true value?"

Fundamental analysis

The first section provided a rather brief description of the fundamental approach, as it mainly served to create ground for this type of analysis. The foremost objective of any fundamental research related to finance is identifying factors driving the price of a certain asset, and possibly using these in order to forecast its future price movements (e.g. Wafi et al., 2015). The underlying assumption being made here is that past and current observations have at least some predictive content, which can only be true if the aforementioned random walk hypothesis does not hold. These factors could range from obvious traditional ones such as supply and demand to more undisclosed relationships such as the aforementioned co-movement between prices for gold and crude oil resulting from hedging activities (Narayan et al., 2010). The respective contribution to the asset price of each factor is expected to vary in absolute terms, but could also fluctuate over time (Fleten et al., 2015; Huisman & Kiliç, 2015). The latter feature in particular is of utmost importance to this type of analysis, since an assumption of constant contribution over time of each factor could lead to significantly incorrect predictions if fundamentals exhibit time-varying significance. Related to the crude oil market, for instance, it has been argued that, as a result of globalization and financialization of the crude oil market since the early 2000s, market prices are no longer only determined by forces of supply and demand but rather by a range of fundamentals as a financial concept (Fan & Xu, 2011). A more recent development on the crude oil market is the significant drop in the production costs of shale oil, which has led to a sharp increase in terms of supply of this oil type. Fan and Xu (2011), in their search for what has driven oil prices since the early 2000s, merely classify fundamental drivers of the crude oil price in three categories: 1) *the basic supply-demand relationship*, and frictions related to this; 2) *geopolitical events*; and 3) the *financialization factor*: market factors which have become increasingly related to the oil market such as stocks, gold, speculation and exchange rates.

Supply and demand

Hamilton (2008) notes that the WTI price behaviour over the period between 1970 and 1997 could be broadly explained by supply and demand forces and their respective low price-elasticity. Sornette et al. (2009) even argue that supply and demand were the main price drivers of crude oil until 2006; beyond this period, speculation cannot be ruled out and the predictive power of the supply-demand relationship sharply decreases. Nevertheless, even though more recent oil prices show significantly different characteristics: high levels of volatility and sensitivity to new information, as well as great variations in the short term (Askari & Krichene, 2008) and their importance seems to have declined, supply-demand shocks appear to remain the main drivers of oil price swings (Lombardi & van Robays, 2011). In line with Fan and Xu (2011), the Baltic Dry Index (BDI) will be chosen to capture fundamentals of supply and demand, short term variations in particular. BDI is preferred over other data sources since this index is published on a daily basis, thus providing enough opportunity to capture any short-term fluctuation on the crude oil market.

Geopolitical events

A significant amount of the oil production takes place in relatively unstable geographic regions. A conflict situation there could lead to drops in the crude oil supply, hence increasingly instable oil prices.

Financialization

Prior to the early 2000s, commodity returns showed negligible correlations with stock markets (Gorton & Rouwenhorst, 2006), little co-movement among each other (Erb & Harvey, 2006), and even provided a risk premium for idiosyncratic risk (Bessembinder, 1992; de Roon et al., 2000). This is in sharp contrast with the price dynamics of the typical financial asset: solely a premium for systematic risk, and high levels of correlation with other indices and each other (Tang & Xiong, 2012). Since then, the crude oil market, like most commodities, has become increasingly interconnected with other financial markets as a result of financialization, thereby increasingly behaving like a typical financial asset (Büyükşahin & Robe, 2011). Indeed, Hamilton and Yu (2011) provide evidence that the risk premia on crude oil futures have dropped significantly since 2005. It has also been argued that the increased participation of hedge funds and institutional investors accelerated the aforementioned interconnectedness (Teo, 2009).

Knowing there is a certain degree of interconnectedness between crude oil and other markets, what are their dynamics, and could these be used in a fundamental analysis in order to predict future oil price movements? Fan and Xu (2011) provide two theories regarding the presumed connection between oil prices and the stock market, which at the same time unambiguously emphasize the aforementioned time variation in fundamentals. (1) In line with Campbell et al. (1998), the effect of the stock market on the oil price depends on the macroeconomic situation. When stock markets become a proxy for the state of the world economy, a positive relation between stocks and crude oil will exist; a higher state of the world economy being associated with more demand for fossil fuels. This has been tested by Kyrtsou et al. (2016), who conclude that the variations in the S&P 500-index drove short-term expectations on the crude oil market since the early 2000s. (2) On the contrary, if stocks lose value, investors seeking safe havens could turn to the crude oil market, which will increase the crude oil price, hence creating a negative relation between both asset types. Further, there is the natural relationship between the crude oil price and the US dollar (USD) exchange rate, since the oil market is settled in USD. This creates a negative relationship between them, a depreciation of the USD being associated with an increase in the crude oil price; in the short run, but also in the long run (e.g. Reboredo et al., 2014).

Lastly, and perhaps one of the most significant developments on the crude oil market since the early 2000s is the presumed speculative behaviour which has entered the market, thereby significantly influencing price dynamics. Following, Li et al. (2016), the rationale behind speculation is as follows: if prices are on the rise (in decline), speculators tend to take long (short) positions in order to profit, thereby amplifying price movements. The authors find an increase in speculative behaviour in more recent periods. Kaufmann (2011) attributes the sharp increase in the crude oil price in 2007-2008 to both

market fundamentals as well as speculative pressure. The author provides three reasons for this proposition: 1) private US crude oil inventories sharply increased after 20 years of steady decline; 2) cointegration between spot and far-month prices broke down, inconsistent with the law of one price; and 3) the failure of predictive models solely based on fundamentals. Speculative behaviour could be based on certain expectations regarding the developments of other fundamentals, thereby leveraging the effects these other factors presumably have on the crude oil price, leading to more significant price shocks.

Based on the findings of Kaufmann (2011), the third one in particular, this research will use the complementary power of a fundamental analysis and the technical approach. This is in line with Bettman et al. (2009), who conclude that while the fundamental approach and the technical approach respectively could explain parts of an asset's price dynamics, a combination of both produces superior results. It is also supported by a questionnaire among financial traders in Hong Kong by Lui and Mole (1998), who find that more than 85 percent of the participants rely on a combination of both types of analysis. Their questionnaire also reveals for short-term decisions there is a sheer skew towards the technical approach. On the contrary, technical analysis has been held accountable for excess volatility on stocks and bond markets, as well as being single-handedly responsible for the 1987 stock market crash (Shiller, 1989). Wong (1993) argues it can create excess reactions on the market in general, without any fundamental reason. Based on these two findings, it could very well be that technical analysis is, at least, a contributing factor to crude oil prices driving away from the fundamentals, as it introduces non-fundamental input to the market. Since this is suggested to be quite a recent trend, given financialization and the changes it brought to the market since the early 2000s, it seems very relevant to add technical factors to the model, thereby combining the explanatory powers for both a fundamental and a technical approach to price developments on market for crude oil.

<u>Technical analysis</u>
Recall that the technical approach uses factors like momentum and the spot-futures spread of an asset in order to predict future movements rather than the aforementioned exogenous – fundamental – factors (Lam, 2003). Traditionally three theories exist regarding the relationship between the current and future price of a commodity, the spread which could exist between the two in particular. The first theory points at the *risk premium* for non-diversifiable risk. In its classic form, under normal backwardation (contango), an arbitrageur who takes the long (short) position may require a futures price below (higher than) the expected future spot price, hence a premium, from a commercial producer, the hedger (Keynes, 1930). As mentioned earlier, these risk premiums used to be relatively large on the market for crude oil but have experienced a dramatic drop since 2005, as per Hamilton and Yu (2011). The authors point at the increasing involvement, and hence buying pressure, of financial investors in oil futures, which has caused the receiving side of the risk premium to shift from the long to the short side of a futures contract.

Specific to a commodity market like the one for crude oil, one could think of another theory regarding the spot-futures spread base on the fact that crude oil is an exhaustible resource. Hotelling

(1931) proposes a theory where the net price of such a resource should rise over time at the rate of interest; in other words, a natural contango market. However, as pointed out by, among others, Litzenberger and Rabinowitz (1995), crude oil futures are often below spot prices, i.e. in backwardation; which is the exact opposite of the aforementioned situation. The authors, approaching oil wells as if it were call options, show that (1) crude oil is only produced if discounted futures prices < spot prices; and 2) strong backwardation materializes if future price volatility is sufficiently high.

The last theory describes the spread between spot and future delivery prices as a function of the opportunity cost related to the commodity's storage, the actual storage costs, and the convenience yield of holding the commodity for a certain period of time (Kaldor, 1939; Working, 1948). The basis for this theory of storage, as discussed by the authors, is that, at time t, one has two options in order to end up with one unit of a certain commodity at t+1: buy the commodity unit on the spot market and store it for one period, or buy a futures contract for delivery of one unit at t+1. As a result of forces of arbitrage, the price of the futures contract then has to equal the spot price raised with the implied liabilities and benefits of commodity storage. The liability in this case is the opportunity costs of holding the commodity, which roughly equal the interest rate; the benefit is derived from the added flexibility of additional commodity stock which could help the storing party to more easily adjust to sudden demand-supply volatility. Hence, the theory of storage arguably breaks the spot-futures spread down to merely a function of the interest rate and the convenience yield; an increase (a decrease) in the spread resulting from a rise (fall) in the interest rate and/or a declining (ascending) convenience yield. The findings of Knetsch (2006) confirm this latter proposition. The author even concludes that, between 1993 and 2005, the convenience yield outperformed the rolling-year futures prices in forecasting prices for delivery one year from now.

The aforementioned dramatic drop in risk premiums since 2005 (Hamilton & Yu, 2011) has been accompanied by the earlier-mentioned high levels of volatility on the crude oil market (e.g. Askari & Krichene, 2008). This seems rather contradictory. Furthermore, the exhaustible resource proposition seems to have become less relevant as a result of the discovery of new oil sources such as shale oil as well as cheaper production methods (Huisman & Kiliç, 2015). Hence, this paper will take the theory of storage as the basis of the technical analysis. It will herewith follow the line of thought as postulated by Ng and Pirrong (1994), in the sense that if: 1) inventories fall relative to demand, and 2) the elasticity of the industry marginal production cost decreases, this will result in an increase in the interest-and-storage adjusted spread. If both 1) and 2) hold, this will lead to a more inelastic supply curve, which should, according to microeconomic theory, in turn, lead to higher commodity price variation. Hence, *an increase in the spot-futures spread should lead to higher price volatility*, and vice versa.

