FORECASTING OIL PRICE VOLATILITY

Namit Sharma

Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

MASTER OF ARTS
in
Economics

Approved :

Roger N. Waud   Nancy A. Wentzler

Brian K. Reid   Thomas Lutton

May, 1998
Falls Church, Virginia

Keywords:
Crude Oil
GARCH
Implied Volatility

FORECASTING OIL PRICE VOLATILITY

Namit Sharma

(ABSTRACT)

This study compares different methods of forecasting price volatility in the crude oil futures market using daily data for the period November 1986 through March 1997. It compares the forward-looking implied volatility measure with two backward-looking time-series measures based on past returns – a simple historical volatility estimator and a set of estimators based on the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) class of models.

Tests for the relative information content of implied volatilities vis-à-vis GARCH time series models are conducted within-sample by estimating nested conditional variance equations with returns information and implied volatilities as explanatory variables. Likelihood ratio tests indicate that both implied volatilities and past returns contribute volatility information. The study also checks for and confirms that the conditional Generalized Error Distribution (GED) better describes fat-tailed returns in the crude oil market as compared to the conditional normal distribution.

Out-of-sample forecasts of volatility using the GARCH GED model, implied volatility, and historical volatility are compared with realized volatility over two-week and four-week horizons to determine forecast accuracy. Forecasts are also evaluated for predictive power by regressing realized volatility on the forecasts. GARCH forecasts, though superior to historical volatility, do not perform as well as implied volatility over the two-week horizon. In the four-week case, historical volatility outperforms both of the other measures. Tests of relative information content show that for both forecast horizons, a combination of implied volatility and historical volatility leaves little information to be added by the GARCH model.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Robert Mackay for his encouragement and support throughout the time this thesis was written. His guidance and comments helped me greatly in developing a first-principles, intuitive approach to the topic.

Equally important to this thesis was the invaluable help I received from Dr. Andrew Thompson, in assembling a set of ready-to-use data from a mass of NYMEX text files, as well as in understanding the oil market and its dynamics.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

Forecasting volatility is fundamental to the risk management process in order to price derivatives, devise hedging strategies and estimate the financial risk of a firm's portfolio of positions. In recent years, Autoregressive Conditional Heteroscedasticity (ARCH) type models have become popular as a means of capturing observed characteristics of financial returns like thick tails and volatility clustering. These models use time series data on returns to model conditional variance. An alternative way to estimate future volatility is to use options prices, which reflect the market's expectation of volatility. Analytical option pricing models can be used to back out implied volatility over the remaining life of the option given the observed market price.

In the construction of volatility forecasts, energy market participants would like to know which model produces the most accurate forecasts, as well as, whether the complex time series models add any significant volatility information beyond that contained in option prices. Day and Lewis (1993) compare the relative information content and predictive power of implied volatility and ARCH-type forecasts for crude oil futures. A similar study by Xu and Taylor (1996) examines the informational efficiency of the PHLX currency options market in predicting volatility. Duffie and Gray (1995) compare the forecasting accuracy of ARCH type models, Markov switching models, and implied volatilities for crude oil, heating oil and natural gas markets.

This study seeks to test the hypothesis that implied volatilities subsume information contained in returns and to determine which method provides the best month-ahead volatility forecasts for crude oil. The results can also be compared with earlier research to check whether the inclusion of more recent data improves ARCH forecasts (longer time series), indicates increased options market efficiency in forming expectations (higher liquidity) or yields a different ranking of forecasts (constructed over a different time horizon).

CHAPTER 2.AN OVERVIEW OF THE OIL MARKET

2.1 The Economic Importance of Oil

The world oil market is a capital-intensive environment characterized by complex interactions deriving from the wide variety of products, transportation/ storage issues and stringent environmental regulation.
Worldwide consumption of oil exceeds $500 billion, roughly 10% of the US GDP. Crude oil is also the world's most actively traded commodity, accounting for about 10% of total world trade.[1]

The economic importance of oil derives not only from the sheer size of the market, but also from the crucial, almost strategic, role it plays in the economies of oil-exporting and oil-consuming countries. Oil prices drive revenues to oil-exporting countries in a large number of which, oil exports comprise over 20% of the GDP. On the other hand, costs of oil imports (typically over 20% of the total import bill) have a substantial impact on growth initiatives in developing countries. Energy price shocks have often been cited as causing adverse macroeconomic impacts on aggregate output and employment, in countries across the world.

2.2 A History of Oil Prices (1947-1998)

Oil prices have averaged $19.27 per barrel, in 1996 dollars, over the period 1947-1997. Prices have only exceeded $22.00 per barrel in response to war or conflict in the Middle East [2]. The major oil price shocks during this period were the 1973 Arab oil embargo, the 1979-80 events in Iran and Iraq, and the 1990 invasion of Kuwait.

As a result of the Yom Kippur war, crude oil price, which had stayed between $2.50 and $3 since 1948, quadrupled from $3 per bbl in 1972 to $12 per bbl by the end of 1974.

[1] Verleger (1993)
[2] WTRG Economics (1998)

The revolution in Iran and the subsequent Iran-Iraq war more than doubled prices from $14 per bbl in 1978 to $35 per bbl in 1981. The ensuing world recession and development of alternative energy sources, led to a decrease in demand and falling prices for most of the 1980s. Efforts by the OPEC to set production quotas, in an effort to shore up prices were largely unsuccessful, as member nations routinely violated limits. For example, in 1986, Saudi Arabia increased production from 2 MMBPD to 5 MMBPD causing crude oil prices to plummet below $10 per barrel.

Prices surged again in 1990-91, in response to the uncertainty created by the Iraqi invasion of Kuwait, but retreated in the face of an US-led military resolution of the conflict and increased supply from other nations.

Recession in the US saw prices decline, until in 1994, inflation-adjusted prices attained their lowest level since 1973. Subsequently, a strong US economy and growth in Asia led to a firming up in demand. World petroleum demand grew 2.8% in 1995 and 2.2% in 1996. Oil prices increased by approximately $6 per bbl over the course of 1996[3]. Despite Iraq's re-entry into the world market in Dec 1996, the price recovery continued well into 1997, until the recent sharp downturn brought about by the Asian economic crises and the mild 1997-98 winter.

2.3 Fundamental Drivers of Supply and Demand

Fluctuations in energy prices are caused by supply and demand imbalances arising from events like wars, changes in political regimes, economic crises, formation/ breakdown of trade agreements, unexpected weather patterns etc. Forward and futures prices imbed the expectations of the market participants about how demand will evolve and how quickly the supply side can react to events, to restore balance.

A dynamic market model based on expectations would predict that prices for immediate delivery will exceed prices for longer delivery horizons, when stocks are low or are

[3] Energy Information Administration (1997)

anticipated to be insufficient to meet short-term needs. This pattern of prices is characteristic of a market in *backwardation*. In contrast, when stocks are high and the probability of stockout is low, forward prices exceed spot prices, a situation which describes a market in *contango*. A fundamental driver of volatility in oil prices is the fact that current stocks can be stored for consumption in the future but future production cannot be "borrowed" to meet immediate needs. This market asymmetry implies that the magnitude of a price increase in a given period due to a disruption in current supplies is likely to be larger as compared to a price drop in response to oversupply. Storage limitations cause energy markets to display volatile day-to-day behavior in spot and nearby futures prices. Volatility decreases for longer futures expirations reflecting the expectation that supply and demand balance in the long run, to reach a relatively stable equilibrium price. This long run price builds in expectations of market production capacity and cost in the long run.

On the demand side, fundamental price drivers are *convenience yield* and seasonality. Convenience yield is directly related to the probability of a disruption in supplies. Depending on the prevailing supply versus demand situation, industrial users may be willing to pay a premium for "immediate energy", reflected in higher near-term forward prices relative to longer-term forward prices. Convenience yield is measured as the net benefit (value of uninterrupted production) minus the cost (including storage costs). Demand for heating oil is seasonal, peaking in winter, while gasoline demand is higher in summer. The seasonal demand for these and other distillates affects the pattern of crude oil prices, although the effect is much less pronounced.

The impact of these fundamental supply and demand drivers must be viewed against the changing backdrop of new developments in the oil market. To quote TIME (April 6, 1998 issue)

> "The industry has changed fundamentally since previous oil shocks that seemed to portend ever higher prices. Oil is coming into the market from every corner of the globe. Current exploration hot spots include the newly independent nations around the Caspian Sea and offshore West Africa.

This diversification acts as an insurance policy against supply disruptions. The growing role of natural gas in the overall energy mix provides a further buffer. Information technology has also allowed the industry to search for oil and make a profit at $15 per bbl., about half the threshold of just a decade ago."

CHAPTER 3.ESTIMATING OIL PRICE VOLATILITY

Energy markets are characterized by extremely high levels of price volatility. In recent years there has been an explosive growth in the use of exchange traded and OTC derivatives in these markets as a means of managing exposure to energy prices. Besides basic or "plain vanilla" swaps and options, the unique nature of price risk in energy markets has given rise to new "exotic" instruments tailored to the needs of end-users. Any flexibility in an energy contract, in terms of price, volume, point of delivery, delivery timing, etc., can be thought of in terms of a financial option or a portfolio of options. Volatility estimates are required for efficient pricing of options and other derivative securities as well as for the effective use of these securities in managing and hedging risk.

3.1 Statistical Features of Oil Prices

Figure 1 shows a time series plot of the first nearby futures price on West Texas Intermediate (WTI) crude, traded at the New York Mercantile Exchange (NYMEX). An examination of the pattern reveals particular tendency for the price to increase or decrease, often called "random walk" behavior

If P_t denotes the price of crude oil on a given business day, the relative price change or percent return R_t is defined as

$$R_t = (P_t - P_{t-1}) / P_{t-1} \quad ; \qquad (1)$$

On a continuous compounding basis, the price return over a given period can be computed as the logarithm of the ending price less the logarithm of the beginning price.[4]

$$\text{Price Return, } r_t = \ln(1+R_t) = \ln(P_t / P_{t-1}) = p_t - p_{t-1} \quad ; \qquad (2)$$

[4] Lim $(1+r/m)^m = e^r$ as the frequency of compounding $m \rightarrow \infty$

where $p_t = \ln P_t$.

To begin with, the log price p_t can be modeled as a standard random walk

$$p_t = \mu + p_{t-1} + \varepsilon_t \quad , \tag{3}$$

i.e. $$r_t = p_t - p_{t-1} = \mu + \varepsilon_t \quad , \tag{4}$$

where $\varepsilon_t = \sigma z_t$, and $z_t \sim \text{IID } N(0,1)$;

The above model implies that the returns are normally distributed with mean μ and constant variance σ.[5] The assumption of normally distributed returns, for modeling purposes, implies a lognormal price distribution (which guarantees that prices will never be negative). Returns series are preferred over prices in analysis of financial time series because they have attractive statistical properties like stationarity.[6] [7]

As Fig 3 and 4 show, however, the actual distribution of daily returns for crude oil nearby futures has fatter tails (and a narrower waist) compared to a fitted Normal distribution (0,2.33).[8] The Kurtosis value for daily oil price returns over the period November 14, 1986 - March 31, 1997 is 38.68 demonstrating presence of fat tails. Thick tails can be modeled by assuming a "conditional" normal distribution for returns; where conditional normality implies that returns are normally distributed on each day, but that the parameters of the distribution change from day to day.[9] Also, as evidenced in Figure 2, the volatility (standard deviation) of oil price returns exhibits "clustering" i.e. bursts of high volatility separated by periods of relative tranquility.

---

[5] In the case of crude oil futures prices, the mean or expected daily price return μ is zero. The finite long-run unconditional variance implies that returns are centered on some value. Negative and positive returns are equally likely to be observed, as Figure 2 shows.

[6] For details on statistical modeling of financial time series see Appendix A.

[7] For crude oil, evidence suggests that log prices are mean-reverting/ stationary (Pilipovic, 1997). The process generating such series would not have a "unit root" i.e. the coefficient for p_{t-1} in the log price random walk equation would be less than one.

[8] Values and Figures based on fitting a Normal distribution to Crude Oil Daily Returns on Nearby Futures (1986-1997) using BestFit®.

Returns are thus not identically distributed with mean 0 and variance σ^2 at each point in time. Instead, it is fair to say that σ_t^2 changes with time t. This time-varying nature of variance is referred to in statistics as *heteroscedasticty*. The persistence of volatility (related to the time it takes for the effects of events in the oil market to dissipate) is an indication of *autocorrelation* in variances.The Box-Ljung statistic BL(p) can be used to test for autocorrelation in variance. Since $\sigma^2 = E[r_t-\mu]^2 = E(r_t^2)$ for $\mu = 0$; squared returns can be used for the test. Under the null hypothesis that a time series is not autocorrelated, BL(p) is distributed χ^2_p, where p denotes the number of autocorrelations used to estimate the statistic. For p=36, the BL(p) statistic for squared oil price returns is 220.91, which rejects the hypothesis that variances of daily returns are not autocorrelated. The presence of ARCH effects can be tested using the original Lagrange Multiplier (LM) test proposed by Engle.[10] The test statistic TR^2 for first order ARCH effects (distributed as χ^2_1) in daily return variances for crude oil is 34.48 , thus rejecting the null hypothesis that variances are homoscedastic.[11] Time varying volatility is modeled statistically by estimating a conditional variance equation in addition to the returns generating process.

