

# AN EMPIRICAL STUDY OF CRUDE OIL MARKET

Yana Roth

Submitted for the degree of Ph.D in Economics at  
Faculty of Economics  
Universita della Svizzera Italiana  
Lugano, Switzerland

Thesis Director:  
Prof. Giovanni Barone Adesi

22 October 2008

## Acknowledgments

I should like to record here my personal sentiments for some defining moments and figures of the time spent researching this thesis.

My PhD studies coincided with my move to Switzerland from Israel, where I first graduated. For me, it was a new departure in more ways than one. As I began the process of settling in a new country and a new language, with Andreas (who was then to become my husband), I was also admitted as a doctoral candidate at the Università della Svizzera italiana, Lugano. It was largely my husband's intelligent encouragement and belief in me that sustained my motivation, and shaped and strengthened my resolve. This thesis is dedicated to him and to my Father: sadly they have both passed away. At different times but with comparable authority, they patiently nurtured and disciplined my young, eager, and often stubborn character. I owe them both lasting gratitude.

I next wish to thank my supervisor, Professor Giovanni Barone-Adesi. It was a great opportunity to be able to study in a prestigious Institute, where I attended advanced lectures and seminars held by top academics and finance specialists; and where I also had the privilege to be in the company of bright, ambitious and stimulating colleagues. Professor Barone-Adesi's brilliant guidance and advice was readily given, towards a satisfactory completion of this doctoral research.

My thanks also go to Professor Patrick Gagliardini, who devoted time and attention to discussing with me the theory and details of my work, making useful and insightful corrections and suggestions.

Among those who made me feel welcome in this Faculty, I wish to remember with particular pleasure the late Professor Pietro Balestra, who was for a time a joint supervisor of my thesis. I also had the good luck to attend his advanced econometrics classes and appreciate the excellence of his teaching.

I am grateful to Dominik Colangelo, who was generous of his time in introducing me to essential software, making sure I became proficient in using it.

All along the way, I have met new colleagues and friends, inside and outside the academic community. I wish to remember them all as I feel they have all somehow played a part in the development of this work, my PhD.

Closer to home, my warmest thoughts go to Peggy Giuliani, whose affectionate and loyal presence has been precious to me at different times, especially of late.

Thank you to Simona Cain, too, who was one of the first friends I made in Ticino and at the University. Her role began when she first taught me to see English as an accessible and enjoyable language, indeed fun to discover.

Finally and naturally, my parents – my family, my roots – deserve pride of place in my recollections, as they are originally responsible for who I am: my most loving and grateful memories are for them.

## **Abstract**

In this thesis I have tried to identify the risks and opportunities crude oil market offers. For this purpose I tested the performance of Univariate and Multivariate GARCH models. The first part of the work describes univariate GARCH models and their application to commodities markets. Physical ownership of the commodity carries an associated flow of services. The net flow of these services per unit of time is called 'convenience yield'. Since convenience yield appears as a factor which cannot be directly observed, the GARCH process was used for its modelling. The GARCH model has been found to provide a good fit for the convenience yield process. It paves the way for new term structure models of commodities prices and its empirical testing.

The second part of my work uses multivariate GARCH models for hedging in oil market. For this reason, bivariate GARCH models have been estimated for cash and future prices. Taking into account the cointegration relation between these variables improves hedging performance. I test performance of cointegrated GARCH model in and out of sample and find it superior to simple bivariate GARCH model.

# Contents

<b>1</b>	<b>Crude Oil Market</b>	<b>6</b>
1.1	Description of oil market . . . . .	7
1.1.1	The Oil Production . . . . .	8
1.1.2	Oil Consumption . . . . .	9
1.2	Global Patterns of Oil Trade . . . . .	10
1.2.1	Oil Trade: Highest Volume, Highest Value . . . . .	10
1.2.2	Distance: The Nearest Market First . . . . .	11
1.2.3	Quality, Industry Structure, and Governments . . . . .	12
1.3	Stocks . . . . .	12
1.4	Costs and Profits . . . . .	14
1.5	Highlights of Oil Market . . . . .	17
1.6	Main Objectives of this Work . . . . .	24
<b>2</b>	<b>Models of commodity markets</b>	<b>25</b>
2.1	Theory of storage . . . . .	25
2.1.1	Convenience Yield . . . . .	26
2.2	Cost of carry . . . . .	28
2.3	A long Term Extension of the analysis . . . . .	29
2.4	Dynamic Analysis of the Term Structure . . . . .	30
2.5	Term Structure Models of Commodity Prices . . . . .	31
2.5.1	One Factor Models . . . . .	31
2.5.2	Two Factor Models . . . . .	35
2.5.3	Three-Factor Models . . . . .	37
2.5.4	Other Models . . . . .	37

<b>3</b>	<b>A GARCH Model of convenience yield</b>	<b>41</b>
3.1	Overview of Univariate GARCH Models . . . . .	41
3.1.1	Families of Univariate GARCH Models . . . . .	44
3.1.2	Nonlinear GARCH . . . . .	46
3.1.3	Integrated and fractionally integrated GARCH . . . . .	51
3.1.4	Semi- and nonparametric ARCH models . . . . .	53
3.1.5	Family of Exponential GARCH models . . . . .	56
3.1.6	Stochastic volatility . . . . .	57
3.1.7	Comparing EGARCH with GARCH . . . . .	58
3.2	Empirical Evidence . . . . .	59
3.3	Convenience Yield under the GARCH Model . . . . .	61
<b>4</b>	<b>A Multivariate GARCH Approach</b>	<b>71</b>
4.1	Multivariate GARCH Models . . . . .	73
4.1.1	Generalizations of the Univariate Standard GARCH model . . . . .	75
4.1.2	Linear combinations of univariate GARCH models . . . . .	80
4.1.3	Nonlinear combinations of Univariate GARCH Models . . . . .	82
4.2	Choosing the Best Model . . . . .	89
4.2.1	Estimation of Univariate Models . . . . .	89
4.2.2	Estimation of Multivariate Models . . . . .	91
4.3	Bivariate Full Bekk GARCH Model of Cash and Future Prices . . . . .	93
4.3.1	Empirical Results . . . . .	96
4.4	Cointegrated VAR-GARCH Model . . . . .	97
4.4.1	Empirical Results . . . . .	99
4.4.2	Hedging Performance In Sample . . . . .	105
4.4.3	Performance out of sample . . . . .	108
4.5	Comparison of Results . . . . .	113
<b>5</b>	<b>Recent Developments and Conclusions</b>	<b>115</b>
5.1	Recent Developments . . . . .	115
5.2	Conclusion . . . . .	118

# Chapter 1

## Crude Oil Market

Commodities constitute the only spot market which has existed nearly throughout the history of humankind. The nature of trading has evolved from barter to more elaborate forward contracting between producers and merchants, then to organized Futures markets with clearing houses. Some key properties of commodity markets contrasting them with stock and bond markets include the following: Commodity spot prices are defined by the intersections of supply and demand curves in a given location, as opposed to the net present value of receivable cash flows. Demand for commodities is generally inelastic to prices, given the indispensable nature of the good. Supply is defined by production and inventory. But in the case of energy commodities, underground reserves also play a role since they have an impact on long-term prices. Commodities represent today a new asset class of its own right. Many institutional investors and funds are increasingly turning to it for diversification benefits and for the returns generated. Because of the growing importance of commodities markets it is important to access the issues of its pricing, risks and hedging.

Commodity price risk is an important element of the world economy at these days as it has an impact on developed and developing countries. The last two decades have experienced dramatic changes in world commodity markets. Political upheavals in some countries, economic mutation, new environmental regulation, a huge rise in the consumption of commodities in countries such as China and other structural changes have contributed to increase the volatility of supply and prices. This had made hedging activities (through forwards, Futures and options) indispensable for many sectors of the economy, the airline industry in particular being an important example.

Our study investigates the oil market and we choose to concentrate on this market because of its importance to global economy and because this market has undergone several changes

in recent years which have not been well documented yet. We would like to identify the risks and opportunities this market offers. For this purpose we would like to test GARCH models performance. The first part of our work will examine univariate GARCH models for convenience yield and its application for oil industry. The second part of our work will use multivariate GARCH models for hedging in oil market. For this reason we will estimate bivariate GARCH model for cash and future prices. From this model we derive optimal hedge ratio (OHR) and build six hedging strategies. We compare performance of portfolio hedged on the basis of traditional hedge ratio with performance of portfolio which uses OHR derived from GARCH model and arrive to conclusion that introducing GARCH dynamics improves hedging performance. In the next step we test and find cointegration between cash and future prices and subsequently build cointegrated GARCH model and from this derive optimal hedge ratio as well. Taking into consideration cointegration relation improves even more hedging performance. We test this performance in and out of sample. Last part of our work describes recent trends and developments in oil market, concludes our work and brings some ideas for future research.

## 1.1 Description of oil market

According to the most widely accepted theory, oil is composed of compressed hydrocarbons, and was formed millions of years ago in a process that began when aquatic plant and animal remains were covered by layers of sediment – particles of rock and mineral. Over millions of years of extreme pressure and high temperatures, these particles became the mix of liquid hydrocarbons that we know as oil. Different mixtures of plant and animal remains, as well as pressure, heat, and time, have caused hydrocarbons to appear today in a variety of forms: crude oil, a liquid; natural gas, a gas; and coal, a solid. Even diamonds are a form of hydrocarbons.

The word "petroleum" comes from the Latin words *petra*, or rock, and *oleum*, oil. Oil is found in reservoirs in sedimentary rock. Tiny pores in the rock allowed the petroleum to seep in. These "reservoir rocks" hold the oil like a sponge, confined by other, non-porous layers that form a "trap."

The world consists of many regions with different geological features formed as the Earth's crust shifted. Some of these regions have more and larger petroleum traps. In some reservoir rock, the oil is more concentrated in pools, making it easier to extract, while in other reservoirs it is diffused throughout the rock.

### **1.1.1 The Oil Production**

The Middle East is a region that exhibits both favorable characteristics – the petroleum traps are large and numerous, and the reservoir rock holds the oil in substantial pools. This region dominance in world oil supply is the clear result other regions, however, also have large oil deposits, even if the oil is more difficult to identify and more expensive to produce. The United States, with its rich oil history, is such a region.

### **Global Oil Supply by Region**

The Mideast remains the largest oil-producing region. Mideast dominance in oil reserves – the estimated amount of oil that can be produced from known reservoirs – is even more pronounced: the region holds about two-thirds of the one trillion barrels of global proved oil reserves, so the region’s critical role in world oil supply will continue and will grow. The United States, by contrast, holds only 4 percent of global proved reserves. Several core developments have shaped the pattern of regional oil production:

The higher oil prices of the 1970s and early 1980s afforded a strong economic incentive to explore for and produce oil, and production rose in many areas. At the same time, oil demand declined -the expected response to the high prices. Saudi Arabia became the "swing supplier," reducing its production as necessary to balance supply and demand. Its rejection of that role in mid-1985 - its output had fallen to about 25 percent of its 1980 peak - brought the full force of the supply/demand imbalance onto markets and resulted in the price collapse of 1986. Prices did not return to the pre-1986 level until the Persian Gulf conflict of 1990-91, and then only briefly. When, in 1998, Asian demand faltered with the region economies, and northern hemisphere demand faltered with the warm winter, the high production levels resulted in another price collapse. The market reaction in 1998, however, was not the same as in 1986:demand did not recover as quickly and supply did not fall as quickly. Hence, the low price period lasted longer and showed lower prices in 1998 than in 1986. In early 2000, oil prices exceeded the levels of the Persian Gulf conflict in nominal terms. Sharp as the price increases were in early 2000, however, crude oil prices remained less than half of the early 1980s peak in terms of real buying power. Saudi Arabia , the market-balancer in the early 1980s, has been the world’s largest producer during the 1990s. Not only did Saudi Arabia increase its production to fill the gap left by the loss of Iraqi and Kuwaiti supplies after Iraq invaded Kuwait in 1990, but production declined in the other two large producers, the United States and the Former Soviet Union. Mideast production would have been higher throughout the 1990s if Iraq’s production had not been constrained by the United Nations sanctions imposed after Iraq invaded Kuwait

in 1990. The so-called "Humanitarian Oil Sales" have provided Iraq only limited and closely controlled reentry into world oil markets. Mideast production also would have been higher at various times if it had not been for the market-balancing role played with varying degrees of success by the Organization of Petroleum Exporting Countries (OPEC). OPEC currently includes Algeria, Indonesia, Iran, Iraq, Kuwait, Libya, Nigeria, Qatar, Saudi Arabia, the United Arab Emirates and Venezuela. Ecuador and Gabon withdrew their membership at the end of 1992 and 1994, respectively. North America is the second largest producing area after the Middle East. The United States, the second largest producing country in the world, accounts for almost 60 percent of the North American region's total. Canada, the United States and Mexico all have long production histories, and production from mature fields has been declining. However, a new surge in technology has benefited both new field development and more complete production from existing fields. North Sea production, off the United Kingdom and Norway, began in the late 1970s. In contrast to predictions from the early 1980s of the imminent decline in the region's production, the North Sea has yet to see its peak. The region's success with new exploration and production technology, and hence its continuing volume growth, has been a central factor in world oil markets for a decade. Production in the Soviet Union peaked at about 12 million barrels a day in the early 1980s, when it was the top world oil producer. The region's demand collapse, in combination with its aggressive production targets set to maintain foreign exchange, masked its rapid production decline in the late 1980s as the Soviet Union broke up. The former Soviet Union has recently been the third-ranked producer, after Saudi Arabia and the United States. One of the most visible new production prospects has been the Caspian Sea in Central Asia, in spite of the enormous logistical and political hurdles involved in getting the oil produced to world markets. (EIA has produced several analysis of these issues).

### **1.1.2 Oil Consumption**

The industrialized countries are the largest consumers of oil but until 1998 had not been the most important growth markets for some years. The countries of the Organization for Economic Cooperation and Development (OECD), for instance, account for almost 2/3 of worldwide daily oil consumption. In contrast, however, oil demand in the OECD grew by some 11 percent over the 1991-97 period, while demand outside the OECD (excluding the Former Soviet Union) grew by 35 percent. The Former Soviet Union presents a special case. The collapse of the Russian economy that accompanied the collapse of Communism led to a decline in oil consumption of more than 50 percent over the 1991-98 period.

The developed economies use oil much more intensively than the developing economies,

and Canada and the United States stand almost alone in their consumption of oil per capita. For instance, oil consumption in the United States and Canada equals almost 3 gallons per day per capita. (The difference is these countries' transportation sectors, with their dependence on private vehicles to travel relatively long distances.) Oil consumption in the rest of the OECD equals 1.4 gallons per day per capita. Outside of the OECD, oil consumption equals 0.2 gallons per day per capita.

Regionally, the largest consuming area remains North America (dominated by the United States), followed by Asia (with Japan the largest consumer), Europe (where consumption is more evenly spread among the nations), and then the other regions. Asia was the region with the fastest demand growth until the 1998 economic crisis in East Asia. The region's economic upheaval is a central reason for the oil price collapse of 1998.

The United States and Canada use oil more for transportation than for heat and power, but the opposite pattern holds for most of the rest of the world: most regions use more oil for heat and power than for transportation. As a result, global demand for oil is highest in the Northern Hemisphere's cold months. There is a swing of 3-4 million barrels per day (some 5 percent) between the 4th quarter of the year, when demand is highest, to the 3rd quarter, when it is lowest. The precise amount varies from year-to-year, depending on weather, economic activity and other factors. While the 4th quarter is not the coldest in any region, estimated demand calculations are swollen by the traditional stock building that occurs during the period.

Demand for crude oil is derived from the demand for the finished and intermediate products that can be made from it. In the short-term, however, demand for crude oil may be mismatched with the underlying demand for petroleum products. This misalignment occurs routinely as a result of stock changes: the need to build stocks to meet seasonal demand, for instance, or the desire to reduce stocks of crude oil for economic reasons. In the longer term, blending non-petroleum additives into petroleum products (such as ethanol or other oxygenating agents into gasoline) can also reduce crude oil demand relative to demand for finished products.

## **1.2 Global Patterns of Oil Trade**

### **1.2.1 Oil Trade: Highest Volume, Highest Value**

There is more trade internationally in oil than in anything else. This is true whether one measures trade by how much of a good is moved (volume), by its value, or by the carrying capacity needed to move it. All measures are important and for different reasons. Volume provides insights about whether markets are over- or under-supplied and whether the infrastructure

is adequate to accommodate the required flow. Value allows governments and economists to assess patterns of international trade and balance of trade and balance of payments. Carrying capacity allows the shipping industry to assess how many tankers are required and on what routes. Transportation and storage play a critical additional role here. They are not just the physical link between the importers and the exporters and, therefore, between producers and refiners, refiners and marketers, and marketers and consumers; their associated costs are a primary factor in determining the pattern of world trade.

### **1.2.2 Distance: The Nearest Market First**

Generally, crude oil and petroleum products flow to the markets that provide the highest value to the supplier. Everything else being equal, oil moves to the nearest market first, because that has the lowest transportation cost and therefore provides the supplier with the highest net revenue, or in oil market terminology, the highest netback. If this market cannot absorb all the oil, the balance moves to the next closest one, and the next and so on, incurring progressively higher transportation costs, until all the oil is placed.

The recent growth in United States dependence on its Western Hemisphere neighbors is an illustration of this "nearer-is-better" syndrome. For instance, Western Hemisphere sources now supply over half the United States import volume, much of it on voyages of less than a week. Another quarter comes from elsewhere in the area called the Atlantic Basin, those countries on both sides of the Atlantic Ocean. This oil, coming especially that which comes from the North Sea and Africa, and takes just 2-3 weeks to reach the United States, boosts the so-called short-haul share of U.S. imports to over three-quarters. Most North Sea and North and West African crude oils stay in the Atlantic Basin, moving to Europe or North America on routes that rarely take over 20 days. In contrast, voyage times to Asia for just the nearest of these, the West African crude oils, would be over 30 days to Singapore, rising to nearly 40 for Japan. Not surprisingly, therefore, most of Asia's oil comes instead from the Middle East, only 20-30 days away.