Now that the presumed price drivers, both from a fundamental as well as from a technical point of view, both in the long run as well as the short run, and certain characteristics of the crude oil market have

become apparent and can be accounted for, appropriate statistical methods for predicting future price movements/volatility have to be chosen.

Forecasting oil price volatility

At this point, this research tries to find out whether the spot-futures spread combined with a set of fundamental factors presumably driving oil price developments could be an *unbiased* estimator for future crude oil price volatility, and the correct statistical method for doing so. It is of great importance that the statistical method meets the data, in the sense that it accurately captures certain known characteristics of the asset it describes.

As noted by Andersen and Bollerslev (1997), volatility is time-varying and predictable. Also, a set of common features of financial asset volatility can be observed. First of all, the aforementioned volatility clustering and mean reversion properties of financial assets in general (Mandelbrot, 1963; Engle & Patton, 2001). Indeed, Charles and Darné (2014) state that WTI prices exhibited significant degrees of volatility clustering in 1985-2011, periods of relatively high volatility being followed by periods of relatively low volatility. However, Ozdemir et al. (2013) do not find a trend of mean reversion for crude oil markets between 1991 and 2011. Two additional characteristics financial assets have in common have been discussed extensively in the literature: the *leverage effect*, and *fat tails* in the distribution. The leverage effect describes the asymmetrical effect on the conditional volatility of positive and negative shocks respectively, as first noted by Black (1976); negative shocks arguably having a more significant impact on the asset's volatility. Wei et al. (2009) find mixed results for the leverage effect on the crude oil market for a dataset covering 1992-2009. Between this period, Brent prices seem to contain significant leverage effects. WTI prices on the other hand only show this effect when an EGARCH model is used. The fat-tails property of financial assets indicates an unconditional volatility distribution where more observations are found in the lower and the upper tail of the distribution than what would be assumed under a normal distribution. Reboredo (2014) indeed finds fat tails in the distribution of multiple crude oils between 1997 and 2010. Pointing at the high observed kurtosis statistic for WTI, Brent, Dubai and Maya, the author states that extreme price changes occurred at a higher frequency when the return distribution showed fat tails.

Knowing that volatility clustering and fat tails are the most significant of the aforementioned asset characteristics on the WTI crude oil market: how could one incorporate this knowledge in a statistical model? Traditionally, the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) has been the bread and butter for a significant amount of the research on modelling conditional volatility. It was first proposed by Bollerslev (1986) after the pioneering ARCH model of Engle (1982). Both models are built on the assumption that financial time series tend to exhibit high degrees of heteroscedasticity (Officer, 1973). By incorporating this feature into the conditional second moment, the ARCH model was the first to actively account for fat tails in the distribution. The model, however, had some significant drawbacks. For instance, the large number of parameters needed in order

for it to work. GARCH was an improvement to the ARCH model, in the sense that it did not have this drawback, while incorporating mean reversion and being relatively easily to modify in order to account for specific other data characteristics (Bollerslev et al., 1993).

The GARCH model, or GARCH(1,1), though, also seems to have some important shortcomings. Engle and Patton (2001), for instance, point at the fact that GARCH(1,1) does not distinguish between positive and negative shocks as it takes the squares of volatility moves rather than their positive and negative amounts, thereby ignoring the aforementioned leverage effect. Lamoureux and Lastrapes (1990), in turn, state that the high predictability of volatility shocks implied by the GARCH model may be spurious as a result of regime shifts or structural breaks. Solutions to both drawbacks respectively have been proposed since the model's establishment. Nelson (1991)'s EGARCH model, for instance, logs the conditional volatility differently across regimes; thereby allowing for the different signs, and hence heterogeneity in volumes, volatility shocks could have. Furthermore, as the impact of regime shifts is arguably the largest in the short-term, Marcucci (2005) advises to add Markov regime-switching features to the GARCH(1,1) in order to account for this in the short-term volatility forecast, while keeping the original model for forecasting long-term volatility - both based on performance results.

Specifically to the crude oil market, a significant amount of research has been conducted on the performance of certain models and the implications of regime shifts (e.g. Wilson et al., 1996; Fong & See, 2002; Vo, 2009; Zhang & Zhang, 2015). These papers indeed show that on the market for crude oil, the significant degree of persistence in volatility shocks implied by the GARCH model is caused by regime shifts. Moreover, Wilson et al. (1996), argue that sudden volatility shifts are among the most significant factors of the market, and therefore definitely should be accounted for. Fong and See (2002) use these arguments as a starting point to test a general regime-switching model without GARCH effects and conclude this significantly reduces volatility clustering for the covered period of 1992-1997. The authors ascribe this to regime shifts dominating ARCH effects on the market for crude oil, and hence, a regime-switching model should be more effective than a GARCH model. These finding are supported by Vo (2009), who compares the predicting powers of four different statistical models for the WTI market between 1986 and 2008, among which the GARCH model and the Markov switching model with constant variance. The author concludes that, although dependent on the chosen evaluation criteria, the Markov switching model outperforms the other models.

In line with these propositions, this paper will use a regime-switching model. Fong and See (2002) have some interesting additional findings from using such a model on the crude oil market. The authors make the probability of a regime switch dependent on the spot-futures spread, as a proxy for supply-demand shocks, and find that the impact of variations in the spread on future volatility is larger in periods of significant demand shocks, thereby supporting the aforementioned assumptions regarding the theory of storage of Ng and Pirrong (1994).

As mentioned earlier, the limitations of the efficient market hypothesis as noted by Fama (1965) mark the starting point of both the fundamental and technical analysis. Indeed the crude oil market seems to contain certain inefficiencies, both in the short-terms as well as the long –term, where prices drift away from their fundamentals. Hence, fundamentals like the supply-demand relationship, and market factors such as the USD exchange rate, the gold price, and the stock market only seem to partially explain the price movements on the WTI market. The increased financialization of the crude oil market since the early 2000s and the presumed speculation which has come with this trend could have played a significant role here. Technical analysis, through the spot-futures spread in particular, could fill this explanatory gap in a volatility forecasting model. In order for such a forecast to be accurate, tomorrow's prices should not only be dependent on today's news; but rather be a function of certain market fundamentals as well as technical factors. Crude oil prices do not seem to follow a random walk, as explained by certain properties of the asset itself; the presumed existence of volatility clustering and fat tails in particular. Volatility clustering seems to be more important than fat tails on the crude oil market, as indicated by regime-switching models outperforming the GARCH model. Hence this research will use a switching-regression model in order to forecast future oil price movements based on a set of fundamentals and technical factors.

Preliminary hypotheses

1. A combination of the fundamental and the technical approach has superior predictive power compared to separate analyses.

2. The data, weekly WTI prices between 2000 and 2015, will contain different volatility regimes. Hence the decision to use a regime-switching model.

3. The fundamentals will show time-varying significance, following Fan and Xu (2011).

4. The rolling-year WTI futures spread will be the most obvious determinant for the probability of a volatility regime switch. The fundamentals (the S&P 500 index, the NYMEX rolling-month gold future, the Baltic Dry Index, and the US Dollar Index) will merely function as leverage for the predictive power of the model all together.

Crude oil data, technics and fundamentals

The preceding section described the theoretical and empirical relationship between the aforementioned set of technical and fundamental factors and the crude oil price. This research will use fifteen years of weekly data over the period 2000-2015, for a total of 832 observations, to further investigate this relationship. Specifically, whether a combination of both types of factors is able to predict future oil price movements. The following chapter will function as a justification as well as an overview of the data which are used for this research, the one thereafter will elaborate on the methodology used throughout this research in order to assess this.

<u>WTI crude oil price</u>

2000-2015 has been a period of significant highs and lows for the crude oil price. Fan and Xu (2011) assigned labels to certain subsets of this time range: 2000–mid-2004 is described as relatively calm, mid 2004–mid-2008 as a period of "bubble accumulation" and mid-2008–early 2009 as the time where the global financial crisis took place. I would like to add two more descriptions in order to partially cover the remaining period between 2009 and 2015 based on the figure below: a period of post-crisis recovery between 2009 and mid-2011, and one characterized by a fierce battle for market share between OPEC and the USA starting mid-2014 and still ongoing by the end of 2015.

<u>Figure 1 and 2 – WTI front-month: weekly prices and returns, 2000-2015</u>
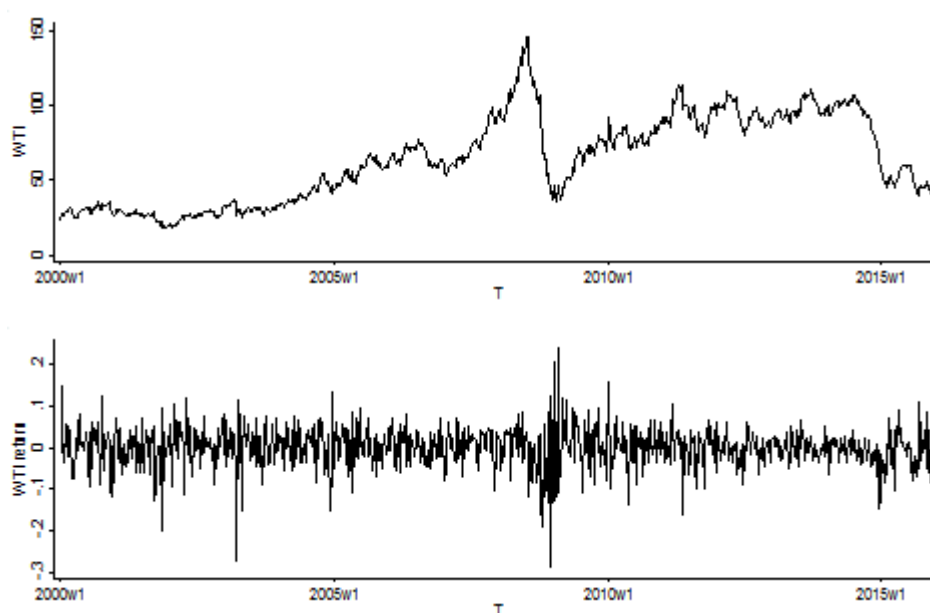
Source: NYMEX, WTI prices in $/bbl.