Often, even when put together, the conditional normality assumption and the simultaneous estimation of conditional variance, do not capture the thick tails entirely. In such cases, estimation may be required to be done with an underlying conditional Student's t or conditional Generalized Error Distribution (GED)[12] which allow for fatter tails in the conditional distribution than the normal distribution.

---

[9] The distribution over longer time horizons is thus a mixture of normal distributions.
[10] See Appendix A
[11] See Appendix B
[12] See Appendix A

Fig 1 : Crude Oil - Nearby Futures Prices
1983 - 1997

Fig 2 : Crude Oil - Daily Returns on Nearby Futures Prices
1983 - 1997

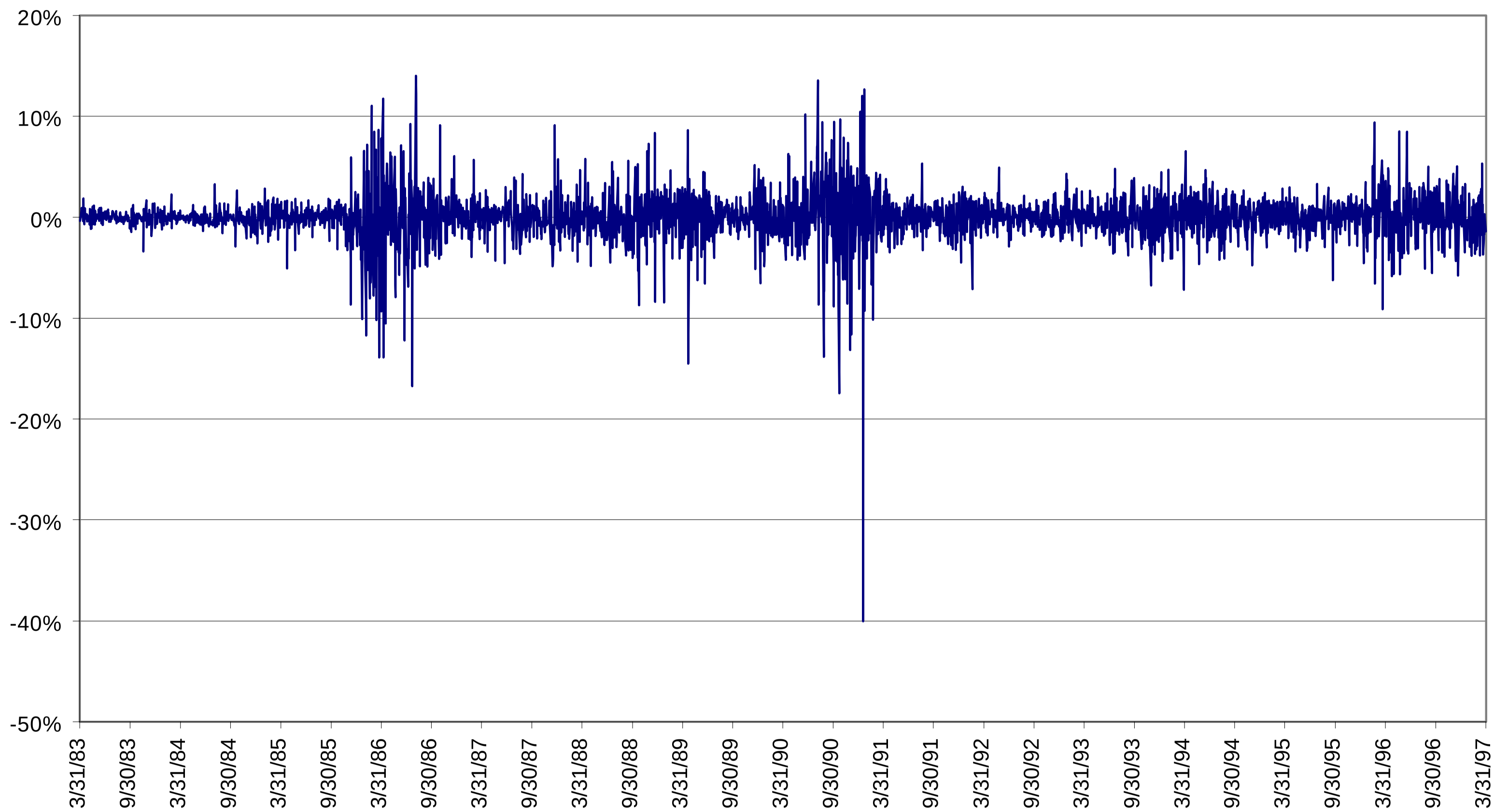


Fig 3: Crude Oil Nearby Futures- Distribution of Daily Returns (1986-1997)

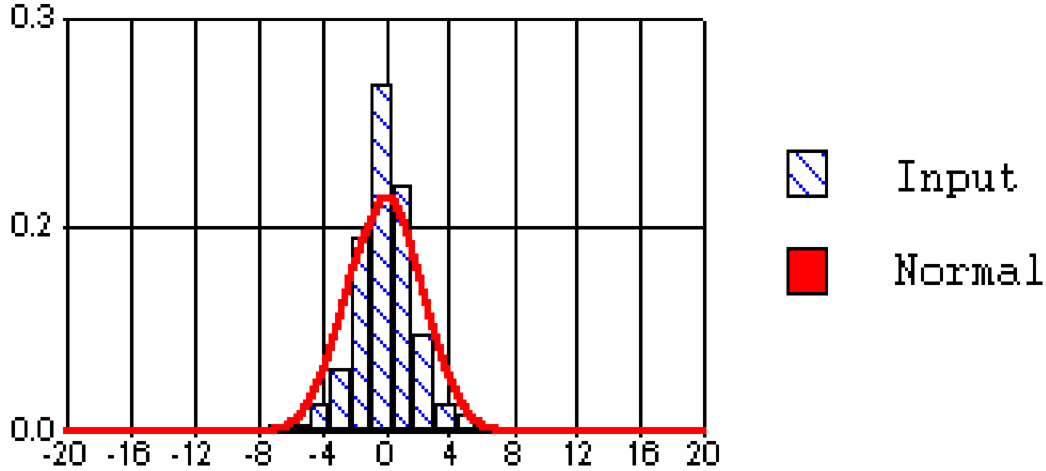
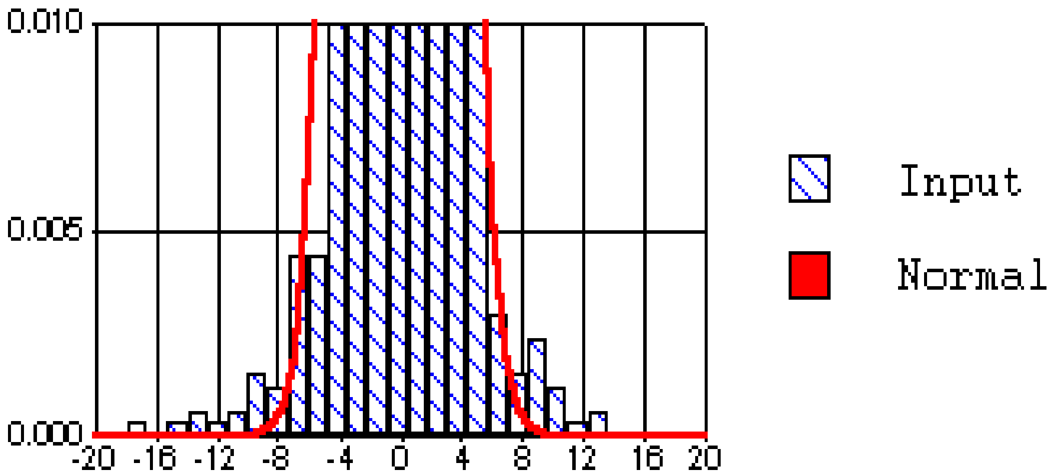


Fig 4: Crude Oil Nearby Futures- Distribution of Daily Returns showing Fat Tails



3.2 Volatility Models

3.2.1 Historical Volatility

Assuming that ε_t is the innovation in mean for energy price returns (defined as log price changes) the simplest estimate of volatility at time t is the given by the historical volatility estimator

$$V_{H,t} = \left[(1/N) \sum_{i=0}^{N-1} \varepsilon^2_{t-i}\right]^{1/2} \quad ; \qquad (5)$$

calculated over the last N days, where N is the forecast period. This historical volatility calculated over the previous N days can be more precisely termed as an N-day simple moving average volatility. The historical volatility estimator assumes that volatility is constant over the estimation period and the forecast period.

Another estimate of volatility derived using historical returns is the long-run or unconditional volatility estimated using all previous returns available at time t. The underlying volatility estimator assumes constant volatility across all periods over time. Variations in the estimate are due to random variation in prices.

3.2.2 ARCH

Autoregressive Conditional Heteroscedasticity (ARCH) type modeling is the predominant statistical technique employed in the analysis of time-varying volatility. In ARCH models, volatility is a deterministic function of historical returns. The original ARCH(q) formulation proposed by Engle (1982) models conditional variance h_t as a linear function of the first q past squared innovations.

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon^2_{t-i} \quad ; \qquad (6)$$

This model allows today's conditional variance to be substantially affected by the (large) square error term associated with a major market move (in either direction) in any of the previous q periods. It thus captures the conditional heteroscedasticity of financial returns and offers an explanation of the persistence in volatility. A practical difficulty with the ARCH(q) model is that in many of the applications a long length q is called for.

3.2.3 GARCH

Bollerslev's Generalized Autogressive Conditional Heteroscedasticity [GARCH(p,q)] specification (1986) generalizes the model by allowing the current conditional variance to depend on the first p past conditional variances as well as the q past squared innovations. That is,

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon^2_{t-i} + \sum_{i=1}^{p} \beta_i h_{t-i} \ ; \qquad (7)$$

By accounting for the information in the lag(s) of the conditional variance in addition to the lagged ε^2_{t-i} terms, the GARCH model reduces the number of parameters required. In most cases, one lag for each variable is sufficient. The GARCH(1,1) model is given by :

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^{\ 2} + \beta_1 h_{t-1} \ ; \qquad (8)$$

GARCH can successfully capture thick tailed returns and volatility clustering. It can also readily be modified to allow for several other stylized facts of asset returns.

3.2.4 EGARCH

The Exponenetial Generalized Autoregressive Conditional Heteroscedasticity (EGARCH) model introduced by Nelson (1991) builds in a directional effect of price moves on conditional variance. Large price declines, for instance can have a larger impact on volatility than large increases. The EGARCH(1,1) specification, for example, is given by

$$\ln(h_t) = \omega + \beta_1 \ln(h_{t-1}) + \theta z_{t-1} + \gamma (|z_{t-1}| - (2/\pi)^{1/2}) ; \quad (9)$$

where $z_t = \varepsilon_t / h_t^{1/2}$ is i.i.d , $E(z_t) = 0$, var $(z_t) = 1$. If θ is < 0 the variance tends to rise (fall) when ε_{t-1} is negative (positive) in accordance with the empirical evidence for returns in some financial markets, especially stock and stock index returns.[13]

3.2.5 Conditional Distribution

In many financial time series the standardized residuals from the estimated models $z_t^{\wedge} = \varepsilon_t^{\wedge} / h_t^{1/2\wedge}$ display excess kurtosis which suggests departure from conditional normality. In such cases, the fat-tailed distribution of the innovations driving an ARCH process can be better modeled using the Student's-*t* or the Generalized Error Distribution (GED).

Taking the square root of the conditional variance and expressing it as an annualized percentage yields a time-varying volatility estimate. A single estimated model can be used to construct forecasts of volatility over any time horizon.

3.2.6 Implied Volatility

An alternative method to infer the conditional variance, or volatility, is to back-out an implied volatility estimate from the observed market price of an option and a particular option valuation model. These are fundamentally different measures than the historical measures since they should reflect, at least to some extent, the market's expectation of volatility. In this sense, they are forward-looking estimates in contrast to the backward-

[13] In equity markets this phenomenon is termed the "leverage effect", where a decrease in the stock price makes the firm more leveraged (in terms of the debt-to-equity ratio) and thus "riskier". The theoretical rationale for an assymetric effect in other markets, such as currencies or commodities is less well developed.

looking measures based exclusively on historical returns data. Option prices are calculated using no-arbitrage techniques like the Black-Scholes formula (or a generalized form of the B-S). For a listed option, all inputs to the formula are observable, except for the constant instantaneous variance, which can be backed out using the observed market price of the option. This yields an implied forecast for volatility of the underlying asset price, over the remaining life of the option. Certain analytical option pricing models, such as the Black- Scholes model, assume a constant variance, and are inconsistent, in a strict sense, with time-varying conditional variance. New models like stochastic volatility, implied volatility trees and implied binomial trees, allow for varying volatility.[14] For near-term, at-the-money (ATM) options, however, the effect of stochastic volatility is less pronounced and options-implied volatilities may provide statistically significant information, about future volatility, in addition to the information from time-series models.[15] Prices for traded options are determined by the market, using an estimate of volatility over the remaining life of the option. It is this estimate that is backed-out when implied volatility is calculated from observed market price of the option. If the implied volatility estimates are close to the actual volatility that prevails, then the options market is effectively using all available information in pricing the option; or, in other words, is informationally efficient.

