DEPARTMENT OF ECONOMICS – NTNU, DECEMBER 2011

Implied Risk-Neutral Densities

An application to the WTI Crude Oil market

Morten Bergendahl Nitteberg

Acknowledgments

I would like to show my gratitude to my supervisor Sjur Westgaard, whose guidance and support have helped me through this valuable process. I would also like to thank my classmates for making the time at NTNU an enjoyable experience. Lastly, I would like to offer my gratitude to all of those who have supported me in any respect during the realization of this project.

Morten Bergendahl Nitteberg, November 2011

Table of Contents

Table of Figures

1. Introduction

In this paper, I'm going to use Options on West Texas Intermediate (WTI) Crude Oil Futures, to obtain information on market sentiments regarding the WTI Oil price distribution at the options maturity. This will be done by extracting risk-neutral density functions (RNDs) for the future oil price, implied by the observed Option prices. Although the concept of estimating implied risk-neutral densities from Option prices has become more popular the last fifteen years, there have not been done much published research on the Oil market. Working on this paper, I have been lucky enough to have access to a unique and complete data set consisting of daily price observations for European WTI Crude Oil Options traded at the New York Mercantile Exchange (NYMEX). This has made it possible for me to extract and analyze implied RNDs from the WTI Oil market during two different time periods, characterized by considerably different market conditions. I find that the estimated densities are consistent with probability theory, and able to adapt to market conditions at the time the Options were traded. Furthermore, all densities are positively skewed, which can be explained by an inverse leverage effect in the oil market.

The main motivation for extracting risk-neutral densities from option prices is the unique information it might give us about market sentiments, and can therefore be a nice supplement to other sources of information. Financial market participants and decision makers want information that can help them gain insight into future financial and economic development. One way to get this kind of information is to use asset and derivative prices as market indicators. Prices depend on the expected future value of the asset, therefore, by observing market prices of assets and put these in the context of the current market situation and historical economic development, we get a general overview of the market's expectations of the future. Analyzing the development in spot and futures prices for oil might tell us something about expectations of the future. If we want to know more about the uncertainty in the future distribution of the Oil price, we could estimate the future oil price volatility. The traditional approach to this problem would be to calculate the historic standard deviation of returns. A disadvantage with this method is that the past volatility is not necessary the same as the future volatility. As we will see later in this paper, the return volatility can vary quite much under different market conditions. By extracting probability densities from Option prices, we are able to obtain information about the markets expected future oil price distribution. We get information about the expected uncertainty in the future Oil price and by analyzing changes in the characteristics (moments) of the implied density we might be able to catch changes in market expectations. This is helpful when we want to estimate the future volatility and it gives us another perspective on the uncertainty in the market. Implied probability densities can also be a useful tool to help gain insight into market views on political or economic events, to help manage risk, or to price exotic derivatives. In a note from Bank of England (2000), they write that the Monetary Policy Committee is provided with information from option markets to quantify market uncertainty about the future course of financial asset prices, and information extracted using implied RND analysis is used in the banks Inflation Report (Clews, Panigirtzoglou, & Proudman, 2000). They point out that implied RNDs have proved useful in estimating the markets assessment of the balance of risk associated with future price movements. On their homepage, Bank of England write that the Macro Financial Analysis Team Division estimate probability density functions from the prices of options on both equity futures and interest rate futures contracts (Bank of England, 2011). Another way implied RND analysis can be valuable is that it quantifies market assessments implied by option prices. This way, changes in the implied density might give us insight into how the market reacts to major news, e.g. political or economic events.

Prices of Options prove to be a rich source of information about market sentiments. Because of their forward-looking structure, these contracts naturally capture information about market expectations in the price. One intuitive way to see this is to look at prices of two European call options on the same asset, with the same time to maturity, but with different strike prices. The price-difference between the two options reflects among other things how likely the market think it is that the price of the underlying asset lies between the two strikes at maturity. In order to price an Option one has to make assumptions about how the underlying asset is distributed in the future. In the famous Black & Scholes Option pricing model, this is done by assuming that the asset price follows a Geometric Brownian Motion with constant drift. Thus it is assumed that the underlying asset is log-normally distributed over discrete periods of time, and that the continuously compounded returns are normally distributed. Furthermore, the volatility of the asset`s continuously compounded returns is assumed to be constant. Together with an assumption of perfect capital markets, Black & Scholes proved that continuous-time dynamic delta hedging allows us to value an Option as if the word is "risk-neutral" (McDonald 2006). The implication is that by using this model, we do not need to take into account attitude towards risk, and this is why we talk about risk-neutral probability densities. Probability densities for the underlying asset implied by this model are always lognormal, and are not affected by risk preferences. Given that the assumptions of the model were true, the shape of the assets probability density should only be affected by the underlying asset`s return volatility and expected value. However, it is well documented that the real world is not consistent with the assumptions in the Black & Scholes model. Capital markets are not perfect, and asset prices do not behave like a geometric Brownian motion. Empirical studies show among other things that financial asset return distributions are characterized by high peaks and fat tails relative to the normal distribution. Still, the Black & Scholes framework is widely used. Derivative traders have known about this inconsistency for a long time and take this into account when pricing Options. By tweaking the input parameters, traders are able to adjust the price until it reflects their best estimate for the risk associated with the underlying asset. One could say that traders use the incorrect input, in a faulty model, to get what they believe is the correct price. And if not the Black & Scholes model is used, more advanced pricing frameworks are able to do the same job. The implication is that we have Option prices that are inconsistent with the assumptions behind the Black & Scholes framework, which leads to implied RNDs that differ from the lognormal distribution.

One way we can observe this inconsistency is through the Black & Scholes implied volatility. Black & Scholes implied volatility is defined as the volatility that yields a Black & Scholes price equal to the observed market price. If we extract implied volatilities from observed Option prices we will usually get something called a volatility smile, or volatility skew. This means that the implied volatility is shaped as a smile in the volatility-strike space. This is a well-known characteristic of option prices and has been observed in the option markets since the crash of 1987 (Hull, 2008). In fact, when implied volatilities are not constant across all strikes, the implied risk-neutral density will be different from the lognormal distribution (Taylor, 2007). The approach used to extract implied RNDs in this paper take advantage of this link. Based on a theoretical result by Breeden and Litzenberger (1976), we are able to convert implied volatilities into implied risk-neutral probabilities. The main obstacle we face is the discrete intervals for which Options are traded. One way or another we must estimate the probability density between the discrete observations. Different empirical methods have been used in the literature to overcome this problem. This paper will focus on a technique that solves this issue by smoothing the implied volatility smile curve to estimate the volatility for all possible strikes. In this paper we use the Stochastic Volatility Inspired (SVI) parameterization of volatility surfaces introduced by Gatheral (2004). This method is chosen partially because of its good track-record for being able to fit observed volatility smiles and because it is based on a solid theoretical fundament that is supposed to ensure absence of arbitrage in the results. Another motivation for using the SVI parameterization is that we have not seen this method used for this purpose in other papers.

I have two goals writing this paper. First, this paper should be viewed upon as an attempt to gain an understanding of both the theoretical and practical aspect of extracting predictive densities from option prices. Secondly, I want to see if I am able to find evidence of market sentiments in the implied densities extracted from WTI Crude Oil Options. This means that the implied densities should be able to reflect the general attitude towards the future price development in the oil market. As mentioned earlier, the price which options are trading at is affected by traders beliefs of the future, and these beliefs is affected by a variety of factors, including economic news, supply and demand, and world events. Market sentiments is the accumulation of all these factors, and should result in implied densities which differ from the lognormal distribution. Further we expect the extracted densities to be able to adapt to characteristics specific to the oil market, and changes in the attitude towards to the future oil price. There have not been done much research on RNDs in the Oil market which makes this contribution relevant. I will continue this paper by first presenting earlier research on the topic, and explain why the SVI based approach to extracting implied RNDs was chosen. Part two will present theory and methodology, by deriving the Breeden and Litzenberger result, and prove the link between this result and the risk-neutral density. In the last part I will discuss the implementation and performance of the SVI approach, and discuss its ability to capture information that is implicit in Crude Oil Option prices.

2. Earlier Research

There have not been done much research on risk-neutral densities in the Oil market. The main contribution to this topic was made by Melick and Thomas (1997). They used a three-lognormal mixture method to extract implied RNDs from Crude Oil Option prices observed during the Persian Gulf crisis of 1991. They found that sentiments in the Option markets at the time were consistent with the media commentary, in that they reflected at significant uncertainty in future Oil prices. The implied densities extracted did provide evidence that suggested a relatively high probability for a large increase in the future Oil price were included in the observed prices, which can be explained by fear of a major disruption in the Oil market. Positively skewed probability densities with a long right tail are also a sign of an inverse leverage effect in the Oil Market. Melick and Thomas also found evidence that confirmed relatively large shifts in market expectations at days when significant crisis-related events struck. More research has been done on implied volatility in the oil market, e.g. Doran and Ronn (2006). They looked at, among other commodities, ten years of Crude Oil price data from NYMEX, and found evidence suggesting that the implied volatility is a biased, but efficient predictor of future realized volatility (Doran & Ronn, 2006).

Following Melick and Thomas there are no major published research papers on RNDs extracted from Oil markets as far as I know. This makes new research on the topic relevant. Having access to daily data of European WTI Crude Oil Option prices traded at NYMEX in the period from 2004 to the spring of 2011 makes it interesting to see if we are able to find evidence of market sentiments and an inverse leverage effect in this data.

2.1 Methods used

Different techniques have been used in earlier research to extract implied probability densities from Option prices. As Taylor (2007) points out, the common goal for all of them is that the problem they must solve is to find a RND whose corresponding Option prices are an acceptable approximation of the prices observed in the market. At the same time they must solve the critical issue of discrete strikes by one way or another interpolate inside and extrapolate outside the range of traded strikes. Although many different approaches have been used, the most popular ones can roughly be divided into two branches. The first branch is methods that assume a particular parametric form for the distribution of the future asset price and solve for the unknown parameters. The other branch consists of methods where the implied volatility smile is interpolated and extrapolated and directly converted into a RND.

Methods in the first branch assume a parametric form for the future asset price. This is a general approach since no assumptions are made about the stochastic process of the underlying. The chosen parametric form is then estimated from the observed prices. The most popular structural form used in the literature seems to be a mixture of lognormal distributions. Two papers which uses this approach is Melick and Thomas (1997), which use a mixture of three lognormal distributions to extract RNDs from crude oil Options, and Bahra (1997) which prefer to use a mixture of two lognormal distributions on LIFFE equity and interest rate options. Other types of parametric specifications of the RND used in the literature are the general beta distribution of the second kind (GB2) or densities from stochastic volatility processes (Taylor, 2007).

The approach used in this paper is among the methods in the latter branch, which involves smoothing the implied volatility function. The implied volatility function is interpolated and extrapolated in some way, e.g. by fitting a parametric form. Then a theoretical result by Breeden & Litzenberger allows us to convert the implied volatility function into an implied risk-neutral density. Shimko (1993) was among the first to use this approach. He assumed a quadratic form to interpolate the implied volatility function. Outside the observed strikes, he attached lognormal tails. Some years later started a trend of using splines to smooth the implied volatility function. Campa et al. (1998) chose to use a cubic spline, while Bliss and Panigirtzoglou (2000) use a natural spline. Splines are harder to use, but in turn they will most often give a better fit to the data compared to the quadratic function.

From the literature, the two most popular methods seems to be the two-lognormal approach of the former branch, and the method of smoothing the IV function with a spline from the latter branch. There are a few papers that compare these two approaches on their stability and ability to capture the wanted features implied in the data. Cooper (1999) compares the two lognormal-mixture to a implied volatility smile approach with a cubic spline. He uses the Heston Stochastic Volatility model to simulate Option prices based on a known distribution, and is thereby able to test how well the true RND is reflected in the extracted implied RND. He finds that the mixture lognormal approach is especially unstable with short maturities, and overall the smile based approach seems to better capture the first two moments of the true RND. Another comparison of the two lognormal mixture approach and a smoothed IV smile approach using a natural spline is done in Bliss and Panigirtzoglou (2000). Their results provide evidence that the former method is by far the most stable and robust of the two. In a note from Bank of England in 2005 they point out that the smile based technique is an improvement upon the parametric one (lognormal mixture) that has been used at the Bank earlier (Clews, Panigirtzoglou, & Proudman, 2000).

2.2 Why chose the SVI smile based approach?

The method we choose should be sufficiently flexible to capture the features of the density implicit within Option prices, e.g. leptokurtosis properties. Further it should be in compliance with theoretical properties of probability densities; it must never be negative, it should sum up to one, and must be defined for all possible strikes. Based on my impression from the earlier research mentioned above, it seems that a smile based approach is the best choice. Depending on the chosen method for smoothing the smile, this approach can be easy to implement, and research provides evidence that it can be both stable and robust (Cooper, 1999). It does not assume anything about the process of the underlying asset and should not be restrictive in possible functional forms.

As mentioned earlier, the critical part of this approach is the interpolation and extrapolation of the implied volatility smile. Shimko (1993) use a quadratic function to solve this problem. An advantage with this parameterization is that with only three parameters to estimate, it is very easy to implement. A disadvantage is that densities can become negative and thus not in compliance with standard properties of probability. It is also possible to obtain a better fit to observed volatilities with other methods (Taylor, 2007). Bahra (1997) suggests that a cubic spline give a good fit, and finds the quadratic function to be somewhat restrictive. Overall, splines seem to yield a better fit than the quadratic function, and they do not have a problem with negative probabilities (Taylor, 2007). But in turn they are harder to implement.