Although the returns on the rolling-month's WTI contract play a significant role in this paper, the main focus will be on predicting its price volatility. Specifically, whether the earlier-mentioned fundamentals and technics are able to predict oil's second moment. For instance, could a deviation in one or more of these factors help to foresee a switch from a relatively low-volatility state to a high-volatility one? Before this paper turns to the methodology behind this assessment of volatility movements in the next chapter,

and the actual analysis in the one thereafter, the remainder of this section will be used to both establish definitions with respect to oil price volatility and the input parameters as well as a set of respective hypotheses.

WTI crude oil price volatility

This research takes the variation in the returns of the WTI contract for front-month delivery as a measure of crude oil price volatility. The concerning contract has been chosen due to its relative proximity to the fundamental and technical data at any time t. Consequently, this paper follows the proposition of Fong and See (2002) and Marcucci (2005) that volatility switches are expected to be the most significant in the short-term.

Since the focus is on predicting future volatility, at any point in time, the standard deviation in returns related to the front-month contract for the forthcoming 5 trade days will be used in order to calculate the volatility measure:

$$\sigma_{t+5} = \sqrt{\frac{1}{5}\sum_{i=1}^{5}(r_{t+i} - \bar{r})^2} \quad \text{where} \quad r_t = ln\left[\frac{P_{front-month_t}}{P_{front-month_{t-1}}}\right] \quad \text{and} \quad \bar{r} = \frac{1}{5}\sum_{i=1}^{5}r_{t+i} \quad (1)$$

The selected range should be both long enough in order to accurately measure price volatility while also being short enough to still be able to capture the assumed short-term effects of the input on the volatility. The method above results in the following figure:

Figure 3 – WTI front-month: weekly forthcoming price volatility, 2000-2015
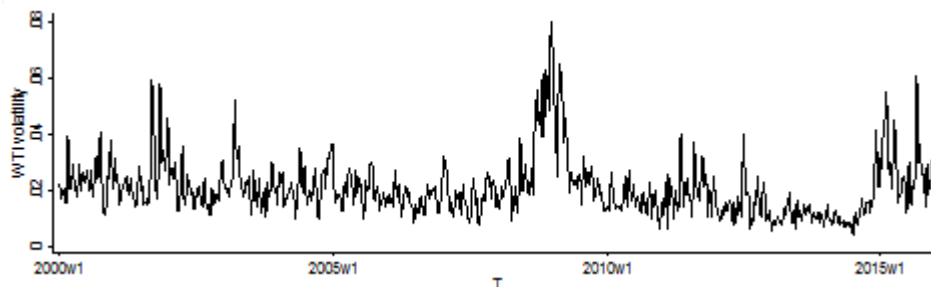
Source: NYMEX

The highs and lows in the WTI price as shown in figure 1, indeed as expected, result in a similar volatility pattern. Moreover, there seems to be great variability in volatility between periods; high and low states of volatility seem relatively easy to distinguish. Two periods of high volatility stand out: the global financial crisis and the price war between OPEC and non-OPEC. Another striking feature is the peak around mid-2002, and another by the end of 2003, within a period which Fan and Xu (2011) label as relatively calm.

With the definition of WTI volatility being established, the next section will elaborate on the remaining factors to be used in the switching-regression model.

Input parameters

First of all, this paper will use the following definition with respect to the aforementioned technical factor, the one-year spread:

$$Spread = ln\left[\frac{P_{next-year_t}}{P_{front-month_t}}\right] \qquad (2)$$

Specifically, this is the spread between the rolling WTI futures contract for front-month delivery and the one for delivery next year. For instance, for any day in January 2000, the former contract will be the February 2000 contract, while in this case the latter is the January 2001 contract. Given this definition, intuitively a similar picture emerges for the spread as for the volatility above:

Figure 4 – WTI rolling-year spread, 2000-2015
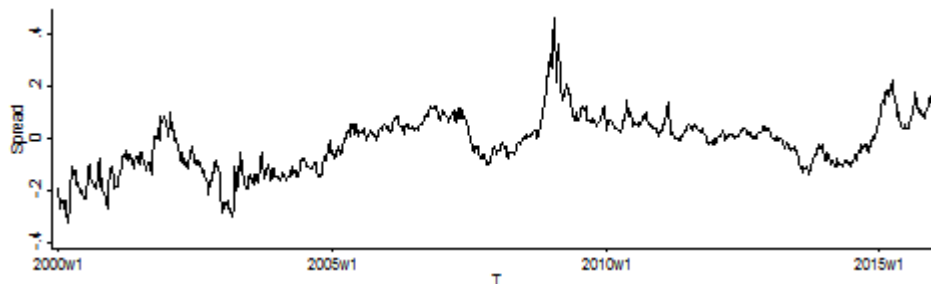
Source: NYMEX

Indeed, peaks in the spread can be seen around mid-2002, during the global financial crisis and by the end of the covered period, possibly resulting from the supply power struggles on the crude oil market. Will the fundamental data show a similar pattern?

Based on the theoretical relations with the crude oil price as established in the previous section, four fundamental factors have been selected for this research. First of all the S&P 500 Index, which is taken as a proxy for the world stock market. Since the index is composed of the 500 largest companies in terms of market capitalization, deviations in this parameter could provide useful information about the state of the economy. Secondly, the NYMEX front-month contract for gold in 100 oz/$. As mentioned earlier, certain investors use gold contracts in order to hedge their oil investments. Hence, this parameter could be an indicator of an upcoming volatility switch. Further, there is the weekly Baltic Dry Index (BDI), which is used to incorporate the supply-demand fundamental. This parameter could both signal the state of the world economy as well as speculative behaviour. Lastly, the US Dollar Index (USDx), which indexes the relative value of the US dollar against a basket of foreign currencies. This parameter can be used in order to track the relative cost of crude oil, since oil contracts are denoted in US dollars.

Ideally, an outcome of this research would be that *any* movement of these fundamental parameters, either positive or negative, raises or lowers the probability of a certain volatility state of the WTI front-month price. In this sense, it would intuitively be more appealing to use the absolute fundamental returns as input for this research, rather than the regular measures. As the absolute spread return contains relatively little information, the spread will still be taken in the earlier-described way. The absolute returns of the fundamentals are displayed below:

Figure 5 and 6 – weekly absolute returns of S&P 500 and front-month gold, 2000-2015

Source: S&P 500

Source: NYMEX
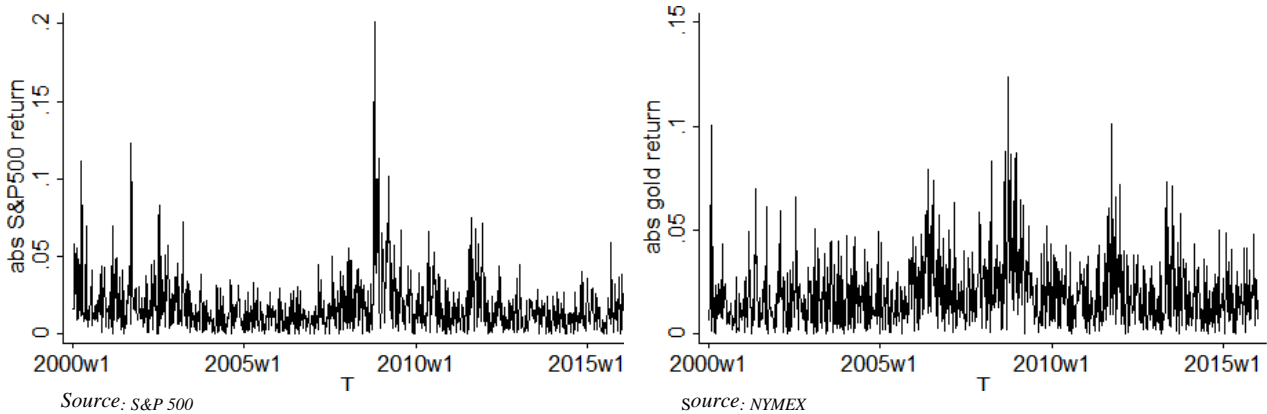
Both parameters show to some degree the same pattern as the crude oil price: higher return fluctuations around mid-2002 and during the global financial crisis. Furthermore, a greater deal of variability in returns can be observed in the post-crisis recovery period around 2011-2012. Gold shows one additional peak by the end of the time range, possibly due to an increase in hedging behaviour resulting from the aforementioned market power tensions between OPEC and non-OPEC.

Figure 7 and 8 – weekly absolute returns of Baltic Dry Index and US Dollar Index, 2000-2015

Source: BDI

Source: ICE
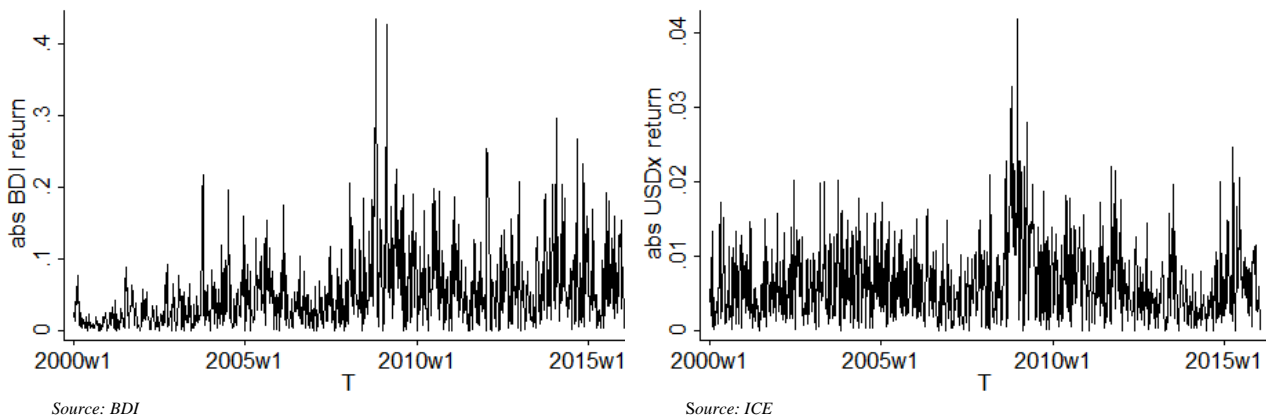
The indices shown above exhibit a somewhat different pattern. The financial crisis period again clearly stands out, but other periods of relatively higher variability are harder to distinguish. BDI seems to experience a trend of increasing volatility over time, possibly due to the aforementioned increase in speculative behaviour. The US Dollar Index, in turn, seems to be relatively constant over the analysed period; with an exception for the financial crisis and the price war period starting from around 2015. This last feature may be due to an increase in demand for crude oil resulting from the lower price, which effectively increases the demand for dollars as well.