3.3 Previous Studies

Day and Lewis (1993) compare forecasts of crude oil volatility from GARCH(1,1), EGARCH(1,1), implied volatility and historical volatility, based on daily data from November 1986 to March 1991. Using OLS regressions of realized volatility on out-of-sample forecasts, they check for unbiasedness of the forecasts (from the coefficient estimates) and for relative predictive power (from the R^2 figures). The accuracy of out-of-sample forecasts is compared using Mean Forecast Error (ME), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). They also check for the within-sample information content of implied volatility, by including it as predictor in the GARCH and

---

[14] Chriss (1997) pp. 361, 411
[15] Jorion (1997) pg. 180

EGARCH models and using Likelihood Ratio (LR) tests[16] on nested equations. They find that both implied volatilities and GARCH/ EGARCH conditional volatilities contribute incremental volatility information. The null hypothesis that implied volatilities subsume all information contained in observed returns is rejected, as is the hypothesis that option prices have no additional information. This would indicate that a composite forecast made using implied volatility and GARCH would yield better results since each would contribute unique information not contained in the other. However, in out-of-sample tests for incremental predictive power, results indicate that GARCH forecasts and historical volatility do not add much explanatory power to forecasts based on implied volatilities. Test for accuracy of forecasts based error criteria also support the conclusion that implied volatilities alone are sufficient for market professionals to predict near-term volatility (up to two months).

Duffie and Gray (1995) construct in-sample and out-of-sample forecasts for volatility in the crude oil, heating oil, and natural gas markets over the period May 1988 to July 1992. Forecasts from GARCH(1,1), EGARCH(1,1), bi-variate GARCH[17], regime switching[18], implied volatility, and historical volatility predictors are compared with the realized volatility to compute the criterion RMSE for forecast accuracy. They find that implied volatility yields the best forecasts in both the in-sample and out-of-sample cases, and in the more relevant out-of-sample case, historical volatility forecasts are superior to GARCH forecasts.

Xu and Taylor (1996) apply the methodology used by Day and Lewis (1993) to test for the informational efficiency of the Philadelphia Stock Exchange (PHLX) currency options market. They also construct volatility forecasts for British Pound, Deutsche Mark, Japanese Yen and Swiss Franc exchange rates quoted against the US dollar using data from January 1985 to January 1992. They improve on the Day and Lewis

[16] See section “Data and Methodology” for a description of the LR test

[17] The bi-variate GARCH model includes volatility information (returns, conditional variance) from a related market and the conditional covariance between the returns in the two markets. Duffie and Gray (1995) use the heating oil market as the related market.

[18] A form of stochastic volatility model in which volatility is allowed to attain a certain number of values with associated conditional probablilities.

methodology, however, by testing GARCH models with underlying Generalized Error Distribution (GED), to better account for the possibility of fat-tailed, non-normal conditional distribution of returns. In addition to using implied volatilities from options with short times to maturity, they also include an implied volatility predictor based on the term structure of volatility expectations. Based on in-sample tests they find that historical returns add no further information beyond that contained in implied volatility estimates. That is, they find that the PHLX currency options market is informationally efficient, and that the choice of the implied volatility predictor (term structure or short maturity) makes no difference to the conclusions. For out-of-sample volatility forecasting they find that implied volatility outperforms both the GARCH and historical volatility predictors.

CHAPTER 4. PROBLEM STATEMENT & APPROACH

The purpose of this study is to compare different methods of forecasting the volatility of crude oil futures prices using daily data from November 14,1986 to March 31,1997.

The forecasting models to be considered include:

- GARCH(1,1) ,
- EGARCH(1,1) ,
- Implied volatility , and
- Historical volatility (as a benchmark) ;

The evaluation of models can be done in two ways:

> *Within-Sample Tests*: These tests examine the information content of implied volatilities. This is done by adding the implied volatility from the option having the shortest time to expiration (but > 5 days) as an exogenous variable in the GARCH specification. Using nested equations with only implied volatility, only past returns, and implied volatility plus past returns, Likelihood Ratio (LR) tests with appropriate χ^2 statistics can be used to determine incremental predictive power.
>
> *Out-of-Sample Tests*: These tests evaluate forecasting accuracy relative to the realized volatility over a portion of the original sample. A four-week forecast horizon is used with the parameters for the GARCH formulation being recalculated each time with a rolling constant sample size. The four week forecasts are evaluated mid-month when the nearest expiry option has about four weeks to expiry. A shorter forecast horizon is also examined with two week GARCH forecasts constructed on the last day of the month, so as to match the horizon for implied volatility of the nearest expiry option with about two weeks to expiry.

Forecasting performance can be evaluated by using the mean forecast error (ME), the mean absolute error (MAE) and the root mean square error (RMSE) calculated from the forecast and the *ex-post* realized volatility. The relative predictive power of the alternative forecasts of volatility can also be evaluated by regressing realized volatility on its forecast and interpreting the coefficient of determination as a measure of forecast efficiency.

This study applies the rigorous approach of the Day and Lewis (1993) paper, adding GARCH with GED distribution as used in the Xu and Taylor (1996) analysis, to crude oil futures prices, over a longer observation period. The results of this study can be used to analyze the following issues:

(1) Does the use of a GED distribution in GARCH estimation provide a better representation of the fat-tailed nature of returns on crude oil futures, as found by Xu and Taylor (1994) for currency futures?

(2) Does the inclusion of more recent data, with presumably more liquid options markets (based on open interest) improve the information content of implied volatilities? Does increased liquidity of the options markets make them more informationally efficient in setting prices?

(3) Does the availability of a longer time series improve the estimates and hence the forecasting power of GARCH forecasts?

(4) Is the quick response advantage of implied volatilities to major market events (being based on market expectations) accentuated in the Day and Lewis (1993) and the Duffie and Gray (1995) studies where out-of-sample forecasts are calculated for the period following the Iraqi invasion of Kuwait? Would the difference between accuracy of forecasts be less pronounced or even completely different in an expanded data set?

CHAPTER 5. DATA AND METHODOLOGY

5.1 Datasets

The data consists of closing prices for the Light Sweet Crude Oil futures contract and the associated options on futures, traded on the New York Mercantile Exchange (NYMEX)[19] between Novermber 14, 1986 to March 31, 1997.[20]

Daily futures returns are calculated using the ratio of log prices on successive trading days. The options prices used are the nearest expiry call options, with more than 5 days to expiry. All implied volatilities are calculated for nearest-to-the-money options (i.e. those options most sensitive to changes in volatility and least noise from the "smile" effect)[21]. The estimates for the closest in-the-money and out-of-the-money call options are averaged to derive the final implied volatilities The one-month Eurodollar rate is chosen as the risk free interest rate.[22] Implied volatility estimates are generated using the Cox-Rubinstein binomial model function (aaBIN_ iv) in Financial CAD®, Version 1.0 b.

In-sample comparison of volatility models is made using the entire data set. Out-of-sample forecasts are constructed over the period January 1, 1994 through March 31, 1997. The parameters are estimated over a constant rolling sample size of roughly 1800 trading days (just over seven years), adding four weeks of the most recent data and dropping four weeks at the beginning of the sample each time.

5.2 Methodology

Returns defined as log price changes are given by equation (2)

---

[19] See Appendix C for contract specifications

[20] November 14, 1986 is the date crude oil futures options started trading on the NYMEX. The end date was chosen based on availability of data from NYMEX.

[21] The term "volatility smile" derives from the fact that a plot of strike price versus implied volatility is a smile shaped curve centered around the at-the-money option. See Xu and Taylor (1996) pg. 184, Day and Lewis (1993) pg. 34, Hotopp (1997) pg. 22

[22] Xu and Taylor (1996) pg. 183

Price Return, $r_t = \ln (P_t / P_{t-1})$;

assuming that t represents one business day and that all close-to-close period are of the same length. [23]Storage costs and convenience yields being small and predictable over short time horizons have a relatively minor effect on volatility in energy markets, and can thus be ignored.[24]

Thus,

$$\text{Innovation,} \quad \varepsilon_t = r_t - \mu \tag{10}$$

$$= r_t \quad \text{if conditional mean } \mu = 0$$

(a reasonable assumption for futures data)[25] ;

The GARCH(1,1) model then defines the conditional variance as

$$h_t = \omega + \alpha_1 \varepsilon^2_{t-1} + \beta_1 h_{t-1} ; \tag{11}$$

The GARCH model is estimated with an underlying GED distribution and with an underlying Normal distribution (GED with tail thickness parameter equal to 2). Twice the difference between the optimal value of the log likelihood functions can be used to run a Likelihood Ratio (LR) test with χ_1^2, to check which characterization better describes the returns distribution.[26]

On each day the implied volatility σ_{It-1} is calculated for the option closest to maturity (but > 5 days). To test the hypothesis that options prices give optimal one-period ahead forecasts, two more GARCH models are estimated

$$h_t = \omega + \alpha_1 \varepsilon^2_{t-1} + \beta_1 h_{t-1} + \delta\, \sigma^2_{It-1} , \tag{12}$$

and

$$h_t = \omega + \delta\, \sigma^2_{It-1} ; \tag{13}$$

[23] See Heynan and Kat (1994) pg. 55

[24] Duffie and Gray (1995) pg. 40

[25] Xu and Taylor (1996) pg. 188

[26] See Appendix A for a discussion of this point

Likelihood ratio tests of the null hypothesis that returns contain no volatility information in addition to that already conveyed by options prices (i.e. $\alpha_1 = \beta_1 = 0$) can be evaluated by comparing LR= 2 (L_1 - L_0) with χ_2^2 (L_0 and L_1 are the maximum log-likelihoods for the equations (13) and (12), respectively). Similarly, since equation (11) is nested in equation (12), the LR test can be used to check the hypothesis that options prices have no incremental information content i.e. $\delta = 0$. The χ^2 statistic in this case has one degree of freedom.[27]

For the EGARCH (1,1) model Equations (11)-(13) can be restated as

$$\ln(h_t)= \omega + \beta_1 \ln(h_{t-1})+ \theta z_{t-1} + \gamma (|z_{t-1}| - (2/\pi)^{1/2}) \quad , \tag{14}$$

$$\ln(h_t)= \omega + \beta_1 \ln(h_{t-1})+ \theta z_{t-1} + \gamma (|z_{t-1}| - (2/\pi)^{1/2}) + \delta \ln(\sigma^2_{It-1}) \text{ , and} \tag{15}$$

$$\ln(h_t = \omega + \delta \ln(\sigma^2_{It-1}) \quad ; \tag{16}$$

To compare the *ex ante* forecasting ability of historical volatility predictors -- forecasts from GARCH(1,1), EGARCH (1,1) – and implied volatility forecasts; the sample can be split into two parts with roughly 2/3 of the data points used to estimate the parameters in the GARCH model, while the remaining data is used to test the forecasts. At each forecast date the following volatility forecasts can be calculated ;

Realized Volatility defined as $$V_{R,t} = [(1/N) \sum_{i=1}^{N} \varepsilon^2_{t+i}]^{½} \quad , \tag{17}$$

where N is the number of days to expiry of the nearest expiry option.

Historical Volatility defined as $$V_{H,t} = [(1/N) \sum_{i=0}^{N-1} \varepsilon^2_{t-i}]^{½} = V_{R,t-N} \quad , \tag{18}$$

[27] Xu and Taylor (1996) pg. 186, Day and Lewis (1993) pg. 40

is the benchmark forecast calculated over the previous N days.

GARCH Volatility defined as $V_{G,t} = [(1/N) \sum_{i=1}^{N} h^{\wedge}_{t+i}]^{1/2}$, (19)

where[28] $h_{t+1} = \omega + \alpha_1 \varepsilon^2_t + \beta_1 h_t$, (20)

and $h^{\wedge}_{t+i} = \omega /(1- \alpha_1 - \beta_1) + (\alpha_1 + \beta_1)^{i-1} [h_{t+1} - \omega /(1- \alpha_1 - \beta_1)]$; (21)

The parameters in the GARCH model are re-estimated on each forecast date (mid-month and month-end) by adding the latest four weeks of observations and deleting the first four weeks of observations in the previous sample. The forecast from the options market $V_{M,t}$ is obtained from a matched maturity option as the implied volatility for an option whose time to maturity is approximately 28 calendar days (for four week forecasts) and 14 calendar days (for two week forecasts).