Mexico and Venezuela have consciously helped the trend toward short-haul shipments. They pro-actively took the strategic decision to make a market for poor quality crudes as large and as profitable as possible, since their reserves are unusually biased toward those hard-to-place grades. Both countries therefore targeted their nearest markets, the U.S. Gulf Coast and the Caribbean, for joint venture refinery investments. They began with refineries that had traditionally run their crudes, and then with refineries that might be upgraded to do so. This policy has turned poor quality crudes into the preferred crude at these sites, significantly increasing the crude oil self-sufficiency of the Western Hemisphere.

A change in trade flow patterns can also be of critical importance to the shipping industry. For example, the Suez crisis of 1957 forced tanker owners back to using the much longer route around the Cape of Good Hope, and resulted in the development of Very Large Crude Carriers (VLCC's) to reduce that voyage's higher costs. The shift to short-haul routes in the 1990's was also critical. Using the growth in world trade volumes as a proxy for demand, tanker owners had been expecting a return to a strong tanker market. But the combination of the surge in short haul imports in the Atlantic Basin and the shift of Middle East exports from the longer United States to shorter Asian voyages led to a sharp decline in average voyage length. This decline was accelerated by the return of Iraqi crude exports, many of which move on the extremely short route from the Black Sea end of the Iraq-Turkey pipeline to the Mediterranean. The tanker owners' outlook was thus fading even before world trade volumes were undermined by the Asian crisis.

### **1.2.3 Quality, Industry Structure, and Governments**

In practice, trade flows do not always follow the simple "nearest first" pattern. Refinery configurations, product demand mix, product quality specifications -all three of which tie into quality - and politics can all change the rankings.

Different markets frequently place different values on particular grades of oil. Thus, a low sulfur diesel is worth more in the United States, where the maximum allowable sulfur is 0.05 percent by weight, than in Africa, where the maximum can be 10 to 20 times higher. Similarly, African crudes - low in sulfur - are worth relatively more in Asia, where they may allow a refiner to meet tighter sulfur limits in the region without investing in refinery upgrades. Such differences in valuing quality can be sufficient to overcome transportation cost disadvantages, as the relatively recent establishment of a significant trade in long-distance African crudes to Asia shows. The cost of moving oil into a particular market can be further distorted from the principle of nearest first by government policies such as tariffs.

## **1.3 Stocks**

According to the Energy Intelligence Group's 1997 report, "How Much Oil Inventory is Enough?," there are 7-8 billion barrels of oil tied up worldwide in industry and government stocks (inventories) at any given time. (The estimate excludes consumer stocks.) Why so much? Mostly, because stocks are needed to keep the global supply system operating. They can be thought of as a huge pipeline stretching from the wellhead to the consumer, filling the tankers, the pipelines, the railcars, and the road tankers, and linking all the markets and

all the segments of the industry together. They are thus the key to the oil industry's proven ability to deliver the right product to the right location at the right time.

Only around 10 percent of this vast stockpile is typically available to the industry to use as and when it pleases. Although minor in volume terms, these stocks – sometimes described as "discretionary" – can affect the industry in major ways, because this subset of total stocks indicates whether markets have too little, too much, or just the right amount of oil. Thus, when stocks are low in a particular market, prices there are likely to be relatively high, encouraging extra supply or reducing demand (whether through fuel switching or other means). Vice versa, when a region's stocks are high, prices are likely to be relatively low in that market. For example, if distillate is in short supply on the U.S. East Coast so that distillate stocks there are low, then the price for East Coast distillate rises relative to distillate prices in other markets, like Europe; relative to other products, like gasoline; and also relative to crude oil. Stocks, particularly projected stocks, are thus viewed as a leading indicator of prices and are one of the most closely watched aspects of the oil market.

It is hard for the industry to follow global stocks as closely as it would like, because the data have large gaps. The United States is the only country to publish comprehensive weekly stock data. The data's uniqueness and timeliness make them market movers, albeit temporary ones, nearly every week, in spite of the fact that they provide only a snapshot, based on preliminary information.

The United States, with its huge and widely dispersed oil market, has by far the largest commercial stocks, some one billion barrels. The Gulf Coast holds the greatest part of the crude oil stocks, but the East Coast, with its high consumption and limited local supply, has the greatest finished product inventories.

The Arab Oil Embargo of 1973 initiated efforts among oil import-dependent nations to build government-controlled stocks of oil-known as strategic stocks -as a buffer against severe supply interruptions. These strategic stockpiles, the largest of which is the U.S. Strategic Petroleum Reserve, or SPR , now account for a significant fraction of the world's inventories. There has been only one strictly emergency use of strategic stocks, and it was small: during the 1991 Gulf Conflict, the United States sold 4 percent of its SPR stocks.

The International Energy Agency's oil-sharing rules, designed to share the burden of an oil supply shortage, require that each participating nation hold stocks equal to 90 days of imports. Most of the participants meet the requirement with industry-owned stocks that can be commandeered in an emergency. Only the United States, Japan, and a few other nations also hold government-owned stockpiles, stored separately.

## 1.4 Costs and Profits

Holding inventory costs money. How much it costs varies, depending on the type of oil being stored, how much storage is available, whether the storage is owned or has to be rented, the price of the oil, and the cost of borrowing money. In all cases, the cost of holding inventory can rapidly become significant compared to the average margins achieved by refiners, marketers, distributors, and other oil industry participants that might need or want to hold inventory. Based on average prices in the first half of the 1990's, holding crude oil for a year would cost a company about 1.50 usd per barrel if it had its own storage and 4usd per barrel if it had to rent storage tank space. For gasoline, the corresponding costs would be 2 and 6 usd per barrel. Thus, storing gasoline in rented tank space costs roughly 1 cent per gallon per month. Companies, therefore, try to operate their supply and distribution systems in ways that keep their inventories as efficient as possible.

The trend in the more mature economies, like that of the United States, toward consolidation of the industry through mergers and acquisitions has helped in this regard. Every gas station, terminal, refinery, etc., must have some oil in inventory. As consolidation has led to facilities being closed, the minimum amount of oil needed to keep the system operating has fallen.

But stocks should not be viewed just as a cost of doing business. Stocks can also be a way to make money; they represent a profitable investment. Such stocks are truly discretionary stocks. They are built or drawn in response to prices, and particularly in response to the difference between today's prices and expectations about where prices will be in the future – the forward price curve. The widespread availability of financial instruments, like futures contracts, has greatly encouraged discretionary stock movements, partly by making the economic signals inherent in the forward price curve easy to see, but especially by reducing the risk of building stocks in a surplus market.

When prices for oil today are lower than prices for oil in the future- a sign of oversupply -the market is said to be in contango. If the contango is wide enough to cover the costs of holding stocks, namely storage and working capital, then a company can lock in a profit on the stocks if it, first, sells oil in the futures market while simultaneously putting the same volume of oil into storage in the futures contract's delivery area, and then, subsequently either delivers the stored oil against the contract or sells the stored oil and buys an offsetting futures contract. Discretionary stockbuilding occurs disproportionately in the U.S. Northeast, particularly around New York, and in Northwest Europe, especially in the Antwerp-Rotterdam-Amsterdam (ARA) area. That is because the world's two active families

of product futures contracts are based on these delivery areas: the NYMEX on New York Harbor and the International Petroleum Exchange on the ARA area.

The opposite of a contango is backwardation. A backwardated market has prices for oil today that are higher than prices for oil in the future - a sign that supplies are tight. Backwardation implies that oil in storage will be worth less later, even if holding it were cost-free. The situation, therefore, creates an incentive for companies to reduce their stocks, which adds supply to the market and helps to correct the indicated shortfall.

There are many other situations that also cause companies to adjust their discretionary stocks because the risk, although not as low as it can be with building on a contango, is judged to be much lower than the potential reward. Three examples: when prices are at unusual levels by historical standards; when prices are moving fast; and when governments' oil-related fiscal policies are expected to change. In all three cases, stocks can be viewed as a buffer that enables a company to change the timing of its purchases, with the high probability that this will lower its costs and, therefore, improve its bottom line. Consumers sometimes do the same thing. For example, if the tax on gasoline at the pump is expected to increase on January 1st, motorists rush in on December 31st and buy early.

The price of crude oil, the raw material from which petroleum products are made, is established by the supply and demand conditions in the global market overall, and more particularly, in the main refining centers: Singapore, Northwest Europe, and the U.S. Gulf Coast. The crude oil price forms a baseline for product prices. Products are manufactured and delivered to the main distribution centers, such as New York Harbor, or Chicago. Product supplies in these distribution centers would include output from area refineries, shipments from other regions (such as the Gulf Coast), and for some, product imports. Product prices in these distribution centers establish a regional baseline. Product is then re-distributed to ever more local markets, by barge, pipeline, and finally truck. The fact the oil markets are physically inter-connected, with supply for a region coming from another region, means that of necessity even local gasoline prices feel the impact of prices abroad.

Oil prices are a result of thousands of transactions taking place simultaneously around the world, at all levels of the distribution chain from crude oil producer to individual consumer. Oil markets are essentially a global auction – the highest bidder will win the supply. Like any auction, however, the bidder doesn't want to pay too much. When markets are "strong" (when demand is high and/or supply is low), the bidder must be willing to pay a higher premium to capture the supply. When markets are "weak" (demand low and/or supply high), a bidder may choose not to outbid competitors, waiting instead for later, possibly lower priced, supplies.

There are several different types of transactions that are common in oil markets. Contract arrangements in the oil market in fact cover most oil that changes hands. Oil is also sold in "spot transactions," that is, cargo-by-cargo, transaction-by-transaction arrangements. In addition, oil is traded in futures markets. Futures markets are a mechanism designed to distribute risk among participants on different sides (such as buyers versus sellers) or with different expectations of the market, but not generally to supply physical volumes of oil. Both spot markets and futures markets provide critical price information for contract makers.

Prices in spot markets - cargo-by-cargo and transaction-by-transaction - send a clear signal about the supply/demand balance. Rising prices indicate that more supply is needed, and falling prices indicate that there is too much supply for the prevailing demand level. Furthermore, while most oil flows under contract, its price varies with spot markets. Futures markets also provide information about the physical supply/demand balance as well as the market's expectations.

Seasonal swings are also an important underlying influence in the supply/demand balance, and hence in price fluctuations. Other things being equal, crude oil markets would tend to be stronger in the fourth quarter (the high demand quarter on a global basis, where demand is boosted both by cold weather and by stock building) and weaker in the late winter as global demand falls with warmer weather. As a practical matter, however, crude oil prices reflect more than just these seasonal factors; they are subject to a host of other influences. Likewise, product prices tend to be highest relative to crude as they move into their high demand season - late spring for gasoline, late autumn for heating oil. The seasonal pattern in actual product prices, again, may be less obvious, because so many other factors are at work.

Price change patterns can vary between regions, depending on the prevailing supply/demand conditions in the regional market, especially in the short-term. Refinery outages or logistics problems in Chicago will lead to rapid price increases in the Midwest without matching increases on the East Coast. Both geography and the unique quality of the gasoline required by the California Air Resources Board contribute to the volatility of gasoline prices there. Sources for additional supply are limited and distant, so any unusual increase in demand or reduction in supply gets a large price response in the market.

That price response, and the differences in regional price movements, are critical to the way the oil market redistributes products to re-balance after an upheaval. The price increase in one area calls forward additional supplies. These new supplies might come from other markets in the United States, or from incremental imports. They may also be augmented by increased output from refineries. The volume and source of the relief supplies are interwoven. The farther away the necessary relief supplies are, the higher and longer the likely price spike

## 1.5 Highlights of Oil Market

In the early 1970s a fourfold rise in the price of oil almost brought a world to a standstill. Thirty five years on, oil prices have quadrupled again briefly soaring to a peak of just over 135 USD a barrel. But so far, this has been a slow motion shock. If the Arab oil-weapon felt like a hammer -blow, this time stagnant oil output and growing emerging market demand have squeezed the oil market like a vice. For almost five years a growling world shrugged it off. Only now it is recoiling in pain. It is clear that high oil prices are hurting many economies-especially in the rich world. Goldman Sachs reckons consumers are handling over 1.8 trillion USD a year to oil producers. In America falling house prices have left consumers resentful-and short of money. Energy tax rebate, voted this year to help people cope with the credit crunch, has in effect been taken right away again. Beset by scarce credit, falling asset prices and costly food, developed-country households are hardly well equipped to foot the oil bill. Stuck for answers, politicians are looking for scapegoats. Top of the list are speculators profiting from other people's hardship. Some 260 billion USD is invested in commodity funds, 20 times the level of 2003. The number of transactions involving oil futures on the NYMEX, the biggest oil market for oil, has almost tripled since 2004. The government of India is so sure that speculation makes commodities dearer that it has banned the trading of future contracts for some of them( although not oil). Germany's Social Democratic Party proposes an international ban on borrowing to buy oil futures, on the same grounds. But this blame is wrong. Such speculators do not own real oil. Every barrel they buy in the futures market they sell back again before the contract ends. That may rise the price of "paper barrels", but not of the oil itself. It is true that high futures prices could lead someone to hoard oil today in the hope of a higher price tomorrow. But inventories are not especially full just now and there are few signs of hoarding. Jeffrey Harris, the chief economist of the Commodity Futures Trading Commission doesn't see any evidence that the growth of speculation in oil has caused the price to rise. Rising prices, after all, might have been stimulating the growing investment, rather than the other way around. There is no clear correlation between increased speculation and higher prices in commodities markets in general. Despite a continuing flow of investment in nickel, for example, its price has fallen by half over the past year. On the other hand, the prices of several commodities that are not traded on any exchange, and therefore much harder for speculators to invest in, have risen even faster than that of oil. Deutsche Bank calculates that cadmium, rare metal, has appreciated twice as much as oil since 2001, for example, and the price of rice has risen fractionally more. Others assume the reverse: that the price is bound to keep rising indefinitely since supplies of oil are running short. As different as these

theories are, they share a conviction that something has gone badly wrong with the market of oil. Yet the evidence suggests that, to the contrary, the rising price is beginning to curb demand and increase supply, just as textbooks say it should.

1. Global oil product demand is expected to rise by a robust 2.5 percent to 88.2 mb/d in 2008, largely due to a weather-related rebound in the OECD and strong demand in non-OECD countries. This represents an increase of 2.2 mb/d, from the slightly revised (-0.1 mb/d) 2007 level of 86.0 mb/d.
2. Non-OPEC supply in 2008 is forecasted to reach 51.0 mb/d (+1.0 mb/d), plus 5.5 mb/d of OPEC gas liquids (+0.7 mb/d). Key growth drivers include the FSU, Latin America and global biofuels. OECD Europe and North America continue to see production decline, despite strong growth from the US Gulf of Mexico and Canadian oil sands.
3. OPEC capacity rises by 1 mb/d in 2008 to average 35.4 mb/d, with the implication that spare capacity will post a modest rise. In reality, the spare capacity comparison will depend to a large extent on OPEC production levels both in the second half of 2007 and next year.
4. Global refinery crude throughput increased by 0.2 mb/d in May to 72.7 mb/d and is 0.4 mb/d higher year-on-year. Crude throughput is forecast to increase rapidly to an August peak of 75.2 mb/d, on the back of higher runs in the OECD and the Middle East.
5. Preliminary end-June OECD stock data show an increase of 7.8 MB, as increases in the US and Japan offset a sharp downturn in Europe. Together with an average 21.2 MB rise in stocks in April and May, this implies a second-quarter stock build around 550 kb/d and forward demand cover at end-June unchanged at 53.6 days.

China's growing thirst for oil is often put forward as one of the main factors behind today's higher oil prices. Demand for diesel there, for example, rose by over 9 percent in the year to April 2008. In the short run, neither demand nor supply of oil is very elastic.

## **Volatility**

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 mean 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. Among commodities traded in financial markets oil is one of the most important, and has extensively been studied in the literature. Its relevance induces innovations in financial markets, generating new contingent -claims that should be used in the estimation of the stochastic behavior of oil prices. One important source of information for the study of oil prices is the futures market. Oil futures markets have included in recent years new futures contracts with longer maturities up to 7 years.

Oil prices are very volatile, have a high degree of mean reversion (Besseminder et.al 1995), and exhibit a complex dynamics. Thus, it is important to analyze the number of risk factors required to model this stochastic behavior (Cortazar and Schwartz, 1994). Several models of the stochastic process followed by commodity prices have been proposed in the literature. They differ in how they specify spot price innovations and how they model the cost-of-carry. The cost-of-carry represents the storage cost plus the interest paid to finance the asset minus the net benefit that accrues to the asset holder, if any. In the commodities literature, the benefit received by the commodity owner, but not by the future owner is called the **convenience yield** which is commonly represented as a dividend yield.

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. Duffie and Gray(1995) compare the forecasting accuracy of ARCH-type models, Markov switching models, and implied volatilities for crude oil and natural gas markets.

Fluctuations in energy prices are caused by supply and demand imbalances arising from events like wars, changes in political regimes, economic crises, unexpected weather patterns etc. Forward and future prices imbed the expectation 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 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 expiration reflecting the expectation that supply and demand balance in the long run, to reach a relatively stable equilibrium price. This work would like to examine the temporal behavior of convenience yield of weekly returns for crude oil futures and its relation to the basis. We will test the hypotheses that convenience yield volatility is time dependent and can be predicted by GARCH(1,1) model. By prediction of convenience yield it is possible to forecast oil prices and this is of course a valuable feature for risk management. Oil futures exhibit backwardation most of the time. Litzenberg and Rabinovitz(1995) explain this phenomenon in terms of option pricing theory. Specifically, oil wells are viewed as call options with exercise price corresponding to the extraction costs. In this context, persistent backwardation is viewed as an inducement for current extraction. Costs of storing oil above ground are high and it severely hinders the ability of oil refiners to cope with situations of sudden high demand, giving rise to inelasticity of oil supply especially in the short run. Due to inelastic supply spot prices must rise significantly to equilibrate the market whenever there is a big increase in demand. Thus, the theory of storage predicts that backwardation is more likely to occur when oil stocks are low. The purpose of this work is to attempt to model these complex features of volatility. Modeling and forecasting the conditional variance, or the volatility, of financial time series is one of the major topics in financial econometrics nowadays.