Based on the visual assessment above, one could say the returns of the aforementioned technical and fundamental factors exhibit the time-varying significance as proposed by Fan and Xu (2011). In other words, different subsets of the period 2000-2015 seem to contain different oil price drivers. Later on in this research, this relationship will be assessed more in-depth.

16

What this research *does* aim to assess is whether the aforementioned set of parameters contains a certain degree of predictability with respect to the crude oil price volatility. The remainder of this chapter will be used to formulate a set of hypotheses related to this predictability feature. To get a sense of how the fundamentals and oil volatility seem to move together, their respective data points are plotted against each other in the figures below:

<u>Figures 9-12 – plots of weekly absolute fundamentals returns against WTI volatility, 2000-2015</u>
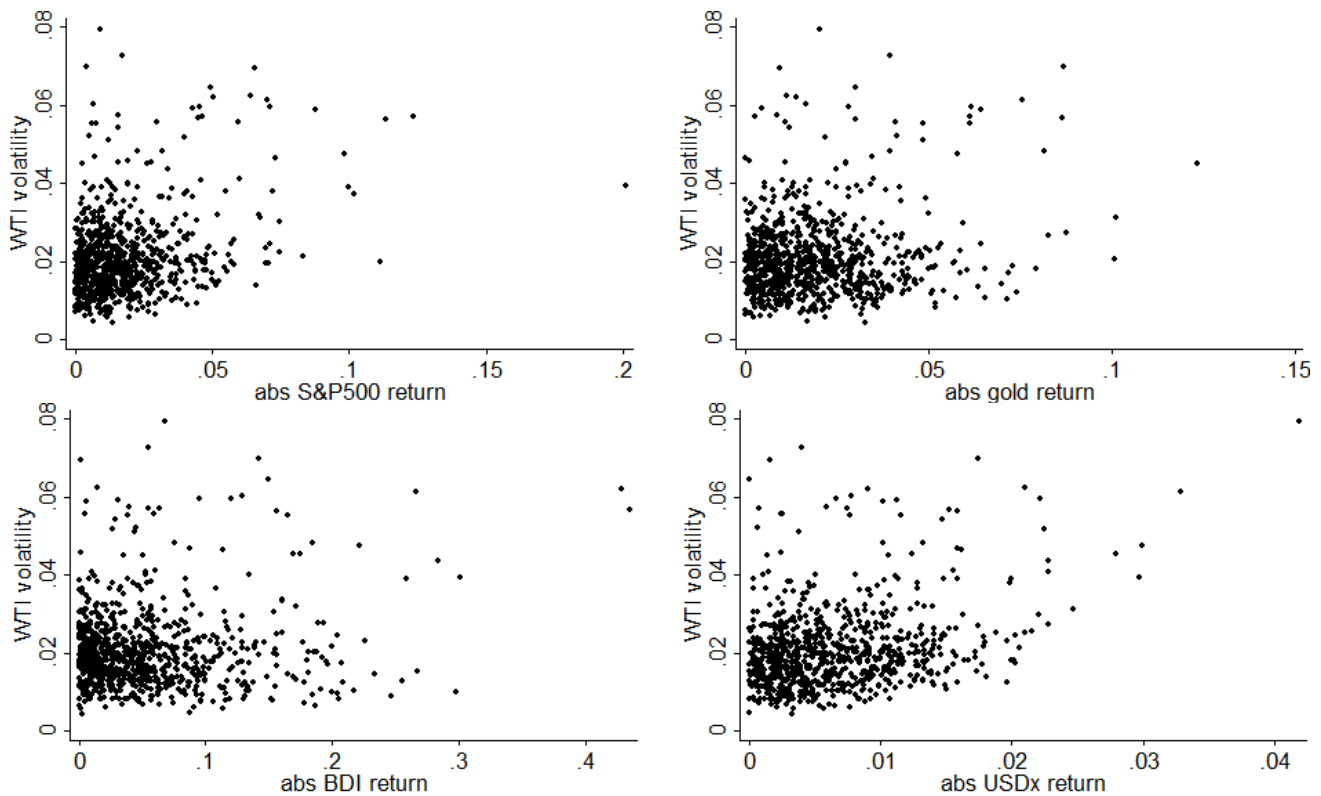
Although it is relatively hard to distinguish a respective trend for each fundamental factor related to WTI's volatility, two things stand out:

1) *There seems to be a general trend for all fundamentals where higher absolute returns go hand-in-hand with higher oil volatility. Being more conservative: higher absolute returns seem to raise the probability of a higher degree of oil volatility.*

2) *For the S&P 500 and the US Dollar Index, this potential relation seems to be relatively more pronounced than for the rolling-month NYMEX gold future and/or the Baltic Dry Index.*

Now that certain expectations about the result of this paper have been established, before turning to the actual assessment, the next chapter will first lay the methodological foundations.

Methodology, a switching-regression model

Recall from the preceding sections that the crude oil market is a market marked with significant fluctuations in terms of volatility over the covered period of 2000-2015. Moreover, the drivers of these price movements seem to exhibit time-varying significance. The type of model most suitable for such a data set according to, among others, Marcucci (2005) is one containing regime-switching elements. Thereby outperforming the more spurious results obtained from comparable GARCH-type of models.

As stated in the previous chapter: the crude oil's return is an important feature of this research, but the emphasis will be put on the developments with respect to the oil price volatility. If we define the return on oil as

$$R_{oil_t} \;=\; ln\left[\frac{P_{front-month_t}}{P_{front-month_{t-1}}}\right] \quad , \tag{3}$$

oil's volatility as (1), and constant factor α as the mean of oil returns as expressed in (2), the very basic model of this research looks like this:

$$Return_t \;=\; \alpha + \begin{cases} volatility_t \\ -volatility_t \end{cases} \cdot \varepsilon_t \tag{4}$$

Which implies that at time t, the return is essentially a volatility process around the return's mean. The basic model will be used to establish a preliminary model of input parameters which raise the probability of currently being in a high-, or low-volatility state, and is expressed as follows:

$$P\langle\theta_1\rangle_t = exp^{(\alpha+\,\beta\Delta spread_t+\gamma\Delta s\&p500_t+\delta\Delta gold_t+\epsilon\Delta bdi_t+\vartheta\Delta usdx_t+\varepsilon_t)} \tag{5}$$

$$\text{with} \quad \sum_{i=1}^{2} P\langle\theta_i\rangle = 1 \tag{6}$$

Then, a one-period lag will be added to the parameters, where volatility is considered to be a probabilistic two-regime process, with regime Θ, the two states being a high-volatility and low-volatility one, and the probability of being in a certain volatility state in the next period being dependent on the aforementioned set of input parameters. Here, $P\langle\theta_1\rangle_t$, the probability of a certain volatility state at time $t > t - 1$ is calculated through:

$$P\langle\theta_1\rangle_t = \; exp^{(\alpha+\,\beta\Delta spread_{t-1}+\gamma\Delta s\&p500_{t-1}+\delta\Delta gold_{t-1}+\epsilon\Delta bdi_{t-1}+\vartheta\Delta usdx_{t-1}+\varepsilon_{t-1})} \tag{7}$$

$$\text{with} \quad \sum_{i=1}^{2} P\langle\theta_i\rangle = 1 \tag{8}$$

Where the deltas in (5) are the absolute returns respectively for each parameter at time t, and at time t-1 in (7). Moreover, as there are two possible states, the probability of being in one regime is equal to one minus the probability of being in the other. This leads to the following two dynamic oil return formulas:

$$R_{\theta_1} \;=\; \mu + \left[P\langle\theta_1\rangle_t \cdot \begin{cases} exp^{\sigma_{\theta_1}} \\ -exp^{\sigma_{\theta_1}} \end{cases}\right] \cdot \varepsilon_t \tag{9}$$

$$R_{\theta_2} \;=\; \mu + \left[(1 - P\langle\theta_1\rangle_t) \cdot \begin{cases} exp^{\sigma_{\theta_2}} \\ -exp^{\sigma_{\theta_2}} \end{cases}\right] \cdot \varepsilon_t \tag{10}$$