Forecasting performance is evaluated by calculating the following measures of forecast error :

Mean Forecast Error: $ME = (1/n) \sum_{t \in S} (V_{F,t} - V_{R,t})$. (22)

Mean Absolute Error: $MAE = (1/n) \sum_{t \in S} |V_{F,t} - V_{R,t}|$. (23)

Root Mean Square Error: $RMSE = [(1/n) \sum_{t \in S} (V_{F,t} - V_{R,t})^2]^{1/2}$. (24)

In the above expressions S indicates the set of times at which ex ante forecasts are produced and n denotes the number of forecasts made using each method[29]. The ME criterion can be expected to have the lowest value amongst the three measures, since positive and negative deviations would average out. The other two criteria avoid that by taking an absolute value (MAE) / squared valued (RMSE) of the deviation. The MAE

[28] Xu and Taylor (1996) pg. 195, Day and Lewis (1993) pg. 40, Heynen and Kat (1994) pg. 52, Duffie and Gray (1995) pg. 47
[29] Xu and Taylor (1996) pg. 195, Day and Lewis (1993) pg. 46

treats large and small deviations equally while the RMSE criterion penalizes large deviations more severely. Note that an optimal forecast will not have MAE = RMSE = 0 because $V_{R,t}$ is only a point estimate of the asset's price volatility which is unobservable.[30]

Further insight into the nature of the different forecast models can be gained by estimating regressions of the form

$$V_{R,t} = b_0 + b_1 V_{F,t} + \eta_t \quad ; \tag{22}$$

where η_t represents the forecast error.[31] Consistent estimates of the regression coefficients in the above equation can be obtained using ordinary least squares (OLS). If the forecasts of volatility are unbiased, the estimate of b_0 will be approximately zero, and the estimate of b_1 will be close to one. The R^2 values can be used as a measure of explanatory power of the different forecasts. The joint null hypothesis that the forecasts are unbiased (i.e. $b_0 = 0$ and $b_1 = 1$) can be examined using the F statistic.

The relative information content of the various forecasts (in terms of incremental predictive power) can be evaluated by running regressions of the form

$$V_{R,t} = b_0 + b_1 V_{M,t} + b_2 V_{G,t} + b_3 V_{H,t} + \eta_t \quad ; \tag{23}$$

By including alternative out-of-sample forecasts one-by-one, and comparing coefficients and R^2 values for composite forecasts with the individual forecast equations (22), the incremental predictive power of the forecasts can be gauged.

[30] Xu and Taylor (1996) pp.199. Lack of a statistical basis to use the RMSE criterion is discussed in Alexander (1996) pg. 257
[31] Day and Lewis (1993) pg. 41, Cumby, Figlewski and Hasbrouck (1993) pg. 55, Heynan and Kat (1994)

CHAPTER 6. EMPIRICAL RESULTS WITHIN –SAMPLE

6.1 Conditional Distribution

To evaluate which conditional distribution better describes the observed characteristics of daily oil price returns, the GARCH and EGARCH models in equations (11) and (14) are estimated using an underlying GED distribution and with an underlying Normal distribution (GED with tail thickness parameter $\nu = 2$). The results are reported in Table 1. T-statistics appear in italics below the respective coefficient estimates. The existence of fat tails is seen from the statistically significant estimates of the tail thickness parameter υ (< 2) in the GED specifications. The difference in maximum log-likelihoods between the GED and the normal conditional distribution is around 90 for both GARCH and EGARCH models. Since the normal distribution is the GED with one restriction (υ=2), the specifications can be compared using the LR test. Computing 2(L1-L0) and comparing with χ^2_1, shows statistical significance at the 0.1 % level. The GED conditional distribution, therefore, better captures the thick-tailed nature of daily returns for crude oil.

6.2 Information Content of Implied Volatilities vs. ARCH models

The information content of implied volatilities versus returns information in ARCH type models is examined using maximum likelihood estimates of equations (11)-(13) for GARCH models and equations (14)-(16) for EGARCH models. Both GARCH and EGARCH models are estimated using the GED conditional distribution.[32]
Results reported in Table 2a show that for the GARCH(1,1) model, adding implied volatility as a predictor leads to an increase of over 40 in the estimated value of maximum log-likelihood. Since equation (11) is equation (12) with the restriction $\delta = 0$, the results can be compared using the LR test statistic distributed as χ^2_1 . The expression

[32] See Appendix D for sample Regression Analysis of Time Series (RATS) code and results

2(L1-L0) evaluates to 80.96 rejecting the null hypothesis that implied volatilities do not add any information to the GARCH model, at the 0.1% level.

The hypothesis that implied volatilities subsume all information contained in the time series of returns can be tested by comparing the L values of equations (12) and (13). Equation (13) is equation (12) with the restrictions $\alpha_1 = 0$ and $\beta_1 = 0$. The value of the LR statistic in this case is 35.33 which exceeds 13.82, the critical χ^2_2 value at the 0.1% significance level. Both implied volatilities and historical returns thus contribute statistically significant information about future volatility. Virtually similar log-likelihood differences are obtained in the case of EGARCH (1,1) models, displayed in Table 2b, echoing the results and conclusions for the GARCH(1,1) models.

For implied volatilities to act as an unbiased estimator of day-ahead volatility, the coefficients ω and δ in equations (13) and (16) must equal 0 and 1 respectively. The joint hypothesis $\omega = 0$ and $\delta = 1$ can again be tested using the LR test, by estimating a restricted version of equations (13) and (16). The χ^2_2 value is 6.22 for GARCH and 7.90 for EGARCH, which rejects the null hypothesis that implied volatilities provide unbiased day-ahead forecasts at the 5% level.

6.3 Correlation between Returns and Volatility

The parameter θ in the EGARCH specification allows for correlation between the conditional variance and the standardized residuals z_t. Unlike in equity markets the "leverage effect" may not be present in the crude oil market. For the EGARCH GED model, θ is not significantly different from zero. An LR test comparing the EGARCH(1,1) specification with the symmetric EGARCH model (with the restriction $\theta = 0$), yields a χ^2_1 of 0.95. The null hypothesis that $\theta = 0$ cannot, therefore, be rejected at the 10% level. A comparison of the maximum log-likelihoods for GARCH GED and the symmetric EGARCH GED shows a difference of only 4.41. There is no clear-cut difference between the two models and thus the EGARCH model can be dropped from the list of models to be tested in the forecasting section.

Table 1

Conditional Distribution - Normal vs GED

GARCH (1,1) $h_t = \omega + \alpha_1 \varepsilon^2_{t-1} + \beta_1 h_{t-1}$ (11)

EGARCH(1,1) $\ln(h_t) = \omega + \beta_1 \ln(h_{t-1}) + \theta z_{t-1} + \gamma (|z_{t-1}| - (2/\pi)^{1/2})$ (14)

Variance Specification	ω	α_1	β_1	θ	γ	ν	Log L	χ^2
GARCH(1,1) GED	0.000	0.092	0.895			1.26	6732.30	
	3.691	*7.721*	*67.779*			*32.088*		
GARCH(1,1) Normal	0.000	0.115	0.877			2	6640.42	183.76
	5.675	*16.433*	*110.035*					
EGARCH(1,1) Normal	-0.101		0.986	0.012	0.185	1.27	6737.19	
	-3.042		*232.214*	*1.013*	*8.931*	*32.021*		
EGARCH(1,1) GED	-0.093		0.987	0.032	0.213	2	6647.34	179.71
	-4.872		*401.032*	*5.056*	*17.681*			

Table 2 a

Information Content of Implied Volatilities vs. ARCH Models

GARCH (1,1) with Conditional GED Distribution

GARCH(1,1) - Returns & Imp Vol	$h_t = \omega + \alpha_1 \varepsilon^2_{t-1} + \beta_1 h_{t-1} + \delta \sigma^2_{It-1}$	(12)
GARCH (1,1) - Returns only	$h_t = \omega + \alpha_1 \varepsilon^2_{t-1} + \beta_1 h_{t-1}$	(11)
GARCH(1,1) - Imp Vol only	$h_t = \omega + \delta \sigma^2_{It-1}$	(13)

Variance Specification	$\omega \times 10^4$	α_1	β_1	δ	ν	Log L	χ^2
GARCH(1,1) - Returns & Imp Vol	-0.089	0.103	0.461	0.432	1.35	6772.78	
	-1.156	*4.683*	*4.237*	*4.139*	*30.299*		
GARCH (1,1) - Returns only	0.071	0.092	0.895		1.26	6732.30	80.96
	3.691	*7.721*	*67.778*		*32.088*		
GARCH(1,1) - Imp Vol only	-0.227			1.005	1.31	6755.12	35.33
	-1.505			*16.401*	*32.071*		

Table 2 b

Information Content of Implied Volatilities vs. ARCH Models

EGARCH (1,1) with Conditional GED Distribution

EGARCH(1,1) - Returns & Imp Vol	$\ln(h_t) = \omega + \beta_1 \ln(h_{t-1}) + \theta z_{t-1} + \gamma (\lvert z_{t-1} \rvert - (2/\pi)^{1/2}) + \delta \ln(\sigma^2_{It-1})$	(15)
EGARCH (1,1) - Returns only	$\ln(h_t) = \omega + \beta_1 \ln(h_{t-1}) + \theta z_{t-1} + \gamma (\lvert z_{t-1} \rvert - (2/\pi)^{1/2})$	(14)
EGARCH(1,1) - Imp Vol only	$\ln(h_t) = \omega + \delta \ln(\sigma^2_{It-1})$	(16)

Variance Specification	ω	β_1	θ	γ	δ	ν	Log L	χ^2
EGARCH(1,1) - Returns & Imp Vol	0.290	0.642	0.011	0.223	0.398	1.35	6772.70	
	1.817	*7.235*	*0.426*	*5.686*	*4.139*	*30.347*		
EGARCH (1,1) - Returns only	-0.102	0.986	0.012	0.185		1.27	6737.19	71.03
	-3.053	*231.311*	*-3.053*	*8.931*		*32.019*		
EGARCH(1,1) - Imp Vol only	0.681				1.096	1.31	6755.96	33.49
	1.845				*23.616*	*32.167*		

CHAPTER 7 : EMPIRICAL RESULTS: OUT-OF-SAMPLE FORECASTING

Using results from the in-sample testing, the GARCH GED model is selected as the representative ARCH type model in order to compare out-of-sample forecasting performance with implied volatilities and historical volatility.

Forecasts of average volatility over a four-week horizon are constructed from the three estimators $V_{G,t}$, $V_{M,t}$ and $V_{H,t}$ on the 14th of each month (or the previous trading day) from January 1994 to Feburary 1997 – a total of 38 forecasts. The nearest maturity option contract has an average of 29 days to expiry on these dates. Thirty nine two-week forecasts are generated on the last trading day of each month from Dec 1993 to February 1997. On each forecast date, the updated GARCH forecast $V_{G,t}$ is constructed using parameters estimated over the previous 1800 odd trading days. The horizon for the GARCH estimate of volatility is matched to the exact number of days the nearest maturity option contract has to expiry[33]. The same number of days is used as the value of N in equations (17) and (18) to calculate $V_{H,t}$ and $V_{R,t}$. Figures 5a and 5b plot the three forecasts against realized volatility for four-week and two-week horizons respectively. Also plotted is the long-run unconditional volatility calculated as

$$V_{U,t} = \omega / (1-\alpha_1 - \beta_1) \quad , \qquad (24)$$

with the parameter values taken from the updated GARCH estimates on every forecast date.

The Mean Forecast Error (ME), Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) for four-week forecasts are reported in Table 3a. The naïve historical volatility estimator performs best on all three error criteria, with GARCH (1,1) GED taking second place, and implied volatility trailing close behind (except on the MAE criterion where implied volatility performs marginally better than GARCH).

[33] Using equations (19)-(21)

The picture changes when forecasts are derived over the shorter two-week horizon. As results in Table 3b show, except for the ME criterion, the ranking of forecasts is reversed with implied volatility coming in first, followed by GARCH and then historical volatility. Also, in the two-week case, the gap between the quality of forecasts is more clearly pronounced than in the four-week comparison. In the four-week case the RMSE for historical volatility is only 4% lower than the RMSE for implied volatility. For two -week forecasts, however, the RMSE for implied volatility is 21% lower than that for historical volatility.

Predictive power can be compared by regressing realized volatility on each of the forecasts, for a given time period, as in equation (22). Comparing the coefficients of determination in Table 4a shows that the historical volatility predictor has higher predictive power ($R^2 = 0.344$) than implied volatility ($R^2 = 0.234$) and GARCH ($R^2 = 0.233$), over a four-week horizon. The joint hypothesis, $b_0 = 0$ and $b_1 = 1$, however, can be rejected at the 5% level, in the case of historical volatility, indicating that the historical volatility forecasts have significant statistical bias. On the other hand, the hypotheses that GARCH and implied volatility forecasts are unbiased forecasts of realized volatility over a four-week horizon cannot be rejected at the 10% level. For two-week forecasts, the ranking based on predictive power (R^2) is the same as that based on MAE and RMSE – implied volatility at the top followed by GARCH and historical volatility. However, none of the forecasts can be said to be unbiased predictors of average realized volatility over a two-week horizon. The $F(2, 37)$ values for all three predictors are above 7, rejecting the null hypothesis, $b_0 = 0$ and $b_1 = 1$, at the 1% level. The fact that all three forecasts have significant statistical biases, over the shorter two-week horizon, ties in with the observed magnitudes of error criteria, all of which are higher for the two-week case compared to the four-week values. As can be seen in Figure 5b, the forecasts tend to over-predict the level of volatility.