The stochastic volatility inspired parameterization requires only five parameters to be estimated, thus it should be relatively easy to implement. It is based on a nice theoretical fundament which relies on the absence of arbitrage. According to Gatheral (2004) the SVI approach will in practice very seldom produce results that imply vertical arbitrage, negative probabilities should therefore not occur. Gatheral also points out that this method usually yields a good fit to observed smiles. Based on my impressions from earlier research, the smile based approach, combined with the arbitrage-free theoretical fundament of the SVI parameterization makes this method seem like a good approach for extracting implied RNDs from option prices.

3. Theory and Method

In this part I will present theory related to the topic. I will start by defining the risk-neutral density function. Then I will explain the concept of how implied RNDs can be extracted from implied volatility smiles. To do this, we first need to derive the Breeden and Litzenberger result which all RND approaches are based on and show how this result is connected to the risk-neutral density. At last we will take a look at the SVI Parameterization by Gatheral.

3.1 The Risk-Neutral Density Function

This section is based on chapter 16 in Taylor (2007). The definition of a risk-neutral density function f_Q in this paper is the density for which European option prices are the discounted expectations of final payoffs; thus

$$C(X) = e^{-r_f T} \int_0^\infty \max(x - X, 0)\, f_Q(x) dx \qquad \text{Equation 1}$$

for all exercise prices X≥0. The standard properties for densities are required:

- f_Q is defined for all x≥0
- $f_Q \geq 0$
- $\int_0^\infty f_Q(x) dx = 1$

Taylor points out that the expected value of the underlying asset under the risk-neutral measure must be the forwards price F. Proof:

Assume we have a call option with an exercise price X equal to zero. Equation 1 becomes:

$$C(0) = e^{-r_f T} \int_0^\infty x f_Q dx \qquad \text{Equation 2}$$

This option is guaranteed to be "in the money". At maturity we exercise the option and buy the underlying. This contract has the same properties as a forward contract. With both contracts we buy the asset at maturity of the contract. The main difference is that we pay the option premium C(0) today instead of paying the forward price F at maturity. A no-arbitrage condition must be:

$$C(0) = e^{-r_f T} \int_0^\infty x f_Q dx = e^{-r_f T} E^Q(S_T) = e^{-r_f T} F \qquad \text{Equation 3}$$

$$F = E^Q(S_T) \qquad \text{Equation 4}$$

3.2 The Breeden and Litzenberger result

All methods for extracting implied risk-neutral densities mentioned in this paper is based on the paper *Prices of State-Contingent Claims Implicit in Option Prices* by Breeden and Litzenberger (1976). They managed to derive a valuation formula for elementary claims using European Options. This result can be used to show a link between the price of European Options and an implied RND function for the underlying asset.

The elementary claim was introduced in the time-state preference framework which can be traced back to the work of Arrow (1964) and Debrau (1959). While developing the theory of complete markets, which is central to the time-state preference framework, they introduced a state contingent security called an elementary claim. This security pays $1 contingent upon the realization of a particular state of the world at a given future date, and pays out nothing if that particular state is not realized. Breeden and Litzenberger show in their paper from 1976 that under an assumption of perfect capital markets, one are able to replicate the payoff of an elementary claim by entering into a Butterfly Spread using Options on the underlying asset. This means, if you want to replicate an elementary claim paying $1 if state X is realized, then you can do so by going long two Call Options, one with strike X-1 and one with strike X+1, and going short two Call Options with strike X. A simple illustration of this strategy for X equal 8 can be seen in the Figure 1 underneath.

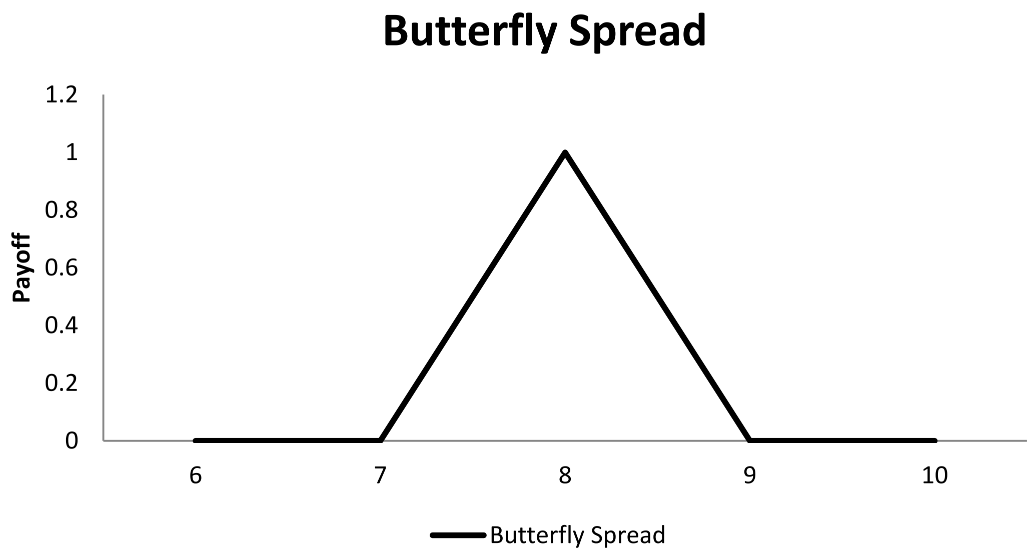

Figure 1: Butterfly Spread

Breeden and Litzenberger show that the payoff of an elementary claim for any given level of the asset at time T can be replicated in a similar manner using a Butterfly Spread strategy. A generalized expression for this portfolio on "state" X when the step size between possible asset prices at time T is ΔX becomes:

$$\frac{1}{\Delta X} * \{[C(X-\Delta X) - C(X)] - [C(X) - C(X+\Delta X)]\} \qquad \text{Equation 5}$$

A no-arbitrage condition must then be that the price of an elementary claim paying $1 if state X is realized is given by Equation 5. The price of an elementary claim divided by the step size ΔX is then given by:

$$\frac{p_X}{\Delta \mathrm{X}} = \frac{1}{(\Delta X)^2} * \{[C(X-\Delta X) - C(X)] - [C(X) - C(X+\Delta X)]\} \qquad \text{Equation 6}$$

In the limit as the step size tends to zero, Breeden and Litzernberger show that this expression becomes:

$$\lim_{\Delta X \to 0} \frac{p_X}{\Delta X} = \frac{\partial^2 C(X)}{\partial X^2} \qquad \text{Equation 7}$$

With Equation 7, Breeden & Litzenberger proved that in a continuous setting the price of an elementary claim is equal to the second derivative of the Call pricing function with respect to the strike price. To establish the link between Equation 7 and the risk-neutral density we can take advantage of the fundamental theorem of asset pricing.

3.3 The risk-neutral measure

Suppose we have a market consisting of a number of securities and a risk-free bond. Then the first and second fundamental Theorem of Asset Pricing states (Harrison & Pliska, 1981):

1) The market is arbitrage-free[1] if and only if there exist at least one equivalent martingale measure (EMM) **Q**.
2) The market is complete if and only if the equivalent martingale measure (EMM) **Q** is unique.

A process is a Martingale if the expected future value, conditional on the past, is its current value. An equivalent martingale measure[2] is a probability measure Q**,** which is equivalent[3] to the real probability measure P**,** under which the bond discounted securities are all Q-martingales. Under the EMM Q all assets have the same expected growth rate as the risk-free bond, independent of the risk associated with the asset. Thus, the Q-measure is neutral with respect to risk (Baxter & Rennie, 1996). In other words, the expected value of any asset under the EMM Q, discounted at the risk-free rate, is equal to today`s asset value. The equivalent martingale measure is just a more accurate name for the well-known risk-neutral measure.

---

[1] A market is arbitrage-free if there is no way of making riskless profits. Riskless profit would be an investment opportunity with zero outlay today, and a positive value at termination with probability one (Baxter & Rennie, 1996).

[2] A Martingale Measure is a measure under which a process is a Martingale (Baxter & Rennie, 1996).

[3] Two measures **P** and **Q** are equivalent if they operate on the same sample space and agree on what is possible (Baxter & Rennie, 1996).

Using the fundamental theorem of asset pricing we know that in a complete market, the price of an elementary claim paying $1 in state S_T is given by the expected payoff under the risk-neutral measure, discounted at the risk-free interest rate, see Equation 8.

$$P_{S_T} = e^{-r_f T} E^Q[\$1] = e^{-r_f T} f_Q(S_T) \qquad \text{Equation 8}$$

Using Equation 7 and Equation 8 we see that the second derivative of the call pricing function with respect to the exercise price is equal to the discounted risk-neutral density (RND) function.

$$\frac{\partial^2 C(X)}{\partial X^2} = e^{-r_f T} f_Q(S_T) \qquad \text{Equation 9}$$

$$f_Q(S_T) = e^{r_f T} \frac{\partial^2 C(X)}{\partial X^2} \qquad \text{Equation 10}$$

Equation 10 verifies the link between the risk-neutral density function and the call pricing function. It states that the risk-neutral density function for the underlying asset is equal to second derivative of the Call pricing function with respect to the strike price, adjusted with the risk free rate of return. Since the Black & Scholes Option pricing function is twice differentiable with respect to the strike price, we can use this model combined with Equation 10 to translate implied volatilities into probability densities. The only assumption made deriving this result is that there are perfect capital markets. Equation 10 does not rely on any assumptions about the process of the underlying asset.

To illustrate how Equation 10 is used in practice, I will in the following example use it to calculate the risk-neutral density for a given strike, assuming the assumptions behind the Black & Scholes model are true. The parameters used are based on WTI Crude Oil data from 1 April 2010, but assuming a constant volatility of 28%, which is the “at the money” implied volatility. We see from Equation 10, that in order to find the density, we need the second derivative of the Black & Scholes Call pricing function with respect to the strike, also called strike gamma. Assuming constant volatility, the strike gamma is given by (Haug, 2006):

$$\frac{\partial^2 C}{\partial X^2} = \frac{n(d_2)e^{-rT}}{X\sigma\sqrt{T}} \qquad \text{Equation 11}$$

Combining Equation 10 and 11, we get the following expression for the RND assuming constant volatility:

$$f_Q(S_T) = \frac{n(d_2)}{X\sigma\sqrt{T}} \qquad \text{Equation 12}$$

The RND is given by the standard normal probability mass function of $d_{2,}$ divided by the strike multiplied by the volatility and the square root of time to maturity. The parameter d_2 is the standard parameter from the Black and Scholes formula. When the underlying is a Futures contract, d_2 is given by (Haug, 2006):

$$d_2 = \frac{\ln\left(\frac{F}{X}\right) - (b + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \qquad \text{Equation 13}$$

On 1 April 2010, we have the following parameters:

Date	4/1/2010	**Risk-free Interest**	0.2915%
Maturity	5/18/2010	**Cost of Carry**	0
Futures Price	$85.34	**Implied Volatility**	28%
Time to maturity	0.12877		

Table 1

Using these parameters amd Equation 12, we get the following level of the risk-neutral density for a future oil Futures price of $100:

d2	-1.628	**S_T**	$100
n(d2)	0.106019	**B&S RND**	0.010552

Table 2

Increasing the implied volatility to 35%, yields the following result:

d2	-1.32501	**S_T**	$100
n(d2)	0.165835	**B&S RND**	0.013204

Table 3

We see that with a volatility of 28%, the density for a Futures price of 100 at maturity is 0,010552. Increasing the volatility to 35%, we get a density of 0,013204. Higher volatility means that larger price fluctuations are more likely, which results in a higher probability for the Futures price to be $100 at maturity.

3.4 RNDs from Implied Volatility Functions

We are now going to see how we can extract implied RNDs from the implied volatility smile. Based on the Breeden and Litzenberger result and the fundamental theorem of asset pricing, we were able to derive a solution for the implied RND, given by Equation 10. The only requirement we need to find the implied density is that the call pricing function is twice differentiable. Assuming a constant volatility, which we used in the example above, we used the second derivative of the Black & Scholes Call pricing formula and got an expression for the RND given by Equation 12. From the discussion earlier in this paper, we know that this is an unrealistic assumption. We know that implied volatilities differ across the traded strikes, and the result is risk-neutral densities that deviate from the lognormal distribution. We could use Equation 12 to calculate the implied density for the traded strikes by using the implied volatility observed at each particular strike. The problem is that to get a continuous probability density, we need to calculate the RND at every possible strike level, and thus we need to estimate the implied volatility between and outside the discrete observations.

To solve this issue, we assume that the implied volatility follows a functional form. Fitting a parametric form to the observed smile lets us interpolate and extrapolate the implied volatility smile. Taylor (2007) shows that if we extend the Black & Scholes Call pricing formula to include a volatility that is dependent on the strike, then the second derivative with respect to the strike, combined with Equation 10, yields the following expression for the RND:

$$f_Q(X) = n(d_2)\left\{\frac{1}{\sigma(x)\sqrt{T}} + \left(\frac{2d_1}{\sigma(x)}\right)\frac{\partial\sigma}{\partial X} + \left(\frac{d_1 d_2 X\sqrt{T}}{\sigma(x)}\right)\left(\frac{\partial\sigma}{\partial X}\right)^2 + \left(X\sqrt{T}\right)\frac{\partial^2\sigma}{\partial X^2}\right\} \quad \text{Equation 14}$$

Where $\sigma(x)$ is the volatility given the strike price x, and d_1 and d_2 are given by:

$$d_1(X) = \frac{\ln\left(\frac{F}{X}\right) + \frac{1}{2}\sigma(X)^2 T}{\sigma(X)\sqrt{T}} \qquad \text{Equation 15}$$

$$d_2(X) = d_1(X) - \sigma(X)\sqrt{T} \qquad \text{Equation 16}$$

With Equation 14, we have a solution for the implied RND assuming that the volatility is dependent on the strike. This expression can be used to transfer an implied volatility function into implied risk-neutral probability densities. The last piece of the puzzle is to choose a parametric form to smooth the smile. In this paper, I have chosen to use the stochastic volatility inspired (SVI) parameterization by Gatheral.