Optimizing the model

In order to ensure the fit of these probabilities with respect to the volatility measure, and, in turn, the volatility measure related to the oil returns, Broyden-Fletcher-Goldfarb-Shanno (BFGS)-optimization

techniques will be used supported by a Marquardt-steps system. BFGS is a Quasi-Newton-category method, which constructs a quadratic model of the objective function in such a way that superlinear convergence is realised; in the case for this model convergence between both the set of input parameters, the probabilities of both regimes, oil's second moment and its returns. Its derivation starts with the quadratic model:

$$m_k(\bar{p}) = f(\bar{x}_k) + \nabla f(\bar{x}_k)^T \bar{p} + \frac{1}{2}\bar{p}^T \beta_k \bar{p} \tag{11}$$

with current iterate $\bar{x}_k$ and symmetric positive definite matrix β_k which will be updated with each iteration. Allowed by the convexity feature of the quadratic model, minimizer $\bar{p}_k$ can be expressed as follows:

$$\bar{p}_k = -\beta_k^{-1} \nabla f(\bar{x}_k) \tag{12}$$

Given this definition of the minimizer, the next iteration then equals:

$$\bar{x}_{k+1} = \bar{x}_k + \alpha_k \bar{p}_k \tag{13}$$

Where the step length α_k is determined by a process originally established by Marquardt (1963). Optimized for this paper, it works as follows: when, at the kth iteration, the equation

$$\varphi_k^* = (\vartheta_k^* + \alpha_k I) \cdot \delta_k^* \tag{14}$$

is constructed and solved for δ_k^* , then we use another outcome from the same author

$$\delta_k = \left. \delta_k^* \middle/ \sqrt{\gamma_{kk}} \right. \tag{15}$$

in order to obtain δ_k. This value will be used as input for the new trial vector

$$b_{k+1} = b_k + \delta_k \qquad , \tag{16}$$

which, in turn, results in a new sum of squares ϕ_{k+1}. Hence, the sum of squares, or the fit of the model, is directly linked to the Marquardt process. It is therefore essential to select step length α_k such that

$$\phi_{k+1} < \phi_k \tag{17}$$

is satisfied. Unless b_k is already at a minimum, a sufficiently large α_k to satisfy (17) should, at least in theory, always exist; this should then lead to a rapid convergence of each iteration, and hence the model, to least sum of squares. Marquardt (1963) provides the following strategy for finding an optimal α_k :

let $v > 1$, α_{k-1} be equal to the previous value for α from the preceding iteration, and the initial value, α_0, arbitrarily be equal to 10^{-2}. Proceed with computing $\phi(\alpha_{k-1})$ and $\phi(\alpha_{k-1}/v)$. Then,

if $\qquad \phi(\alpha_{k-1}/v) \leqq \phi_k$, $\qquad\qquad\qquad\qquad$ **let $\alpha_k = \alpha_{k-1}/v$** $\qquad$ (18)

if $\qquad \phi(\alpha_{k-1}/v) > \phi_k$, and $\phi(\alpha_{k-1}) \leqq \phi_k$ $\qquad\qquad$ **let $\alpha_k = \alpha_{k-1}$** $\qquad$ (19)

if $\qquad \phi(\alpha_{k-1}/v) > \phi_k$, and $\phi(\alpha_{k-1}) > \phi_k$, increase α_k by v until

$\qquad$ for some smallest value of ω $\phi(\alpha_{k-1}v^{\omega}) \leqq \phi_k$. $\qquad\qquad$ **let $\alpha_k = \alpha_{k-1}v^{\omega}$** $\qquad$ (20)

The iteration k is said to have converged when

$$\frac{|\delta_k|}{\tau + |b_k|} < \epsilon \tag{21}$$

for all k and some small values for ϵ and $\tau > 0$, Marquardt (1963) mentions 10^{-5} and 10^{-3} respectively. In practice $v = 10$ seems to work, which is also the default number of iterations used in this model. After doing this, we end up with a new model:

$$m_{k+1}(\bar{p}) = f(\bar{x}_{k+1}) + \nabla f(\bar{x}_{k+1})^T \bar{p} + \frac{1}{2} \bar{p}^T \beta_{k+1} \bar{p} \tag{22}$$

To ensure the quality of each $m_{k+1}(\bar{p})$, it must match the gradient of the objective function in $\bar{x}_k$ and $\bar{x}_{k+1}$. The latter requirement is true by construction, since

$$\nabla m_{k+1}(\bar{0}) = \nabla f(\bar{x}_{k+1}) \tag{23}$$

For $\bar{x}_k$, in turn

$$\nabla m_{k+1}(-\alpha_k \bar{P}_k) = \nabla f(\bar{x}_{k+1}) - \alpha_k \beta_{k+1} \bar{P}_k = \nabla f(\bar{x}_k) \tag{24}$$

Additionally, after a sufficient amount of iterations has resulted in a least squares fit of the model and its parameters, the Akaike information criterion (AIC) will be calculated for the entire model in order to assess the quality of the model and its predictions. Introduced by Akaike (1969), the model will be chosen if it minimizes the Kullback-Leibner distance between the actual situation and the one as predicted by the model. AIC combines the goodness of fit with the number of parameters for its verdict, and is expressed as follows:

$$-2L_m + 2m \tag{25}$$

where L_m is the maximum log-likelihood and m the number of parameters. In his pioneering work, Akaike (1969) suggests to take the model with the lower AIC value, hence, due to the positive impact on the measure, AIC imposes a 'penalty' on the use of parameters.

Each time an acceptable model in terms of fit has been calculated, its input parameters will be checked for significance. In case of insignificant parameters, the least significant parameter will be omitted and the model will be recalculated until all input is significant in terms of their respective contribution to the probability of a certain crude oil volatility state. Besides the basis model without lags, an additional model will be analysed where the input parameters contain a one-period lag. The latter model is particularly interesting in terms of practical use, since it is founded on the idea that a certain absolute return in one or more of the input parameters this week will increase or decrease the probability of a certain volatility state next week.

Now that the assumed theoretical relationships have been established, a visual assessment of the data has led to the formation of some hypotheses and the methodology has been explained in-depth, this paper continues with the actual analyses and their respective results in the next section.

Analysis and discussion

The section below will show and analyse the results of the simple switching regression as established in the last chapter. As mentioned earlier, two versions of the model will be tested, one basic model for all parameters at the same point in time, and one with input parameters containing a one-period lag. This latter model is particularly interesting as it could have significant predictive power with respect to crude oil's second moment in the next period. Before a general conclusion is drawn, the two respective models will be analysed separately. Recall that the basic current-state model has the structure as in (9) and (10), with volatility states being subject to conditions (5) and (6). For the lagged model the basic structure is the same, but the volatility states, and the probabilities of a certain volatility state next period in particular, are subject to (7) and (8).

<u>The basic model</u>

The basic current-state version of the volatility model aims to establish the general relation between the return, the two presumed volatility states and the input parameters. Essentially, it will provide an answer to the question how deviations in one or more parameters at time t will influence the probability of currently being in a high- or low-volatility state.

Single-parameter regressions

Before turning to the actual analysis of this model, first the singular regressions per input parameter will be assessed below in order to establish the ceteris paribus-conditions.

<div align="center">Table 1 – the basic model: single-parameter two-state switching regressions</div>

n	μ	σ_1	σ_2	c	Δspread	Δs&p500	Δgold	Δbdi	Δusdx	AIC
1	0.0031	-2.427	-3.305	-1.525						-3.251
	(-0.0016)	(0.116)	(0.064)	(0.425)						
2	0.0034	-2.288	-3.239	-2.552	13.483					-3.290
	(0.0015)	(0.115)	(0.046)	(0.574)	(3.564)					
3	0.0030	-2.429	-3.317	-2.648		54.529				-3.293
	(0.0015)	(0.105)	(0.056)	(0.494)		(12.350)				
4	0.0034	-2.419	-3.301	-2.411			38.352			-3.266
	(0.0016)	(0.111)	(0.063)	(0.576)			(11.277)			
5	0.0030	-2.404	-3.292	-2.118				8.025		-3.257
	(0.0016)	(0.112)	(0.058)	(0.503)				(3.176)		
6	0.0028	-2.435	-3.322	-2.663					158.436	-3.286
	(0.0015)	(0.106)	(0.060)	(0.534)					(35.015)	

Regressions 2-6 each contain one probability parameter plus a constant c, a mean and two volatility states (σ_1, σ_2). The former two are related to the volatility states, as in (5) and (6), in each regression the coefficients correspond to σ_1. Further, the mean and volatility states define the returns as in (9) and (10). Based on their respective coefficient-standard error combinations, we conclude that each parameter is significant at the one-percent level, and hence, c.p., positively contributes to being in a state-one (σ_1)

volatility. For instance, in regression 2, a one-percent deviation in the spread leads to a probability of being in σ_1 of $e^{(-2.552+13.483\cdot0.01)} = 8.91$ percent. σ_1 is the high-volatility regime in this case, since its coefficient is lower than the one for σ_2, and volatility is determined by $e^{(\sigma_i)}$. Continuing for regression 2, the outcomes lead to the following return formulas:

$$R_{\theta_1} = 0.0034 + \left[P\langle\theta_1\rangle_t \cdot \begin{Bmatrix} exp^{(-2.288)} \\ -exp^{(-2.288)} \end{Bmatrix} \right] \cdot \varepsilon_t \tag{26}$$

$$R_{\theta_2} = 0.0034 + \left[(1 - P\langle\theta_1\rangle_t) \cdot \begin{Bmatrix} e^{(-3.239)} \\ -e^{(-3.239)} \end{Bmatrix} \right] \cdot \varepsilon_t \tag{27}$$

With annual volatility in regime one equal to $e^{(-2.288)} \cdot \sqrt{52} = 73.2$ percent, and, likewise, an annual measure of 28.3 percent for regime one. In table 2, the remaining results for each regression are visualized.

Table 2 – the basic model: full singular two-state switching regression results (Δ 1%)

	$P\langle\theta_1\rangle_t$	σ_1	$R_{\theta_1} =$	$R_{\theta_2} =$	σ_1 p.a.	σ_2 p.a.
Δspread	8.9%	high	$0.003434 \pm P\langle\theta_1\rangle_t \cdot 0.1015 \cdot \varepsilon_t$	$0.003434 \pm (1 - P\langle\theta_1\rangle_t) \cdot 0.0392 \cdot \varepsilon_t$	73.2%	28.3%
Δs&p500	12.2%	high	$0.002984 \pm P\langle\theta_1\rangle_t \cdot 0.0881 \cdot \varepsilon_t$	$0.002984 \pm (1 - P\langle\theta_1\rangle_t) \cdot 0.0363 \cdot \varepsilon_t$	63.5%	26.2%
Δgold	13.2%	high	$0.003379 \pm P\langle\theta_1\rangle_t \cdot 0.0890 \cdot \varepsilon_t$	$0.003379 \pm (1 - P\langle\theta_1\rangle_t) \cdot 0.0369 \cdot \varepsilon_t$	64.2%	26.6%
Δbdi	13.0%	high	$0.002994 \pm P\langle\theta_1\rangle_t \cdot 0.0903 \cdot \varepsilon_t$	$0.002994 \pm (1 - P\langle\theta_1\rangle_t) \cdot 0.0372 \cdot \varepsilon_t$	65.1%	26.8%
Δusdx	34.0%	high	$0.002811 \pm P\langle\theta_1\rangle_t \cdot 0.0876 \cdot \varepsilon_t$	$0.002811 \pm (1 - P\langle\theta_1\rangle_t) \cdot 0.0361 \cdot \varepsilon_t$	63.1%	26.0%

Each probability of being in state one is based on a benchmark deviation of one percent in the concerning parameter. The higher probability resulting from the USDx may be attributable to the fact that a one-percent deviation is relatively large for this parameter. On the right side of the table, the estimated annual crude oil volatilities per parameter for each state are shown.