Since, both implied volatility and GARCH models add volatility information distinct from each other, composite forecasts are expected to have enhanced predictive power. The relative information content of each of the forecast models can be estimated by comparing regression results for two-model composite forecast equations with the individual forecast regressions. Results reported in Table 5a, for the four-week horizon, indicate that neither GARCH nor implied volatility adds much information to that contained in historical forecasts. The increase in R^2 is limited, even when all three predictors are used in the composite forecast. Also, only the historical volatility coefficient is significant in all regressions, which suggests that historical volatility subsumes information contained in other forecasts over the four-week horizon.

In the two-week case, the maximum increase in predictive power (> 25%) is obtained by combining the historical volatility and implied volatility predictors ($R^2 = 0.231$) as compared to 0.172 for historical volatility alone, and 0.183 for implied volatility alone. Adding the GARCH predictor to the composite historical-implied volatility predictor does not contribute additional predictive power.

Figure 5a : Comparison of Out-of-Sample Forecasts with Realized Volatility
Four - Week Forecasts

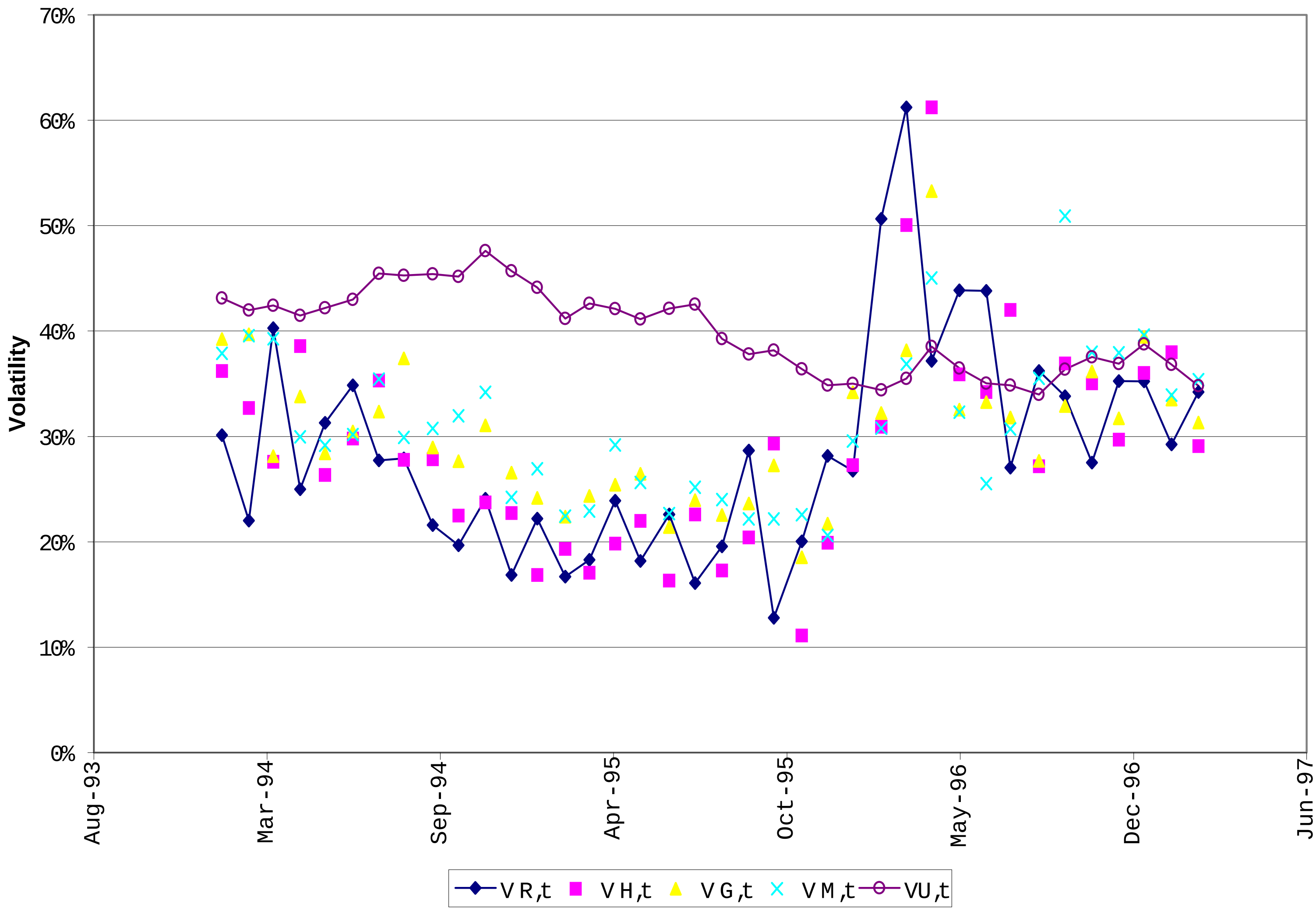


Figure 5b : Comparison of Out-of-Sample Forecasts with Realized Volatility
Two - Week Forecasts

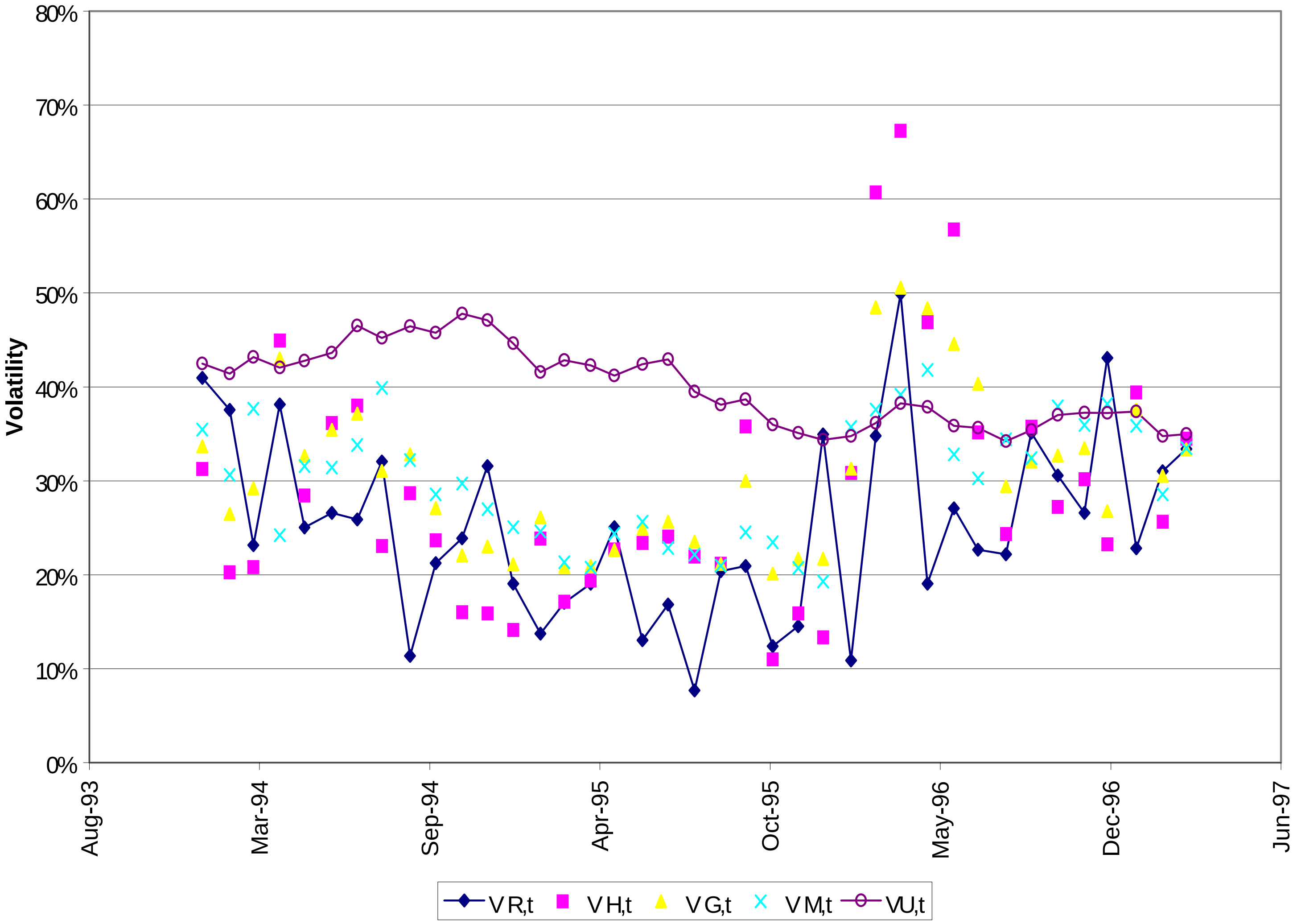


Table 3 a

Comparison of Out-of-Sample Forecast Accuracy

Four-Week Forecasts

Average Unconditional Volatility = 39.90%

Average Realized Volatility = 28.70%

Error	Historical Volatility	GARCH (1,1) GED	Implied Volatility
ME	0.16%	1.66%	2.38%
MAE	7.32%	7.60%	7.47%
RMSE	9.01%	9.13%	9.35%

Table 3 b

Comparison of Out-of-Sample Forecast Accuracy

Two-Week Forecasts

Average Unconditional Volatility = 40.02%

Average Realized Volatility = 25.16%

Error	Historical Volatility	GARCH (1,1) GED	Implied Volatility
ME	3.79%	5.44%	4.89%
MAE	9.98%	8.79%	8.39%
RMSE	12.90%	11.06%	10.21%

Table 4 a

Comparison of Out-of-Sample Predictive Power

Four-Week Forecasts

$$V_{R,t} = b_0 + b_1 V_{F,t} + \eta_t$$

Forecasting Model $V_{F,t}$	b_0	b_1	R^2	F(2,36)
Historic Volatility	0.112	0.606	0.344	4.02 *
	2.646	*4.353*		
GARCH(1,1) GED	0.063	0.737	0.233	1.34
	0.916	*3.312*		
Implied Volatility	0.068	0.702	0.234	2.31
	1.018	*3.319*		

* Significant at the 5% level

Table 4 b

Comparison of Out-of-Sample Predictive Power

Two-Week Forecasts

$$V_{R,t} = b_0 + b_1 V_{F,t} + \eta_t$$

Forecasting Model $V_{F,t}$	b_0	b_1	R^2	F(2,37)
Historic Volatility	0.160	0.315	0.172	21.72**
	4.477	*2.775*		
GARCH(1,1) GED	0.099	0.497	0.181	11.40**
	1.603	*2.860*		
Implied Volatility	0.055	0.652	0.183	7.04**
	0.798	*2.679*		

** Significant at the 1% level

Table 5 a

Comparison of the Relative Information Content for Out-of-Sample Forecasts

Four-Week Forecasts

$$V_{R,t} = b_0 + b_1 V_{F,t} + b_2 V_{G,t} + b_3 V_{H,t} + \eta_t$$

Forecast Combination	b_0	b_1	b_2	b_3	R^2
Imp Vol + GARCH	0.033	0.405	0.421		0.268
	0.463	*1.292*	*1.278*		
Hist Vol + GARCH	0.148		-0.280	0.774	0.352
	2.044		*-0.62*	*2.53*	
Imp Vol + Hist Vol	0.075	0.216		0.499	0.356
	1.206	*0.794*		*2.578*	
Imp Vol + GARCH + Hist Vol	0.120	0.332	0.493	0.739	0.375
	1.567	*1.123*	*-1.008*	*2.41*	

Table 5 b

Comparison of the Relative Information Content for Out-of-Sample Forecasts

Two-Week Forecasts

$$V_{R,t} = b_0 + b_1 V_{F,t} + b_2 V_{G,t} + b_3 V_{H,t} + \eta_t$$

Forecast Combination	b_0	b_1	b_2	b_3	R^2
Imp Vol + GARCH	0.044	0.393	0.292		0.217
	0.629	*1.283*	*1.247*		
Hist Vol + GARCH	0.114		0.360	0.094	0.182
	1.477		*0.680*	*0.275*	
Imp Vol + Hist Vol	0.063	0.438		0.197	0.231
	0.918	*1.651*		*1.49*	
Imp Vol + GARCH + Hist Vol	0.077	0.510	-0.232	0.319	0.233
	0.965	*1.519*	*-0.357*	*0.866*	

CHAPTER 8. SUMMARY AND CONCLUSIONS

Within-sample tests of the forecasting models yield results that are in conformance with earlier studies and other financial time series forecasting literature.

The GARCH GED model is found to increase the maximum log-likelihood function value significantly, demonstrating that the GED conditional distribution is better able to capture the fat-tailed distribution characteristic of financial price returns, confirming the applicability of a similar conclusion reached by Xu and Taylor (1994) for currency returns.

This study confirms the findings of Day and Lewis (1993) about the relative information content of implied volatilities and historical returns information contained in ARCH type models, for oil futures, over a larger sample period from November 14, 1986 to March 31, 1997. Implied volatilities and time series returns information are found to contribute incremental volatility information beyond that contained in the other. The results are different from those reported by Xu and Taylor, who show that the implied volatility information from the Philadelphia currency options market subsumes information contained in historical returns. The expansion in the volume of options trading in the last six years, since the Day and Lewis (1993) study does not seem to have improved the informational efficiency of the NYMEX crude oil options market in predicting volatility.