Forecasted conditional variances are used in portfolio selection, derivative pricing and hedging, risk management and market timing. Among solutions to tackle this problem, the ARCH model proposed by Engle (1982) and the GARCH specification introduced by Bollerslev(1986)are certainly among the most widely used. A GARCH model states that the one step ahead conditional variance of return is a deterministic linear function of lagged squared values of the series and past conditional variances. One drawback of the GARCH model is the symmetry in the response of volatility on past shocks, failing to accommodate sign asymmetries. Researchers, beginning with Black (1976), pointed out the asymmetric response of the conditional variance of the series to unanticipated news, represented by shocks. In particular, asset prices tend to exhibit price clustering in which large (small) price changes tend to be followed by large (small) price changes. This clustering of volatility strongly suggests that conditional volatility of asset returns is time varying. Evidence suggest, that to obtain more robust estimates of conditional volatility requires a more general class of GARCH models that allow for regime shifts as part of the data generating process. Regime switching GARCH

models were introduced recently by Hamilton and Susmel (1994), Gai (1994) and Gray (1996). There are several common features in these models. First, the conditional volatility process is allowed to switch stochastically between a finite numbers of regimes. Second, the timing of regime switch is usually assumed to be governed by a first order Markov process. Fong and See (2003) apply a regime switching model to examine the temporal behavior of the conditional volatility of crude oil futures. More recently, the Tree Structured GARCH developed by Audrino and Bühlmann (2001) also considers multiple regimes but no statistical framework is adopted to test if the regimes are statistically significant. Audrino and De Giorgi (2007) propose a new term structure model in which short term interest rate and the market price of risk are subjected to discrete regime shifts. Authors believe that probability to be at any given time in a specific regime has to be time varying and directly related to some relevant state variables.

This model is different from the regime switching models introduced in the literature above. The underlying regime variable doesn't follow a Markov process. Instead, regimes are directly characterized by multiple thresholds on the regime variable. I would like to probe this model for modeling volatility of daily returns of WTI crude oil futures. Commodities are perceived as a distinct asset class because of their countercyclical nature: over the last 45 years, both commodity spot prices and commodity Futures have outpaced inflation. Moreover, in the current period of historically low interest rates world wide and sluggish stock markets, commodity-related investments perform remarkably well. Hedge funds, which used to be profitable through such strategies as buying and selling volatility, while remaining neutral to the stock market, are looking for new styles as well as new assets, including real assets and commodities.

Oil market is not a free market and the price is not set only by forces of supply and demand. The most promising territory for extracting oil lies in unstable places such as Middle East and Russia. What is more, the oil is controlled by state-owned firms, which often seem blind to the signals sent by the markets. The recent rise of the oil prices is different from the previous hikes in that in contrary to the previous supply driven shocks this time we have a demand driven shock and the supply capacity is tight and there is uncertainty in the market. Three other factors will add to uncertainty. The first is the flood of new speculative investment in commodities, especially oil. Speculators are thought to have put more than 100 usd billion into commodities markets in the past few years, helping to propel the price of oil even higher. But this new hot money could quit the oil market in an instant, causing prices to plunge (and throwing the energy industry's investment plans into disarray). The second unknown is how expensive oil will affect the world economy. Past surges in the oil price have led to rises

in inflation and interest rates that have triggered recessions. This time might be different, partly because growing demand, rather than a reduction in supply, has underpinned the price rise. That means it has been steady and gradual, giving consumers more time to adjust. In addition, inflation remains low and oil exporters are supporting consumption in the United States, the biggest importer. Dearer oil will eventually curb demand, but at what price, and at what cost to the world economy, nobody knows.

**Making the oil** What of the notion that oil scarcity will lead to economic disaster? Majority of experts in the field insist that the key is to avoid price controls and monetary policy blunders of the sort that turned the 1970s oil shocks into economic disasters. Prof. Kenneth Rogoff, a Harvard professor and the former chief economist of the IMF, thinks concerns about peak oil are greatly overblown. The extraction of more oil is the matter of its price (real options theory). There is still much of the black gold underground but its extraction is costly and only at the certain price oil producers will enter into this projects. This occurred during the Gulf war when a number of oil fields in Texas and Southern California (where the deposits are such that the cost of extraction is relatively high) began operations when the price of oil rose. The problem is that these projects and implementing of new technology need time and some stability in the future price of oil. In the oil industry deferral options are critical. The option to choose when to start a project is an initiation or deferment option. Initiation options are particularly valuable in natural resource exploration where a firm can delay mining a deposit until market conditions are favorable. If natural resource companies were committed to producing all resources discovered, they would never explore in areas where the estimated extraction cost exceeded the expected future price at which the resource could be sold. Expansion option is an option to build production capacity in excess of expected level of output (so it can produce at higher rate if needed). Management has the right (not the obligation to expand). If project conditions turn out to be favorable, management will exercise this option.

In oil industry production could be phased in over time. Conventional NPV will significantly undervalue these assets. Two operating options are important: The option to defer and the option of deferring expansion program.

#### **Oil and Current Account imbalances.**

From an American point of view, the rise in oil prices has explained half of the widening of the current account deficit since 2003, a bigger share than that accounted for by China. A chapter in the International Monetary Fund's latest World Economic Outlook, published recently, explores the implications of these gushing petrodollars for the world's current account imbalances. Thanks to higher prices and increased production, the revenue of oil exporters-

members of OPEC, plus Russia and Norway and a few others -reached almost 800 USD billion in 2005. Adjusted for inflation, this beats the previous peak in 1980; and in the past three years real oil revenues have risen by almost twice as much as in 1973-76 or 1978-81. In relation to world GDP, however, revenues have increased by only a little more than during those earlier price shocks. The oil exporters' current-account surpluses reached almost 400 USD billion last year, more than four times the total for 2002, and are forecast to rise to 480 USD billion in 2006. The combined surpluses of China and other emerging Asian economies were only 240 USD billion last year. Oil exporters' surpluses are also much bigger relative to the size of their economies, averaging 18 percent of GDP last year, twice as much as at the previous peak in 1980, and much more than China's surplus of 7 percent of GDP.

Petrodollars can be recycled to oil -importing economies in one of two ways. Either oil exporters can import more goods and services, or they can invest their windfalls in global capital markets, thereby financing the current account deficits of America and other oil importers. Whether the exporters spend or save their revenues determines the path of global imbalances, including America's vast deficit.

After previous increases in oil prices, current-account imbalances have tended to adjust fairly swiftly, through three main channels. First, higher oil prices reduced real incomes and profits in importing countries, squeezing domestic spending and hence imports of both oil and other goods. Second, as oil prices pushed up inflation, central banks raised interest, which further slowed domestic demand and so trimmed the external deficit. And third, importers' real exchange fell, helping to bring their current accounts back towards balance. Conversely, exporters' surpluses were reduced by stronger domestic demand (and hence imports) and rising real exchange rates. The IMF suggests that this time may be different, with the oil producers' surpluses persisting for longer than in the past. Oil exporters have spent a smaller share of their latest windfall on imports of goods and services than during previous oil shocks. Since 2002 they have spent barely half of their extra revenues, compared with three-quarters in the 1970s. Past windfalls typically encouraged wasteful budgetary blow-outs, leaving government finance in trouble once oil prices fell back. This time governments are being more cautious, even though the futures markets expect oil prices to stay high.

America gains little, in terms of its current-account balance, even from the imports that oil exporters do buy. It now accounts for only 8 percent of OPEC countries total imports; the European Union has 32 percent. So even if the exporters spent all their extra revenue, America's current account deficit would increase as oil prices rise- This partly explains why in recent years the EU's trade balance barely changed even as American's deficit has grown sharply.

## 1.6 Main Objectives of this Work

In this work I would like to describe commodity markets and its importance in today's economy. In particular, I would like to concentrate on crude oil market, to identify the risk and opportunities it contains. For this purpose I will apply GARCH models for modeling and risk management in crude oil market. The first part of my work will model convenience yield of crude oil as univariate GARCH process. At the beginning we will provide short survey of univariate GARCH models. After fitting different models to the data we will come up with the model which has the best performance in and out of sample.

The second part of my work will address the issue of hedging within the oil industry. To derive optimal hedging ratio I will use Multivariate GARCH approach. After survey of Multivariate GARCH models I will estimate bivariate GARCH model for spot and futures prices. I will also test and find that spot and futures prices are cointegrated and I will incorporate this result into my analysis. I will build different hedging strategies and compare their performance on the basis of bivariate GARCH and bivariate GARCH with cointegration. Estimating the last one will allow us to build Error-Correction GARCH model. We will subsequently find that this model outperforms the others. On its basis we will derive optimal hedge ratio and build different hedging strategies which performance we will test in and out of sample. This last model will outperform the previous models and will allow us to hedge more efficiently in and out of sample. We will report RMSE and Sharp ratio results to support this claim. Conclusions, survey of recent developments and directions for future work will follow.

## Chapter 2

# Models of commodity markets

### 2.1 Theory of storage

The theory of storage aims to explain the difference between spot and Future prices by analyzing the reasons why agents hold inventories. The models proposed over the last 50 years for price determination of storable commodities have emphasized the importance of the knowledge of quantities produced and stored for the derivation of testable predictions about price trajectories. The theory of storage illuminates the benefit of holding the physical commodity: inventories have a productive value since they allow us to meet unexpected demand, avoid the cost of frequent revisions in the production schedule and eliminate manufacturing disruption. In order to represent the advantages attached to the ownership of the physical good, Kaldor (1939) and Working (1948) define a notion of a convenience yield as a benefit that "accrues to the owner of the physical commodity but not to the holder of a forward contract." Brennan (1958) and Telser (1958) view the convenience yield as an "embedded timing option attached to the commodity" since inventory (e.g. a gas storage facility) allows us to put the commodity on the market when prices are high and hold it when prices are low.

Another approach directly analyzes the role of inventory in explaining commodity spot price volatility. A statistical study performed by Fama E. F. (1987) on commodity Futures including metals, wood and animals shows that the variance of prices decreases with inventory level.

In markets for storable commodities such as oil, inventories play a crucial role in price formation. As in manufacturing industries, inventories are used to reduce costs of changing production in response to fluctuations (predictable or otherwise) in demand, and to reduce marketing costs by helping to ensure timely deliveries and avoid stockouts. Producers must

determine their production levels jointly with their expected inventory drawdowns or buildups. These decisions are made in light of two prices - a spot price for sale of the commodity itself, and a price for storage. As I will explain, although the price of storage is not directly observed, it can be determined from the spread between futures and spot prices. This price of storage is equal to the marginal value of storage, i.e., the flow of benefits to inventory holders from a marginal unit of inventory, and is termed the marginal **convenience yield**. Thus there are two interrelated markets for a commodity: the cash market for immediate, or "spot", purchase and sale, and the storage market for inventories held by both producers and consumers of the commodity.

### 2.1.1 Convenience Yield

Similar to the term structure of interest rates, commodity price curves exist which convey information about future expectations. In addition they reflect the prevailing yield curve (cost of carry) and storage costs. Energy prices are said to be in *contango* when the forward prices are greater than spot prices; prices are said to be in *backwardation* when spot prices exceed forward prices. The term structure has little forecasting power, however. Forward prices have not been proven to be accurate forecasts of future spot prices. The notion of "**convenience yield**" Kaldor (1939), Working (1948) is often used as an explanation for backwardation in future prices of storable commodities. Convenience yield, as defined by Brennan and Schwartz (1985), is "the flow of services that accrues to an owner of the physical commodity but not to the owner of a contract for future delivery of the commodity." The nature of these services and benefits depends on the type of commodity, the identity of the holder and the location of storage (Brennan (1958)). Brennan and Schwartz (1985) as well as Gibson and Schwartz (1990) argue that backwardation is equal to the present value of the marginal convenience yield of the commodity inventory. These papers determine the convenience yield exogenously. Litzenberger and Rabinowitz (1995) determine backwardation endogenously and it is associated with the option feature of reserves. Since oil reserves are spanned by a forward contract and an at-the-money-forward put option, convenience yield can be interpreted as the expected present value of the price protection services associated with storing oil in the ground rather than holding a forward contract. We can now turn to the calculation of convenience yield from futures and spot prices. Let  $\psi_{t,T}$  denote the (capitalized) flow of marginal convenience yield over the period  $t$  to  $t + T$ . Then, to avoid arbitrage opportunities,  $\psi_{t,T}$  must satisfy:

$$\psi_{t,T} = (1 + r_T)P_t - F_{t,T} + k_t \tag{2.1}$$

where  $P_t$  is the spot price at time  $t$ ,  $F_{t,T}$  is the futures price at time  $t$  for delivery at time  $t+T$ ,  $r_T$  is the risk-free  $T$ -period interest rate, and  $k_t$  is the per-unit cost of physical storage. To see why (2.1) must hold, note that the (stochastic) return from holding a unit of the commodity from time  $t$  to time  $t+T$  is  $\psi_{t,T} + (P_{t+T} - P_t) - k_T$ . Suppose that one also shorts a futures contract at time  $t$ . The return on this futures contract is  $F_{t,T} - F_{t+T} = F_{t,T} - P_{t+T}$ , so one would receive a total return by the end of the period that is equal to  $\psi_{t,T} + F_{t,T} - P_t - k_T$ . No outlay is required for the futures contract, and this total return is non-stochastic, so it must equal the risk-free rate times the cash outlay for the commodity, i.e.,  $(r_t \times P_t)$ , from which (2.1) follows. Note that the convenience yield obtained from holding a commodity is very much like the dividend obtained from holding a company's stock. (The ratio of the net convenience yield to the spot price,  $(\psi_{t,T} - k_T)/P_t$  is referred to as the percentage net basis, and is analogous to the dividend yield on a stock.). In fact, if storage is always positive, one can view the spot price of a commodity as the present value of the expected future flow of convenience yield, just as the price of a stock can be viewed as the present value of the expected future flow of dividends.

The future price of an energy product is determined by many factors. The no-arbitrage, cost and carry model predicts that futures prices will differ from spot prices by the storage and financing costs relevant to inventory. The spot price is the only source of uncertainty in the basic model. From (2.1) we can see that the futures price could be greater or less than the spot price, depending on the magnitude of the net (of storage costs) marginal convenience yield. If marginal convenience yield is large, the spot price will exceed the futures price,  $F_{t+T}$ , and in this case we say that the futures market exhibits **strong backwardation**. If net marginal convenience yield is precisely zero, we see from (2.1) that the spot price will equal the discounted future price:  $P_t = F_{t,T}/(1 + r_T)$ . If net marginal convenience yield is positive but not large, the spot price will be less than the futures price, but greater than the discounted future price:  $F_{t,T} > P_t > F_{t,T}/(1 + r_T)$ . In this case we say that the futures market exhibits **weak backwardation**. (We say that the futures market is in **contango** when  $P_t < F_{t,T}$ . Thus contango includes weak backwardation and zero backwardation. For an extractive resource commodity like crude oil, we would expect the futures market to exhibit weak or strong backwardation most of the time, and this is indeed the case. The reason is that owning in-ground reserves is equivalent to owning a call option with an exercise price equal to the extraction cost, and with a payoff equal to the spot price of the commodity. If there were no backwardation, producers would have no incentive to exercise this option, and there would be no production (just as the owner of a call option on a non dividend paying stock would wait to exercise it until just before the option expiration date). If spot price

volatility is high, the option to extract and sell the commodity becomes even more valuable, so that production is likely to require strong backwardation in the futures market. During periods of high volatility, convenience yield is high, so that the spot price exceeds the futures price. This "real option" characteristic of an extractive resource creates a second reason for the convenience yield to depend positively on the level of volatility. As explained earlier, convenience yield will increase when volatility increases because greater volatility increases the demand for storage; market participants will need greater inventories to buffer fluctuations in production and consumption. But in addition, greater volatility raises the option value of keeping the resource in the ground, thereby raising the spot price relative to the futures price.

## 2.2 Cost of carry

Cost of carry is the sum of the riskless interest rate and the marginal cost of storage. Because carry is always positive, the cost and carry model predicts that energy prices will always be in contango. Empirical evidence suggests, however, that the term structure of energy is not fully explained by carry. The term structure of energy prices is not always in contango. Oil and natural gas are often become backward dated due to external factors or supply concerns. Further, the market rarely shows full carrying charges. In other words, futures prices as predicted by a cost of carry model generally exceed those observed in the market, even when prices are in contango.

Several theories have been advanced to explain why market prices are less than full carry. **Keynes** introduced the theory of **normal backwardation**, Keynes (1930). He proposed that for most commodities there are natural hedgers who desire to shed risk. In the oil markets, producers would act as net sellers of forward contracts in order to insulate themselves from the potential for future adverse price movements. They require the presence of speculators in the market, willing to assume the long side of their hedges, and must entice them with an expectation of profit. The speculators will only be willing to buy forward if the forward price is below the expected spot price to give them an expected profit of the difference between prices at the two tenors. Their expected profit is equivalent to the producer's expected loss (pay a positive insurance premium) to guard against unforeseen price movements. This theory is fundamentally flawed, largely because it ignores the hedging activities of end-users. However, it does provide a framework for describing price curve behavior which reflects the volatility of prices, the degree of risk aversion of producers and cost of trading and inventories.

As we know, commodity markets are not always in contango or backwardation. They shift between the two types of behavior as a result of many forces supplemental to carrying costs and

hedging activity. The concept of convenience yield is another element that sheds important light on the behavior of commodity price curves. The relationship between spot and future prices allows identification of at least three variables influencing the futures price: the spot price, the convenience yield and the interest rate, which is implicitly included in the financing costs. Convenience yield and the spot are positively correlated: both of them are the inverse function of the stock level. Very important is to define the basis of commodity futures. This is the spread between commodity spot price and future price. If the commodity is storable, changes in the spot price of the underlying commodity are somewhat smoothed by fluctuations in inventories. Thus, the level of inventories also affects the size and sign of the commodity futures basis. Brennan (1958), Telser (1958), Fama and French (1987), Brennan (1991), and others have used variations of storage cost, inventory level, and seasonal supply and demand variables to explain variation of the basis. The examination of the arbitrage relationships between physical and paper markets shows that the basis has an asymmetrical behavior: in contango, its level is limited to the storage costs. This is not the case for backwardation. Furthermore, the basis is stable in contango, and volatile in backwardation, since in this situation stocks cannot absorb price fluctuations. This asymmetry has implications on the dynamic of convenience yield.