3.5 The SVI Parameterization

The stochastic volatility inspired parameterization of the implied volatility surface was first introduced in Gatheral (2004). With only five parameters, this approach is relatively easy to calibrate to the observed volatility smile, and it will according to Gatheral almost never in practice produce results that implies vertical[4] arbitrage (Gatheral, 2004). Gatheral based this method on the implications made by the moment formula for implied volatility at extreme strikes, introduced in Lee (2003). Under a no-arbitrage condition Lee proves a connection between the number of finite moments in the process of the underlying asset and the behavior of the implied volatility function at extreme strikes. Using this result Lee is able to derive no-arbitrage bounds for the tails of the Black & Scholes implied volatility smile. In order to prevent arbitrage, Lee shows that implied variance must always be linear in "log moneyness" when the absolute value of "log moneyness" goes to infinity. "Log moneyness" is defined as the natural logarithm of the exercise price divided by the underlying asset price.

$$k = \ln\left(\frac{X}{F}\right) \qquad \text{Equation 17}$$

Thus, to ensure absence of arbitrage, a parameterization of the implied variance skew must in the variance-"log moneyness" space be linear in the tails and curved in the middle. Gatheral proposes the following expression for the variance as a function of "log moneyness":

[4] Vertical arbitrage is an arbitrage opportunity using options on the same underlying asset with the same time to maturity but with different strikes. E.g. bull, bear and butterfly spreads.

$$var(k) = a + b\left\{\rho(k-m) + \sqrt{(k-m)^2 + \sigma^2}\right\} \quad \text{Equation 18}$$

With the left and right asymptote given by:

$$v_{left} = a - b(1-\rho)(k-m) \quad \text{Equation 19}$$

$$v_{right} = a + b(1+\rho)(k-m) \quad \text{Equation 20}$$

Figure 2 illustrates how the SVI variance behaves in the log-moneyness space. “Near the money”, the function is curved to fit the implied skew. Moving away from the money the variance function converges to the left and right asymptote.

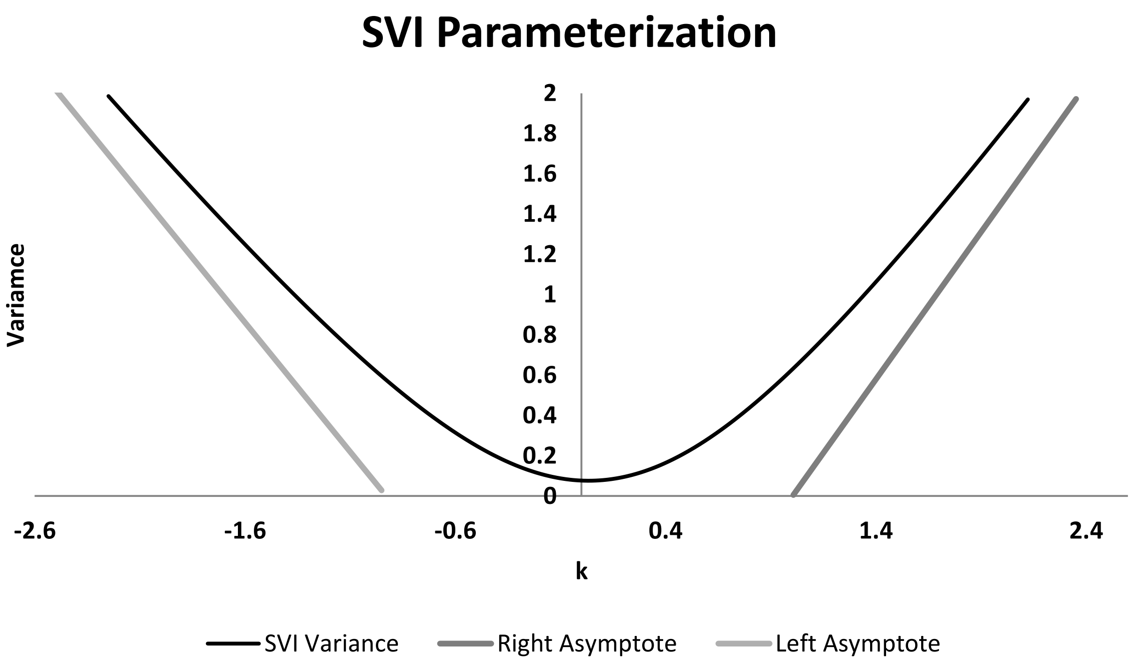

Figure 2: SVI Parameterization

The SVI parameters can be given the following interpretation (Gatheral, 2004):

- a: The overall level of the variance
- b: The angle between the left and right asymptote
- σ: Smoothness of the middle curvature
- ρ: Orientation of the graph

- m: Level along the x-axis

3.6 Summary of the method derived in this section

Based on the Breeden and Litzenberger result and the fundamental theorem of asset pricing, we found that the implied risk-neutral density for the underlying asset is given by the second derivative of the call pricing function with respect to the exercise price, adjusted by the risk free rate, see Equation 10:

$$f_Q(S_T) = e^{r_f T} \frac{\partial^2 C(X)}{\partial X^2} \qquad \text{Equation 10}$$

Assuming that the volatility is dependent of the strike level, then the second derivative of the Black & Scholes Option pricing formula with respect to the strike price combined with Equation 10 becomes:

$$f_Q(X) = n(d_2)\left\{\frac{1}{\sigma(x)\sqrt{T}} + \left(\frac{2d_1}{\sigma(x)}\right)\frac{\partial\sigma}{\partial X} + \left(\frac{d_1 d_2 X\sqrt{T}}{\sigma(x)}\right)\left(\frac{\partial\sigma}{\partial X}\right)^2 + \left(X\sqrt{T}\right)\frac{\partial^2\sigma}{\partial X^2}\right\} \qquad \text{Equation 14}$$

Using Equation 14, we can calculate the implied RND directly, given that we have a twice differentiable function for the volatility with respect to the strike. We solve this problem by fitting the stochastic volatility inspired parameterization by Gatheral to the observed volatility smile, se Equation 21. The first and second derivative of Equation 21 with respect to the strike can be seen in the appendix.

$$\sigma(k) = \sqrt{a + b\left\{\rho(k-m) + \sqrt{(k-m)^2 + \sigma^2}\right\}} \qquad \text{Equation 21}$$

To calculate the implied RND, we follow almost the same approach as in the simple example discussed earlier. The main difference is that we first have to calibrate Equation 21 to the implied volatility smile. Then we use the calibrated parameters from Equation 21, together with Equation 14, to solve for the implied RND.

4. Implementation and Discussion

I will start this section by presenting the data used in this paper. Then I will take a closer look at how the method of extracting implied RNDs by smoothing the implied the volatility smile with the SVI parameterization performs. This will involve discussing how well we are able to interpolate and extrapolate the implied volatility smile found in the WTI Crude Oil data. A quadratic function, following the approach in Taylor (2007) will be used as comparison. At last I will discuss implied RNDs extracted from the Oil data from two time periods characterized by different market conditions.

4.1 The WTI Crude Oil Option Data

Working on this paper I have had access to a unique data set consisting of settle prices for European Light Sweet Crude Oil (WTI) Options traded at the New York Mercantile Exchange (NYMEX). The underlying contract of these options is the Light Sweet Crude Oil (WTI) Futures traded at the same exchange. The WTI-Futures contract is the most liquid Oil-trading instrument available in the industry today. About 600,000 WTI Options or Futures contracts are traded daily, which translates to a volume of 600 Million barrels of Oil. This makes WTI Crude Oil an important benchmark in the Oil industry, and serves as a hedging tool for hundreds of commercial oil companies (CME Group). Over the last hundred years, Oil has become a commodity we are completely dependent on, and large fluctuations in the Oil price affect the whole world economy.

The data set consists of settle prices for all European WTI Option contracts traded at NYMEX between 2004 and the spring of 2011. The settle price is decided by a settlement committee and is based on the last minutes of trading. The settlement committee establish the at the money volatility and create the volatility surface for out of the money Options based on traded outrights and spreads (CME Group). The observed price is therefore affected by a minimal of noise. In addition to data on Oil, the 3 Month LIBOR rate were used as the “risk-free” interest rate. Option Data was obtained from CME Group, while close prices for the underlying Futures contract and 3 Month LIBOR rate were obtained from Reuters EcoWin.

Trade Date	Contract	Call / Put	Month	Year	Strike Price	Settle Price	Open Interest	IV	Exchange
4/1/2010	LC	C	6	2010	60.0000000	25.36000000	850.000000	.411600	NYM
4/1/2010	LC	C	6	2010	62.0000000	23.37000000	25.000000	.391100	NYM
4/1/2010	LC	C	6	2010	66.0000000	19.42000000	600.000000	.363500	NYM
4/1/2010	LC	C	6	2010	68.0000000	17.47000000	375.000000	.348300	NYM
4/1/2010	LC	C	6	2010	70.0000000	15.54000000	550.000000	.337600	NYM
4/1/2010	LC	C	6	2010	72.5000000	13.16000000	100.000000	.323200	NYM
4/1/2010	LC	C	6	2010	73.5000000	12.23000000	500.000000	.315200	NYM
4/1/2010	LC	C	6	2010	74.5000000	11.32000000	300.000000	.310700	NYM
4/1/2010	LC	C	6	2010	76.0000000	9.97000000	50.000000	.301200	NYM
4/1/2010	LC	C	6	2010	76.5000000	9.53000000	410.000000	.298100	NYM
4/1/2010	LC	P	6	2010	80.0000000	1.38000000	477.000000	.287800	NYM
4/1/2010	LC	P	6	2010	82.0000000	1.97000000	50.000000	.282400	NYM
4/1/2010	LC	P	6	2010	83.0000000	2.33000000	44.000000	.279900	NYM
4/1/2010	LC	P	6	2010	85.0000000	3.18000000	300.000000	.274600	NYM
4/1/2010	LC	P	6	2010	86.0000000	3.67000000	14.000000	.271500	NYM
4/1/2010	LC	P	6	2010	87.0000000	4.22000000	100.000000	.269500	NYM
4/1/2010	LC	P	6	2010	120.0000000	34.67000000	350.000000	.366500	NYM
4/1/2010	LC	P	6	2010	125.5000000	40.16000000	400.000000	.380900	NYM
4/1/2010	LC	P	6	2010	130.0000000	44.65000000	300.000000	.411600	NYM

Table 4

Table 4 shows a small extraction from the raw data obtained from CME Group. This is only a small part of the total data, which consists of about 600 000 price observations. In the first column we see the trade date, the day these prices were observed. The next column tells us that these are European Options on WTI Crude Oil Futures, and the third column distinguishes between Call and Put Options. The next two columns tells us when the underlying Futures contract expires. Then the strike and settle price follows before the open interest in each contract. The second last column is implied volatility calculated by CME Group. Since implied volatility is only included in the newest part of the data, and because CME Group assumes zero interest in their calculation, I have calculated implied volatility myself for the data used.

Light Sweet Crude Oil (WTI) Future:

- Underlying Commodity: WTI Light Sweet Crude Oil
- Settlement Type: Physical
- Expiration: Trading ceases on the third business day prior to the twenty-fifth calendar day of the month preceding the delivery month.

European Light Sweet Crude Oil (WTI) Options:

- Underlying Contract: Light Sweet Crude Oil (WTI) Future
- Settlement Type: Financial
- Expiration: The Option expires three business days before trading ceases in the underlying Futures contract

To ensure a best possible result in the analysis, the following "data wash" were done to clean up the data used. On any given day, out of the money call and put Options with an open interest of less than 100 contracts were removed. All obviously misquoted observations were removed. All observations implying vertical arbitrage were removed, e.g. call options with different strikes quoted with the same settle price. Minimum price fluctuation for WTI options is $0,01 per barrel, thus all observations with quotes lower than $0,05 were removed. The precision in quotes below this threshold is deemed too low. Put-Call parity is used to convert out of the money put prices into the corresponding call prices. After this cleanup was done, a typical day used in the analysis consisted of Call prices for about 25 different strikes.

4.2 Smoothing the Implied Volatility Smile

To extract risk-neutral densities from the implied volatility smile, we need to somehow estimate the implied volatility for all possible strikes. In this section I will use a set of WTI Crude Oil Option prices observed 1 April 2010 as an example. These Options were on the underlying WTI Futures contract with delivery in June 2010, and the exercise date for the Option contracts were on the 17 of May, with 47 days until maturity. I am now going to take a closer look at how the SVI parameterization solves the problem of smoothing the IV smile, by fitting it to the smile extracted from the prices observed 1 April. The next section will continue on the same example, discussing the corresponding risk-neutral density for the WTI Crude Oil price at the 17 of May. To help measure the performance of the SVI Parameterization, I will also use a quadratic function following the approach in Taylor (2007).