Multi-parameter regressions
Even more interesting is to see how these parameters work together in order to explain the current state of the volatility. As mentioned earlier, the first regression will consist of the complete set of parameters. New regressions will follow until the parameters in the model are at least significant at the ten-percent level, each time omitting the least significant parameter from the preceding regression. The results from this process are visualized in table 3.

Table 3 – the basic model: two-state switching regressions

n	μ	σ_1	σ_2	c	Δspread	Δs&p500	Δgold	Δbdi	Δusdx	AIC
1	0.0036 (0.0015)	-2.279 (0.100)	-3.252 (0.043)	-4.969 (1.058)	12.187 (4.022)	52.260 (15.015)	17.616 (14.250)	**4.997** **(5.041)**	80.102 (44.471)	-3.251
2	0.0036 (0.0015)	-3.250 (0.045)	-2.277 (0.105)	4.727 (1.016)	-12.336 (4.155)	-52.141 (15.239)	**-19.443** **(14.105)**		-82.506 (43.700)	-3.290
3	0.0035 (0.0015)	-3.250 (0.047)	-2.280 (0.106)	4.427 (1.010)	-12.923 (4.393)	-53.195 (15.627)			-95.157 (44.123)	-3.293

Absolute deviations in BDI and gold do not seem to make a statistically significant contribution to the current WTI price volatility. For regression 1, since $\sigma_1 > \sigma_2$, this current state is a high-volatility one; for regressions 2-3, the model describes the effect of the probability parameters on a low-volatility state. After omitting the two factors, and following (5)-(6), (9)-(10), three dynamic regressions remain:

$$P\langle\theta_1\rangle_t = exp^{(\ 4.427-12.923\ \Delta spread_t - 53.195\Delta s\&p500_t - 95.157\Delta usdx_t + \varepsilon_t)} \tag{28}$$

$$R_{\theta_1} = 0.0035 + \left[P\langle\theta_1\rangle_t \cdot \begin{Bmatrix} exp^{(-3.250)} \\ -exp^{(-3.250)} \end{Bmatrix} \right] \cdot \varepsilon_t \tag{29}$$

$$R_{\theta_2} = 0.0035 + \left[(1 - P\langle\theta_1\rangle_t) \cdot \begin{Bmatrix} exp^{(-2.280)} \\ -exp^{(-2.280)} \end{Bmatrix} \right] \cdot \varepsilon_t \tag{30}$$

For simultaneous absolute deviations in each parameter of up to 2.74 percent, $P\langle\theta_1\rangle_t$, the probability of currently being in a low-volatility state is 100 percent. For higher deviations this probability decreases rapidly, essentially making a high-volatility state more likely. For instance, a simultaneous three-percent absolute deviation results in $P\langle\theta_1\rangle_t = 0.663$, for five percent this number decreases to only two percent. Following (29) and (30), this could have significant impact on the returns on crude oil investments.

Hence, the current state of WTI's rolling month futures price volatility seems to be driven by deviations in the spread, and the markets for stocks and currencies. This is confirmed by the AIC of the final model, which has the lowest value of all basic current-state regressions. Adding lags to the parameters, will this relation still hold for oil's second moment in the next period?

The lagged model

Except for the lagged parameters, the methodology for the lagged model remains identical to the one used to describe the current volatility state. First, single-parameter regressions will be performed in order to establish the general relation between the lagged parameters and the current volatility state, i.e. the relation between the current parameters and next week's volatility.

Single-parameter regressions

Figure 4 - the lagged model: single-parameter two-state switching regressions

n	μ	σ_1	σ_2	c	Δspread(-1)	Δs&p500(-1)	Δgold(-1)	Δbdi(-1)	Δusdx(-1)	AIC
1	0.0031	-2.427	-3.305	-1.525						-3.251
	(-0.0016)	(0.116)	(0.064)	(0.425)						
2	0.0027	-2.361	-3.277	-1.931	8.542					-3.274
	(0.0015)	(0.108)	(0.051)	(0.439)	(2.406)					
3	0.0031	-2.410	-3.308	-2.659		49.986				-3.292
	(0.0015)	(0.105)	(0.053)	(0.492)		(11.455)				
4	0.0027	-2.374	-3.270	-2.472			30.702			-3.263
	(0.0016)	(0.123)	(0.056)	(0.587)			(11.683)			
5	0.0029	-2.351	-3.262	-2.253				6.153		-3.257
	(0.0016)	(0.132)	(0.058)	(0.591)				(3.086)		
6	0.0027	-2.377	-3.273	-2.764					128.960	-3.272
	(0.0016)	(0.124)	(0.059)	(0.661)					(37.129)	

Again, regressions 2-6 each contain one probability parameter plus a constant c, a mean and two volatility states (σ_1, σ_2). The former two are related to the volatility states, for this model as in (7) and (8); in each regression coefficients correspond to σ_1. Returns, defined by the mean and volatility states, are still defined as in (9) and (10). Based on their respective coefficient-standard error combinations, we conclude that each parameter is at least significant at the five-percent level, and hence, c.p., could lead to an increased probability of a σ_1-state next week. For instance, in regression 2, a one-percent deviation in the spread leads to a probability of being in σ_1 next period of $e^{(-1.931+8.542\cdot0.01)} = 15.8$ percent, which is significantly higher than what we observed for the basic model. Also, for the lagged-spread regression σ_1 is the high-volatility regime in this case, as $\sigma_1 < \sigma_2$, and volatility is determined by $e^{(\sigma_i)}$. Continuing for regression 2, the outcomes lead to the following return formulas:

$$R_{\theta_1} = 0.0027 + \left[P\langle\theta_1\rangle_t \cdot \left\{ \begin{matrix} e^{(-2.361)} \\ -e^{(-2.361)} \end{matrix} \right) \right] \cdot \varepsilon_t \tag{31}$$

$$R_{\theta_2} = 0.0027 + \left[(1 - P\langle\theta_1\rangle_t) \cdot \left\{ \begin{matrix} e^{(-3.277)} \\ -e^{(-3.277)} \end{matrix} \right] \cdot \varepsilon_t \tag{32}$$

With annual volatility under σ_1 equal to $e^{(-2.361)} \cdot \sqrt{52} = 68.0$ percent, and, likewise, an annual measure of 27.2 percent for σ_2. In table 5, the remaining results for each regression are visualized.

Table 5 – the lagged model: full singular two-state switching regression results (∆ 1%)

	$P\langle\theta_1\rangle_t$	σ_1	$R_{\theta_1} =$	$R_{\theta_2} =$	σ_1 p.a.	σ_2 p.a.
∆spread(-1)	15.8%	high	$0.002687 \pm P\langle\theta_1\rangle_t \cdot 0.0943 \cdot \varepsilon_t$	$0.002687 \pm (1 - P\langle\theta_1\rangle_t) \cdot 0.0378 \cdot \varepsilon_t$	68.0%	27.2%
∆s&p500(-1)	11.5%	high	$0.003057 \pm P\langle\theta_1\rangle_t \cdot 0.0898 \cdot \varepsilon_t$	$0.003057 \pm (1 - P\langle\theta_1\rangle_t) \cdot 0.0366 \cdot \varepsilon_t$	64.8%	26.4%
∆gold(-1)	11.5%	high	$0.002651 \pm P\langle\theta_1\rangle_t \cdot 0.0931 \cdot \varepsilon_t$	$0.002651 \pm (1 - P\langle\theta_1\rangle_t) \cdot 0.0380 \cdot \varepsilon_t$	67.1%	27.4%
∆bdi(-1)	11.2%	high	$0.002877 \pm P\langle\theta_1\rangle_t \cdot 0.0952 \cdot \varepsilon_t$	$0.002877 \pm (1 - P\langle\theta_1\rangle_t) \cdot 0.0383 \cdot \varepsilon_t$	68.7%	27.6%
∆usdx(-1)	22.9%	high	$0.002664 \pm P\langle\theta_1\rangle_t \cdot 0.0929 \cdot \varepsilon_t$	$0.002664 \pm (1 - P\langle\theta_1\rangle_t) \cdot 0.0379 \cdot \varepsilon_t$	67.0%	27.3%

Also for the lagged model, each parameters on its own positively contributes to being in a high volatility state. Though, except for the spread, the probability coefficients are slightly lower than for the basic model. This could mean that the spread's current deviations have a higher impact on future oil volatility relative to current volatility.

Multi-parameter regressions
With the singular relation established, this research will again also assess whether a simultaneous deviation in multiple parameters could significantly influence the probability of a certain volatility state in the next period. As for the basic model, the first regression will consist of the complete set of parameters. New regressions will follow until the parameters in the model are at least significant at the ten-percent level, each time omitting the least significant parameter from the preceding regression. The results from this process are visualized in table 6.