The out-of-sample comparisons of forecasting accuracy are in conformance with forecasting literature, which suggests that over longer time periods, historical volatility often outperforms other estimators.[34] Historical volatility beats GARCH forecasts in both the Day and Lewis (1993) study as well as the Duffie and Gray (1995) forecasts for crude oil. The same result is achieved for four-week forecasts in this study, but importantly, the two-week forecasts indicate greater accuracy for GARCH forecasts over historical volatility. GARCH models produce damped forecasts, which being an average of the

[34] Hotopp (1997) pg. 23

recent volatilities and the underlying mean volatility, tend to mean-revert the further ahead the model is asked to forecast, which could explain the better performance over the shorter horizon. The increased forecast efficiency could also be a result of using a longer sample period to estimate GARCH parameters, as compared to the other studies.

Improved forecasts from implied volatility over the shorter two-week period matches a similar finding by Day and Lewis (1993) in their comparison of near-term and distant-term volatility forecasts, though they use a different number of days from the 12 and 29 days used in this study.

The result that stands in sharp contrast to earlier studies is the month-ahead volatility forecast performance of the implied volatility predictor. One reason could be the fact that implied volatilities lose their competitive advantage, when asked to forecast over a period of moderate shocks, instead of periods with large price shocks. Both the Day and Lewis (1993) study and the Duffie and Gray (1995) study use the post-Gulf War period to test forecasts, which may bias the results in favor of the forward looking implied volatility measure. The other reason could be the implied volatility estimates used. This study uses an average implied volatility derived from two near at-the-money call options with > 5 days to expiry. Further refinements are possible such as including put options; increasing the number of options used to derive the average and excluding implied volatilities more than five standard deviations from their mean, or values from days when fewer than 100 contracts were traded. Also, the nearest expiry option contract has variable time to expiry on successive dates (between 7 to 36 days)- the term structure of volatility expectations – which impacts implied volatility estimates. The Day and Lewis (1993) paper adjusts for the variable length of forecast intervals, and the Xu and Taylor (1994) paper estimates a time-varying term structure model. This study attempts to solve the problem by working backwards from the number of days the nearest maturity option contract has to expiry on each forecast date and matching the horizon of GARCH and historical forecasts, on each forecast date. Though this ensures matched horizon forecasts on each date, it also means that the error criteria and regressions are computed using forecasts made over 8 to 16 days (average 12 days) for the two-week forecasts and 24 to 32 days (average 29 days) for the four-week forecasts. All forecasting results must also be viewed with some caution since they are calculated using less than 40 forecasts.

The computationally intensive GARCH model is outperformed by implied volatility for two-week forecasts and by historical volatility for four-week forecasts in terms of forecast accuracy and predictive power. Also, as the comparison of relative information content shows, GARCH forecasts do not add significant incremental predictive power beyond a combination of implied volatility and historical volatility forecasts. Practitioners in the oil industry, could thus avoid complicated time series analysis and instead base two-week and four-week forecasts on a composite implied volatility + historical volatility predictor. Greater rigor can be applied in the generation of implied volatility estimates to further improve forecast quality.

GARCH forecasts can be improved by explicitly modeling the observed mean reversion in log prices[35] and including a measure of backwardation/ contango as reflected in the price differential between the first nearby and the second nearby futures prices. A composite predictor based on a truer characterization of observed return characteristics in the flexible GARCH models and more sophisticated implied volatility estimates could significantly improve volatility forecasts in the crude oil market.

[35] Pilipovic (1997) pg. 78

BIBLIOGRAPHY

Conceptual Issues

Verleger, Philip K. *Adjusting to Volatile Energy Prices*, Institute for International Economics, Washington DC, 1993.

Pilipovic, D. *Energy Risk*, McGraw Hill, New York, 1998.

Petroleum 1996 – Issues and Trends, Energy Information Administration, U.S. Department of Energy, Washington DC, 1997.

Oil Price History and Analysis, WTRG Economics Website, 1998.

Hotopp, S. " Practical Issues Concerning Volatility and Its Measurement, Past and Predicted ", *Volatility in Capital Markets*, (ed. I. Nelken), Glenlake Publishing Co. Ltd., Chicago, 1997.

Jorion, P. "Forecasting Risks and Correlations", *Value At Risk - The New Benchmark for Controlling Market Risk*, The McGraw-Hill Companies Inc., 1997, pp. 165 - 182.

RiskMetrics - Technical Document, Fourth Edition, J.P. Morgan & Co. Inc., New York, 1996.

Chriss, Neil A, *Black Scholes and Beyond*, Irwin, New York, 1997.

Volatility Modeling

Mills, Terence C. *The Econometric Modelling of Financial Time Series*, Cambridge University Press, 1993.

"Modeling, estimating, forecasting and comparing monthly volatility: A comprehensive study", *RiskMetrics Monitor,* Fourth Quarter 1995, J.P. Morgan & Co. Inc., New York, 1995.

Enders, W. "Modeling Economic Time Series: Trends and Volatility", *Applied Econometric Time Series*, John Wiley & Sons, 1995, pp. 135 - 210

Campbell J.Y., A.W. Lo and A.C. MacKinlay. "Nonlinearities in Financial Data", *The Econometrics of Financial Markets,* Princeton University Press, Princeton, 1997.

Alexander, C. " Volatility and Correlation Forecasting", *The Handbook of Risk Management and Analysis*, (ed. C.Alexander), John Wiley & Sons, Chichester, 1996.

ARCH - Survey articles

Engle R.F. and D.L. McFadden (eds.). "ARCH Models", *Handbook of Econometrics, Vol IV*, Elsevier Science, 1994, pp. 2961 - 3038.

Bollerslev T., R.Y.Chou and K.F. Kroner. "ARCH Modeling in Finance - A review of theory and empirical evidence", *Journal of Econometrics 52*, 1992, pp. 6 - 59.

Applications

Day T.E. and C.M. Lewis. "Forecasting Futures Markets Volatility", *The Journal of Derivatives*, Winter 1993, pp. 33 -50.

Duffie D. and S. Gray. " Volatility in Energy Prices", *Managing Energy Price Risk,* Risk Publications, London ,1995, pp. 39 - 55.

Hsieh D.A. "Modeling Heteroscedasticity in Daily Foreign-Exchange Rates", *Journal of Business & Economic Statistics*, July 1989, pp. 307 - 317.

Xu X. and S.J. Taylor. “Conditional Volatility and Informational Efficiency of the PHLX Currency Options Market”, *Forecasting Financial Markets,* (ed C. Dunis), John Wiley & Sons, Chichester, 1996.

Nelson, D.B. “Modeling Stock Market Volatility Changes”, *Modeling Stock Market Volatility,* (ed P.E. Rossi), Academic Press, San Diego, 1996.

Nelson, D.B. “Conditional Heteroskedasticity in Asset Returns : A New Approach”, *Modeling Stock Market Volatility,* (ed P.E. Rossi), Academic Press, San Diego, 1996.

Cumby R., S. Figlewski and J. Hasbrouck . “Forecasting Volatilities and Correlations with EGARCH Models”, *The Journal of Derivatives,* Winter 1993, pp. 51 - 63.

Heynen R.C. and H.M. Kat. “Volatility Prediction : A Comparison of the Stochastic Volatility, GARCH (1,1) and EGARCH (1,1) Models”, *The Journal of Derivatives*, Winter 1994,pp. 50 -65.

APPENDIX A

STATISTICAL MODELING OF FINANCIAL TIME SERIES

Introduction

Analysis of risk and uncertainty in financial markets has given rise to techniques which allow for modeling of temporal dependencies in variances and covariances. The key insight offered by the ARCH model lies in the distinction between the *conditional* and *unconditional* second order moments. While the unconditional covariance matrix for the variables of interest may be time invariant, the conditional variances and covariances often depend non-trivially on the past states of the world. Neglecting heteroscedasticity can lead to a large loss in asymptotic efficiency and better evaluation of economic forecasts can be done by conditioning information on the current information set.

To analyze a financial time series using formal statistical methods, it is useful to regard the observed series, $(y_1, y_2, ..., y_t)$ often denoted as $\{y_t\}_1^T$ as a particular realization of a stochastic process.

The stochastic process can be thus described by a T-dimensional probability distribution, the first and the second moments of which are the

T means : $E(y_1)$, $E(y_2)$, ..., $E(y_t)$;

T variances : $\sigma^2(y_1)$, $\sigma^2(y_2)$, ..., $\sigma^2(y_t)$;

and the ${}^TC_2$ covariances : $Cov(y_i, y_j)$, $i < j$.

To infer the above T+(T+1)/2 parameters from a single realization (T values) we need to make the simplifying assumptions of *ergodicity* (sample moments approach population moments as T --> infinity) and *stationarity* (the mean and variance of the process are constant and the autocovariances/ autocorrelations depend only the lag and not on time t).

i.e. $E(y_1) = E(y_2) = ... = E(y_t) = \mu$;

$\sigma^2(y_1) = \sigma^2(y_2) = ... = \sigma^2(y_t) = \sigma_y^2$;

Autocovariance $\gamma_k = Cov(y_t, y_{t=k}) = E[(y_t - \mu)(y_{t-k} - \mu)]$;

Autocorrelation function (ACF) $\rho_k = Cov(y_t, y_{t-k}) / [\sigma^2(y_t). \sigma^2(y_{t-k})]^{1/2} = \gamma_k / \gamma_o$;

The ACF indicates the length and strength of the "memory" of the process and plays a major role in modeling dependencies among observations.

A fundamental theorem in time series analysis known as *Wold's decomposition* states that every weakly stationary, purely non-deterministic (any component that can be perfectly predicted from past values has been subtracted), stochastic process $(y_t - \mu)$ can be written as a linear combination (or linear filter) of a sequence of uncorrelated random variables.

This linear filter representation is given by

$$y_t - \mu = a_t + \Psi_1 a_{t-1} + \Psi_2 a_{t-2} + ... \quad ; \qquad (1)$$

The a_t are a sequence of uncorrelated random variables, often known as *innovations*, drawn from a fiyed distribution with $E(a_t) = 0$, $V(a_t) = \sigma^2$ < infinity and $Cov(a_t, a_{t-k}) = E(a_t, a_{t-k}) = 0$ for all k <>0.Such a sequence is a *white-noise* process $a_t \sim WN(0, \sigma^2)$. $\Sigma|\Psi_j|$ < infinity is the condition for y_t to be stationary.

Taking $\mu = 0$ and choosing $\Psi_j = \phi^j$ in equation (1) gives

$y_t - \phi\, y_{t-1} = a_t$ or $(1 - \phi B)\, y_t = a_t$; a *first-order autoregressive* process AR(1)

In general,

$$y_t = (1+ \phi B + \phi^2 B^2 + ...)a_t$$
$$= a_t + \phi a_{t-1} + \phi^2 a_{t-2} + ... \quad ;$$

This linear filter representation will converge as long as $|\phi| < 1$, which is therefore the stationarity condition.

Taking $\mu = 0$ and choosing $\Psi_1 = -\theta$ and $\Psi_j = 0$ in equation (1) gives

$y_t = a_t - \theta a_{t-1}$ or $(1- \theta B)^{-1} y_t = a_t$ a *first-order moving average* process MA(1)

which will converge if $|\theta| < 1$ i.e. if the model is *invertible*. Implies that the effect of past observations decreases with age).

$(1-\phi B) y_t = (1- \theta B)a_t$ is a *first-order autoregressive moving average* process (ARMA)

More general ARMA(p,q) processes are obtained by combining AR(p) and MA(q) processes :

$$y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - ... - \phi_p y_{t-p} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - ... - \theta_\theta a_{t-q} ,$$

or $(1-\phi_1 B - \phi_2 B^2 - ... - \phi_p B^p) y_t = (1-\theta_1 B - \theta_2 B^2 - ... - \theta_q B^q) a_t ,$

i.e. $\phi(B) y_t = \theta(B) a_t ;$

In the AR (1) process, if $\phi=1$

$y_t = y_{t-1} + a_t ,$ then y_t is said to follow a *random walk.*

If a constant θ_0 is included

$y_t = y_{t-1} + \theta_0 + a_t$, then y_t is said to follow a *random walk with drift.*

The random walk is an example of a class of non-stationary processes (ϕ not < 1) known as *integrated processes*.

First differencing y_t leads to a stationary model $\Delta y_t = \theta_0 + a_t$.

Generally, a series may need first differencing d times to attain stationarity, and the series so obtained may itself be autocorrelated. If this autocorrelation is modeled by an ARMA(p,q) process, then the model for the original series is of the form

$$\phi(B)\, \Delta y_t = \theta_0 + \theta(B)\, a_t \ ;$$

which is said to be an *autoregressive-integrated-moving average* process of orders p,d and q, or ARIMA (p,d,q) and y_t is said to be integrated of order d, I(d).

Univariate non-linear stochastic models

In the model

$$y_t = y_{t-1} + a_t \ ;$$

a_t can be

(1) white noise : stationary and uncorrelated sequence drawn from a fiyed distribution (strict white noise if independent as well), in which case y_t is said to follow a random walk or

(2) a martingale difference : uncorrelated but not necessarily stationary : this implies that there could be dependence between higher conditional moments, most notably conditional variances.