4.2.1 Interpolation

When extracting RNDs by smoothing the implied volatility smile, the quality of the results is largely dependent on how we smooth the smile. The first part of this problem is to interpolate inside the range of traded strike prices. Most of the time, almost all the probability mass is located inside this range, making it critical to get a good fit in this part of the process. Thus, we want the fitted function to be as close to the discrete observations as possible. I have solved this problem by minimizing the sum of squared deviations from the market implied volatilities.

$$MIN \sum (\sigma_{mkt} - \sigma_{SVI})^2 \qquad \text{Equation 22}$$

To get an idea what an interpolated smile looks like I have calibrated both the SVI and quadratic function to the smile extracted from the set of WTI Options traded at April 1 2010. The result is displayed in Figure 3 underneath. We see that the quadratic function has problems getting a close fit on the right side of the smile, where the implied volatilities deviate from the quadratic form. On the left side of the smile the quadratic function obtains a better fit. The SVI function, although not perfect on the right side, get a pretty good fit for the whole smile, especially good on the left side. From this example we see that the SVI approach benefits from being able to adapt to different kinds of shapes. It is also interesting to see that we observe a clearly defined smile extracted from our set of Crude Oil Options. Both “out of the money” Call and Put Options with maturity 17 May traded at 1 April had a significantly higher Black & Scholes implied volatility compared to “at the money” Options. This suggests a market expectation of a higher probability for large price movements in both directions compared to the lognormal distribution. This means that we expect to find fat probability tails on both sides of the corresponding risk-neutral density we will look at a little bit later.

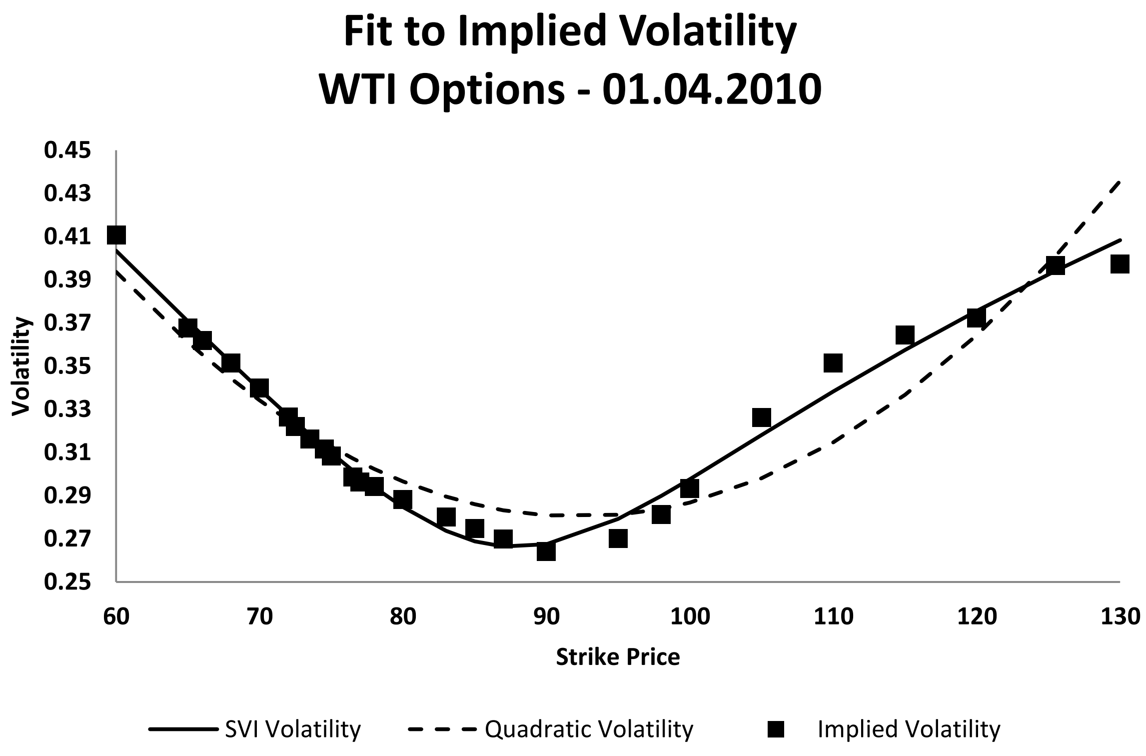

Figure 3: Interpolated Volatility Smile

As mentioned above, the problem of smoothing the smile inside the range of traded strikes are critical. Most of the time, this range will cover a large portion of the implied density. In Figure 4 underneath, I have extracted the implied density based on the SVI function we have fitted in Figure 3. In this case, using strikes ranging from 60 to 130, we have already found 99.4% of the total probability mass. We see that only a very small part of the left tail needs to be estimated before we are very close to 100%. It is a great advantage when the traded range covers most of the implied density, because this part of the density is then based on observed information in the market. Other times we might not be this lucky and a bigger part of the probability mass may be missing. To obtain the missing part of the density, we have to extrapolate the smile outside the range of traded strikes.

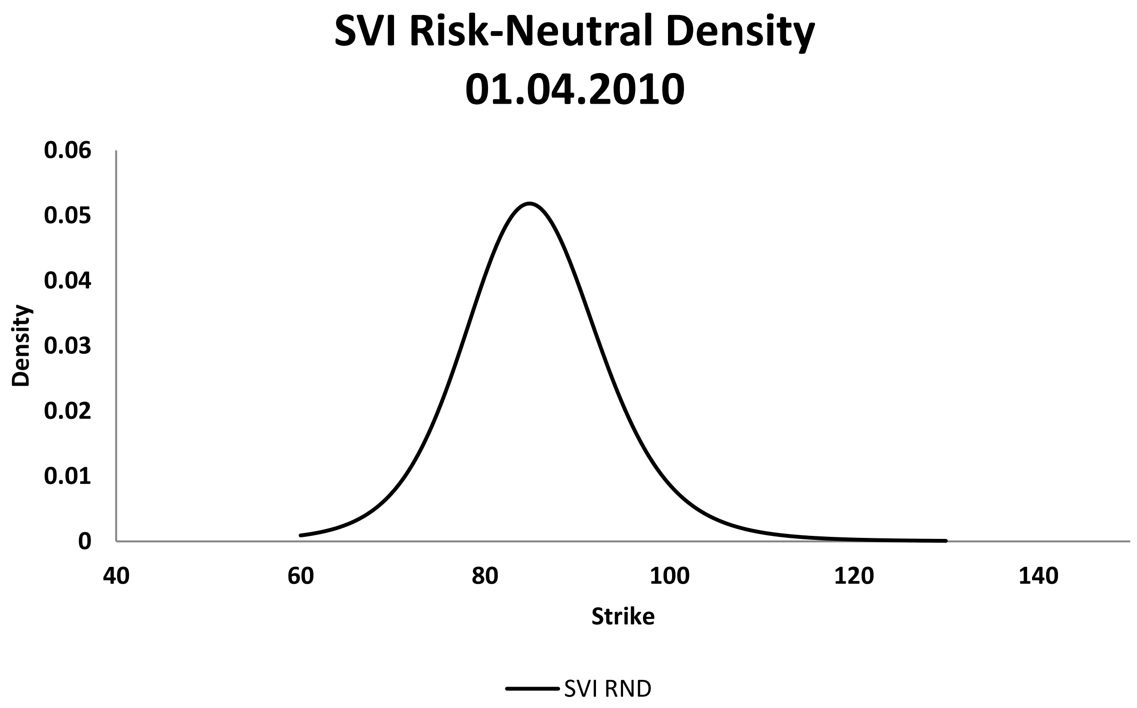

Figure 4: RND - Range of traded strikes

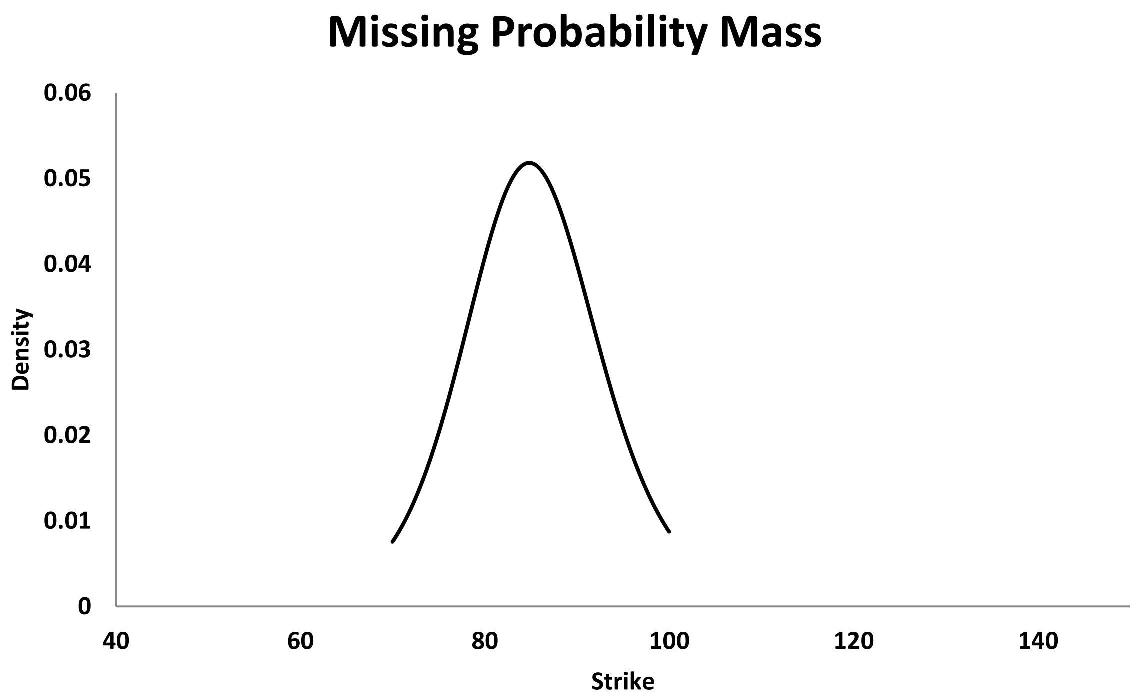

Figure 5: Missing probability mass

4.2.2Extrapolation

Figure 5 illustrates the problem of missing probability mass. For some reason, the available market data only covers a smaller portion of the total RND. The missing part in Figure 5 is exaggerated on purpose to illustrate the problem. In this case we have to extrapolate the implied volatility smile to get an estimate of the missing tails of the distribution. This is a more difficult task because we have little information about market expectations outside the range of traded strikes.

Using the parameter values we obtained when we interpolated the smile, we can continue to estimate the smile outside the range of traded strikes. Doing this for the same smile we interpolated above, we get the result seen in Figure 6. If we look at how the two functions behave outside the observed smile, we see that they estimate the volatility quite differently. On the right side, the SVI function follow the curvature of the smile by flattening out slightly. The quadratic function on the other hand increases fast. The difference is smaller on the left side, but also here we see that the quadratic function increases a bit faster. However, even though we observe that the two functions behave different outside the implied smile, it is difficult to say anything about which method behaves best. To help analyze the behavior, we can display the same information in the implied variance – "log-moneyness" space, see Figure 7. Log-moneyness is as before defined as the natural logarithm of the strike, divided by the forward price. In Figure 7, the curvature we see around log-moneyness equal to zero, is where both functions are fitted to the smile.