### **2.3 A long Term Extension of the analysis**

The Keynesian analysis can be extended rather simply. When the whole price curve is taken into account, the eventual simultaneous presence of contango and backwardation along the curve can be explained by a surplus in the supply or demand of futures contracts for specific maturities. In order to palliate these imbalances, a risk premium is offered to the speculators, provided they accept to take a position in the futures market that compensates for the net position of the other operators. In the way of the preferred habitat theory developed for interest rates Modigliani and Sutch (1966), the term structure of commodity prices is then regarded as a succession of segments having different maturities. Market participants select their segment in accordance with their economic needs. Therefore, all the categories of operators do not necessarily intervene on all the maturities. The level of the premium they are willing to pay and the sense of market's imbalance can be different for each segment. Thus, the risk premium is a function of the maturity. Lastly, in order to take into account the eventual distortions of the price curves, this premium must be able to vary with the period, as the operators expectations and risk aversion change. Gabillon (1995) is the first to claim that in the long run there are other factors than storage costs and convenience yield to explain the behavior of the

prices' term structure. He proposes to separate the term structure of crude oil prices into two distinct segments. The first segment, corresponding to the shorter maturities (from the 1st to 18th month), is mostly used for hedging purposes. As a result, production, consumption, stock level, and the fear of inventory disruptions are the most important explanatory factors of the price relationship. However, for the longer maturities the explanatory factors change: interest rates, anticipated inflation, and the prices for competing energies determine the future prices. In that case, the information provided by the prices is used for investment purposes. In this analysis, agents have preferred habitats: they are specialized in the holding of certain subsets of maturities and they are reluctant to alter their portfolio to take advantage of arbitrage opportunities. The latter is, therefore left unexploited. Later, Lautier (2003), showed that there is a segmentation of a crude oil price curve, explained by liquidity factors, and that segmentation evolves as the futures market matures. The segmentation in 2003 is situated around the 28th month, and not the 18th.

## 2.4 Dynamic Analysis of the Term Structure

The most important feature of the commodity price curves dynamic is probably the difference between the price behavior of first nearby contracts and deferred contracts. The movements in the prices of the prompt contracts are large and erratic, while the prices of long-term contracts are relatively still. This results in the decreasing pattern of volatilities along the price curve. Indeed, the variance of futures prices and the correlation between the nearest and subsequent futures prices decline with maturity. This phenomenon is usually called "the Samuelson effect". Intuitively, it happens because a shock affecting the nearby contract price has an impact on succeeding prices that decreases as maturity increases Samuelson (1965). Indeed, as futures contracts reach their expiration date, they react much more strongly to information shocks due to ultimate convergence of future prices to spot prices upon maturity. These price disturbances influencing mostly the short-term part of the curve are due to the physical market, and to demand and supply shocks. Anderson (1985), Milonas (1986), and Fama E. F. (1987) have provided empirical support for this hypothesis for a large number of commodities and financial assets. Deaton (1992), Deaton (1996) and Chambers and Bailey (1996) showed that the Samuelson effect is a function of storage costs. More precisely, a high cost of storage leads to relatively little transmission of shocks via inventory across periods. As a result, future price's volatility declines rapidly with the maturity. Violations of the Samuelson effect might occur at a shorter horizon when inventory is high. In particular, price volatilities can initially increase with the maturity of the contract, because with enough

inventories, stock-outs may not be possible for the nearest delivery months.

Computing of volatility of futures in our data set shows that next to expiration futures are more volatile than the following to expiration contract.

Principal component analysis presents another strategy for dealing with the dynamic of futures prices Cortazar and Schwartz (1994), Tolmasky (2002), and Lautier (2005). This statistical method reduces the dimensionality of a data set by collapsing the information it contains. In data sets including many variables, groups of variables often move together because they are influenced by the same driving forces. In many systems, there are only a few such driving forces. Applying a principal component analysis to crude oil futures curves gives rise to three conclusions. Firstly, it leads to the identification of the type of price curve movements, which are quite simple to describe. Three different kind of movements can indeed be distinguished: a parallel shift in the curve(level factor), a relative shift of the curve(steepestness factor), and the curvature factor. Secondly, the principal component analysis makes it possible to calculate the contribution of each component to volatility. In case of a crude oil, the first two factors account for 99 % of the total variance of the futures prices. Therefore, one can consider that most of the risk associated with futures price moves is accounted for two factors, instead of all future prices.

## 2.5 Term Structure Models of Commodity Prices

### 2.5.1 One Factor Models

Brennan and Schwartz (1985) and Gibson and Schwartz (1990) use a geometric Brownian motion in their one-factor models to model spot price of crude oil. Among these models, Brennan and Schwartz's model[1985] is the most well known. It has been extensively used in subsequent research on commodity prices(see, e.g., Schwartz[1998], Schwartz and Smith (2000), Nowman and Wang (2001), Cortazar, Schwartz, and Casassus (2001), and Veld-Merkoulova and de Roon (2003)):

$$dS(t) = \mu S(t)dt + \sigma_s(t)dz \quad (2.2)$$

where  $S$  is the spot price,  $\mu$  is the drift of the spot price,  $\sigma_s(t)$  is the spot price volatility and  $dz$  is an increment to a standard Brownian motion associated with  $S$ .

However, the Geometric Brownian motion is probably not the best way to represent the price dynamic. Indeed, the storage theory and Samuelson effect show that the mean-reverting process is probably more relevant.

Schwartz(1997), Cortazar and Schwartz(1997), and Routledge, Seppi, and Spatt(2000) retain a mean-reverting process for their one-factor models. Among these models, that of Schwartz(1997), inspired by Ross(1995), is probably the most well known. In that case, the dynamic of the spot price is the following:

$$dS = S\kappa(\mu - \ln S)dt + \sigma_s S dz_s \quad (2.3)$$

where  $S$  is the spot price,  $\kappa$  is the speed of adjustment of the spot price,  $\mu$  is the long-run mean log price,  $\sigma_s$  is the spot price volatility and  $dz_s$  is an increment of a standard Brownian motion associated with the spot price. In this situation, the spot price fluctuates around its long-run mean. The presence of a speed of adjustment insures that the state variable will always return to its long-run mean  $\mu$ . Therefore, two characteristics describe the spot price behavior. It has a propensity to return to its long-term mean, but simultaneously, random shocks can move it away from  $\mu$ . The use of a mean reversion process for the spot price makes it possible to take into account the behavior of the operators in the physical market. When the spot price is lower than its long-run mean, the industrials, expecting a rise in the spot price, reconstitute their stocks, whereas the producers reduce their production rate. The increasing demand and the simultaneous reduction of supply have a rising influence on the spot price. Conversely, when the spot price is higher than its long-run mean, industrials try to reduce their surplus stocks and producers increase their production rate, thus pushing the spot price to lower levels. This formulation of the spot price behavior is preferable to the geometric Brownian motion, but it is not without flaws. For example, the storage theory shows that in commodity markets, the basis does not behave similarly in backwardation and in contango. Initially, the mean-reverting process does not allow taking into account that characteristic. The mean-reverting process was also used by Cortazar and Schwartz in 1997 in a more sophisticated model. The authors introduce a variable convenience yield that depends on the deviation of the spot price to a long-term average price.

Among the other one-factor models, the models studied by Brennan (1991) are quite interesting because they all rely on a specific hypothesis concerning the convenience yield, which is an endogenous variable. The first model was developed with Schwartz in 1985. In that case, the convenience yield is a simple linear function of the spot price. The second model expresses the convenience yield  $C$  as a non-linear function of a price  $S$ :

$$C(S) = a + bS + cS^2 \quad (2.4)$$

This formula is chosen because it is more flexible than the one used in 1985. However, it was

afterwards forgotten, as the third model, which is also non-linear, was preferred. The third model underlines that when the convenience yield is low, it cannot be lower than the opposite of the storage cost. The latter is supposed to be constant for a large spread of stock levels as long as the storage capacities are not saturated.

**Hotelling Principle** This is an important concept in commodity price theory. The Hotelling Principle, Hotelling (1931), developed under the conditions of perfect competition and certainty, states that the net price of an exhaustible resource, such as crude oil, should rise over time at the rate of interest. Denoting the spot price of oil at time  $t$  by  $S_t$  and the extraction cost per unit at time  $t$  by  $x_t$  this may be expressed as:

$$S_t - x_t = (S_0 - x_0)e^{rt} \quad (2.5)$$

This principle is based on the condition that in an interior equilibrium (in which oil is produced in every period), each producer must be indifferent between present and future production of the oil. Under certainty the future spot price in period  $t$  is known and is equal to the current  $t$ -period future price. If the extraction per unit is proportional to the price of oil, the principle implies that the discounted futures price will equal the spot price. If the extraction cost per unit has a fixed component, then discounted futures prices will equal spot and there will be no backwardation only if the fixed extraction cost rises at the rate of interest. In fact, whenever total per unit extraction costs rise by less than the interest rate, the market will exhibit weak backwardation. Strong backwardation, however will occur only if extraction costs decline sufficiently fast over time. Litzenberger and Rabinowitz (1995) document that between February 1984 and April 1992 the nine months future price was strongly backwardated 77 % of the time and weakly backwardated 94 % of the time. Weak backwardation means that discounted futures prices are below the current spot price. Their paper introduces uncertainty into the theory of commodity pricing and production and shows that it can account for backwardation, even when extraction costs increase at the rate of interest. In my work I would like to test whether the predictions of the theory they have developed were valid also during the period after 1992 until today. Litzenberger and Rabinowitz's theory is based on the call characteristic of oil reserves, predicts a negative association between backwardation and volatility. It is useful to note that the alternative explanations of backwardation are not mutually exclusive.

### **Heaney's Theory**

Heaney (2002) offers an alternative explanation for holding stocks-the existence of a sales timing option that accrues to stockholder. The firm can always choose between selling ex-

isting commodity stocks or using the commodity in the production of finished goods. The producer(stockholder) holds a put option on the stored commodity at a price at least equal to the marginal cost of production at some future time. The combination of this put option and the underlying stock holding creates a call option whose value is increasing in commodity price, much like the convenience yield discussed in the earlier literature. If the impact of convenience yield is to be identified then the cost of carry pricing model is a useful starting point for analysis. In its basic form this model captures the storage value noted by Keynes, with the future price,  $F_{t,T}$ , quoted at time  $t$  for a contract maturing at time  $T$ , expressed in terms of the underlying commodity price,  $P_t$ , quoted at time  $t$ , and the costs of storage which include,  $r$ , the continuously compounding risk-free rate of return for the period  $t$  to  $T$ , the physical costs of storage,  $s$ , continuously compounding for the period  $t$  to  $T$  and the exponential function term,  $exp(\cdot)$ . The convenience yield,  $cy$ , is included for the period from time  $t$  to  $T$  in the model below. The cost of carry model, adjusted for convenience yield, takes the form:

$$F_{t,T} = P_t \exp((r + s - cy)(T - t)) \quad (2.6)$$

This can be rearranged to give the interest adjusted basis:

$$IAB_t = \ln(P_t/F_{t,T}) + r = cy - s \quad (2.7)$$

Drawing on the cost of carry model and extending it to deal with the impact of stock outs, Scheinkman and Schechtman (1983) show that it is possible to model commodity prices in terms of two pricing regimes, value in consumption and value in storage, much like the process that Keynes described. The process driving the underlying commodity price is written as:

$$P_t = \max(E_t(P_T) \exp(-r - s), P_t(x)) \quad (2.8)$$

where  $P_t(x)$  is the commodity price in the market for immediate consumption given  $x$  units of commodity are available in the market. If the risk neutral world is assumed then the future price,  $F_{tT}$ , is equal to the expected spot price,  $E_t(P_t)$  and so it is possible to rewrite the relationship as:

$$P_t = \max(F_{tT} \exp(-r - s), P_t(x)) \quad (2.9)$$

## 2.5.2 Two Factor Models

Most of the time, the second state variable is the convenience yield. However, models based on long-term price or on volatility of the spot price have also been developed.

### The Convenience Yield as the Second State Variable.

Schwartz (1997) model is the most famous model of commodity prices. It was used as a reference to develop several models that are more sophisticated (Hilliard (1998), Smith and Schwartz (2000)).

Inspired by the one proposed by Gibson and Schwartz in 1990, the latest model is more tractable than its former version: it has an analytical solution. The two factor model supposes that the spot price  $S$  and the convenience yield  $C$  can explain the behavior of the futures price  $F$ . The dynamic of this state variable is:

$$\begin{cases} dS = (\mu - C)Sdt + \sigma_s S dz_s \\ dC = [\kappa(\alpha - C)]dt + \sigma_C dz_C \end{cases} \quad (2.10)$$

with  $k, \sigma_s, \sigma_C > 0$  where  $\mu$  is the drift of the spot price,  $\sigma_s$  is the spot price volatility,  $dz_s$  is an increment to a standard Brownian motion associated with  $S$ ,  $\alpha$  is the long run mean of the convenience yield,  $\kappa$  is the speed of adjustment of the convenience yield,  $\sigma_C$  is the volatility of the convenience yield,  $dz_C$  is an increment to a standard Brownian motion associated with  $C$ .

The idea of mean reverting process is retained in this model. However it is applied to the convenience yield. When applied to the convenience yield, the Ornstein-Uhlenbeck process relies on the hypothesis that there is a regeneration property of inventories, namely that there is a level of stocks which satisfies the needs of industry under normal conditions. When the convenience yield is low, the stocks are abundant and the operators sustain a high storage cost compared with the benefits related to holding raw materials. Therefore, if they are rational, they try to reduce these surplus stocks. Conversely, when the stocks are rare the operators tend to reconstitute them. Moreover, as the storage theory shows, the two state variables are correlated. Both the spot price and the convenience yield are indeed an inverse function of the inventory level. Nevertheless, as Gibson and Schwartz[1990] have demonstrated, the correlation between these two variables is not perfect. Therefore, the increments to standard Brownian motions are correlated, with:

$$E[dz_s \times dz_C] = \rho dt \quad (2.11)$$

where  $\rho$  is the correlation coefficient.

This model, which is quite tractable, presents a limit. Indeed, it ignores that in commodity markets, the price's volatility is positively correlated with the degree of backwardation. The basis has an asymmetrical behavior: in contango, its level is limited to storage costs, but this is not the case in backwardation. Furthermore, the basis is stable in contango, and volatile in backwardation, since in the later situation stocks cannot absorb price fluctuations. This phenomenon can lead to the assumption that the convenience yield is an option. (Brennan 1991) introduces an asymmetric convenience yield in his model because he takes into account a non-negativity constraint on inventory. However, he supposes that convenience yield is deterministic. In the model presented by Routledge, Seppi, and Spatt (2000) the asymmetry in the behavior of the convenience yield is introduced in the correlation between the spot price and the convenience yield. This correlation is higher in backwardation than in contango. The convenience yield is endogenous, and it is determined by the storage process. However, it is stochastic. The two factors are the spot price and exogenously transitory shocks affecting supply and demand. Lautier and Galli (2001) propose a two-factor model inspired by the Schwartz model(1997), in which the convenience yield is also mean reverting and acts as a continuous dividend. However, the asymmetry is introduced in the convenience yield dynamic: it is high and volatile in backwardation, when inventories are rare. it is conversely low and stable when inventories are abundant. The asymmetry is measured by the parameter  $\beta$ . When the later is set to zero, the asymmetrical model is reduced to Schwartz's model.

### **The long-Term Price as a Second State Variable**

Another approach of the term structure of commodity prices consists in considering the decreasing pattern of volatilities along the price curve. In that situation, it is possible to infer that the two extremities of the price curve, the spot and the long term prices. This kind of approach was followed by Gabillon (1992) and Smith and Schwartz (2000). Gabillon (1992) uses the spot and long term prices as state variables. In this model, the convenience yield is an endogenous variable, dependent on the two factors. The use of the long term price as a second state variable is justified by the fact that this price can be influenced by elements that are exogenous to the physical market, such as expected inflation, interest rates, or prices for renewable energies. The author proposes a geometric Brownian motion to represent the behavior of the long term price. Moreover, the two state variables are assumed to be positively correlated. Schwartz and Smith propose a two-factor model where the state variables come from decomposition of the spot price- This decomposition authorizes the distinction between short term variations and long term equilibrium.

$$\ln(S_t) = \chi_t + \xi_t \tag{2.12}$$

where:  $\chi_t$  is the short-term deviation in prices; and  $\xi_t$  is the equilibrium price level. These two factors are not directly observable, but they are estimated from spot and futures prices. The movements in prices for long-maturity futures contracts provide information about the equilibrium price level, whereas the difference between the short and long term futures prices provide information about short term variations in prices. The most important advantage of this model is that it avoids the questions concerning the convenience yield, its estimation and its economic significance. The idea of a long term equilibrium is also in line with recent works on long memory in the commodity futures markets. However, using the long term price as a state variable introduces a new problem: is it feasible to represent a stable equilibrium with a stochastic variable?

**Seasonality** Apart from the nature of state variables, research has been conducted on the seasonality of commodity prices. In this field, Gabillon once again forged new territory in 1992 with a model including a seasonal function as a composite of sine and cosine functions. The same type of formalization was retained by Richter and Sorensen (2002). However, the later model takes the spot price and the convenience yield as state variables, whereas Gabillon(1992) retains the spot and the long term prices.

### 2.5.3 Three-Factor Models

Until 1997, every term structure model of commodity prices assumed the interest rate constant. Such a hypothesis amounts to say that the term structure of interest rates is flat, which is all the more reductive as the horizon of analysis is remote. Schwartz (1997) proposes a model including three state variables: the spot price, the convenience yield, and the interest rate. The latter has a mean reverting behavior. In 2000, Schwartz and Smith proposed an extension of their short-term/long-term model in which the growth rate for the equilibrium price level is stochastic. Such an extension improves the model's ability to fit long-term futures prices. (Yan 2002) introduces a model which incorporates stochastic convenience yield, interest rates and volatility. He also introduced simultaneous jumps in the spot price and volatility. Yan finds that stochastic volatility and jumps do not alter the futures price at a given point in time. Nevertheless, they play important roles in pricing options on futures.