Figure 6 – the lagged model: two-state switching regressions

n	μ	σ_1	σ_2	c	Δspread(-1)	Δs&p500(-1)	Δgold(-1)	Δbdi(-1)	Δusdx(-1)	AIC
1	0.0027	-2.305	-3.268	-3.634	6.151	37.350	15.550	**1.323**	43.481	-3.302
	(0.0015)	(0.116)	(0.047)	(0.772)	(2.373)	(10.359)	(12.013)	**(3.456)**	(38.317)	
2	0.0027	-2.317	-3.273	-3.512	6.080	37.478	**14.798**		48.307	-3.305
	(0.0015)	(0.112)	(0.047)	(0.703)	(2.350)	(10.372)	**(11.873)**		(37.990)	
3	0.0027	-3.283	-2.339	3.162	-6.429	-37.675			-57.728	-3.305
	(0.0015)	(0.048)	(0.109)	(0.605)	(2.395)	(10.599)			(37.283)	

For the lagged model, the same omissions as the ones performed in order to enhance the basic model have to be executed. Again, the first regression shows the lowest significance for BDI, followed by the gold parameter in the second regression. Hence, we conclude that neither of these two factors significantly contributes to a certain crude oil volatility state, neither currently nor for the next-period. For regressions 1 and 2, due to $\sigma_1 > \sigma_2$, this next-week state is a high-volatility one; for the third and most significant regression, the model describes the relations of the probability parameters with a low-volatility state. As the AIC for this third regression shows the lowest value for any lagged-parameter regression, we conclude this to be the best-fit model.

After omitting BDI and gold, following (7)-(10), three dynamic regressions remain:

$$P\langle\theta_1\rangle_t = exp^{(3.162 - 6.429\,\Delta spread_{t-1} - 37.675\Delta s\&p500_{t-1} - 57.728\Delta usdx_{t-1} + \varepsilon_{t-1})} \tag{33}$$

$$R_{\theta_1} = 0.0027 + \left[P\langle\theta_1\rangle_t \cdot \begin{cases} exp^{(-3.283)} \\ -exp^{(-3.283)} \end{cases}\right] \cdot \varepsilon_t \tag{34}$$

$$R_{\theta_2} = 0.0027 + \left[(1 - P\langle\theta_1\rangle_t) \cdot \begin{cases} exp^{(-2.339)} \\ -exp^{(-2.339)} \end{cases}\right] \cdot \varepsilon_t \tag{35}$$

For simultaneous absolute deviations in each parameter of up to 3.10 percent, $P\langle\theta_1\rangle_t$, the probability of currently being in a low-volatility state is 100 percent. For higher deviations this probability decreases rapidly, essentially making a high-volatility state more likely. For instance, a simultaneous 3.5-percent absolute deviation results in $P\langle\theta_1\rangle_t = 0.669$, for five percent this number decreases to only 14.5 percent. Following (34) and (35), this could have significant impact on oil investment returns.

Hence, up to a certain extent, the lagged model and the basic one show the same results. The impact of each respective as well as simultaneous movement in the lagged parameters has a somewhat lower impact on the crude oil volatility of next week, compared to the effect on the current volatility state. Perhaps there is a direct effect of the parameters on crude oil's second moment, which in the lagged model is smoothened out over a week's time. Regardless, the analysis above shows significant evidence for a relation between the rolling-year WTI spread, the S&P 500, the US Dollar Index and the state of the volatility of the rolling-month WTI futures contract, both for the current period as well as the one for next week. This last feature in particular could be of great value for any oil market participant, be it an investor, (speculative) trader, producer or Finance minister of an oil-producing country.

Conclusion

This research has provided the theoretical foundations for the impact of movements in a set of four fundamental and one technical factor on the WTI crude oil price volatility. Moreover, it has statistically assessed and proved the relationship for a subset of these parameters. The four fundamentals are the S&P 500 index, the rolling-month NYMEX gold future, the Baltic Dry Index, and the US Dollar Index. While separate analyses of the impact of fundamental or technical factors on the crude oil price respectively have been performed before, this research is, to the author's best knowledge, the first in combining both types of factors into one model related to crude oil volatility.

The limitations of the efficient market hypothesis as noted by Fama (1965) mark the starting point of both the fundamental and technical analysis. Indeed the market for WTI crude oil seems to contain certain inefficiencies, both in the short-terms as well as the long –term, where prices drift away from their fundamentals. Hence, fundamentals like the supply-demand relationship, and market factors such as the USD exchange rate, the gold price, and the stock market only seem to partially explain the price movements on the WTI market. The increased financialization of the crude oil market since the early 2000s and the presumed speculation which has come with this trend could have played a significant role here. Technical analysis, through the spot-futures spread in particular, could fill this explanatory gap in a volatility forecasting model. In order for such a forecast to be accurate, tomorrow's prices should not only be dependent on today's news; but rather be a function of certain market fundamentals as well as technical factors. Crude oil prices do not seem to follow a random walk, as explained by certain properties of the asset itself; the presumed existence of volatility clustering and fat tails in particular. Volatility clustering seems to be more important than fat tails on the crude oil market, as indicated by regime-switching models outperforming the GARCH model. Moreover, expectedly a model based on fundamental and technical factors should outperform separate analyses.

The data showed clustering of volatility on the market for crude oil is accompanied with significant high and lows in the WTI price over the covered period between 2000 and 2015, containing structural breaks. The movements in the assessed fundamental factors as well as the rolling-year spread show similar patterns. The model focuses on the absolute returns of these factors, hence working towards a conclusion based on *any* movement in the concerning parameters. Based on the data, the parameters were expected to have a positive impact on the probability of a certain volatility state; the most significant contributions being attributable to the S&P 500 index and the US dollar index.

The patterns in the data, the structural breaks in particular, induced the choice for the use of a switching-regression model. Essentially two models were estimated, one for the current state of the volatility based on current deviations in the assessed input factors and another for next week's impact on WTI's second moment resulting from current changes in the fundamentals and rolling-year spread.

Based on the results the following can be concluded:

- Simultaneous changes in BDI or the rolling-month NYMEX gold future, as part of a combined model with the rolling-year WTI spread, the US Dollar Index and the S&P 500, do not seem to have a statistically significant impact on the volatility state of the WTI rolling-month future, neither the current one nor the one for next period.

- On the other hand, simultaneous deviations in the rolling-year WTI futures spread, S&P 500 and US Dollar Index seem to increase the probability for a certain WTI volatility state, both the current one as well as the one for next week.

- For the current state, a simultaneous absolute deviation in these three parameters larger than 2.74 percent significantly increases the chance of a volatility switch in the current period. Furthermore, a simultaneous return larger than 3.10 percent now will significantly increase the chance of a volatility switch in the next period. These measures are at least significant at the ten-percent level, most even at the one-percent level.

These results could significantly improve risk management on the crude oil market. The outcome from the lagged model is particularly useful as it provides market participants with the probability of risks affecting their investments in the near-future based on the current developments on connected markets.

Future research could focus on the maximum amount of lags in the parameters which still provides significant impact on WTI's volatility. Moreover, more fundamentals such as trader's net investments in crude oil and commodities with strong correlations to oil such as natural gas could be added to the model. Lastly, certain corrections to the model such as seasonalities in the data could be assessed. For now, the first steps towards a better understanding of the combined effect of different analyses have led to a significantly pragmatic set of outcomes which could serve the oil market well.

Bibliography

Akaike, H. (1969). *Fitting Autoregressive Models for Prediction. Annals of the Institute of Statistical Mathematics, 21,* 243-247.

Alvarez-Ramirez, J., Cisneros, M., Ibarra-Valdez, C., & Soriano, A. (2002). *Multifractal Hurst Analysis of Crude Oil Prices. Physica, 313,* 651-670.

Alvarez-Ramirez, J., Alvarez, J., & Rodriguez, E. (2008). *Short-Term Predictability of Crude Oil Markets: a Detrended Fluctuation Analysis Approach. Energy Economics, 30,* 2645-2656.

Alvarez-Ramirez, J., Alvarez, J., & Solis, R. (2010). *Crude Oil Market Efficiency and Modelling: Insights from the Multi-Scaling Autocorrelation Pattern. Energy Economics, 32,* 993-1000.

Andersen, T.G., & Bollerslev, T. (1997*). Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns. The Journal of Finance, 52,* 975-1005.

Askari, H., & Krichene, N. (2008). *Oil Price Dynamics (2002-2006). Energy Economics, 30,* 2134-2153.

Bęben, M., & Orlowski, A. (2001). *Correlations in Financial Time Series: Established versus Emerging Markets. The European Physical Journal B, 20,* 527-530.

Bessembinder, H. (1992). *Systematic Risk, Hedging Pressure, and Risk Premiums in Futures Markets. Review of Financial Studies, 5,* 637-667.

Bettman, J.L., Sault, S.J., & Schultz, E.L. (2009). *Fundamental and technical analysis: substitutes or complements. Accounting and Finance, 49,* 21-36.

Black, F. (1976). *Studies of Stock Price Volatility Changes. Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section,* 177-181.

Bollerslev, T. (1986). *Generalized Autoregressive Conditional Heteroscedasticity. Journal of Econometrics, 31,* 307-327.

Bollerslev, T., Engle, R.F., & Nelson, D.B. (1993). *ARCH Models* (University of San Diego, California).

Brown, G.W., & Cliff, M.T. (2004). *Investor Sentiment and the Near-Term Stock Market. Journal of Empirical Finance, 11,* 1-27.

Büyükşahin, B., & Robe, M.A. (2011). *Does 'Paper Oil' Matter? Energy Markets' Financialization and Equity-Commodity Co-Movements. SSRN Electronic Journal:* DOI: 10.2139/ssrn.1855264.

Campbell, J.Y., Lo, A.W., & MacKinlay, A.C. (1998). *The Econometrics of Financial Markets.* Princeton: Princeton University Press.

Charles, A., & Darné, O. (2014). *Volatility Persistence in Crude Oil Markets. Energy Policy, 65,* 729-742.

Di Matteo, T., Aste, T., & Dacorogna, M. (2003). *Scaling Behaviors in Differently Developed Markets. Physica A, 324,* 183-188.

Dorner, D. (1996). *The Logic of Failure.* Cambridge: Perseus.

Engle, R.F. (1982). *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of the United Kingdom Inflation. Econometrica, 50,* 987-1008.

Engle, R.F., & Patton, A.J. (2001). *What Good is a Volatility Model?. Quantitative Finance, 1,* 237-245.