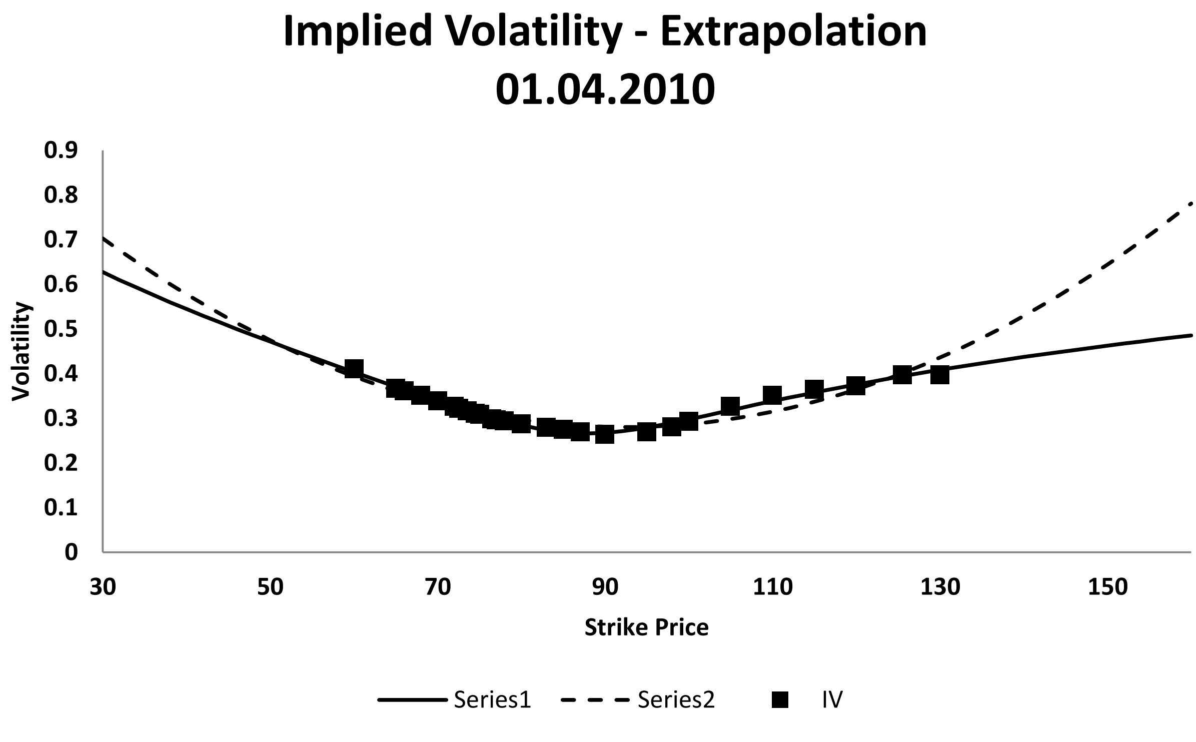

Figure 6: Implied volatility extrapolation

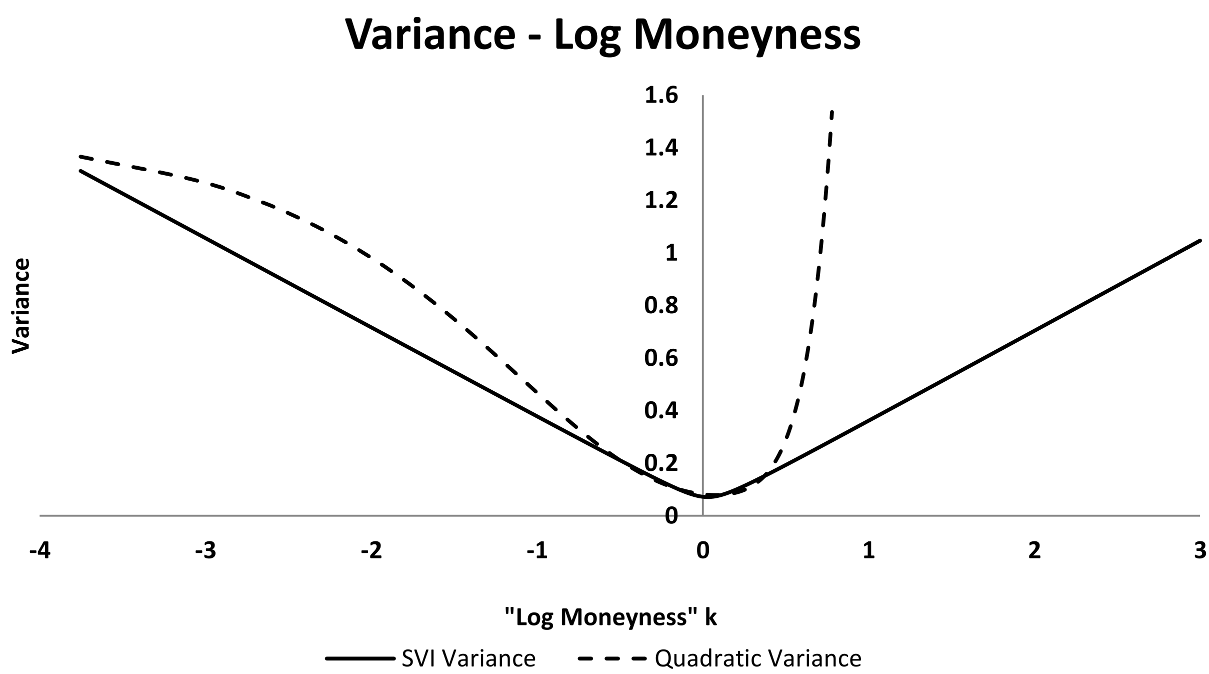

Figure 7: Implied variance - "Log-moneyness"

Figure 7 show us that if we continue to estimate the volatility at more extreme strikes, the two functions behave very different. Outside this curvature in the middle, the implied SVI variance continues in a straight line, while the implied quadratic variance increases very fast on both sides. We already know from the theory section that a required condition to secure absence of arbitrage at extreme strikes is that the slope of the variance function is not steeper than linear in the log-moneyness space (Lee, 2003). In Figure 7 above, the quadratic variance clearly violates this condition. The SVI variance on the other hand adapts to the curvature in the smile, and then converges to the two straight asymptotes on either side. This is likely to be the reason earlier research points out problems with negative probabilities when a quadratic function is used (Taylor, 2007).

Based on what we have seen in this example, I think the SVI parameterization look like a promising solution to the problem of smoothing the implied volatility smile. Since most of the probability mass in the implied density is derived from range of traded strikes, this is the area of the smile that it is most important to get a good fit. On the smile used in the example above, which is a typical smile seen from the data used in this paper, we achieved a relatively good fit using the SVI function. Outside the range of traded strikes, we have seen that the estimated volatility-tails are in compliance the no arbitrage condition at extreme strikes introduced in Lee (2003). This means that the little part of the density we were not able to recover from the interpolated smile will be based on volatilities which should ensure absence of arbitrage. This should ensure that we do not have problems with negative probabilities in the estimated densities.

4.3 The SVI implied Risk-Neutral Density

Having obtained a parameterization of the implied volatility smile observed in the Oil data, we are able to convert the estimated volatility at each possible strike directly into risk-neutral probabilities using Equation 14. Doing this using the SVI parameterization fitted to the smile extracted from Options traded 1 April 2010 which we have discussed above, results in the implied RND seen in Figure 8. The solid black line represents the SVI RND, and for the sake of comparison I have included a lognormal density function using the “at the money” implied volatility of 28% across all strikes.

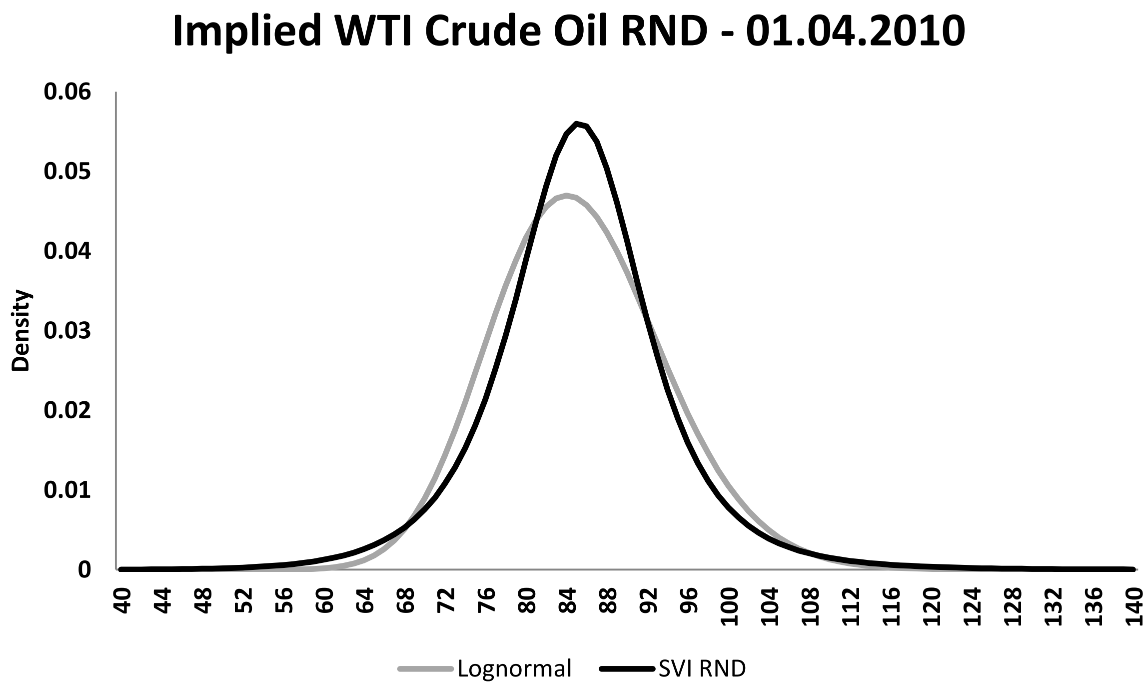

Figure 8: WTI Crude Oil implied RND

	Mean	Volatility	Skewness	Kurtosis
SVI RND	85.34	9.065695779	0.38	6.39
Lognormal RND	85.34	8.574204647	0.30	3.16

Table 5

It is easy to see the leptokurtic properties of the density which is based on the SVI function. Compared to the lognormal density, it got a high peak and fat tails on both sides. The first four moments of the distributions which can be seen in Table 5, confirms what we see in Figure 8. The SVI density has a Kurtosis of 6.39 compared to 3.16 for the lognormal function. Both densities are positively skewed with fatter right tails. The SVI slightly more with a Skewness of 0.38. Both densities are positively skewed compared to the normal distribution, which is symmetrical and thus have a Skewness of 0. This means that the SVI density extracted from Crude Oil Options got a larger part of the probability mass located in the right tail relative to the left. From the second moment we see that the volatility of the SVI density is slightly larger than that of the lognormal, with 9.06 for the former and 8.57 for the latter. The first moment is the same for both densities. We know from the theory presented earlier in the paper that a theoretical constraint for the implied RND, is that the expected value should be equal to the forward price. The price of the Futures contract traded 1 April 2010 with delivery in June was $85.34. Thus, both the SVI density and the lognormal density yield the correct expected value.

It is very interesting to see the high peak and positive skew of the SVI density extracted from our Crude Oil Options. Leptokurtic properties are something we would expect to find, and it is positive that our approach seems to be able to adapt to the leptokurtic properties reflected in our observed prices. Especially interesting is the positive skew. This suggests that a higher chance of large price increases in the Crude Oil price have been included in the traded Option prices. This is in compliance with the theory of an inverse leverage effect in the Oil market. Oil markets have a special characteristic in that volatility and prices are positively correlated, which is opposite of what is observed in equity markets. The theory of a leverage effect in stock markets was first introduced in Black (1976), where he suggested that negative shocks in the stock price usually leads to higher volatility than equally large positive shocks, thus price and volatility are negatively correlated (Black, 1976). The intuition behind the leverage effect is that falling stock prices makes the equity less valuable compared to the debt, increasing the leverage of the firm. This makes the firm a riskier investment, and the result is higher volatility. We could also say it the other way around. Higher volatility increases the risk of the firm, thus we demand a higher expected return to compensate for the increased risk, and prices go down. However, research shows that the opposite relationship between price and volatility is true in energy markets. This could partially be explained by an inelastic short run supply in the Oil market. A positive demand shock results in higher prices due to the time it takes the supply side to adapt to the higher demand. A little available inventory results in higher and more volatile prices. The positive skew observed in the SVI density can therefore be explained by the existence of an inverse leverage effect in the Oil market, and that this effect has been taken into account in the traded Option prices. If the volatility increase when the oil price increase, then the probability for even higher prices increase. This could explain the fat probability tail observed in Figure 8. The same characteristics were found by Melick and Thomas (1997) on the data from the period of the Gulf Crisis in 1991. The implied densities they analyzed suggested a relatively high probability for a large increase in the future Oil price were included in the observed prices, which can be explained by fear of a major disruption in the Oil market. Comparing the Skewness we see in the example above to research done on equity Options supports the theory of an inverse leverage effect. In their research on RNDs from equity markets, Bliss and Panigirtzoglou (2000) found that the extracted densities from FTSE 100 Options had a mean Skewness of -0,54.

This example has showed us that our approach is able to adapt to characteristics that are specific to oil markets. We have found a high peak and positive skew, indicating both leptokurtic properties and an inverse leverage effect in the oil price distribution. We have also seen that the SVI parameterization was able to get a good fit to the implied volatility smile extracted from WTI Crude Oil Options. Furthermore, we found that outside the range of traded strikes, the SVI function estimated a linear variance-tails in the log-moneyness space, which is in compliance with the no-arbitrage condition of Lee (2003). Converting the SVI volatility into a risk-neutral density resulted in a probability density which seems to behave according to probability theory, with positive probabilities for all strikes ranging from 0 to 1000, and the cumulative density did not exceed 1.

4.4 WTI Crude Oil RNDs

In this part I will focus on analyzing RNDs extracted from two different periods where Crude Oil Options were sold under quite different market conditions. Figure 9 displays the WTI Crude Oil price development from June 2007 until April 2011. The two periods used in this analysis in seen as the solid black lines.

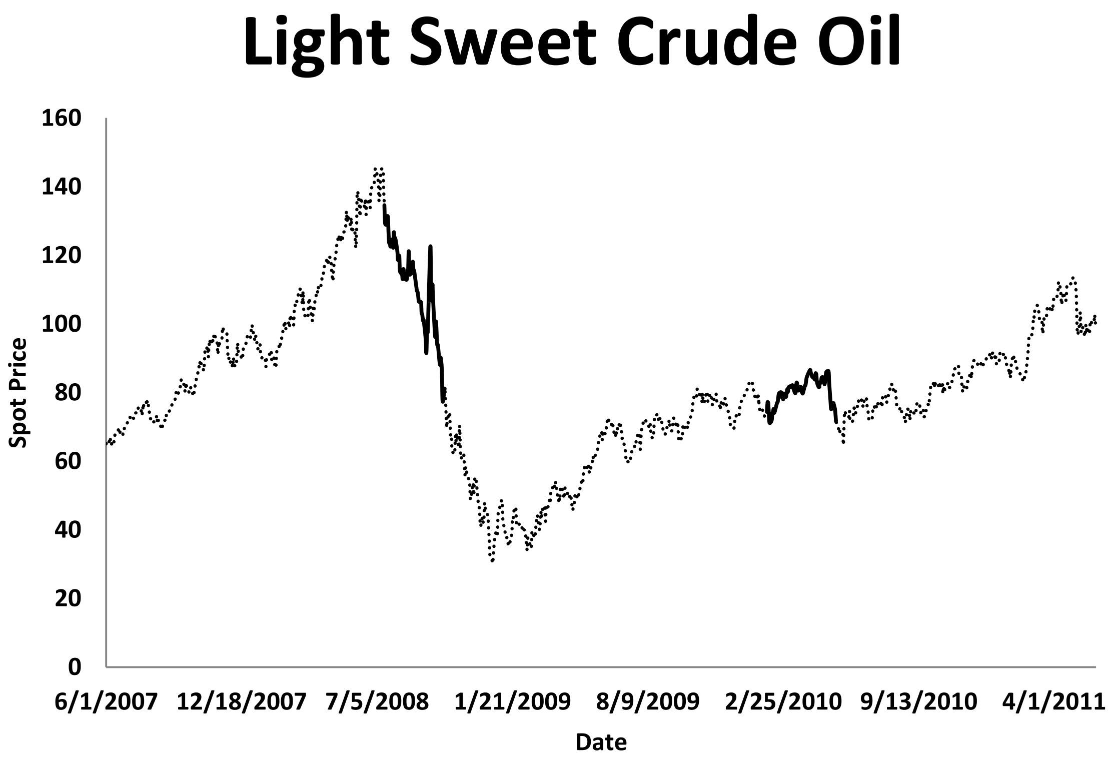

Figure 9: WTI Cude Oil price developtment

The first period starts July 16 2008 and ends on October 13, covering the first half of the huge Oil price plunge in the fall of 2008. In the months before this period starts, the price of Oil had been increasing steadily for quite a while and it reached its all-time high of more than $145 per barrel in the first half of July 2008. The following weeks the price started to drop, partially due to lower demand for energy because of the economic crisis. On 14 July one barrel of Sweet Crude Oil cost $145, and only two days later on the day our period starts, the price had dropped to $129. The down-market continued the entire fall until the price reached $30 on December 23th. The second period I will look at starts on 1 February 2010 and ends 14 May. This period is characterized by very different market conditions compared to the down-market of 2008. Instead, the price had a flat development, starting at about $76 per barrel in February and having decreased to $71.61 in the middle of May.