### 2.5.4 Other Models

**Volatility** Volatility has important implications for the pricing of commodity derivatives and the construction of optimal hedge ratios. For example commodity options will be under

priced if the historical unconditional volatility is assumed during periods when there is a switch from 'low' to 'high' volatility regime. Similarly, highly inaccurate hedge ratios may result when they are computed on the assumption that there are no sudden changes in volatility. Lastly, since the spread between spot and futures prices tend increase during periods of high volatility, commodity traders will face higher basis risks, which compounds the problem of determining an optimal hedge ratio. Fong and See (2003) examine temporal behavior of daily returns for crude oil futures and its relation to the basis, i.e the percentage difference between contemporaneous spot and future price. Similarly to Litzenberg and Rabinovitz they find that oil futures exhibit backwardation 71 % of the time. Moreover, they find that the more volatile are demand and supply shocks, the higher must the spot price be relative to future price to overcome producer's preference to leave oil underground. Authors attempt to model complex features of volatility by using a general regime switching model. Their model builds on the standard GARCH approach by allowing for jumps in the conditional variance between a discreet number of states or regimes. The model is flexible in that all GARCH parameters can switch between regimes. The regimes are treated as a latent variable that can be estimated along with the other parameters of the model using maximum likelihood. The transition between states is assumed to be governed by a Markov process to capture the possible effects of basis on the persistence of volatility regimes, they allow transition probabilities to be a time varying function of the basis. The paper concludes that regime shifts are present in the data and dominate GARCH effects. They also find that conditional on high volatility state, the more negative the basis, the more persistent is the regime.

**Stochastic Convenience Yield** As we have seen convenience yield is an important variable in many term structure models. In the energy markets, oil and natural gas consumers are characterized by relatively inelastic demand. End users, refiners, and distribution companies cannot do business without a supply of oil and gas. Because they cannot afford to be without oil and gas, these firms hold energy inventories. The financial benefit that accrues to holders of inventories is called convenience yield. The value of the convenience yield influences the term structure of energy prices. One theory of backwardation holds that when excess supplies are available, inventories increase and convenience yield declines. Low convenience yields push the market toward full carry, forward prices in steep contango. When gas and oil supplies are short, inventories are drained from the market and end-users are willing to pay more today for an uninterrupted supply of energy. As convenience yield increases, the market can swing into backwardation. Because a commodity can be consumed, its price is a combination of future asset and current consumption values. However, unlike financial derivatives, storage of

energy products is costly and sometimes practically impossible like in the case of electricity. Consequently, physical ownership of the commodity carries an associated flow of services. On the one hand, the agent has the option of flexibility with regards to consumption (no risk of commodity shortage). On the other hand, the decision to postpone consumption implies storage expenses. The net flow of these services per unit of time is termed the convenience yield and we will note it as  $\delta_t$ . From now on we assume that  $\delta_t$  is quoted on a continuously compounded basis. Intuitively, the convenience yield corresponds to dividend yield for stocks.

$$\delta_t = \text{Benefit of direct access} - \text{Cost of carry}$$

Commodity pricing models are obtained via various assumptions on the behavior of  $\delta_t$ . The implicit assumption is that  $S_t$ , the spot price process of the commodity, in fact exists. This is not true for some commodities, such as electricity. Even for mature markets like crude oil where spot prices are quoted daily, the exact meaning of the spot is difficult to pin down. Nevertheless, we will maintain the industry-standard assumption of traded spot asset. By a basic no-arbitrage argument it follows that the price of a forward contract  $F(t; T)$  which has payoff  $S_T$  at future time  $T$  must equal

$$F(t; T) = S_t E_Q \left[ e^{\int_t^T (r_s - \delta_s) ds} \right] \quad (2.13)$$

Indeed, we can replicate the payoff by either entering into the forward contract, or by borrowing  $S_t$  today and holding the commodity itself from now until  $T$ . As usual,  $Q$  is an equivalent martingale measure used for risk neutral pricing. Reformulating (2.13) implies that the risk-neutral drift of the spot commodity must equal to  $(r_t - \delta_t)S_t$ . From a modeling point of view, it remains to specify the stochastic processes for  $S_t$ ,  $r_t$  and  $\delta_t$ . However, the difficulty lies in the fact that the convenience yield is unobserved. Notice that  $\delta_t$  is defined indirectly as a "correction" to the drift of the spot process. Thus, for any model we specify, we must address the issue of estimating or filtering the convenience yield process given the model observables. Gibson and Schwartz (1989) were the first to introduce stochastic convenience yield. They assume that spot price of oil is not unique determinant of price of oil contingent claim. They derive two-factor pricing model and analyze its performance in valuing short as well as long term oil contracts. They define two proxies for  $S_t$  and  $\delta_t$  since neither of these two state variables can actually be observed. There is no true spot market for crude oil and therefore Gibson and Schwartz identify this state variable with the settlement price of the closest to maturity crude oil future trading on the New York Mercantile Exchange.

The procedure which has been used to compute the instantaneous convenience yield of

crude oil relies on the well known relationship between the futures and the spot price of a commodity when there is neither interest rate nor convenience yield uncertainty, namely

$$\delta_{T-1,T} = r_{T-1,T} - 12 \ln \left[ \frac{F(S,T)}{F(S,T-1)} \right] \quad (2.14)$$

where  $\delta_{T-1,T}$  denotes the  $T-1$  periods ahead one month forward convenience yield, and  $r_{T-1,T}$  denotes the  $T-1$  periods ahead one month riskless forward interest rate. The following graph shows weekly convenience yield dynamics during the period 1990-2006.

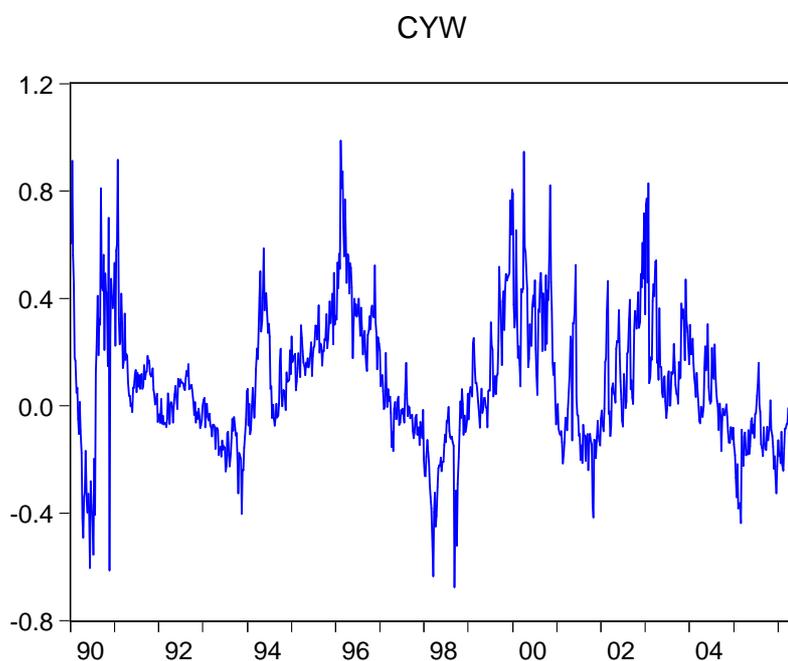


Figure 2.1: Convenience Yield for Crude Oil 1990-2006

However, in this work we will propose a GARCH model for convenience yield dynamic. But before performing calculations we will take a brief overview of univariate GARCH models. Afterwards, we will compare different GARCH models for convenience yield dynamics. In the second part of my work we will address the issue of hedging in the crude oil market. We will propose hedging strategy based on the bivariate GARCH model of spot and futures prices of crude oil. The strategy will directly incorporate the fact that spot and futures prices of crude oil are cointegrated. Conclusions and directions for future work will follow.

## Chapter 3

# A GARCH Model of convenience yield

Financial economists are concerned with modeling volatility in asset returns. This is important as volatility is considered a measure of risk, and investors want a premium for investing in risky assets. Banks and other financial institutions apply so-called value-at-risk models to assess their risks. Modeling and forecasting volatility or, in other words, the covariance structure of asset returns, is therefore important. The fact that volatility of returns fluctuates over time has been known for a long time. Originally, the emphasis was on another aspect of return series: their marginal distributions were found to be leptokurtic. Returns were modeled as independent and identically distributed over time. In a classic work, Mandelbrot (1963) and Mandelbrot and Taylor (1967) applied so-called stable Paretian distributions to characterize the distribution of returns. Rachev and Mittnik (2000) contains an informative discussion of stable Paretian distributions and their use in finance and econometrics. Observations in return series of financial assets observed at weekly and higher frequencies are in fact not independent. While observations in these series are uncorrelated or nearly uncorrelated, the series contain higher order dependence. Models of Autoregressive Conditional Heteroscedasticity (ARCH) form the most popular way of parameterizing this dependence. Its generalization is a GARCH model.

### 3.1 Overview of Univariate GARCH Models

In this section we will present an overview of univariate GARCH models citing Teräsvirta (2006).

**The ARCH Model** The autoregressive conditional heteroscedasticity (ARCH) model is the first model of conditional heteroscedasticity. According to Engle (1982), the original idea was to find a model that could assess the validity of the conjecture of Friedman (1977) that the unpredictability of inflation was a primary cause of business cycles. Uncertainty due to this unpredictability would affect the investors behavior. Pursuing this idea required a model in which this uncertainty could change over time. Engle (1982) applied his resulting ARCH model to parameterizing conditional heteroscedasticity in a wage-price equation for the United Kingdom. Let  $\epsilon_t$  be a random variable that has a mean and a variance conditionally on the information set  $F_{t-1}$ , (the  $\sigma$  field generated by  $\epsilon_{t-j}$  ;  $j \geq 1$ ). The ARCH model of  $\epsilon_t$  has the following properties. First,  $E[\epsilon_t/F_{t-1}] = 0$  and, second, the conditional variance  $h_t = E[\epsilon_t^2/F_{t-1}]$ , is a nontrivial positive-valued parametric function of  $F_{t-1}$ . The sequence  $[\epsilon_t]$  may be observed directly, or it may be an error or innovation sequence of an econometric model. In the latter case:

$$\epsilon_t = y_t - \mu_t(y_t) \tag{3.1}$$

where  $y_t$  is an observable random variable and  $\mu_t(y_t) = E[y_t/F_{t-1}]$ , the conditional mean of  $y_t$  given  $F_{t-1}$ . Engle's (1982) application was of this type. In what follows, the focus will be on parametric forms of  $h_t$  and  $\mu_t(y_t)$  will be ignored. Engle assumed that  $\epsilon_t$  can be decomposed as follows:

$$\epsilon_t = z_t h_t^{1/2} \tag{3.2}$$

where  $z_t$  is a sequence of independent, identically distributed (iid) random variables with zero mean and unit variance. This implies  $\epsilon_t/F_{t-1} \sim D(0, h_t)$  where D stands for the distribution (typically assumed to be a normal or a leptokurtic one). The following conditional variance defines an ARCH model of order q:

$$h_t = \alpha_0 + \sum_{j=1}^{q-1} \alpha_j \epsilon_{t-j}^2 \tag{3.3}$$

where  $\alpha_0 > 0, \alpha_j \geq 0, j = 1, \dots, q - 1$ , and  $\alpha_q > 0$ . The parameter restrictions in (3.3) form a necessary and sufficient condition for positivity of the conditional variance. Suppose the unconditional variance  $E[\epsilon_t^2] = \sigma^2 < \infty$ . The definition of  $\epsilon_t$  through the decomposition (3.2) involving  $z_t$  then guarantees the white noise property of the sequence  $\epsilon_t$  since  $z_t$  is a sequence of iid variables. Although the application in Engle (1982) was not a financial one, Engle and others soon realized the potential of the ARCH model in financial applications

that required forecasting volatility. The ARCH model and its generalizations are thus applied to modeling, among other things, interest rates, exchange rates and stock and stock index returns. Bollerslev and Kroner (1992) already listed a variety of applications in their survey of these models. Forecasting volatility of these series is different from forecasting the conditional mean of a process because volatility, the object to be forecast, is not observed. The question then is how volatility should be measured. Using  $\epsilon_t^2$  is an obvious but not necessarily a very good solution if data of higher frequency are available; see Andersen and Bollerslev (1998) for discussion.

**The Generalized ARCH model.** In applications, the ARCH model has been replaced by the so-called generalized ARCH (GARCH) model that Bollerslev (1986) and Taylor (1986) proposed independently of each other. In this model, the conditional variance is also a linear function of its own lags and has the form:

$$h_t = \alpha_0 + \sum_{j=1}^q \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (3.4)$$

The conditional variance defined by (3.4) has the property that the unconditional autocorrelation function of  $\epsilon_t^2$ ; if it exists, can decay slowly, albeit still exponentially. For the ARCH family, the decay rate is too rapid compared to what is typically observed in financial time series, unless the maximum lag  $q$  in (3.3) is long. As (3.4) is a more parsimonious model of the conditional variance than a high-order ARCH model, most users prefer it to the simpler ARCH alternative. The overwhelmingly most popular GARCH model in applications has been the GARCH(1,1) model, that is,  $p = q = 1$  in (3.4). A sufficient condition for the conditional variance to be positive with probability one is  $\alpha_0 > 0, \alpha_j > 0, j = 1, \dots, q$  and  $\beta_j \geq 0, j = 1, \dots, p$ . The necessary and sufficient conditions for positivity of the conditional variance in higher-order GARCH models are more complicated than the sufficient conditions just mentioned and have been given in Nelson and Cao (1992). The GARCH(2,2) case has been studied in detail by He and Terasvirta (1999b). Note that for the GARCH model to be identified if at least one  $\beta_j > 0$  (the model is a genuine GARCH model) one has to require that also at least one  $\alpha_j > 0$ : If  $\alpha_1 = \dots = \alpha_q = 0$ , the conditional and unconditional variances of  $\epsilon_t$  are equal and  $\beta_1, \dots, \beta_p$  are unidentified nuisance parameters. The GARCH(p,q) process is weakly stationary if and only if  $\sum \alpha_j + \sum \beta_j < 1$ .

### 3.1.1 Families of Univariate GARCH Models

Since its introduction the GARCH model has been generalized and extended in various directions. This has been done to increase the flexibility of the original model. For example, the original GARCH specification assumes the response of the variance to a shock to be independent of the sign of the shock and just be a function of the size of the shock. Several extensions of the GARCH model aim at accommodating the asymmetry in the response. These include the GJR-GARCH model of Glosten, Jagannathan, and Runkle (1993), the asymmetric GARCH models of Engle and Ng (1993) and the quadratic GARCH of Sentana (1995). The GJR-GARCH model has the form (3.2), where

$$h_t = \alpha_0 + \sum_{j=1}^q \{\alpha_j + \delta_j I(\varepsilon_{t-j} > 0)\} \varepsilon_{t-j}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (3.5)$$

In (3.5),  $I(\varepsilon_{t-j} > 0)$  is an indicator function obtaining value one when the argument is true and zero otherwise. In the asymmetric models of both Engle and Ng, and Sentana, the centre of symmetry of the response to shocks is shifted away from zero. For example,

$$h_t = \alpha_0 + \alpha_1 (\varepsilon_{t-1} - \gamma)^2 + \beta_1 h_{t-1} \quad (3.6)$$

with  $\gamma \neq 0$  in Engle and Ng (1993). The conditional variance in Sentana's Quadratic ARCH (QARCH) model (the model is presented in the ARCH form) is defined as follows:

$$h_t = \alpha_0 + \alpha' \varepsilon_{t-1} + \varepsilon_{t-1}' A \varepsilon_{t-1} \quad (3.7)$$

where  $\varepsilon_t = (\varepsilon_t, \dots, \varepsilon_{t-q+1})'$  is a  $q \times 1$  vector,  $\alpha = (\alpha_1, \dots, \alpha_q)$  is a  $q \times 1$  parameter vector and  $A$  a  $q \times q$  symmetric parameter matrix. In (3.7), not only squares of  $\varepsilon_{t-i}$  but also cross-products  $\varepsilon_{t-i} \varepsilon_{t-j}$ ,  $i \neq j$ ; contribute to the conditional variance. When  $\alpha \neq 0$  the QARCH generates asymmetric responses. The ARCH equivalent of (3.6) is a special case of Sentana's model. Constraints on parameters required for positivity of  $h_t$  in (3.7) becomes clear by rewriting (3.7) as follows

$$h_t = \begin{bmatrix} \varepsilon_{t-1} & 1 \end{bmatrix} \begin{bmatrix} A & \alpha/2 \\ \alpha'/2 & \alpha_0 \end{bmatrix} \begin{bmatrix} \varepsilon_{t-1} \\ 1 \end{bmatrix} \quad (3.8)$$

The conditional variance  $h_t$  is positive if and only if the matrix in the quadratic form on the right hand side of (3.8) is positive definite.

Some authors have suggested modeling the conditional standard deviation instead of the conditional variance: see Taylor (1986), Schwert (1990), and for an asymmetric version, Za-

koian (1994). A further generalization of this idea appeared in Ding, Granger and Engle (1993). These authors proposed a GARCH model for  $h_t^k$  where  $k > 0$  is also a parameter to be estimated. Their power GARCH model is (3.2) with

$$h_t^k = \alpha_0 + \sum_{j=1}^q \alpha_j |\varepsilon_{t-j}|^{2k} + \sum_{j=1}^p \beta_j h^k h_{t-j}, \quad k > 0 \quad (3.9)$$

The authors argued that this extension provides flexibility lacking in the original GARCH specification of Bollerslev (1986) and Taylor (1986). The proliferation of GARCH models has inspired some authors to define families of GARCH models that would accommodate as many individual models as possible. Hentschel (1995) defined one such family. The first order GARCH model has the general form

$$\frac{h_t^{\lambda/2} - 1}{\lambda} = \omega + \alpha h_{t-1}^{\lambda/2} f^\nu(z_{t-1}) + \beta \frac{h_{t-1}^{\lambda/2} - 1}{\lambda} \quad (3.10)$$

where

$$f^\nu(z_t) = |z_t - b| - c(z_t - b) \quad (3.11)$$

This family contains a large number of well-known GARCH models. The Box-Cox type transformation of the conditional standard deviation  $h_t^{1/2}$  makes it possible, by allowing  $\lambda \rightarrow 0$  to accommodate models in which the logarithm of the conditional variance is parameterized, such as the exponential GARCH model to be considered in Section (3.2). Parameters  $b$  and  $c$  in  $f^\nu(z_t)$  allow the inclusion of different asymmetric GARCH models such as the GJR-GARCH or threshold GARCH models in this family.