Non-linear stochastic processes are capable of modeling volatility in financial time series which display protracted quiet periods interspersed with bursts of turbulence. A simple way in which non-linearity can be introduced into a time series is to allow the variance (or the conditional variance) of the process to change either at discrete points in time or continuously. For a non-linear stationary process $\{y_t\}$, the variance $\sigma^2(y_t)$ is a constant for all t, but the conditional variance $\sigma^2[y_t | y_{t-1}, y_{t-2}, ...]$ depends on the observations and thus can change from period to period.

ARCH processes

Consider a case where conditional standard deviations are a function of past values of y_t i.e.

$$\sigma_t = F(y_{t-1}) = \{\omega + \alpha_1 (y_{t-1} - \mu)^2\}^{1/2} \;;$$

where ω and α_1 are both positive, and y_t is generated by the process

$$y_t = \mu + \sigma_t z_t \quad ;$$

With $z_t \sim NID(0,1)$ and independent of σ_t, y_t is then white noise and conditionally normal, i.e.

$$y_t | y_{t-1}, y_{t-2}, ... \sim NID(\mu , \sigma_t^2) \; ,$$

so that conditional variance $\sigma^2(y_t | y_{t-1}) = \omega + \alpha_1 (y_{t-1} - \mu)^2$.

This process is stationary if, and only if, $\alpha_1 < 1$, in which case the unconditional variance is

$$\sigma^2(y_t) = \omega / (1- \alpha_1).$$

This model was first introduced by Engle(1982) and is known as a *first order autoregressive conditional heteroscedastic* [ARCH(1)] process.

A more convenient notation is to define $\varepsilon_t = y_t - \mu = \sigma_t z_t$ and $h_t = \sigma_t^2$, so that the ARCH(1) model can be written as

$$\varepsilon_t \mid y_{t-1}, y_{t-2}, \ldots \sim NID(0, h_t) ,$$

$$h_t = \omega + \alpha_1 \varepsilon^2_{t-1} \;;$$

The conditional variance function for a general ARCH(q) process can be written as

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon^2_{t-i} = \omega + \alpha(B) \varepsilon^2_{t-i} \quad ;$$

where B denotes the lag or backshift operator $B^i y_t = y_{t-i}$. For the model to be well defined and the conditional variance to be positive $\omega > 0$ and the α_i s >= 0.

Defining $v_t = \varepsilon^2_t - h_t$, the ARCH(q) model may be re-written as

$$\varepsilon^2_t = \omega + \alpha(B) \varepsilon^2_{t-i} + v_t \quad ;$$

Since $E_{t-1}(v_t) = 0$, the model corresponds directly to an AR(q) model for the squared innovations ε^2_t . The process is covariance stationary if and only if the sum of the positive autoregressive parameters is less than one, in which case the unconditional variance equals $Var(\varepsilon_t) = \sigma^2 = \omega \,/\, (1 - \alpha_1 - \alpha_2 \ldots - \alpha_q)$.

Even though the ε_t's are serially uncorrelated, they are clearly not independent through time. The standardized process, $z_t = \varepsilon_t / \sigma_t$ will have a conditional mean of zero and a time invariant conditional variance of unity. The unconditional distribution for ε_t will have fatter tails than the distribution for z_t. In many financial time series this may not adequately account for the leptokurtosis i.e. the standardized residuals from the estimated models $z_t^{\wedge} = \varepsilon_t^{\wedge} / \sigma_t^{\wedge}$ are leptokurtic which suggests departure from conditional normality. In addition to the potential gains in efficiency, the exact form of the error distribution also plays an important role in several important applications of the ARCH model such as option pricing. Other parametric densities that have been considered in the estimation of ARCH models include the Student-t , normal-Poisson mixture, power exponential, generalized exponential and the normal - lognormal mixture distributions.

The Student-t distribution, used by Bollerslev (1987) is identical to a normal-inverted gamma mixture

$$f(z) = n^{-0.5} \pi^{-0.5} \Gamma (n+1)/2 \, \Gamma (n/2)-1 \text{ x } (1 + z^2 / n-2)^{-0.5 (n+1)} \;;$$

where n is the degree of freedom. As $n \rightarrow \infty$, the distribution becomes the standard normal.

Nelson (1988) suggested the generalized error distribution (GED)

$$f(z) = 0.5 \, \upsilon \, \Gamma (3/ \upsilon) \, \Gamma (1/ \upsilon) -1.5 \exp(-0.5|z/\lambda|^{\upsilon}) \;;$$

where $\lambda = (1/\upsilon)^{0.5} \Gamma (3/ \upsilon)^{-0.5} 2^{-2/\upsilon}$ is used to used to normalize the variance at unity. For υ=2 the GED distribution is equivalent to the normal distribution.

In many of the applications with the linear ARCH(q) model a long length q is called for. A practical difficulty with ARCH models is that with large q, estimation will often lead to violation of the non-negativity constraints on the α_i s that are required to ensure that h_t is always positive.

To obtain more flexibility, the *generalized* ARCH (GARCH) process has been proposed. The GARCH (p,q) process has the conditional variance function

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon^2_{t-i} + \sum_{i=1}^{p} \beta_i h_{t-i} = \omega + \alpha(B)\varepsilon^2_{t-i} + \beta(B) h_{t-i} \quad ;$$

in which $E[h_{t+s} \mid \varepsilon_t, \varepsilon_{t-1} \ldots] \dashrightarrow \sigma^2$ (unconditional variance of ε_t) as s --> infinity i.e. the discrete time (as financial observations tend to be) GARCH (1,1) process converges to a continuous time diffusion model (in which modern finance's stochastic differential equation theories are set) as the sampling interval gets arbitrarily small.

The GARCH(1,1) model

$$h_t = \omega + \alpha_1 \varepsilon^2_{t-1} + \beta_1 h_{t-1} \quad , \alpha_0 > 0, \alpha_1 >=0, \beta_1 >=0 \quad ;$$

can be written as

$$\varepsilon^2_t = \omega + (\alpha_1 + \beta_1)\varepsilon^2_{t-1} + v_t - \beta_1 v_{t-1} \quad \text{by defining } v_t = \varepsilon^2_t - h_t .$$

By ARMA analogue, the process $\{\varepsilon_t\}$ will be stationary if $\alpha_1 + \beta_1 < 1$

Setting $\alpha_1 + \beta_1 = 1$ (the autoregressive polynomial has a unit root), leads to the *integrated* ARCH, or IGARCH(1,1) model which with $\alpha_0 = 0$ is

$$h_t = \alpha_1 \varepsilon^2_{t-1} + (1 - \alpha_1) h_{t-1} \quad ;$$

In the IGARCH model a shock to the conditional variance is persistent in the sense that it remains important for future forecasts of all horizons.

Just as an ARMA model often leads to a more parsimonious representation of the temporal dependencies in the conditional mean than an AR model, the GARCH formulation provides similar flexibility over the ARCH model when parameterizing the conditional variance. In most empirical applications with finely sampled data, the simple

GARCH(1,1) model with $\alpha_1 + \beta_1$ close to one is found to provide a good description of the data.

GARCH successfully captures thick tailed returns, and volatility clustering, and can readily be modified to allow for several other stylized facts of asset returns such as non-trading periods (lower volatility) and predictable information releases (high ex-ante volatility) by making α_0 a function of time. It is not well suited to capture the "leverage effect" - the tendency for asset prices to be negatively correlated with changes in volatility - since in the GARCH formulation the conditional variance is a function only of the magnitudes of the lagged residuals and not their signs.

In the exponential GARCH (EGARCH) model, h_t is an asymmetric function of past ε_t s (depends on both the size and the sign of the lagged residuals). The EGARCH (1,1) model is

$$\ln(h_t) = \omega + \beta_1 \ln(h_{t-1}) + \theta z_{t-1} + \gamma(|z_{t-1}| - (2/\pi)^{1/2}) ;$$

where $z_t = \varepsilon_t / \sigma_t$ is i.i.d , $E(z_t) = 0$, var $(z_t) = 1$. Unlike in the linear GARCH(1,1) model there are no restrictions on the parameters to ensure non-negativity of the conditional variance. If θ is < 0 the variance tends to rise (fall) when ε_{t-1} is negative (positive) in accordance with the empirical evidence for retuns in many financial markets.

Another important class of models is the switching ARCH or SWARCH model. This class of models postulates that there are several different ARCH models and that the asset switches from one to another following a Markov chain.

Many theories in finance (e.g. CAPM) call for an explicit tradeoff between the expected returns and the variance, or the covariance between returns. The (G)ARCH in mean, or (G)ARCH-M models capture such relationships by making the conditional mean an explicit function of the conditional variance in an attempt to incorporate a measure of risk into the returns generating process.

The variety of parametric ARCH models complicates the search for the "true" model but provides flexibility to find the context specific model that makes analysis of time-varying volatility tractable.

Testing for ARCH

The original Lagrange Multiplier (LM) test proposed by Engle involves the computation of R^2 from the regression of ε^2_t on a constant and $\varepsilon^2_{t-1}, \ldots, \varepsilon^2_{t-p}$. Under the null hypothesis that there is no ARCH, the test statistic $T R^2$ is asymptotically distributed as a chi-squared distribution with p degrees of freedom, where T denotes the sample size. The alternative hypothesis is that the errors are ARCH(p).

The intuition behind this test is very clear. If the data are homoscedastic, then the variance cannot be predicted and variations in ε^2_t will be purely random. However, if ARCH effects are present, large values of ε^2_t will be predicted by large values of the past squared residuals.

Estimation

Specifying the ARCH model as all discrete time stochastic processes $\{\varepsilon_t\}$ of the form

$$\varepsilon_t = z_t \sigma_t \quad z_t \text{ i.i.d}, E(z_t) = 0, \text{var}(z_t) = 1;$$

with σ_t a time varying, positive, and measurable function of the time t-1 information set. By definition ε_t is serially uncorrelated with mean zero but the conditional variance of ε_t is σ_t^2

which may be changing through time. In most applications, ε_t will correspond to the innovation in mean for some other stochastic process, say $\{y_t\}$, where

$$\varepsilon_t = y_t - \beta x_t;$$

Let $f(z_t)$ denote the density function for z_t , and let θ be the vector of all the unknown parameters in the model (the dependence of ε_t and σ_t on the parameter vector θ are suppressed for convenience). By the prediction error decomposition, the log likelihood function for the sample $\varepsilon_T ,\varepsilon_{T-1}, \ldots, \varepsilon_1$ becomes, apart from initial conditions

$$L(\theta) = \sum_{t=1}^{T} [\log f (\varepsilon_t \sigma_t^{-1}) - \log \sigma_t] \ ;$$

The second term in the summation is a Jacobian term arising from the transformation from z_t to ε_t. The above function also defines the sample log-likelihood for $y_T ,y_{T-1}, \ldots, y_1$.

Given a parametric representation for $f(z_t)$, maximum likelihood estimates for the parameters of interest can be computed directly from the above log-likelihood function by a number of different numerical optimization techniques. For z_t normally distributed, the conditional density entering the likelihood function takes the form

$$\log f (\varepsilon_t \sigma_t^{-1}) = -0.5 \log 2\pi - 0.5\, \varepsilon_t^2 \sigma_t^{-2} \ ;$$

In terms of the equations

$$\varepsilon_t = y_t - \beta x_t \ ;$$

and $$h_t = \omega + \alpha_1 \varepsilon^2_{t-1} \ ;$$

the appropriate log likelihood function is

$$\text{Log } L = -0.5 \log 2\pi - 0.5 \log h_t - 0.5\, \varepsilon_t^2 h_t^{-1} ;$$

which can be maximized with respect to ω, α_1 and β. Since the first order equations are non-linear, the solution requires some sort of search algorithm. The entire procedure is available as a typical set of programming statements on time series packages like Regression Analysis of Time Series (RATS). The constant term - 0.5 log 2π is excluded, as it has no effect on the optimal solution. The initial guess values for ω and β are obtained from an OLS regression of y_t on x_t, and α_1 is set equal to a small positive number.