In both periods I have extracted RNDs every 3 or 4 days for Options with the same maturity date, with a total of 27 and 32 RNDs from the two periods. All options in period 1 had the WTI Futures contract with delivery in November 2008 as the underlying asset, and all Option contracts in period 2 where on the WTI Futures contract with delivery in June 2010. We will therefore get two sets of risk-neutral predictive densities describing the WTI Futures price at the end of each period. This makes it possible to see how the implied RND change as time passes. These two periods were chosen to see if we are able to find evidence that the SVI approach used are able to produce RNDs that reflect market expectations. Given the very different market conditions the two sets of densities are extracted from, we would expect to find RNDs with characteristics reflecting the Oil market at the time they were traded.

Figure 10 and 11 on the following two pages displays the RNDs from the two periods as probability time-charts. In Figure 10 we see the RND of July 16 as the very wide density in the front, and the RND of October 13 as the tall density all the way in the back. All densities in this chart are predictive densities for the WTI Futures price on October 18. The down-market of the time is easily recognized as the peaks of the densities moving left towards lower prices as time passes. In Figure 11 we see the February 1 RND in front and the May 14 RND in the back, all being predictive densities for the futures price at May 18. In this chart all densities except for the last 3 is located just behind each other, because of the flat price development in this period. The WTI Oil price dropped a bit in the last part explaining the left shift of the 3 last RNDs. From both charts we see that the density starts out very wide, with more than 3 months until the maturity date, the range of likely Oil prices at maturity is very wide. There is still a long time to expiration and the probability of large price fluctuations is relatively high. As time passes and the expiration date draws nearer the densities gradually gets taller and narrower, mostly reflecting the fact that there is less time left for large price fluctuations. The last couple of densities in both charts are narrow and very tall, predicting the futures price only 5 and 4 days later.

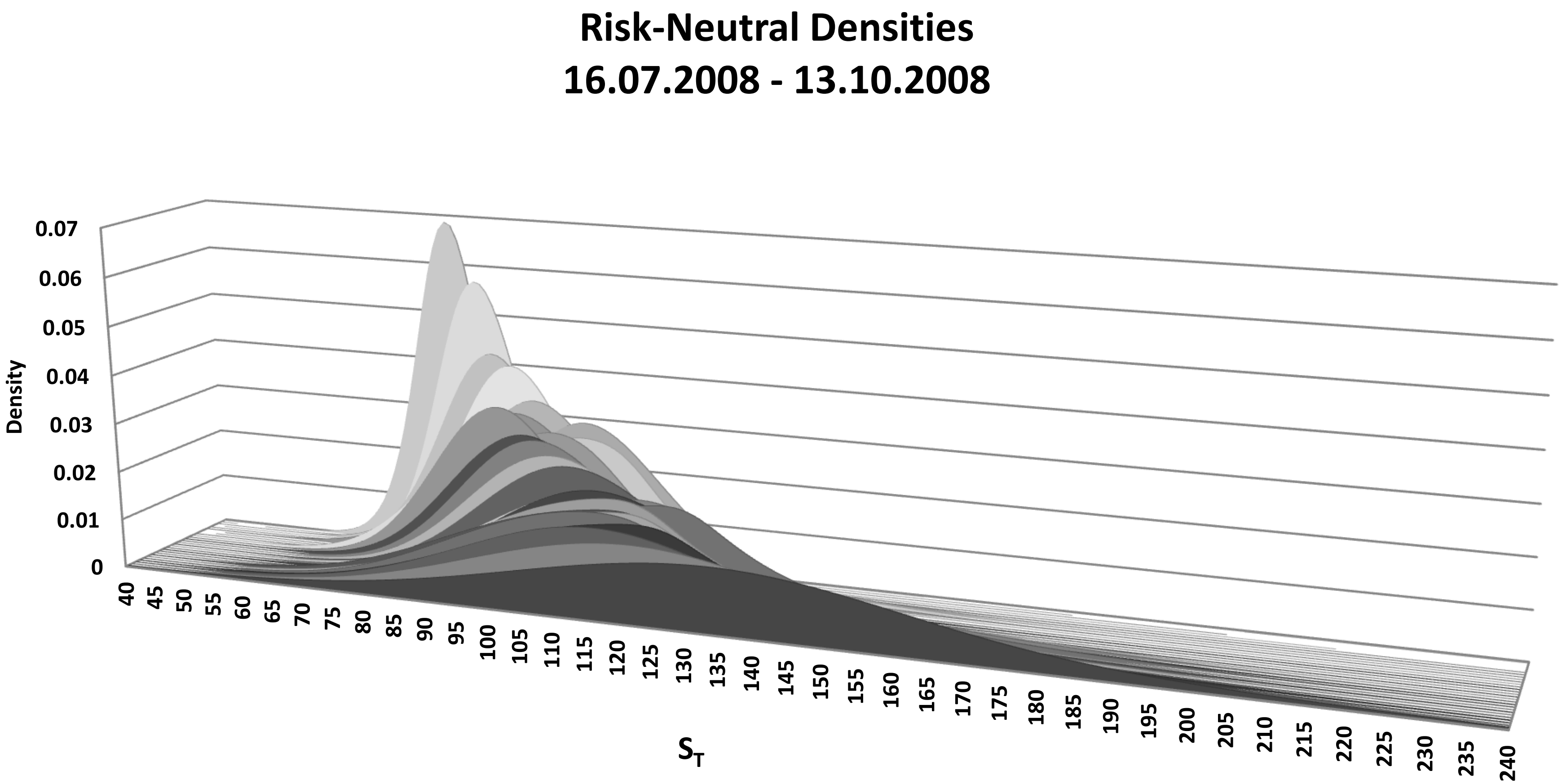

Figure 10: RNDs Fall 2008

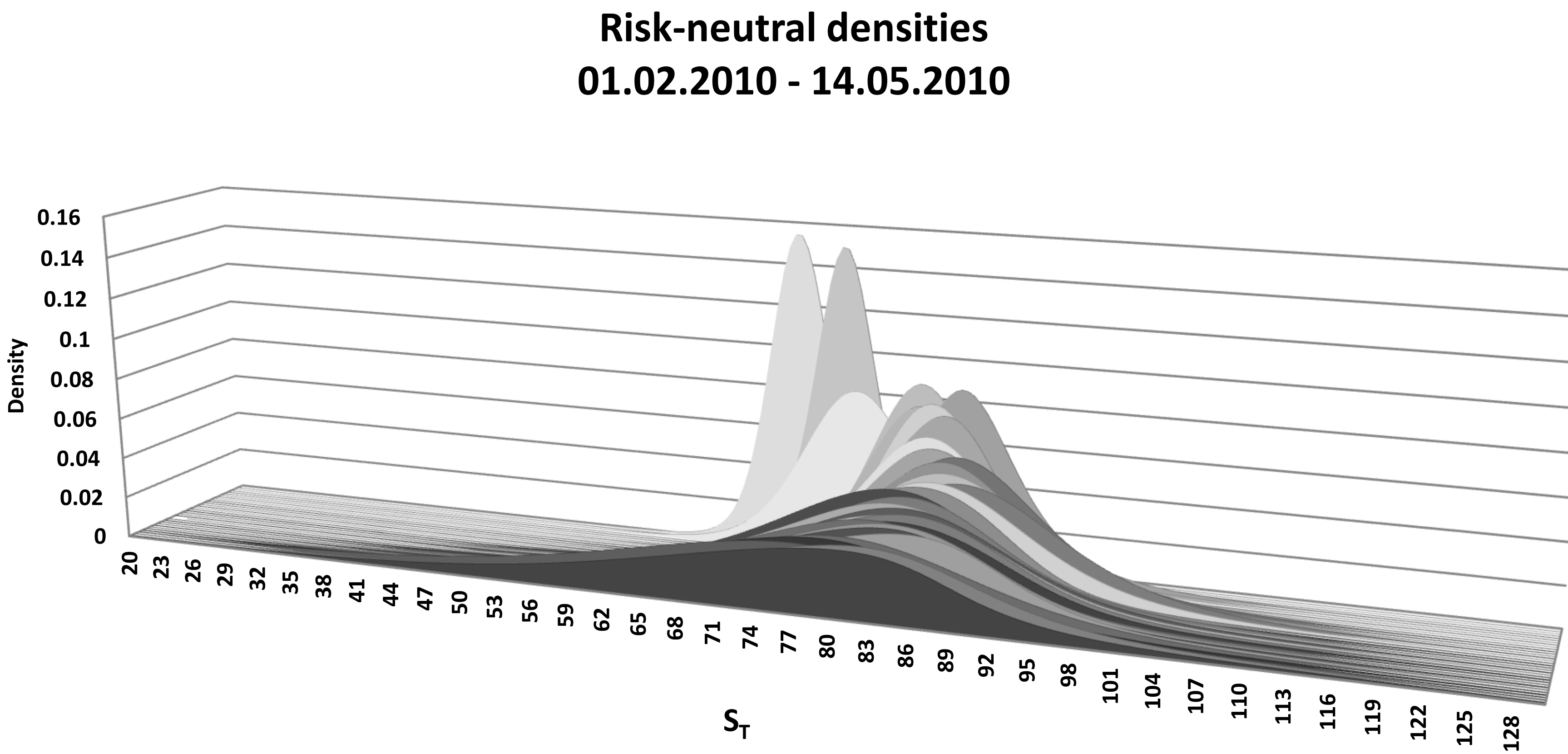

Figure 11: RNDs Spring 2010

The same information seen in Figure 10 and 11 can be displayed in a slightly different way, which better illustrates the risk-neutral prediction for the futures price at maturity. Figure 12 and 13 on the next page, display the RNDs as risk-neutral probability bounds around the expected price. The darkest shaded area covers a 10% risk-neutral probability bound around the expected value. The two lighter grey areas cover respectively a 50% and 90% bound. At each point in time, these bounds are interpreted as a 10%, 50% and 90% risk-neutral probability that the WTI Oil price is within the respective range at maturity. Since the probability we are analyzing is risk-neutral, we know that the expected price equals the Futures price at the time. Thus the WTI Futures price is located in the middle of the darkest shaded bound. We recognize the same time effects as we have discussed above. Starting the period with about 3 months until maturity, especially the 90% bound is very wide. In Figure 12, this bound actually covers the range from about \$100 to about \$215, with the Futures price at the time being \$129.4. Moving closer to the maturity of the Option contracts these bounds gradually get narrower, following the trend of the charts seen above.

From both the time charts seen above in Figure 10 and 11, and in the two time charts seen below in Figure 12 and 13, we see that the densities extracted from Options traded in the market with high and falling Oil prices in 2008, compared to the narrower densities of those extracted from the Oil market in the spring of 2010. As discussed earlier, this is a sign of the inverse leverage effect characteristic of Oil markets. In the couple of years leading up to the Oil price peak in July 2008, the price had been steadily increasing. During this period there were multiple disruptive events affecting the Oil market. Tension between OPEC and the United States grew as the price surged to \$110 in March 2008. On 17 April the same year, it reached a new all-time high of \$117, after an attack on an Oil pipe line in Nigeria. As tension between USA and Israel against Iran grew in the end of April and start of June, prices continued to increase, and by the start of July it had reached a new all-time high of about \$140 per barrel. Based on the theory of an inverse leverage effect, we should expect to find a relatively high uncertainty reflected in densities extracted in the first period. If we look at the development of “at the money” implied volatility and the volatility of our RNDs in the two periods, which are displayed in Figure 14 and 15, we see that volatility were very high in 2008 compared to 2010. The first RND extracted in 2008, when the price only days before had been at its all-time high, had a “at the money” implied volatility of 48%, compared to 34% for the first RND extracted in 2010. The RND volatility follows the same trend, starting at 34 in 2008, compared to 14 in 2010.