Another family of GARCH models that is of interest is the one of Terasvirta (1999) defined as follows:

$$h_t^k = \sum_{j=1}^q g(z_{t-j}) + \sum_{j=1}^q c_j(z_{t-j}) h_{t-j}^k, \quad k > 0 \quad (3.12)$$

where  $\{g(z_t)\}$  and  $\{c(z_t)\}$  are sequences of independent and identically distributed random variables. In fact, the family was originally defined for  $q = 1$ , but the definition can be generalized to higher-order models. For example, the standard GARCH( $q$ ;  $q$ ) model is obtained by setting  $g(z_t) = \alpha_0 = q$  and  $c_j(z_{t-j}) = \alpha_j z_{t-j}^2 + \beta_j, j = 1, \dots, q$ , in (3.12).

Many other GARCH models such as the GJR-GARCH, the absolute-value GARCH, the Quadratic GARCH and the power GARCH model belong to this family. Note that the power GARCH model itself nests several well-known GARCH models; see Ding et al. (1993) for details. Definition (3.12) has been used for deriving expressions of fourth moments, kurtosis and the autocorrelation function of  $\varepsilon_t^2$  for a number of first-order GARCH models and the

standard GARCH(p; q) model.

The family of augmented GARCH models, defined by Duan (1997), is a rather general family. The first-order augmented GARCH model is defined as follows. Consider (3.2) and assume that

$$h_t = \begin{cases} |\lambda\phi_t - \lambda| & \text{if } \lambda \neq 0 \\ \exp\{\phi_t - 1\} & \text{if } \lambda = 0 \end{cases} \quad (3.13)$$

where

$$\phi_t = \alpha_0 + \zeta_{1,t-1}\phi_{t-1} + \zeta_{2,t-1} \quad (3.14)$$

In (3.14),  $(\zeta_1, \zeta_2)$  is a strictly stationary sequence of random vectors with a continuous distribution, measurable with respect to the available information until  $t$ . Duan defined an augmented GARCH(1,1) process as (3.2) with (3.13) and (3.14), such that

$$\begin{aligned} \zeta_{1t} &= \alpha_1 + \alpha_2 |\varepsilon_t - c|^\delta + \alpha_3 \max(0, c - \varepsilon_t)^\delta \\ \zeta_{2t} &= \alpha_4 \frac{|\varepsilon_t - c|^\delta - 1}{\delta} + \alpha_5 \frac{\max(0, c - \varepsilon_t)^\delta - 1}{\delta} \end{aligned} \quad (3.15)$$

This process contains as special cases all the GARCH models previously mentioned, as well as the Exponential GARCH model to be considered in Section 3.2. Duan (1997) generalized this family to the GARCH(p; q) case and derived sufficient conditions for strict stationarity for this general family as well as conditions for the existence of the unconditional variance of  $\varepsilon_t$ . Furthermore, he suggested misspecification tests for the augmented GARCH model.

### 3.1.2 Nonlinear GARCH

**Smooth Transition GARCH** As mentioned above, the GARCH model has been extended to characterize asymmetric responses to shocks. The GJR-GARCH model, obtained as setting  $g(z_t) = \alpha_0$ . A nonlinear version of the GJR-GARCH model is obtained by making the transition between regimes smooth. (Hagerud 1997), Gonzalez-Rivera (1998) and Anderson, Nam, and Vahid (1999) proposed this extension. A smooth transition GARCH (STGARCH) model may be defined as equation with

$$h_t = \alpha_{10} + \sum_{j=1}^q \alpha_{1j} \varepsilon_{t-j}^2 + (\alpha_{20} + \sum_{j=1}^q \alpha_{2j} \varepsilon_{t-j}^2) G(\gamma, c; \varepsilon_{t-j}) + \sum_{j=1}^p \beta_j h_{t-j} \quad (3.16)$$

where the transition function

$$G(\gamma, c; \varepsilon_{t-j}) = (1 + \exp(-\gamma \Pi(\varepsilon_{t-j} - c_k)))^{-1}, \quad \gamma > 0 \quad (3.17)$$

When  $K = 1$  (3.17) is a simple logistic function that controls the change of the coefficient of  $\epsilon_{t-j}^2$  from  $\alpha_{1j}$  to  $\alpha_{1j} + \alpha_{2j}$  as a function of  $\epsilon_{t-j}$ ; and similarly for the intercept. In that case, as  $\gamma \rightarrow \infty$ , the transition function becomes a step function and represents an abrupt switch from one regime to the other. Furthermore, at the same time setting  $c_1 = 0$  yields the GJR- GARCH model because  $\varepsilon_t$  and  $z_t$  have the same sign. When  $K = 2$  and, in addition,  $c_1 = -c_2$  in (3.17), the resulting smooth transition GARCH model is still symmetric about zero, but the response of the conditional variance to a shock is a nonlinear function of lags of  $\varepsilon_t$ . Smooth transition GARCH models are useful in situations where the assumption of two distinct regimes is too rough an approximation to the asymmetric behavior of the conditional variance. Hagerud (1997) also discussed a specification strategy that allows the investigator to choose between  $K = 1$  and  $K = 2$  in (3.17). Values of  $K > 2$  may also be considered, but they are likely to be less common in applications than the two simplest cases. The smooth transition GARCH model (3.16) with  $K = 1$  in (3.17) is designed for modeling asymmetric responses to shocks. On the other hand, the standard GARCH model has the undesirable property that the estimated model often exaggerates the persistence in volatility (the estimated sum of the  $\alpha$  and  $\beta$ -coefficients is close to one). This in turn results in poor volatility forecasts. In order to remedy this problem, Lanne and Saikkonen (2005) proposed a smooth transition GARCH model whose first-order version has the form:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \delta_1 G_1(\theta; h_{t-1}) + \beta_1 h_{t-1} \quad (3.18)$$

In (3.18)  $G_1(\theta; h_{t-1})$  is a continuous bounded function such as (3.17): Lanne and Saikkonen use the cumulative distribution function of the gamma-distribution. A major difference between (3.16) and (3.18) is that in the latter model the transition variable is a lagged conditional variance. In empirical examples given in the paper, this parameterizations clearly alleviates the problem of exaggerated persistence.

**Threshold GARCH and extensions** If (3.16) is defined as a model for the conditional standard deviation such that  $h_t$  is replaced by  $h^{1/2}$ ,  $h_{t-j}$  by  $h_{t-j}^{1/2}$ ,  $j = 1, \dots, p$  and  $\epsilon_{t-j}^2$  by  $|\varepsilon_{t-j}|$ ,  $j = 1, \dots, q$ , then choosing  $K = 1$  and letting  $\gamma \rightarrow \infty$  in (3.17) yields the threshold GARCH (TGARCH) that (Zakoian 1994) considered.

The TGARCH (p;q) model is thus the counterpart of the GJR-GARCH model in the case where the entity to be modeled is the conditional standard deviation instead of the conditional variance. Note that in both of these models, the threshold parameter has a known value (zero). In Zakoian (1994), the conditional standard deviation is defined as follows:

$$h_t^{1/2} = \alpha_0 + \sum(\alpha_j^+ \epsilon_{t-j}^+ - \alpha_j^- \epsilon_{t-j}^-) + \sum \beta_j h_{t-j}^{1/2} \quad (3.19)$$

where  $\epsilon_{t-j}^+ = \max(\epsilon_{t-j}, 0)$  and  $\epsilon_{t-j}^- = \min(\epsilon_{t-j}, 0)$ . Rabemananjara and Zakoian (1993) introduced an even more general model in which  $h_t^{1/2}$  can obtain negative values, but it has not gained wide acceptance. Nevertheless, these authors provide evidence of asymmetry in the French stock market by fitting the TGARCH model to the daily return series of stocks included in the CAC 40 index of the Paris Stock Exchange. The TGARCH model is linear in parameters because the threshold parameter is assumed to equal zero. A genuine nonlinear threshold model is the Double Threshold ARCH (DTARCH) model of Li and Li (1996). It is called a double threshold model because both the autoregressive conditional mean and the conditional variance have a threshold-type structure as defined in Tong (1990). The conditional mean and variance are defined as follows:

$$y_t = \sum(\phi_{0k} + \sum \phi_{jk} y_{t-j}) I(c_{k-1}^m < y_{t-b} \leq c_k^m)_l + \epsilon_t \quad (3.20)$$

$$h_t = \sum(\alpha_{0l} + \sum \alpha_{jl} \epsilon_{t-j}^2) I(c_{l-1}^v < y_{t-d} \leq c_l^v) \quad (3.21)$$

Furthermore,  $k = 1, \dots, K, l = 1, \dots, L$ , and  $b$  and  $d$  are delay parameters,  $b, d \leq 1$ . The number of regimes in (3.20) and (3.21),  $K$  and  $L$ , respectively, need not be the same, nor do the two threshold variables have to be equal. Other threshold variables than lags of  $y_t$  are possible. For example, replacing  $y_{t-d}$  in (3.21) by  $\epsilon_{t-d}$  or  $\epsilon_{t-d}^2$  may sometimes be an interesting possibility. Another variant of the DTARCH model is the model that Audrino and Bühlmann (2001) who introduced it called the Tree-Structured GARCH model. A practical problem is that the tree-growing strategy of Audrino and Bühlmann (2001) does not seem to prevent under identification: if  $K$  is chosen too large, it is not identified. A similar problem is present in the DTARCH model as well as in the STGARCH one. Hagerud (1997) and Gonzalez-Rivera (1998), however, do provide linearity tests in order to avoid this problem in the STGARCH framework.

**Time-varying GARCH** An argument brought forward in the literature, see for instance Mikosch and Starica (2004), is that in applications the assumption of the GARCH models having constant parameters may not be appropriate when the series to be modeled are long. Parameter constancy is a testable proposition, and if it is rejected, the model can be generalized. One possibility is to assume that the parameters change at specific points of time, divide the series into sub series according to the location of the break-points, and fit separate

GARCH models to the subseries. The main statistical problem is then finding the number of break-points and their location because they are normally not known in advance. Chu (1995) has developed tests for this purpose. Another possibility is to modify the smooth transition GARCH model (3.16) to fit this situation. This is done by defining the transition function (3.17) as a function of time:

$$G(\gamma; c; t^*) = 1 + \exp(-\gamma \Pi(t^* - c_k))^{-1}, \quad \gamma > 0 \quad (3.22)$$

where  $t^* = t/T$ . Standardizing the time variable between zero and unity makes interpretation of the parameters  $c_k, k = 1, \dots, K$ , easy as they indicate where in relative terms the changes in the process occur. The resulting time varying parameter GARCH (TV-GARCH) model has the form

$$h_t = \alpha_0(t) + \sum \alpha_j(t) \epsilon_{t-j}^2 + \sum \beta_j(t) h_{t-j} \quad (3.23)$$

where  $\alpha_0(t) = \alpha_{01} + \alpha_{02}G(\gamma, c; t^*), \alpha_j(t) = \alpha_{j1} + \alpha_{j2}G(\gamma, c; t^*), j=1, \dots, q$ , and  $\beta_j(t) = \beta_{j(1)} + \beta_{j(2)}G(\gamma, c; t^*), j = 1, \dots, p$ . This is the most flexible parameterizations. Some of the time-varying parameters in (3.23) may be restricted to constants a priori. For example, it may be assumed that only the intercept  $\alpha_0(t)$  is time-varying. This implies that the 'baseline volatility' or unconditional variance is changing over time. If change is allowed in the other GARCH parameters then the model is capable of accommodating systematic changes in the amplitude of the volatility clusters that cannot be explained by a constant-parameter GARCH model. This type of time-varying GARCH is mentioned here because it is a special case of the smooth transition GARCH model. Other time-varying parameter models of conditional heteroscedasticity, such as non stationary ARCH models with locally changing parameters, are discussed in SPOKOINY (2007).

**Markov-switching ARCH and GARCH** Markov-switching or hidden Markov models of conditional heteroscedasticity constitute another class of nonlinear models of volatility. These models are an alternative way of modeling volatility processes that contains breaks. Hamilton and Susmel (1994) argued that very large shocks, such as the one affecting the stocks in October 1987, may have consequences for subsequent volatility so different from consequences of small shocks that a standard ARCH or GARCH model cannot describe them properly. Their Markov regime-switching ARCH model is defined as follows:

$$h_t = (\alpha_0^i + \sum \alpha_j^i \epsilon_{t-j}^2) I(s_t = i) \quad (3.24)$$

where  $s_t$  is a discrete unobservable random variable obtaining values from the set  $S = 1, \dots, k$ . It follows a (usually first-order) homogeneous Markov chain:  $Pr(s_t = j/s_{t-1} = i) = p(ij)$ ,  $i, j = 1, \dots, k$ . Cai (1994) considered a special case of (3.24) in which only the intercept  $\alpha_0^i$  is switching, and  $k = 2$ . But then, his model also contains a switching conditional mean. Furthermore, Rydén, Terasvirta and Åsbrink (1998) showed that a simplified version of (3.24) where  $\alpha_j^i = 0$  for  $j \geq 1$  and all  $i$ ; is already capable of generating data that display most of the stylized facts that Granger and Ding (1995) ascribe to high-frequency financial return series. This suggests that a Markov-switching variance alone without any ARCH structure may in many cases explain a large portion of the variation in these series. Nevertheless, it can be argued that shocks drive economic processes, and this motivates the ARCH structure. If the shocks have a persistent effect on volatility, however, a parsimonious GARCH representation may be preferred to (3.24). Generalizing (3.24) into a GARCH model involves one major difficulty: a straightforward (first-order) generalization would have the following form:

$$h_t = (\alpha_0^i + \alpha_1^i \varepsilon_{t-1}^2 + \beta^i h_{t-1}) I(s_t = i) \quad (3.25)$$

From the autoregressive structure of (3.25) it follows that  $h_t$  is completely path-dependent: its value depends on the unobservable  $s_{t-j}, j=0,1,2,\dots,t$ . This makes the model practically impossible to estimate because in order to evaluate the log-likelihood, this unobservable have to be integrated out of this function. Simplifications of the model that circumvent this problem can be found in Gray (1996) and Klaassen(2002). A good discussion about their models can be found in Haas, Mittnik and Paoletta (2004). These authors present another Markov-switching (MS-) GARCH model whose fourth moment structure they are able to work out. That does not seem possible for the other models. The MS-GARCH model of Haas et al. (2004) is defined as follow:

$$\varepsilon_t = z_t \Sigma h_{it}^{1/2} I(s_t = i)$$

where  $s_t$  is defined as in (3.22). Furthermore,

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + B h_{t-1}$$

where  $h_t = (h_{1t}, \dots, h_{kt})'$ ,

$$\alpha_i = (\alpha_{i1}, \dots, \alpha_{ik})'$$

,  $i = 0, 1$ , and

$$B = \text{diag}(\beta_{11}, \dots, \beta_{kk})'$$

. Thus, each volatility regime has its own GARCH equation. The conditional variance in a given regime is only a function of the lagged conditional variance in the same regime, which is not the case in the other models. The identification problem mentioned earlier is present here as well. If the true model has fewer regimes than the specified one, the latter contains unidentified nuisance parameters. Liu (2006) provides a number of results, including conditions for strict stationarity and the existence of higher-order moments, for this MS-GARCH model. More information about Markov-switching ARCH and GARCH models can be found in Lange and Rahbek (2007).

### 3.1.3 Integrated and fractionally integrated GARCH

In applications it often occurs that the estimated sum of the parameters  $\alpha_1$  and  $\beta_1$  in the standard first-order GARCH model (3.4) with  $p = q = 1$  is close to unity. Engle and Bollerslev (1986), who first paid attention to this phenomenon, suggested imposing the restriction  $\alpha_1 + \beta_1 = 1$  and called the ensuing model an integrated GARCH (IGARCH) model. The IGARCH process is not weakly stationary as  $E\varepsilon_t^2$  is not finite. Nevertheless, the term "integrated GARCH" may be somewhat misleading as the IGARCH process is strongly stationary. Nelson (1991) showed that under mild conditions for  $z_t$  and assuming  $\alpha_0 > 0$ , the GARCH(1,1) process is strongly stationary if:

$$E[\ln(\alpha_1 + \beta_1 z_t^2)] < 0 \tag{3.26}$$

The IGARCH process satisfies (3.26). The analogy with integrated processes, that is, ones with a unit root, is therefore not as straightforward as one might think. For a general discussion of stationarity conditions in GARCH models, see 2007 LINDNER. Nelson (1991) also showed that when an IGARCH process is started at some finite time point, its behavior depends on the intercept  $\alpha_0$ . On the one hand, if the intercept is positive then the unconditional variance of the process grows linearly with time. In practice this means that the amplitude of the clusters of volatility to be parameterized by the model on the average increases over time. The rate of increase need not, however, be particularly rapid. One may thus think that in applications with, say, a few thousand observations, IGARCH processes nevertheless provide a reasonable approximation to the true data-generating volatility process. On the other hand, if  $\alpha_0 = 0$  in the IGARCH model, the realizations from the process collapse to zero almost surely. How rapidly this happens, depends on the parameter  $\beta_1$ . Although the investigator may be prepared to accept an IGARCH model as an approximation, a potentially disturbing fact is that this means assuming that the unconditional variance of the process to be modeled does not exist. It is not clear that this is what one always wants to do. There exist other explanations to the

fact that the sum  $(\alpha_1 + \beta_1)$  estimates to one or very close to one. First Diebold (1986) and later Lamoureux and Lastrapes (1990) suggested that this often happens if there is a switch in the intercept of a GARCH model during the estimation period. This may not be surprising as such a switch means that the underlying GARCH process is not stationary. Another, perhaps more puzzling, observation is related to exponential GARCH models to be considered later on in this work. Malmsten (2004) noticed that if a GARCH(1,1) model is fitted to a time series generated by a stationary first-order exponential GARCH model,(see section 3.2), the probability of the estimated sum  $\alpha_1 + \beta_1$  exceeding unity can sometimes be rather large. In short, if the estimated sum of these two parameters in a standard GARCH(1,1) model is close to unity, imposing the restriction  $(\alpha_1 + \beta_1) = 1$  without further investigation may not necessarily be the most reasonable action to take. Assuming  $p = q = 1$ , the GARCH(p,q) equation (3.4) can also be written in the "ARMA(1,1) form" by adding  $\epsilon_t^2$  to both sides and moving  $h_t$  to the right-hand side:

$$\epsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1)\epsilon_{t-1}^2 + \nu_t - \beta_1\nu_{t-1} \quad (3.27)$$

where  $\{\nu\} = \{\epsilon_t^2 - h_t\}$  is a martingale difference sequence with respect to  $h_t$ . For the IGARCH process,(3.4) has the "ARIMA(0,1,1) form"

$$(1 - L)\epsilon_t^2 = \alpha_0 + \nu_t - \beta_1\nu_{t-1} \quad (3.28)$$

Equation (3.28) has served as a starting-point for the fractionally integrated GARCH (FIGARCH) model. The FIGARCH(1,d,0) model is obtained from (3.28) by replacing the difference operator by a fractional difference operator:

$$(1 - L)^d \epsilon_t^2 = \alpha_0 + \nu_t - \beta_1\nu_{t-1} \quad (3.29)$$

The FIGARCH equation (3.29) can be written as an infinite-order ARCH model by applying the definition  $\nu_t = \epsilon_t^2 - h_t$  to it. This yields:

$$h_t = \alpha_0(1 - \beta_1)^{-1} + \lambda(L)\epsilon_t^2 \quad (3.30)$$

where  $\lambda(L) = 1 - (1 - L)^d(1 - \beta_1 L)^{-1}\epsilon_t^2 = \sum_{j=1}^{\infty} \lambda_j L^j \epsilon_t^2$ ,  $\lambda_j \geq 0$  for all  $j$ . Expanding the fractional difference operator into an infinite sum yields the result that for long lags  $j$ ,

$$\lambda_j = [(1 - \beta_1)\Gamma(d)^{-1}]^{-(1-d)} = c_j^{-(1-d)}; c > 0 \quad (3.31)$$

where  $d \in (0, 1)$  and  $\Gamma(d)$  is the gamma function. From the above equation it is seen that the effect of the lagged  $\epsilon_t^2$  on the conditional variance decays hyperbolically as a function of the lag length. This is the reason why Baillie, Bollerslev and Mikkelsen (1996) introduced the FIGARCH model, as it would conveniently explain the apparent slow decay in autocorrelation functions of squared observations of many daily return series. The FIGARCH model thus offers a competing view to the one according to which changes in parameters in a GARCH model are the main cause of the slow decay in the autocorrelations. The first-order FIGARCH model (3.30) can of course be generalized into a FIGARCH(p; d; q) model. The probabilistic properties of FIGARCH processes such as stationarity, still an open question, are quite complex, see, for example, Davidson (2004) and GIRAITIS (2007) for discussion. The hyperbolic GARCH model introduced in the first-mentioned paper contains the standard GARCH and the FIGARCH models as two extreme special cases; for details see Davidson (2004).