Appendix B

Tests for Heteroscedasticity and Autocorrelation in Variances

```
CALENDAR(DAILY) 1986 11 14
ALL 1997:3:31
OPEN DATA C:\WINRATS\CL\NFUT_RET.RAT
DATA(FORMAT=RATS)
SET FILTER_MISSING = %VALID(NFUT_RET)
SAMPLE(SMPL=FILTER_MISSING) NFUT_RET / C_NFUT_RET
CAL(IRREGULAR)
STATS C_NFUT_RET
SET Y = C_NFUT_RET
LINREG(NOPRINT) Y / E
# CONSTANT
SET RESSQR = E**2
LINREG RESSQR
# CONSTANT RESSQR{1}
COMPUTE CHISTAT=%NOBS*%RSQUARED
CDF CHISQR CHISTAT 1
```

```
Statistics on Series C_NFUT_RET
Observations   2604
Sample Mean    0.00010543011        Variance             0.000543
Standard Error 0.02331253441        SE of Sample Mean    0.000457
t-Statistic      0.23078            Signif Level (Mean=0) 0.81750490
Skewness        -2.14054            Signif Level (Sk=0)  0.00000000
Kurtosis        38.68568            Signif Level (Ku=0)      NA

Dependent Variable RESSQR - Estimation by Least Squares
Usable Observations   2603      Degrees of Freedom  2601
Centered R**2     0.013247      R Bar **2   0.012868
Uncentered R**2   0.036969      T x R**2      96.229
Mean of Dependent Variable        0.0005434482
Std Error of Dependent Variable  0.0034632936
Standard Error of Estimate        0.0034409395
Sum of Squared Residuals          0.0307960075
Regression F(1,2601)              34.9177
Significance Level of F           0.00000000
Durbin-Watson Statistic           2.017454
Q(36-0)                           220.915162
Significance Level of Q           0.00000000

   Variable              Coeff             Std Error       T-Stat     Signif
*******************************************************************************
1.  Constant          0.0004809060  0.0000682689    7.04429  0.00000000
2.  RESSQR{1}         0.1150949256  0.0194775045    5.90912  0.00000000

Chi-Squared(1)=     34.481651 with Significance Level 0.00000000
```

Appendix C

NYMEX Division Light, Sweet Crude Oil Futures and Options Contract Specifications

Trading Unit

Futures: 1,000 U.S. barrels (42,000 gallons).
Options: One NYMEX Division light, sweet crude oil futures contract.

Trading Hours

Futures and Options: 9:45 A.M. - 3:10 P.M., for the open outcry session. After-hours trading is conducted via the NYMEX ACCESS® electronic trading system starting at 4 P.M. on Monday through Thursday, and concluding at 8 A.M. On Sunday, the electronic session begins at 7 P.M. All times are New York time.

Trading Months

Futures: 30 consecutive months plus long-dated futures which are initially listed 36th, 48th, 60th, 72nd, and 84th months prior to delivery.
Additionally, trading can be executed at an average differential to the previous day's settlement prices for periods of two to 30 consecutive months in a single transaction. These calendar strips are executed during open outcry trading hours.
Options: Twelve consecutive months, plus three long-dated options at 18, 24, and 36 months out on a June - December cycle.

Price Quotation

Futures and Options: Dollars and cents per barrel.

Minimum Price Fluctuation

Futures and Options: $0.01 (1¢) per barrel ($10 per contract).

Maximum Daily Price Fluctuation

Futures: $15.00 per barrel ($15,000 per contract) for the first two contract months. Initial back month limits of $1.50 per barrel rise to $3.00 per barrel if the previous day's settlement price in any back month is at the $1.50 limit. In the event of a $7.50 per barrel move in either of the first two contract months, back month limits are expanded to $7.50 per barrel from the limit in place in the direction of the move.
Options: No price limits.

Last Trading Day

Futures: Trading terminates at the close of business on the third business day prior to the 25th calendar day of the month preceding the delivery month.

Options: Beginning with the August 1997 contract, trading will end three business days before the underlying futures contract. All prior contracts, as well as the December 1997, June 1998, and December 1998 contracts already listed as long-dated options, expire on the Friday before the termination of futures trading, unless there are less than three trading days left to futures termination, in which case, options expire two Fridays before the futures contract.

Exercise of Options

By a clearing member to the Exchange clearinghouse not later than 5:30 P.M., or 45 minutes after the underlying futures settlement price is posted, whichever is later, on any day up to and including the option's expiration.

Option Strike Prices

The first twenty strike prices are in increments of $0.50 (50 cents) per barrel above and below the at-the-money strike price and the next ten are in increments of $2.50 above the highest and below the lowest existing strike price for a total of 61 strike prices. The at-the-money strike price is nearest to the previous day's close of the underlying futures contract. Strike price boundaries are adjusted according to the futures price movements.

Delivery

F.O.B. seller's facility, Cushing, Oklahoma, at any pipeline or storage facility with pipeline access to Arco, Cushing storage, or Texaco Trading and Transportation Inc., by in-tank transfer, in-line transfer, book-out, or inter-facility transfer (pumpover).

Delivery Period

All deliveries are rateable over the course of the month and must be initiated on or after the first calendar day and completed by the last calendar day of the delivery month.

Alternate Delivery Procedure (ADP)

An Alternate Delivery Procedure is available to buyers and sellers who have been matched by the Exchange subsequent to the termination of trading in the spot month contract. If buyer and seller agree to consummate delivery under terms different from those prescribed in the contract specifications, they may proceed on that basis after submitting a notice of their intention to the Exchange.

Exchange of Futures for, or in Connection with, Physicals (EFP)

The commercial buyer or seller may exchange a futures position for a physical position of equal quantity by submitting a notice to the Exchange. EFPs may be used to either initiate or liquidate a futures position.

Deliverable Grades

Specific domestic crudes with 0.42% sulfur by weight or less, not less than 37 degrees API gravity nor more than 42 degrees API gravity. The following domestic crude streams are deliverable: West Texas Intermediate, Low Sweet Mix, New Mexican Sweet, North Texas Sweet, Oklahoma Sweet, South Texas Sweet.

Specific foreign crudes of not less than 34 degrees API nor more than 42 degrees API. The following foreign streams are deliverable: U.K. Brent and Norwegian Oseberg Blend, for which the seller shall receive a 30¢-per-barrel discount below the settlement price; U.K.Forties is delivered at a 35¢ discount; and Nigerian Bonny Light and Columbian Cusiana are delivered at 25¢ and 15¢ premiums respectively to the final settlement price.

Inspection

Inspection shall be conducted in accordance with pipeline practices. A buyer or seller may appoint an inspector to inspect the quality of oil delivered. However, the buyer or seller who requests the inspection will bear its costs and will notify the other party of the transaction that the inspection will occur.

Position Limits

15,000 contracts for all months combined, but not to exceed 1,000 in the last three days of trading in the spot or 7,500 in any one month.

Margin Requirements

Margins are required for open futures or short options positions. The margin requirement for an options purchaser will never exceed the premium.

Futures: CL
Options:LO

Appendix D

RATS Code and Results for GARCH (1,1) GED Estimation

```
CALENDAR(DAILY) 1986 11 14
ALL 1997:3:31
OPEN DATA C:\WINRATS\CL\NFUT_RET.RAT
DATA(FORMAT=RATS)
SET FILTER_MISSING = %VALID(NFUT_RET)
SAMPLE(SMPL=FILTER_MISSING) NFUT_RET / C_NFUT_RET
CAL(IRREGULAR)
STATS(NOPRINT) C_NFUT_RET
SET Y = C_NFUT_RET
COMPUTE GSTART=2,GEND=%NOBS
DECLARE SERIES E
DECLARE SERIES H
DECLARE REAL L
NONLIN W A1 B1 D
FRML HF = W  + A1*E{1}**2 + B1*H{1}
COMPUTE D = 2.0
FRML LAMBDA = EXP((-1./D)*LOG(2) + 0.5*%LNGAMMA(1./D) $
-0.5*%LNGAMMA(3./D))
FRML LOGL = $
(H(T)=HF(T)),(E(T)=Y(T)),$
LOG(D) - 0.5*ABS(E/(LAMBDA*SQRT(H)))**D - $
(LOG(LAMBDA) + (1+1./D)*LOG(2) + %LNGAMMA(1./D)) - $
.5*LOG(H)
LINREG(NOPRINT) Y / E
# CONSTANT
COMPUTE W=%SEESQ,A1=.05,B1=.05
SET H = %SEESQ
MAXIMIZE(METHOD=SIMPLEX,ITERS=5,NOPRINT) LOGL GSTART GEND
MAXIMIZE(METHOD=BHHH,RECURSIVE,ITERS=100) LOGL GSTART GEND
PRINT GEND GEND H
COMPUTE L=%FUNCVAL
OPEN COPY GARCH_GED_RET.TXT
DECLARE VECTOR COEFFS(4) STDERR(4) TSTATS(4) SIGNIF(4)
COMPUTE COEFFS=%BETA
EWISE STDERR(I) = SQRT(%XX(I,I))
EWISE TSTATS(I) = %BETA(I)/STDERR(I)
DO I = 1,4
CDF(NOPRINT) NORMAL TSTATS(I) 2599
DISPLAY(UNIT=COPY) 'C' COEFFS(I) 'S' STDERR(I) 'T' TSTATS(I) '%' %SIGNIF
END DO I
DISPLAY(UNIT=COPY) 'H' H(GEND)
DISPLAY(UNIT=COPY) 'L' L
CLOSE COPY
```

```
Estimation by BHHH
Iterations Taken    14
Usable Observations   2603      Degrees of Freedom  2599
Function Value                  6732.30228890

  Variable                  Coeff              Std Error          T-Stat      Signif
*******************************************************************************
1.  W                   0.0000070944   0.0000019223      3.69069  0.00022365
2.  A1                  0.0915469269   0.0118575662      7.72055  0.00000000
3.  B1                  0.8945105899   0.0131976510     67.77802  0.00000000
4.  D                   1.2610769439   0.0393010122     32.08765  0.00000000

 ENTRY         H
   2604    0.0004882817148
```

Appendix D

RATS Code and Results for EGARCH GED Estimation with Returns and Imp Vol

```
CALENDAR(DAILY) 1986 11 14
ALL 1997:3:31
OPEN DATA C:\WINRATS\CL\RETIV2.RAT
DATA(FORMAT=RATS)
SET FILTER_MISSING = %VALID(NFUT_RET).AND.%VALID(IMP_VOL)
SAMPLE(SMPL=FILTER_MISSING) NFUT_RET / C_NFUT_RET
SAMPLE(SMPL=FILTER_MISSING) IMP_VOL / C_IMP_VOL
CAL(IRREGULAR)
STATS(NOPRINT) C_NFUT_RET
STATS(NOPRINT) C_IMP_VOL
SET Y = C_NFUT_RET
SET S = C_IMP_VOL
COMPUTE GSTART=2,GEND=%NOBS
DECLARE SERIES E
DECLARE SERIES H
DECLARE REAL L
NONLIN W TH B1 G DE D
FRML K = ABS(E(T)/SQRT(H(T))) - SQRT(2.0/%PI)
FRML HF = EXP(W+ G*K{1} + B1*LOG(H{1})+TH*E{1}/SQRT(H{1})+DE*LOG((S{1}**2)/252))
COMPUTE D = 2.0
FRML LAMBDA = EXP((-1./D)*LOG(2) + 0.5*%LNGAMMA(1./D) $
-0.5*%LNGAMMA(3./D))
FRML LOGL = $
(H(T)=HF(T)),(E(T)=Y(T)),$
LOG(D) - 0.5*ABS(E/(LAMBDA*SQRT(H)))**D - $
(LOG(LAMBDA) + (1+1./D)*LOG(2) + %LNGAMMA(1./D)) - $
.5*LOG(H)
LINREG(NOPRINT) Y / E
# CONSTANT
COMPUTE W=LOG(%SEESQ),TH=.05,B1=.05,G=.05,DE=.05
SET H = %SEESQ
NLPAR(SUBITERATIONS=100)
MAXIMIZE(METHOD=SIMPLEX,ITERS=5,NOPRINT) LOGL GSTART GEND
MAXIMIZE(METHOD=BHHH,RECURSIVE,ITERS=100) LOGL GSTART GEND
PRINT GEND GEND H
COMPUTE L=%FUNCVAL
OPEN COPY EGARCH_GED_RETIV.TXT
DECLARE VECTOR COEFFS(6) STDERR(6) TSTATS(6) SIGNIF(6)
COMPUTE COEFFS=%BETA
EWISE STDERR(I) = SQRT(%XX(I,I))
EWISE TSTATS(I) = %BETA(I)/STDERR(I)
DO I = 1,6
CDF(NOPRINT) NORMAL TSTATS(I) 2594
DISPLAY(UNIT=COPY) 'C' COEFFS(I) 'S' STDERR(I) 'T' TSTATS(I) '%' %SIGNIF
END DO I
DISPLAY(UNIT=COPY) 'H' H(GEND)
DISPLAY(UNIT=COPY) 'L' L
CLOSE COPY
```

```
Estimation by BHHH
Iterations Taken    12
Usable Observations   2600      Degrees of Freedom  2594
Function Value             6772.70497873
   Variable                  Coeff              Std Error          T-Stat       Signif
*******************************************************************************
1.  W                 0.2899988786    0.1596084253      1.81694  0.06922635
2.  TH                0.0105346804    0.0247072691      0.42638  0.66983112
3.  B1                0.6417771869    0.0887005090      7.23533  0.00000000
4.  G                 0.2228905355    0.0391969636      5.68642  0.00000001
5.  DE                0.3980256876    0.0961562260      4.13936  0.00003483
6.  D                 1.3548276710    0.0446438500     30.34746  0.00000000

 ENTRY          H
                                                          2601  0.0003106242096
```

VITA

Namit Sharma was born and raised in India, where he acquired a Bachelor's degree in Engineering from the Birla Institute of Technology and Science (BITS) at Pilani and an MBA from the Indian Institute of Management at Calcutta. He worked for Philips Electronics in India from 1993 to 1996 before joining Virginia Tech's Economics Program as a Graduate Assistant in 1997. He was employed at Mackay Consulting from Jan –May 1998, and would be joining National Economic Research Associates (NERA) from July 1998.