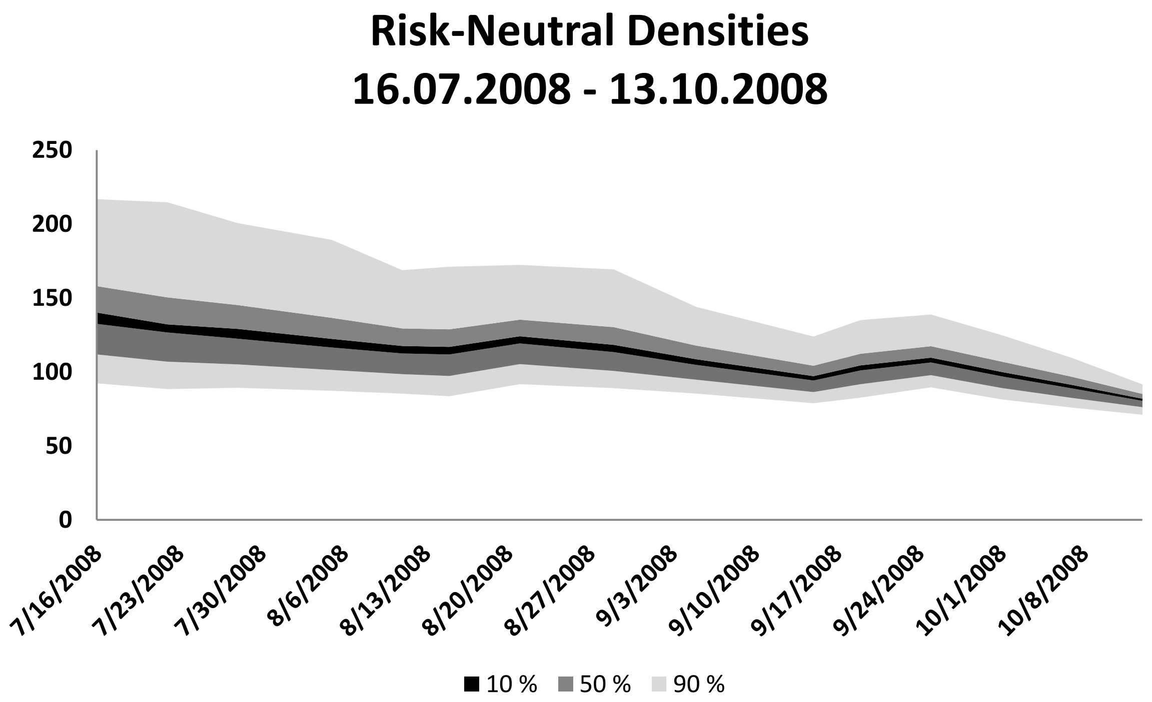

Figure 12: RNDs Fall 2008

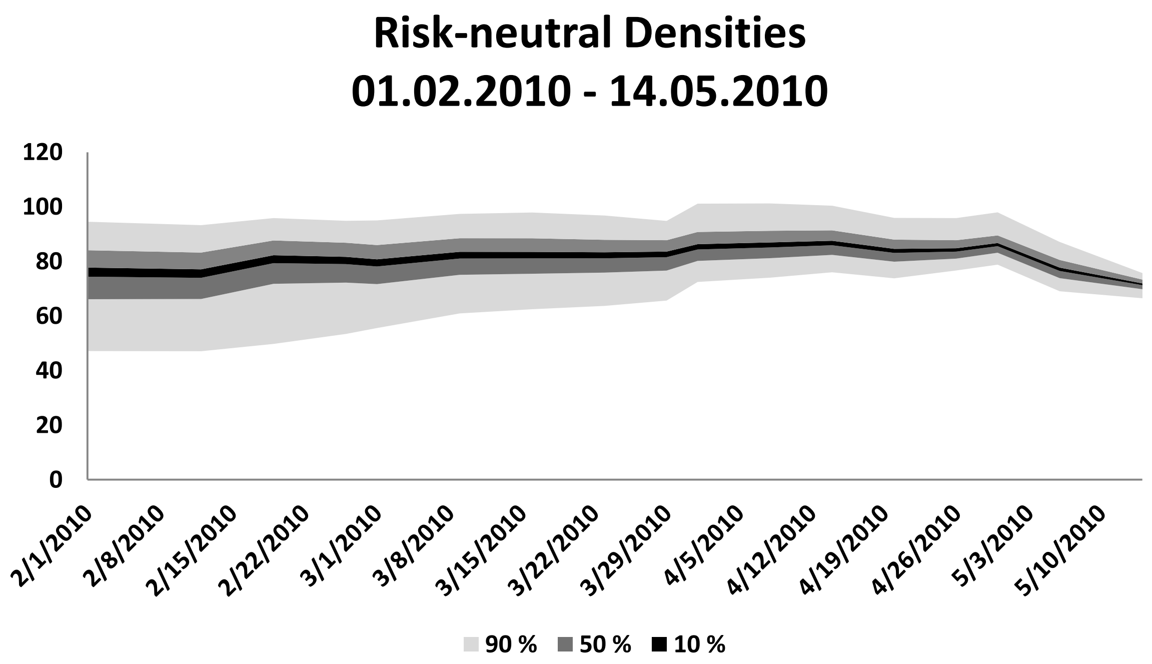

Figure 13: RNDs Spring 2010

The high uncertainty regarding Oil supply in the period leading up to our first implied RND in 2008, could explain the increasing Oil prices and the resulting high volatility, following the theory of an inverse leverage effect in the Oil market. As we see in the two figures on the next page, volatility continued to be high throughout the period of 2008 compared to 2010. This means that, implicit in the observed prices, Oil market participants continued to expect high future volatility throughout the whole first period, relative to the more “normal” period in 2010. The average “at the money” implied volatility from is 50% and 30% in respectively the densities from 2008 and 2010. Calculating the higher moments paints a picture supporting the observations already made. From Figure 16 we see that the Kurtosis, although unstable, were higher or equally high in 2010, compared to the RNDs of 2008. Knowing that the densities of 2008 generally were wider with higher volatility, it makes sense that those densities also were flatter compared to those of 2010, explaining the lower Kurtosis. The much higher “at the money” implied volatility used throughout the first period resulted in flatter volatility smiles, and even though not being the same across all strikes, which results in lognormal densities, a flatter smile will result in lower peaks and less of a leptokurtic look. This can also be explained by the fact that a lower “at the money” volatility, means that the probability for large price movements included in the Option price is lower and more of the probability mass will be centered around the mean, resulting in a higher peak. The average kurtosis is 5.3 and 8.56 in respectively the first and second period. From Figure 16 we see that all densities are positively skewed. The calculated skew is also a bit unstable throughout the two periods, but we see a clear tendency of a greater skew in densities from 2008, with an average skew of 0.71, compared to an average of 0.45 for those extracted from Options traded in 2010.

The implied risk-neutral densities extracted from the two different time periods show multiple signs indicating that our approach is able to adapt to expectations in the Oil market. All RNDs from the quiet period of 2010 are showing leptokurtic properties, with high peaks and fat tails. There were also a positive skew present in all densities, meaning that more of the probability mass was located in the right tail relative to the left. As we have mentioned earlier, these results suggest an inverse leverage effect is included in the traded Option prices, which can be explained by a greater fear in the market of increasing, relative to decreasing, energy prices (Geman, 2005). As Geman points out, this fear is not surprising, given the negative effect on the world economy of higher energy prices.