### 3.1.4 Semi- and nonparametric ARCH models

The ARCH decomposition of returns has also been used in a semi- or nonparametric approach. The semi-parametric approach is typically employed in situations where the distribution of  $z_t$  is left unspecified and is estimated non parametrically. In nonparametric models, the issue is the estimation of the functional form of the relationship between  $\epsilon_t^2$  and  $\epsilon_{t-1}^2, \dots, \epsilon_{t-q}^2$ . Semi- and nonparametric ARCH models are considered in detail in Linton (2007).

**GARCH-in-mean model** GARCH in mean models are often used for predicting the risk of a portfolio at a given point of time. From this it follows that the GARCH type conditional variance could be useful as a representation of the time-varying risk premium in explaining excess returns, that is, returns compared to the return of a riskless asset. An excess return would be a combination of the non-forecastable difference between the ex ante and ex post rates of return and a function of the conditional variance of the portfolio. Thus, if  $y_t$  is the excess return at time t, then

$$y_t = \varepsilon_t + \beta + g(h_t) - E_g(h_t) \quad (3.32)$$

where  $h_t$  is defined as a GARCH process (3.4) and  $g(h_t)$  is a positive valued function. Engle, Lilien and Robins (1987) originally defined

$$g(h_t) = \delta h_t^{1/2}, \quad \delta > 0$$

which corresponds to the assumption that changes in the conditional standard deviation appear less than proportionally in the mean. The alternative  $g(h_t) = \delta h_t$  has also appeared in the literature. Equations (3.32) and (3.4) form the GARCH-in-mean or GARCH-M model. It has been quite frequently applied in the applied econometrics and finance literature. Glosten et al. (1993) developed their asymmetric GARCH model as a generalization of the GARCH-M model. The GARCH-M process has an interesting moment structure. Assume that  $Ez_t^3 = 0$  and  $E(\epsilon_t)^4 < \infty$ . From (3.32) it follows that the k-th order auto covariance

$$E(y_t - E y_t)(y_{t-k} - E y_t) = E[\varepsilon_{t-k} g(h_t)] + \text{cov}(g(h_t), g(h_{t-k})) \neq 0$$

This means that there is forecastable structure in  $y_t$ ; which may contradict some economic theory if  $y_t$  is a return series. Hong (1991) showed this in a special case where  $g(h_t) = \delta h_t$ ,  $[E(\varepsilon)]_t^4 < \infty$ , and  $h_t$  follows a GARCH(p,q) model. In that situation, all autocorrelations of  $y_t$  are nonzero. Furthermore,

$$E[y_t - E(y_t)]^3 = 3E[g(h_t) - E(g(h_t))] + E[g(h_t) - E g(h_t)]^3 \neq 0(3.33)$$

It follows from (3.33) that a GARCH-M model implies postulating a skewed marginal distribution for  $y_t$  unless  $g(h_t) \equiv \text{constant}$ . For example, if  $g(h_t) = \delta h^{1/2}$ ,  $\delta < 0$ , this marginal distribution is negatively skewed. If the model builder is not prepared to make this assumption on the one of forecastable structure in  $y_t$ , the GARCH-M model, despite its theoretical motivation, does not seem an appropriate alternative to use. For more discussion of this situation, see He, Silvennoinen and Terasvirta (2006).

**Stylized facts and the first-order GARCH model** As already mentioned, financial time series such as high-frequency return series constitute the most common field of applications for GARCH models. These series typically display rather high kurtosis. At the same time, the autocorrelations of the absolute values or squares of the observations are low and decay slowly. These two features are sometimes called stylized facts of financial time series. Granger and Ding (1995) listed a few more such features. Among them is the empirical observation that in a remarkable number of financial series, the autocorrelations of the powers of observations,  $|\varepsilon_t|^k$ , peak around  $k = 1$ . Granger and Ding called this stylized fact the Taylor effect as Taylor (1986) was the first to draw attention to it (by comparing the autocorrelations of  $\varepsilon_t^2$  and  $|\varepsilon_t|^k$ ). One way of evaluating the adequacy of GARCH models is to ask how well they can be expected to capture the features or stylized facts present in the series to be modeled. The expressions for kurtosis and the autocorrelation function of absolute values and squared

observations are available for the purpose. They allow one to find out, for example, whether or not a GARCH(1,1) model is capable of producing realizations with high kurtosis and low, slowly decaying autocorrelations. The results of Malmsten and Terasvirta (2004) who have used these expressions, illustrate the well known fact, see, for example, Bollerslev, Engle and Nelson (1994), that a GARCH model with normally distributed errors does not seem to be a sufficiently flexible model for explaining these two features in financial return. Note that  $(\alpha_1 + \beta_1)$  is the decay rate of the autocorrelation function, that is, the  $j$ -th autocorrelation  $p_j = (\alpha_1 + \beta_1)^{j-1} p_1$  for  $j \geq 1$ . Increasing the "baseline kurtosis", that is, the kurtosis of the distribution of  $z_t$ ; the error, helps the GARCH(1,1) model to capture the stylized fact of high kurtosis/low autocorrelation. But then, this does not simultaneously affect the decay rate  $\alpha_1 + \beta_1$  of the autocorrelations. Recently, Kim and White (2004) suggested that the standard estimator of kurtosis exaggerates the true kurtosis and that robust measures yield more reliable results. It follows that high kurtosis values estimated from return series are a result of a limited number of outliers. If this is the case, then the use of a non-normal (heavy-tailed) error distribution may not necessarily be an optimal extension to the standard normal-error GARCH model. However, Terasvirta and Zhao (2006) recently studied 160 daily return series and, following Kim and White (2004), used robust kurtosis and autocorrelation estimates instead of standard ones. Their results indicate that leptokurtic distributions for  $z_t$  are needed in capturing the kurtosis-autocorrelation stylized fact even when the influence of extreme observations is dampened by the use of robust estimates. As to the Taylor effect, He and Terasvirta (1999a) defined a corresponding theoretical property, the Taylor property, as follows. Let  $\rho(|\varepsilon_t|^k, |\varepsilon_{t-j}|^k)$  be the  $j$ -th order autocorrelation of  $|\varepsilon_t|^k$ . The stochastic process has the Taylor property when  $\rho(|\varepsilon_t|^k, |\varepsilon_{t-j}|^k)$  is maximized for  $k = 1$  for  $j = 1, 2, \dots$ . In practice, He and Terasvirta (1999a) were able to find analytical results for the AVGARCH(1,1) model, but they were restricted to comparing the first-order autocorrelations for  $k = 1$  and  $k = 2$ . For this model,  $\rho(|\varepsilon_t|, |\varepsilon_{t-1}|) > \rho(|\varepsilon_t|^2, |\varepsilon_{t-1}|^2)$  when the kurtosis of the process is sufficiently high. The corresponding results for the standard GARCH(1,1) model (3.4) with  $p = q = 1$  and normal errors are not available as the autocorrelation function of  $|\varepsilon_t|^k$  cannot be derived analytically. Simulations conducted by He and Terasvirta (1999a) showed that the GARCH(1,1) model probably does not possess the Taylor property, which may seem disappointing. But then, the results of Terasvirta and Zhao (2006) show that if the standard kurtosis and autocorrelation estimates are replaced by robust ones, the evidence of the Taylor effect completely disappears. This stylized fact may thus be a consequence of just a small number of extreme observations in the series.

### 3.1.5 Family of Exponential GARCH models

**Definition and properties** The Exponential GARCH (EGARCH) model is another popular GARCH model. Nelson (1991) who introduced it had three criticisms of the standard GARCH model in mind. First, parameter restrictions are required to ensure positivity of the conditional variance at every point of time. Second, the standard GARCH model does not allow an asymmetric response to shocks. Third, if the model is an IGARCH one, measuring the persistence is difficult since this model is strongly but not weakly stationary. Shocks may be viewed persistent as the IGARCH process looks like a random walk. However, the IGARCH model with  $\alpha_0 > 0$  is strictly stationary and ergodic, and when  $\alpha_0 = 0$ , the realizations collapse to zero almost surely. The second drawback has since been removed as asymmetric GARCH models such as GJR-GARCH (Glosten et al. (1993)) or smooth transition GARCH have become available. A family of EGARCH( $p$ ;  $q$ ) models may be defined as in (3.2) with

$$\ln(h_t) = \alpha_0 + \sum_{j=1}^q (z_{t-j}) + \sum_{j=1}^p \beta_j \ln h_{t-j} \quad (3.34)$$

When  $g_j(z_{t-j}) = \alpha_j z_{t-j} + \psi_j(|z_{t-j}| - E|z_{t-j}|)$ ,  $j = 1, \dots, q$ , (3.34) becomes the EGARCH model of Nelson (1991). It is seen from (3.34) that no parameter restrictions are necessary to ensure positivity of  $h_t$ . Parameters  $\alpha_j$ ,  $j = 1, \dots, q$ , make an asymmetric response to shocks possible. When  $g_j(z_{t-j}) = \alpha_j \ln(z_{t-j}^2)$ ,  $j = 1, \dots, q$ , (3.2) and (3.34) form the logarithmic GARCH (LGARCH) model that Geweke (1986) and Pantula (1986) proposed. The LGARCH model has not become popular among practitioners. A principal reason for this may be that for parameter values encountered in practice, the theoretical values of the first few autocorrelations of  $\epsilon_t^2$  at short lags tend to be so high that such autocorrelations can hardly be found in financial series such as return series. This being the case, the LGARCH model cannot be expected to provide an acceptable fit when applied to financial series. Another reason are the occasional small values of  $\ln \epsilon_t^2$  that complicate the estimation of parameters. As in the standard GARCH case, the first-order model is the most popular EGARCH model in practice. Nelson (1991) derived existence conditions for moments of the general infinite-order Exponential ARCH model. The conditions imply that if the error process  $z_t$  has all moments and  $\sum_{j=1}^p \beta_j^2 < 1$  in (3.34) then all moments for the EGARCH process  $\epsilon_t^2$  exist. For example, if  $z_t$  is a sequence of independent standard normal variables then the restrictions on  $\beta_j$ ,  $j = 1, \dots, p$ , are necessary and sufficient for the existence of all moments simultaneously. This is different from the family (3.12) of GARCH models considered in Section 3.2. For those models, the moment conditions become more and more stringent for higher and higher even moments. The

expressions for moments of the first-order EGARCH process can be found in He, Terasvirta and Malmsten (2002); for the more general case, see He (2000).

**Stylized facts and the first-order EGARCH model** Before we considered the capability of first-order GARCH models to characterize certain stylized facts in financial time series. It is instructive to do the same for EGARCH models. For the first-order EGARCH model, the decay of autocorrelations of squared observations is faster than exponential in the beginning before it slows down towards an exponential rate; see He et al. (2002). Thus it does not appear possible to use a standard EGARCH(1,1) model to characterize processes with very slowly decaying autocorrelations. Malmsten and Terasvirta (2004) showed that the symmetric EGARCH(1,1) model with normal errors is not sufficiently flexible either for characterizing series with high kurtosis and slowly decaying autocorrelations. As in the standard GARCH case, assuming normal errors means that the first-order autocorrelation of squared observations increases quite rapidly as a function of kurtosis for any fixed  $\beta_1$  before the increase slows down. Analogously to GARCH, the observed kurtosis/autocorrelation combinations cannot be reached by the EGARCH(1,1) model with standard normal errors. The asymmetry parameter is unlikely to change things much. Nelson (1991) recommended the use of the so-called Generalized Error Distribution (GED) for the errors. It contains the normal distribution as a special case but also allows heavier tails than the ones in the normal distribution. Nelson (1991) also pointed out that a t-distribution for the errors may imply infinite unconditional variance for  $\varepsilon_t$ . As in the case of the GARCH(1,1) model, an error distribution with fatter tails than the normal one helps to increase the kurtosis and, at the same time, lower the autocorrelations of squared observations or absolute values. Because of analytical expressions of the autocorrelations for  $k > 0$  given in He et al. (2002) it is possible to study the existence of the Taylor property in EGARCH models. Using the formulas for the autocorrelations of  $|\varepsilon|^k$ ,  $k > 0$ , it is possible to find parameter combinations for which these autocorrelations peak in a neighborhood of  $k = 1$ . A subset of first-order EGARCH models thus has the Taylor property. This subset is also a relevant one in practice in the sense that it contains EGARCH processes with the kurtosis of the magnitude frequently found in financial time series. For more discussion on stylized facts and the EGARCH(1,1) model, see Malmsten and Terasvirta (2004).

### 3.1.6 Stochastic volatility

The EGARCH equation may be modified by replacing  $g_j(z_{t-j})$  by  $g_j(s_{t-j})$  where  $s_t$  is a sequence of continuous unobservable independent random variables that are often assumed

independent of  $z_t$  at all lags. Typically in applications,  $p = q = 1$  and  $g_1(s_{t-1}) = \delta s_{t-1}$  where  $\delta$  is a parameter to be estimated. This generalization is called the autoregressive stochastic volatility (SV) model, and it substantially increases the flexibility of the EGARCH parameterizations. For evidence of this, see Malmsten and Terasvirta (2004) and Carnero, Peña and Ruiz (2004). A disadvantage is that model evaluation becomes more complicated than that of EGARCH models because the estimation does not yield residuals.

### 3.1.7 Comparing EGARCH with GARCH

The standard GARCH model is probably the most frequently applied parameterization of conditional heteroscedasticity. This being the case, it is natural to evaluate an estimated EGARCH model by testing it against the corresponding GARCH model. Since the EGARCH model can characterize asymmetric responses to shocks, a GARCH model with the same property, such as the GJR-GARCH or the smooth transition GARCH model, would be a natural counterpart in such a comparison. If the aim of the comparison is to choose between these models, they may be compared by an appropriate model selection criterion as in Shephard (1996). Since the GJR-GARCH and the EGARCH model of the same order have equally many parameters, this amounts to comparing their maximized likelihoods. If the investigator has a preferred model or is just interested in knowing if there are significant differences in the fit between the two, the models may be tested against each other. The testing problem is a non-standard one because the two models do not nest each other. Several approaches have been suggested for this situation. Engle and Ng (1993) proposed combining the two models into an encompassing model. If the GARCH model is an GJR-GARCH( $p; q$ ) one (both models can account for asymmetries), this leads to the following specification of the conditional variance:

$$\begin{aligned} \ln_t = & \sum_{j=1}^q \alpha_j^* z_{t-j} + \psi_j^* (|z_{t-j}| - E|z_{t-j}|) + \sum_{j=1}^p \beta_j^* \ln h_{t-j} \\ & + \ln \left( \alpha_0 + \sum_{j=1}^q \alpha_j + \omega_j I(\varepsilon_{t-j}) \epsilon_{t-j}^2 + \sum_{j=1}^p \beta_j h_{t-j} \right) \end{aligned} \quad (3.35)$$

Setting  $(\alpha_j, \omega_j) = (0, 0), j = 1, \dots, q$ ; and  $\beta_j = 0, j = 1, \dots, p$ ; in (3.35) yields an EGARCH( $p; q$ ) model. Correspondingly, the restrictions  $(\alpha_j^*, \omega_j^*) = (0, 0); j = 1, \dots, q$ , and  $\beta_j^* = 0, j = 1, \dots, p$ , define the GJR-GARCH( $p; q$ ) model. Testing the models against each other amounts to testing the appropriate restrictions in (3.33) : A Lagrange Multiplier test may be constructed for the purpose. The test may also be viewed as another misspecification test and

not only as a test against the alternative model. Another way of testing the EGARCH model against GARCH consists of forming the likelihood ratio statistic despite the fact that the null model is not nested in the alternative. This is discussed in Lee and Brorsen (1997) and Kim, Shephard and Chib (1998). Let  $M_0$  be the EGARCH model and  $M_1$  the GARCH one, and let the corresponding log-likelihoods be  $L_T(\epsilon, M_0, \theta_0)$  and  $L_T(\epsilon, M_1, \theta_1)$ , respectively. The test statistic is

$$LR = 2L_T(\epsilon, M_1, \hat{\theta}_1) - L_T(\epsilon, M_1, \tilde{\theta}_1) \quad (3.36)$$

The asymptotic null distribution of (3.36) is unknown but can be approximated by simulation. Assuming that the EGARCH model is the null model and that  $\tilde{\theta}_1$  is the true parameter, one generates  $N$  realizations of  $T$  observations from this model and estimates both models and calculates the value of (3.36) using each realization. Ranking the  $N$  values gives an empirical distribution with which one compares the original value of (3.34) : The true value of  $\theta_0$  is unknown, but the approximation error due to the use of  $\tilde{\theta}_0$  as a replacement vanishes asymptotically as  $T \rightarrow \infty$ . If the value of (3.36) exceeds the  $100(1 - \alpha)$  % quantile of the empirical distribution, the null model is rejected at significance level  $\alpha$ . Note that the models under comparison need not have the same number of parameters, and the value of the statistic can also be negative. Reversing the roles of the models, one can test GARCH models against EGARCH ones. Chen and Kuan (2002) proposed yet another method based on the pseudo-score, whose estimator under the null hypothesis and assuming the customary regularity conditions is asymptotically normally distributed. This result forms the basis for a  $\chi^2$ -distributed test statistic; see Chen and Kuan (2002) for details. Results of small-sample simulations in Malmsten (2004) indicate that the pseudo-score test tends to be oversized. Furthermore, the Monte Carlo likelihood ratio statistic seems to have consistently higher power than the encompassing test, which suggests that the former rather than the latter should be applied in practice.