Eom, C., Oh, G., & Jung, W-S. (2008). Relationship between Efficiency and Predictability in Stock Price Change. *Physica, 387,* 5511-5517.

Erb, C., & Harvey, C. (2006). *The Strategic and Tactical Value of Commodity Futures. Financial Analysts Journal, 62,* 69-97.

Fama, E.F. (1965). *The Behavior of Stock-Market Prices. The Journal of Business, 38,* 34-105.

Fama, E.F. (1991). *Efficient Capital Markets II. The Journal of Finance, 66,* 1575-1617.

Fan, Y., & Xu, J-H. (2011). *What Has Driven Oil Prices Since 2000? A Structural Change Perspective. Energy Economics, 33,* 1082-1094.

Fleten, S-E., Huisman, R., Kiliç, M., Pennings, E., & Westgaard, S. (2015). *Electricity Futures Prices: Time Varying Sensitivity to Fundamentals. Journal of Energy Markets, 8,* 1-21.

Fong, W.M., & See, K.H. (2002). *A Markov Switching Model of the Conditional Volatility of Crude Oil Prices. Energy Economics, 24,* 71-95.

Gorton, G., & Rouwenhorst, G. (2006). *Facts and Fantasies about Commodity Futures. Financial Analysts Journal, 62,* 47-68.

Graham, B., & Dodd, D.L. (1965). *The Intelligent Investor.* New York: Harper & Row.

Hamilton, J.D. (2008). *Understanding Crude Oil Prices* (University of California Energy Institute).

Hamilton, J.D., & Yu, J.C. (2011). *Risk Premia in Crude Oil Futures Prices* (University of California, San Diego).

Hotelling, H. (1931). *The Economics of Exhaustible Resources. The Journal of Political Economy, 39,* 137-175.

Huisman, R., & Kiliç, M. (2015). *Managing oil price risk: Dealing with the time-varying relationship between the price of oil and fundamentals.* In V. Kaminsky (Ed.), *Managing Energy Price Risk.*

Kahneman, D., & Tversky, A. (1973). *On the Psychology of Prediction. Psychological Review, 80,* 237-251.

Kaizoji, T. (2000). *Speculative bubbles and crashes in stock markets: an interacting-agent model of speculative activity. Physica, 287,* 493-506.

Kaldor, N. (1939). *Speculation and Economic Stability. The Review of Economic Studies, 7,* 1-27.

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

Keynes, J.M. (1930). *A Treatise on Money.* London: Macmillan.

Kristoufek, L., & Vosvrda, M. (2013). *Measuring Capital Market Efficiency: Global and Local Correlations Structure. Physica A, 392,* 184-193.

Kristoufek, L., & Vosvrda, M. (2014). *Commodity Futures and Market Efficiency. Energy Economics, 42,* 50-57.

Kyrtsou, C., Mikropoulou, C., & Papana, A. (2016). *Does the S&P500 Index Lead the Crude Oil Dynamics? A Complexity-Based Approach. Energy Economics (Accepted Manuscript).*

Lam, M. (2003). *Neural Network Techniques for Financial Performance Prediction: Integrating Fundamental and Technical Analysis. Decision Support Systems, 37,* 567-581.

Lamoureux, C.G., & Lastrapes, W.D. (1990). *Persistence in Variance, Structural Change, and the GARCH Model. Journal of Business and Economic Statistics, 8,* 225-234.

Lean, H.H., McAleer, M., & Wong, W-K. (2010). *Market Efficiency of Oil Spot and Futures: a Mean-Variance and Stochastic Dominance Approach. Energy Economics, 32,* 979-986.

Li, H., Kim, M.J., & Park, S.Y. (2016). *Nonlinear Relationship Between Crude Oil Prices and Net Futures Positions: A Dynamic Conditional Distribution Approach. International Review of Financial Analysis, 44,* 217-225.

Litzenberger, R.H., & Rabinowitz, N. (1995). *Backwardation in Oil Futures Markets: Theory and Empirical Evidence. The Journal of Finance, 50,* 1517-1545.

Lo, A.W. (1991). *Long-Term Memory in Stock Market Prices. Econometrica, 59,* 1279-1313.

Lo, A.W., & MacKinlay, A.C. (1999). *A Non-Random Walk Down Wall Street.* Princeton: Princeton University Press.

Lombardi, M.J., & Robays, I. van (2012). *Do Financial Investors Destabilize the Oil Price?* (ECB Working Paper No. 1346).

Malkiel, B.G. (1973). *A Random Walk Down Wall Street.* New York: W.W. Norton & Company.

Malkiel, B.G. (2003). *The Efficient Market Hypothesis and Its Critics. Journal of Economic Perspective, 17,* 59-82.

Mandelbrot, B. (1963). *The Variation of Certain Speculative Prices. The Journal of Business, 36,* 394-419.

Marcucci, J. (2005). *Forecasting Stock Market Volatility with Regime-Switching GARCH Models. Studies in Nonlinear Dynamics and Econometrics, 9,* 1-55.

Marquardt, D.W. (1963). *An Algorithm for Least-Squares Estimation of Nonlinear Parameters. Journal of the Society for Industrial and Applied Mathematics, 11,* 431-441.

Narayan, P.K., Narayan, S., & Zheng, S. (2010). *Gold and Oil Futures: Are Markets Efficient?. Applied Energy, 87,* 3299-3303.

Nelson, D.B. (1991). *Conditional Heteroscedasticity in Asset Returns: A New Approach. Econometrica, 59,* 347-370.

Ng, V.K., & Pirrong, S.C. (1994). *Fundamentals and Volatility: Storage, Spreads, and the Dynamics of Metals Prices. The Journal of Business, 67,* 203-230.

Officer, R.R. (1973). *The Variability of the Market Factor of the New York Stock Exchange. Journal of Business, 46,* 436-453.

Ozdemir, Z.A., Gokmenoglu, K., & Ekinci, C. (2013). *Persistence in Crude Oil Spot and Futures Prices. Energy, 59,* 29-37.

Reboredo, J.C. (2014). *How Do Oil Prices Co-Move? A Copula Approach. Energy Economics, 33,* 948-955.

Reboredo, J.C., Rivera-Castro, M.A., & Zebende, G.F. (2014). *Oil and US Dollar Exchange Rate Dependence: A Detrended Cross-Correlation Approach. Energy Economics, 42,* 132-139.

Roon, F. de, Nijman, T., & Veld, C. (2000). *Hedging Pressure Effects in Futures Markets. Journal of Finance, 55,* 1437-1456.

Samuelson, P. (1965). *Proof that Properly Anticipated Prices Fluctuate Randomly. Industrial Management Review Spring, 6,* 41-49.

Serletis, A., & Andreadis, I. (2004). *Random Fractal Structures in North American Energy Markets. Energy Economics, 26,* 389-399.

Shiller, R.J. (1989). *Market Volatility.* Massachusetts: The MIT Press.

Shiller, R.J. (2000). *Irrational Exuberance.* Princeton: Princeton University Press.

Sornette, D., Woodward, R., & Zhou, W-X. (2009). *The 2006-2008 Oil Bubble: Evidence of Speculation, and Prediction. Physica A, 388,* 1571-1576.

Summers, L.H. (1986). *Does the Stock Market Rationally Reflect Fundamental Values?. The Journal of Finance, 16,* 591-601.

Tabak, B.M., & Cajueiro, D.O. (2007). *Are the Crude Oil Markets Becoming Weakly Efficient over Time? A Test for Time-Varying Long-Range Dependence in Prices and Volatility. Energy Economics, 29,* 28-36.

Tang, K., & Xiong, W. (2012). *Index Investment and the Financialization of Commodities. Financial Analysts Journal, 68,* 54-74.

Teo, M. (2009). *The Geography of Hedge Funds. Review of Financial Studies, 22,* 3531-3561.

Verma, R., & Verma, P. (2007). *Noise Trading and Stock Market Volatility. Journal of Multinational Financial Management, 17,* 231-243.

Vo, M.T. (2009). *Regime-Switching Stochastic Volatility: Evidence from the Crude Oil Market. Energy Economics, 31,* 779-788.

Wafi, A.S., Hassan, H, & Mabrouk, A. (2015). *Fundamental Analysis Vs Technical Analysis in the Egyptian Stock Exchange – Empirical Study. International Journal of Business and Management Study, 2,* 212-218.

Wang, Y., & Liu, L. (2010). *Is WTI Crude Oil Market Becoming Weakly Efficient over Time?: New Evidence from Multiscale Analysis Based on Detrended Fluctuation Analysis. Energy Economics, 32,* 987-992.

Wang, Y., Wei, Y., & Wu, C. (2011). *Detrended Fluctuation Analysis on Spot and Futures Markets of Western Texas Intermediate Crude Oil. Physica, 390,* 864-875.

Wei, Y., Wang, Y., & Huang, D. (2010). *Forecasting Crude Oil Market Volatility: Further Evidence Using GARCH-Class Models. Energy Economics, 32,* 1477-1484.

Wilson, B., Aggarwal, R., & Inclan, C. (1996). *Detecting Volatility Changes across the Oil Sector. Journal of Futures Markets, 16,* 313-320.

Wong, M.C.S. (1993). *The Impact of Technical Trend Signals on Excess Market Reactions: an Empirical Evidence from the Hong Kong Stock Market* (Asia-Pacific Financial and Forecasting Research Centre).

Working, H. (1948). *Theory of the Inverse Carrying Charge in Futures Markets. Journal of Farm Economics, 30,* 1-28.

Ying, F., & Xu, J-H. (2011). *What Has Driven Oil Prices since 2000? A Structural Change Perspective. Energy Economics, 33,* 1082-1094.

Zhang, Y-J., & Zhang, L. (2015). *Interpreting the Crude Oil Price Movements: Evidence from the Markov Regime Switching Model. Applied Energy, 143,* 96-109.

Zunino, L., Zanin, M., Tabak, B.M., Pérez, D.G., & Rosso, O.A. (2010). *Complexity-Entropy Causality Plane: A Useful Approach to Quantify the Stock Market Inefficiency. Physica, 389,* 1891-1901.