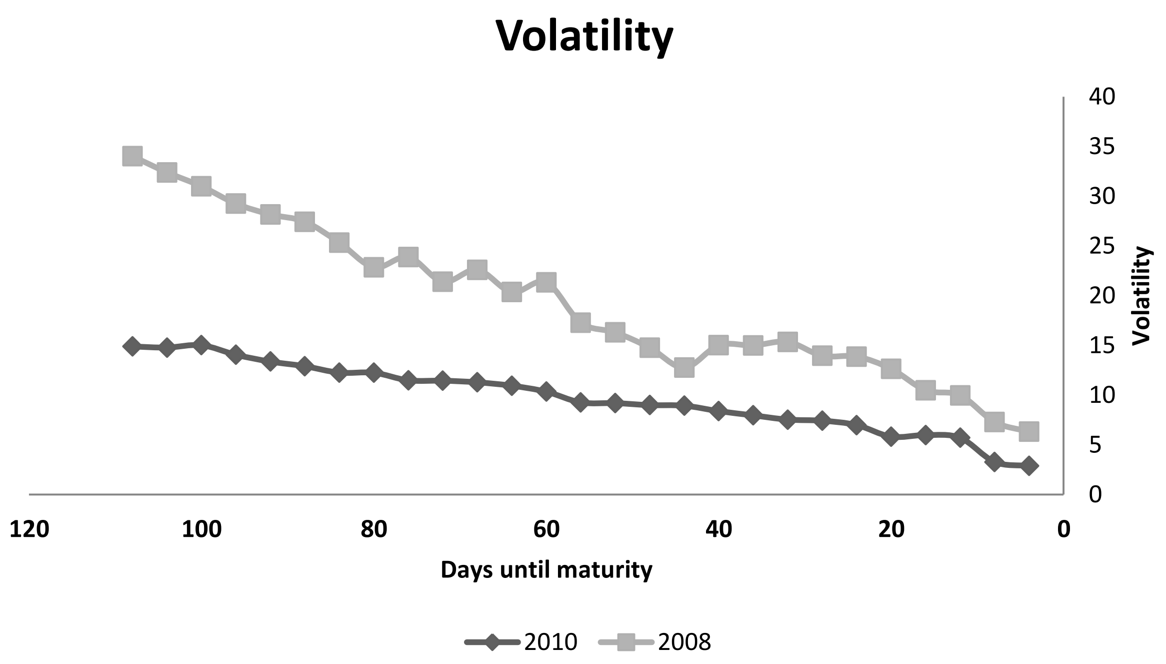

Figure 14: Volatility

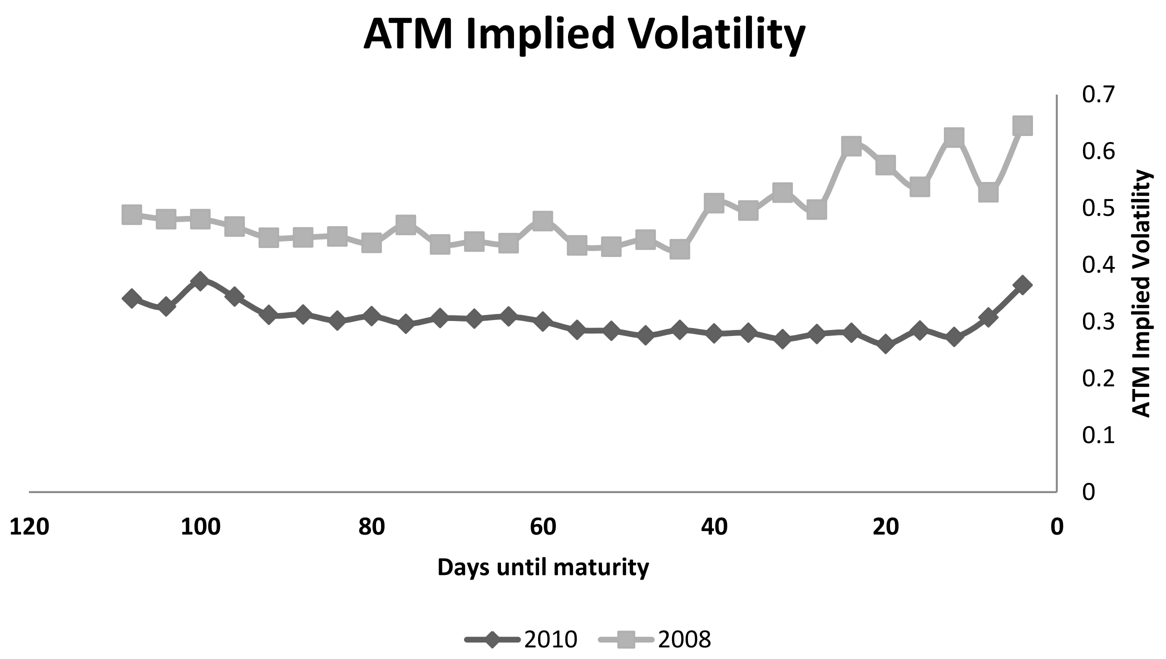

Figure 15: ATM Implied Volatility

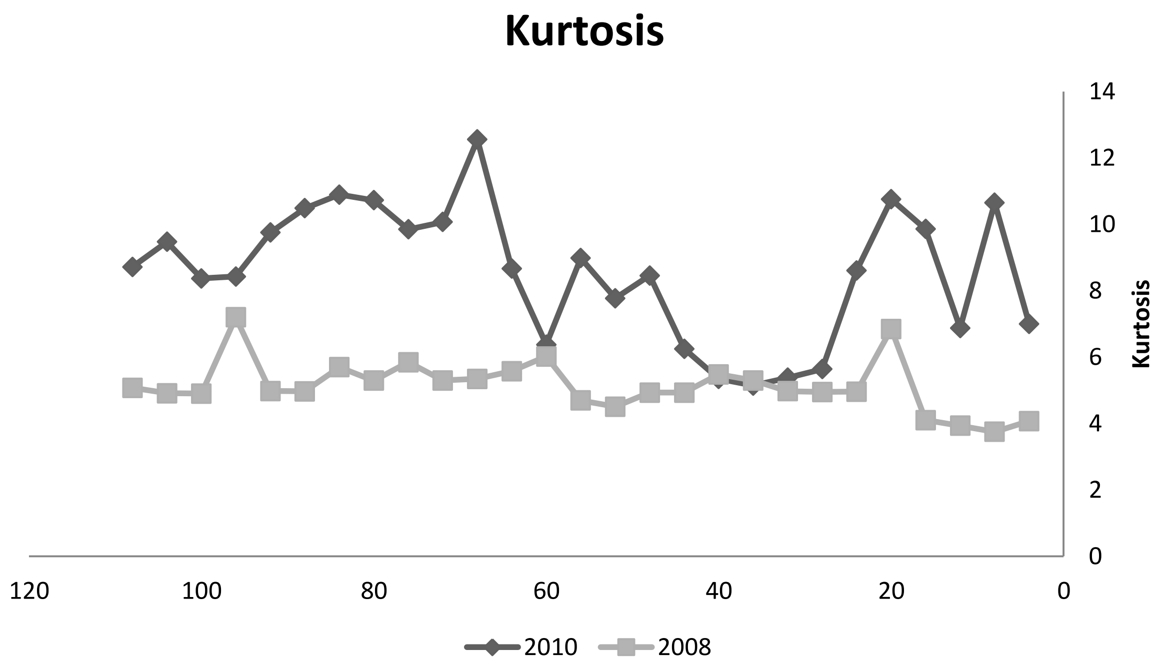

Figure 16: Kurtosis

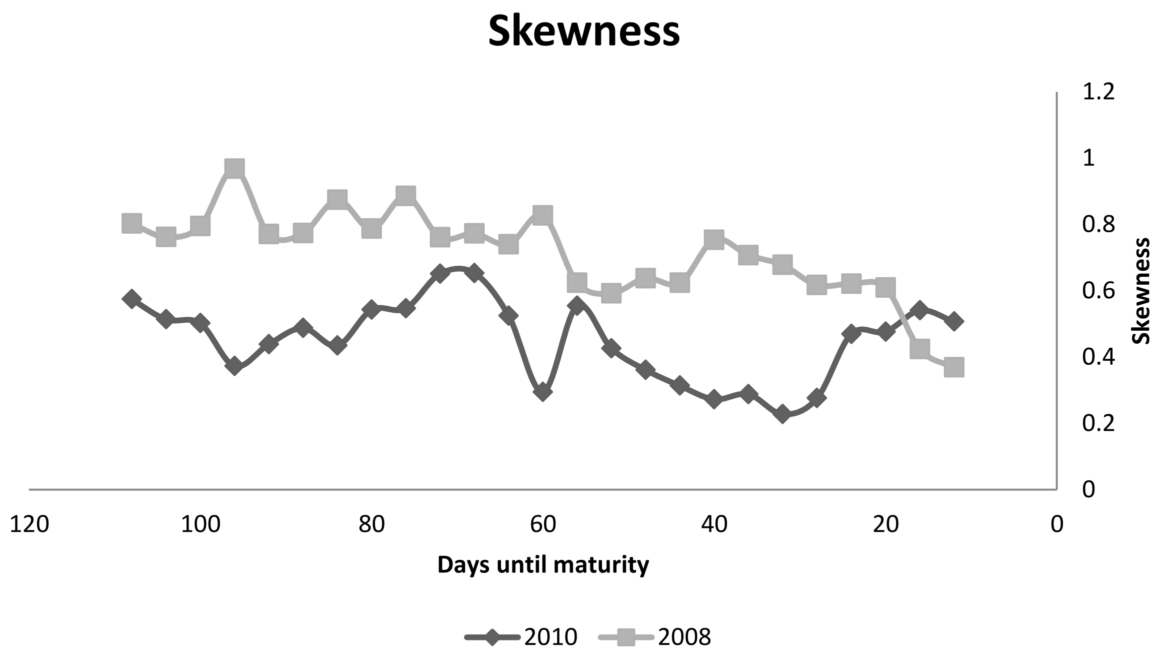

Figure 17: Skewness

These results show us that the implied volatility from Options traded in the period of 2008 was considerably higher compared to the period of 2010. This is consistent with earlier research, which suggests that volatility in oil markets tend to increase sharply when prices spike. I addition to high volatility and wide densities, we have observed that densities from the high-price period of 2008 were more positively skewed compared to those based on options from 2010. If a positive skew indicates, as Geman points out, higher aversion for increasing prices. Then, it makes perfect sense that the fear for higher prices were greater in the volatile oil market of 2008, were prices already were very high, compared to the period in 2010 when the oil price were lower and more stable.

5. Conclusion

This paper has presented and applied a way of extracting the implied risk-neutral density function of the future WTI Crude Oil price from Option prices with different strikes. This method is derived from a theoretical result by Breeden and Litzenberger, which proves the link between prices of Options and the risk-neutral measure. The main assumption made, to overcome the problem of Options being traded at discrete strike intervals, is that the implied volatility smile has a particular functional form. No assumptions are made regarding the stochastic process of the WTI Crude Oil price. Thus, the approach used should be able to adapt to properties and sentiments of the Oil market, as long as they are reflected in the observed Option price. To smooth the implied volatility smile, we have used the stochastic volatility inspired parameterization, introduced by Gatheral. Imposing a minimum of structure, the approach used in this paper proves to perform well at extracting implied RNDs from the Oil market, which is characterized by positively skewed and leptokurtic distributions.

Throughout the paper, the approach of obtaining densities from implied volatility smiles smoothed by the SVI parameterization have worked very well. We have seen that the SVI parameterization yielded a good fit to implied volatility smiles implicit in the Oil data. Further it extrapolated the smile outside the range of traded strikes by creating volatility tails that secure absence of arbitrage in the corresponding option prices. This approach yielded implied RNDs which were consistent with probability theory, avoiding problems with negative probabilities.

Applied to the Oil market, we find that the extracted densities are able to reflect market sentiments of the time the Options were traded. We also find that they are able to adapt to typical characteristics of Oil markets. The densities extracted from WTI Crude Oil Options show leptokurtic properties, and have high peaks and fat tails compared to the lognormal distribution. All densities are positively skewed, with long probability tails towards higher oil prices. Studying two sets of RNDs extracted from different time periods allowed us to study how the shape of the density is affected by the time to maturity and the market condition under which the Option price were observed. In both periods, densities start out very wide, with more than three months to maturity, and then gradually get narrower as time pass.

Densities extracted from the high Oil price period of 2008 had considerably higher volatility and were wider compared to the counterparts from 2010. The densities from 2008 were also flatter, confirmed by a lower kurtosis. All densities extracted were positively skewed, but more so in those from 2008.

The positive skew observed in all densities can be explained by an inverse leverage effect in the Oil market, which means that the price and volatility are positively correlated. Market participants fear a positive shock to the Oil price more than an equally large price drop due to the negative effect a positive price shock have on the world economy. Higher volatility at higher prices would also increase the probability of even higher prices, which is exactly what we observe in the positively skewed densities. Densities extracted from the period of 2008 were more skewed compared to those extracted from 2010, indicating that the inverse leverage effect was stronger in a market with high prices, which can be explained by more fear for increasing prices when the price and volatility already are high. Earlier research on RNDs from equity markets show that the characteristics we have found are specific to energy markets.

It would be interesting to do a more comprehensive study of implied risk-neutral densities from Oil markets. This paper presents results suggesting the existence of an inverse leverage effect in the Crude Oil market, and show that implied RNDs are able to adapt to the sentiments of the market for which the Option was traded. One aspect further research could look into is the development of RNDs from the Oil market over a longer time interval. It would make it possible to better see how the inverse leverage effect is reflected in the RND by analyzing the development of the Skewness over time. It would also be interesting to compare the performance of the SVI parameterization to the more popular splines used in other research papers. Another direction further research could take, is to test if implied densities extracted from the oil data is able to predict the realized oil price.

Bibliography

Bank of England. (2011, November 25). Retrieved November 25, 2011, from http://www.bankofengland.co.uk/statistics/impliedpdfs/

Bahra, B. (1997). "Implied risk-neutral probability density functions from option prices: theory and application". Bank of England.

Baxter, M., & Rennie, A. (1996). *Financial Calculus: An Introduction to Derivative Pricing*. Cambridge University Press.

Bertrand, T. (2011). Implied distribution as a function of the volatility smile.

Black, F. (1976). Studies of Stock Price Volatility Changes. *American Statistical Association*.

Bliss, R. R., & Panigirtzoglou, N. (2000). Testing the stability of implied probability density functions. *Bank of England*.

Breeden, D. T., & Litzenberger, R. H. (1978). Prices of State-Contingent Claims Implicit in Option Prices. *Journal of Business*, 621-651.

Clews, R., Panigirtzoglou, N., & Proudman, J. (2000). Recent Developments in Extracting Information from Options Markets. *Bank of England*.

CME Group. (n.d.). Retrieved July 30, 2011, from http://www.cmegroup.com/market-data/files/CME_Group_Settlement_Procedures.pdf

CME Group. (n.d.). Retrieved July 2011, from http://www.cmegroup.com/trading/files/WTI-Fact-Card.pdf

Cooper, N. (1999). Testing techniques for estimating implied RNDs from the prices of European-style Options. *Bank of England*.

Copeland, T. E., Weston, J. F., & Shastri, K. (2004). *Financial Theory and Corporate Policy*. Addison Wesley.

Doran, J. S., & Ronn, E. I. (2006). "The Bias in Black-Scholes/Black Implied Volatility: An Analysis of Equity and Energy Markets".

Gatheral, J. (2004). A parsimonious arbitrage-free implied volatility parametrization with the application to the valuation of volatility derivatives.

Geman, H. (2005). *Commodities and commodity derivatives: modelling and pricing for agriculturals, metals, and energy*. John Wiley and Sons.

Harrison, M., & Pliska, S. (1981). *Martingales and Stochastic integrals in the theory of continuous trading.* Stochastic Processes and their Applications.

Haug, E. G. (2006). *The Complete Guide to Option Pricing Formulas.* McGraw-Hill.

Hull, J. C. (2008). *Options, Futures, and Other Derivatives.* Prentice Hall.

Lee, R. W. (2003). The Moment Formula for Implied Volatility at Extreme Strikes. *Mathematical Finance*.

McDonald, R. L. (2006). *Derivatives markets.* Addison-Wesley.

Melick, W. R., & Thomas, C. P. (1997). Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during during. *The Journal of Financial and Quantitative Analysis*.

Taylor, S. J. (2007). *Asset Price Dynamics, Volatility, and Prediction.* Princeton University Press.

Appendix

Calculation of Moments

Given a random variable X, the first four moments is defined as:

Mean:

The mean is a measure of the central tendency of a distribution. It is defined as the weighted average of a random variable X, where all outcomes are weighted with the corresponding probability of each outcome occurring.

$$m = E[X] \qquad \text{Equation 12}$$

Variance:

The variance is a measure of the dispersion of a distribution. It is defined as the weighted average squared deviation from the mean.

$$v = E[X - m]^2 \qquad \text{Equation 24}$$

Skewness:

The Skewness is a measure of asymmetry of a distribution. If the distribution has a longer tail to the right, it has a positive skew. A longer tail to the left means the distribution is negatively skewed.

$$s = E\left[\left(\frac{X - m}{\sqrt{v}}\right)^3\right] \qquad \text{Equation 25}$$

Kurtosis:

The Kurtosis is a measure of peaked-ness of a distribution. It is measured relatively to the Gaussian distribution. A kurtosis of more than 3 has a high peak relative to the Gaussian curve, and a kurtosis of less than 3 is flatter than the Gaussian curve. Kurtosis equal to 3 is normally distributed.

$$k = E\left[\left(\frac{X - m}{\sqrt{v}}\right)^4\right] \qquad \text{Equation 26}$$

Given that we know the density function of a random variable X, we can solve the following integral to help find the first four moments (Bertrand, 2011):

$$I_z = E[X^z] = \int_0^{\infty} x^z f_X(X)dx \qquad Z = 1,2,3,4 \qquad \text{Equation 27}$$

This integral must be solved numerically. In this paper I used a method based on the Simpson`s rule for numerical integration. Using Equation 27 the moments are given by:

$$m = I_1 \qquad \text{Equation 28}$$

$$v = I_2 - m^2 \qquad \text{Equation 29}$$

$$s = v^{-\frac{3}{2}}(I_3 - 3mv - m^3) \qquad \text{Equation 30}$$

$$k = v^{-2}(I_4 - 4msv^{\frac{3}{2}} - 6m^2v - m^4) \qquad \text{Equation 31}$$

SVI Derivatives

The first and second derivatives of the SVI volatility function with respect to the strike price were needed for the approach used in this paper. These two expressions were found using the Wolfram Alpha mathematical calculator.

The first derivative of the SVI volatility function with respect to the strike:

$$\frac{b\left(\frac{\log\left(\frac{k}{f}\right)-m}{\sqrt{\left(m-\log\left(\frac{k}{f}\right)\right)^2+\sigma^2}}+p\right)}{2\,k\sqrt{a+b\left(\sqrt{\left(m-\log\left(\frac{k}{f}\right)\right)^2+\sigma^2}+p\log\left(\frac{k}{f}\right)-m\,p\right)}}$$

The second derivative of the SVI volatility function with respect to the strike:

$$\left(b\left(\left(2\left(-p\left(\left(m-\log\left(\frac{k}{f}\right)\right)^2+\sigma^2\right)^{3/2}+m\left(\left(m-\log\left(\frac{k}{f}\right)\right)^2+\sigma^2\right)-\right.\right.\right.\right.$$

$$\left.\log\left(\frac{k}{f}\right)\left(\left(m-\log\left(\frac{k}{f}\right)\right)^2+\sigma^2\right)+\sigma^2\right)$$

$$\left.\left(a+b\sqrt{\left(m-\log\left(\frac{k}{f}\right)\right)^2+\sigma^2}+b\,p\log\left(\frac{k}{f}\right)-b\,m\,p\right)\right)\Big/$$

$$\left.\left.\left(\left(m-\log\left(\frac{k}{f}\right)\right)^2+\sigma^2\right)^{3/2}-b\left(\frac{\log\left(\frac{k}{f}\right)-m}{\sqrt{\left(m-\log\left(\frac{k}{f}\right)\right)^2+\sigma^2}}+p\right)^2\right)\right)\Big/$$

$$\left(4\,k^2\left(a+b\sqrt{\left(m-\log\left(\frac{k}{f}\right)\right)^2+\sigma^2}+b\,p\log\left(\frac{k}{f}\right)-b\,m\,p\right)^{3/2}\right)$$