## 3.2 Empirical Evidence

**Data description** Crude Oil Futures trade in units of 1,000 U.S. barrels(42,000 gallons). The option is on one NYMEX light, sweet crude oil future contract. Crude oil futures trade 30 consecutive months plus long-dated futures initially listed 36,48,60,72 and 84 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:12 consecutive months, plus three long-dated options at 18,24, and 36 months out on a June/December

cycle. Minimum price fluctuation is 10 USD per contract. Maximum daily price fluctuation for futures initial limit of 3.00 USD per barrel are in place in all but the first two months an rise to 6.00 USD per barrel if the previous day's settlement in any back months is at the 3.00 USD limit. In the event of a 7.50 USD per barrel move in either of the first two contract months, limits on all months become 7.50 USD per barrel from the limit in place in the direction of the move following a one-hour trading halt. On options there are no price limits. 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 of the future contract. If the 25th calendar day of the month is the non-business day, trading shall cease on the third business day prior to the last business day preceding the 25th calendar day. Trading of options ends three business days before the underlying futures contract.

**Description of future contract** Open outcry trading is conducted from 10:00 AM until 2:30 PM New York time. Electronic trading is conducted from 6:00 PM until 5:15 PM via the CME Globex trading platform, Sunday through Friday. There is a 45 min break each day between 5:15PM (current trade date) and 6:00 PM(next trading day).

**Trading Months** The current year and the next five years. A new calendar year will be added following the termination of trading in the December contract of the current year. 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. Settlement type is physical.

**Delivery**

F.O.B.seller's facility. Cushing, Oklahoma, at any pipeline or storage facility with pipeline access to TEPPCO, Cushing storage, or Equilon Pipeline Co., by in-tank transfer, in-line transfer, book-out or inter-facility transfer(pump over). Complete delivery rules and provisions are detailed in Chapter 200 of the Exchange Rulebook. 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. An alternative 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. The commercial buyer or seller may exchange a future position for a physical position(EFP) 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 crude with 0.42 per cent

sulfur by weight or less, not less than 37° API gravity nor more than 42° 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° API nor more than 42° API. The following foreign streams are deliverable: UK Brent and Forties, for which the seller shall receive a 30 per cent per barrel discount below the final settlement price; Norwegian Oseberg Blend is delivered at 55 cent per barrel discount; Nigerian Bonny Light, Qua Lboe, and Colombian Cusiana are delivered at 15 cent premiums.

### 3.3 Convenience Yield under the GARCH Model

Consider a discrete time economy. Let  $\delta_t$ , convenience yield, to be defined as proposed by the formula above. We derive convenience yield from the weekly prices of next to expiration future contract and the following month to expiration futures trading on NYMEX during the period from January 1990 until March 2006. As a proxy for riskless rate we use 3 month forward T-bill rate quoted weekly. Next figure shows weekly observations of convenience yield time series.

We assume that convenience yield,  $cy_t$  follows a GARCH process.

$$cy_t = \mu + \varepsilon_t \quad (3.37)$$

$$\sigma_t^2 = K + G_1 \sigma_{t-1}^2 + A_1 \varepsilon_{t-1}^2 \quad (3.38)$$

However, we see that after estimation of this type of model standardized residuals have distribution which is far from White Noise and they exhibit high degree of autocorrelation. Closer look at the time series of convenience yield reveals that it has a long-term trend component. In order to get rid of this trend we will use a Hodrick-Prescott Filter. The Hodrick-Prescott Filter is a smoothing method that is widely used among macroeconomists to obtain a smooth estimate of the long term trend component of a series. Technically, the Hodrick-Prescott(HP) filter is a two-sided linear filter that computes the smoothed series  $s$  of  $y$  by minimizing the variance of  $y$  around  $s$ , subject to a penalty that constrains the second difference of  $s$ . That is, the HP filter chooses  $s$  to minimize:

$$\sum_{t=1}^T (y_t - s_t)^2 + \lambda \sum_{t=2}^{T-1} ((s_{t+1} - s_t - (s_t - s_{t-1}))^2) \quad (3.39)$$

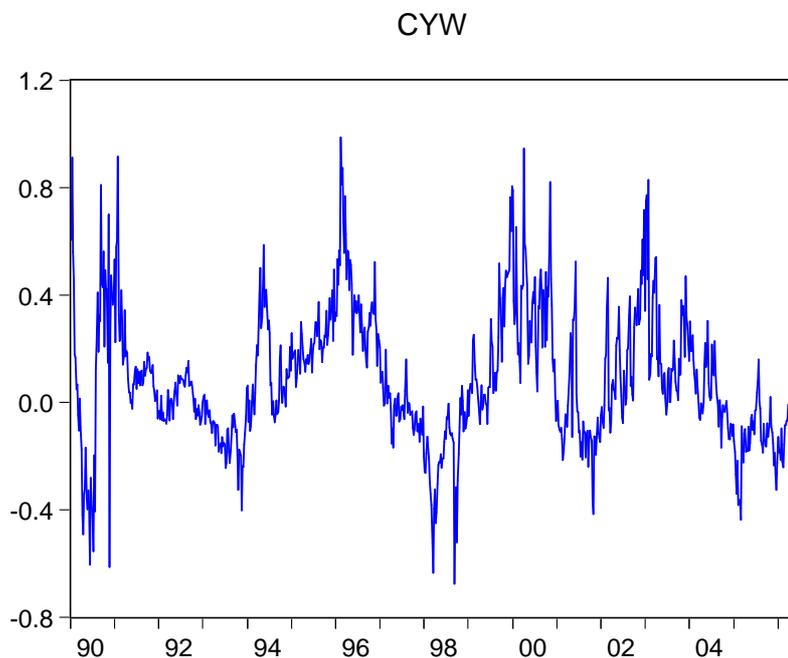


Figure 3.1: Convenience Yield for Crude Oil 1990-2006

The penalty parameter  $\lambda$  controls the smoothness of the series  $\sigma$ . The larger the  $\lambda$ , the smoother the  $\sigma$ . As  $\lambda \rightarrow \infty$ ,  $s$  approaches a linear trend.

The next chart shows detrended convenience yield dynamics. We can visually notice that this series is more appropriate for GARCH modeling than original convenience yield series.

After detrending convenience yield series we would like to test if GARCH modeling is an appropriate one. For this purpose we perform Engle's hypotheses test for the presence of ARCH/GARCH effects on the series  $dcy_t$ . Under the null hypothesis that a time series is a random sequence of Gaussian disturbances (i.e., no ARCH effects exist), this test statistic is asymptotically Chi-Square distributed. The result of this test indicates presence of ARCH/GARCH effects. We can also see that series has autoregressive component and according to this we will specify our mean equation with autoregressive component of order two. For the variance specification, we employ a GARCH(1,1) model. To model the thick tail in residuals, we will assume that the errors follow a Student's  $t$ -distribution. We come up with the following model:

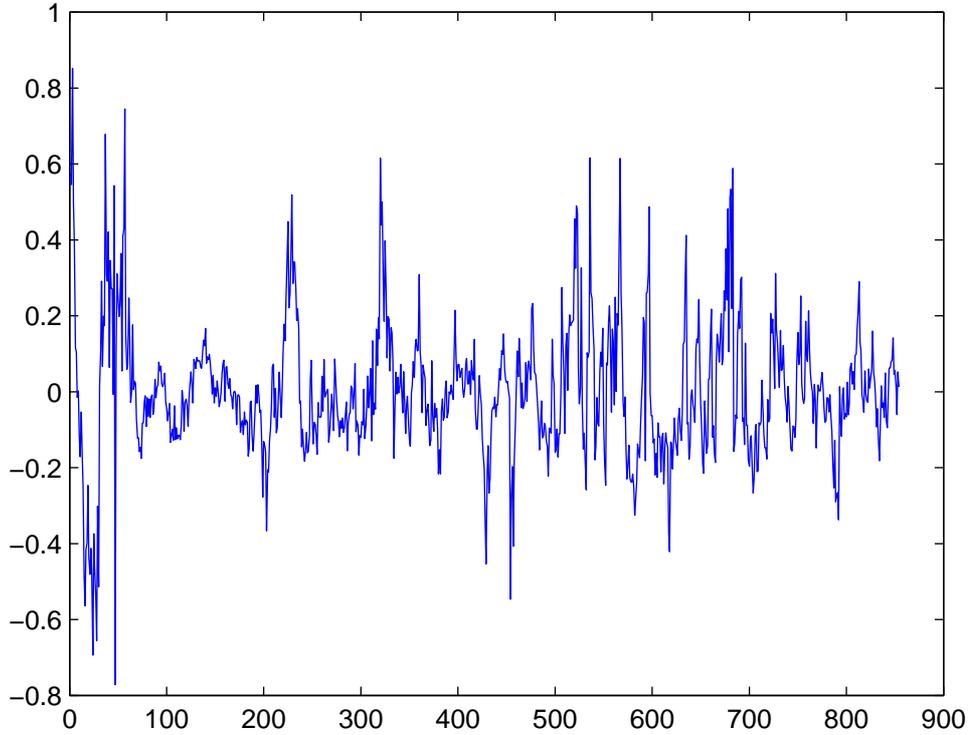


Figure 3.2: Detrended Convenience Yield for Crude Oil 1990-2006

$$cy_t = dcy_t + trend \quad (3.40)$$

$$dcy_t = 0.64488dcy_{t-1} + 0.13278dcy_{t-2} + \epsilon_t \quad (3.41)$$

$$\sigma_t^2 = 0.0005 + 0.1933\sigma_{t-1}^2 + 0.7864\epsilon_{t-1}^2 \quad (3.42)$$

Next we will plot quantiles of standardized residuals of our model against the quantiles of normal distribution. The plot indicates that it is primarily large negative shocks that are driving the departure from normality.

The plot above compares quantiles of empirical distributions of residuals obtained from estimation of GARCH models with quantiles of Normal distribution. It can be seen that empirical distribution has fatter tails, it is typical for financial data and support the fact that we chose to estimate our model with t-distribution assumption.

Plot of innovations show some volatility clustering but when we plot standardized residuals

Parameter	Value	SE	T Statistic
AR(1)	0.6449	0.0372	17.3342
AR(2)	0.1328	0.0384	3.4577
K	0.0005	0.0002	2.7289
GARCH(1)	0.78636	0.0398	19.7696
ARCH(1)	0.1933	0.0447	4.3226
DoF	4.7175	0.7047	6.6941

Table 3.1: Estimation of GARCH Model for Convenience Yield

(the innovations divided by their conditional standard deviation) they appear generally stable with little clustering.

When we plot the ACF of the squared standardized innovations, they also show no correlation.

Now we perform ARCH and Q-test on standardized innovation and in contradiction to the pre-estimation analysis we accept the null hypothesis that a time series is a random sequence of Gaussian disturbances (i.e., no ARCH effects exist). Therefore, our model has significant explanatory power.

We also perform forecasting and Monte Carlo simulations based on our GARCH model. Afterwards we compare between them and we can see that they predict very close values.

To see how the model might fit real data, we examine static forecasts for out-of-sample data. We estimate our model for the period 05/01/1990-14/3/2005. After it we make a forecast for the post-estimation period 14/3/2005-14/3/2006. Since the actual volatility of convenience yield is unobserved, we will use the squared detrended convenience yield series  $(dcy_t)^2$  as a proxy for the realized volatility. A plot of the proxy against the forecasted GARCH volatility of convenience yield provides an indication of the model's ability to track variations in market volatility.

It can be seen that GARCH volatility resembles well realized volatility.

We also perform rolling dynamic forecast of our model. We estimate our model on sub sample of 500 observations and perform a forecast of 50 observations out of sample. Next we move 50 observations and perform the same procedure again. For each forecast we estimate the error between forecasted and observed values. Next we calculate the root mean squared error. The RMSE is a quadratic scoring rule which measures the average magnitude of the error. Expressing the formula in words, the difference between forecast and corresponding observed values are each squared and then averaged over the sample. Finally, the square root of the average is taken. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. This means the RMSE is most useful when large

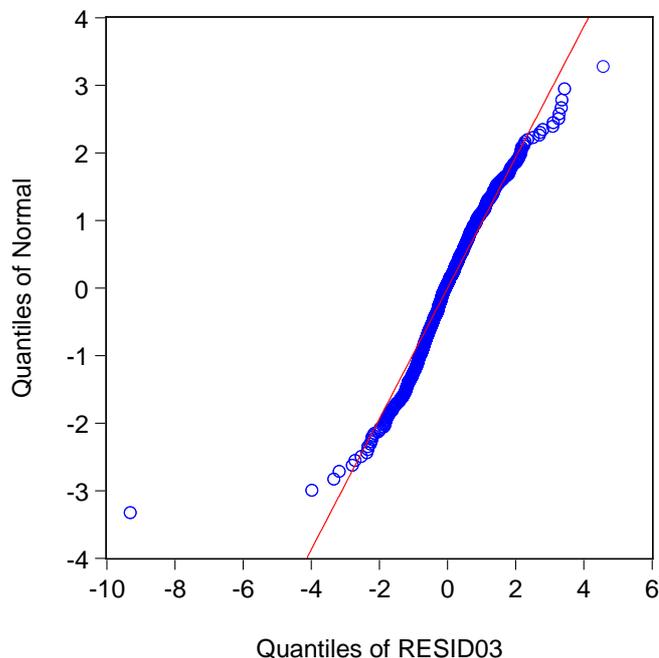


Figure 3.3: QQ plot

errors are particularly undesirable.

$$\text{RMSE}(\hat{\theta}) = \sqrt{E((\hat{\theta} - \theta)^2)}$$

From the box plot of RMSE we see that the average is 0.2 and it indicates that our model provides a good forecast.

The average of 0.2 means that our error is 20 % for the week. This is not so low value. However, one should bear in mind that we use close to expiration futures to carry our results. This is well known fact in oil risk management industry that longer term futures prices produce more accurate estimation.

The autocorrelation function of standardized residuals after estimation of GARCH model looks different from autocorrelation function estimated for squared returns before estimation. We see that autocorrelation is efficiently removed and it means that our model explained well the convenience yield dynamic.

Interesting to notice is the fact that convenience yield takes negative value significant amount of time(42.8 % of the time). Negative convenience yield means that one can make

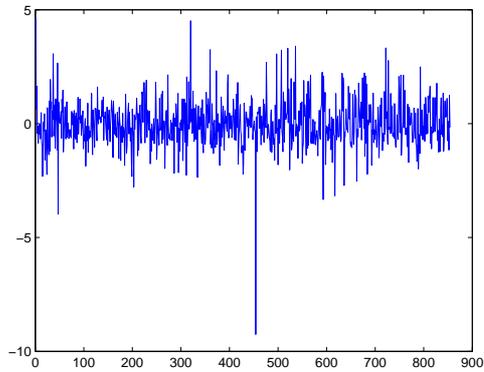


Figure 3.4: Standardized Residuals of GARCH Model

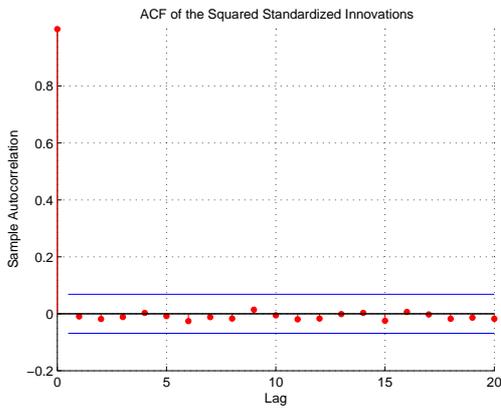


Figure 3.5: ACF of the Squared Standardized Innovations

profits simply by buying and storing oil and economic significance of this is that arbitrage opportunities exist. We doubt that this happens only for short period of time as we calculate convenience yield using next to expire and the following to next to expire future contracts. We repeat calculation of convenience yield using second and third contacts and the outcome is surprising-convenience yield is negative 51.64 % of the time - even larger proportion than in the case of closer to maturity future contracts.

From the residuals graph we can see that GARCH model provides a good fit for convenience yield dynamic. One can use this result in explaining and modeling of oil prices term structure. It can also be useful for modeling of price and inventory dynamics within petroleum industry.

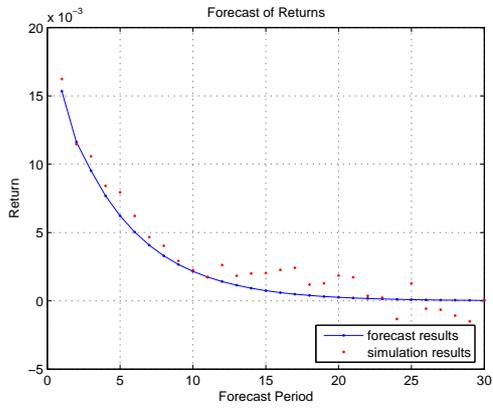


Figure 3.6: Forecast and Simulation Results

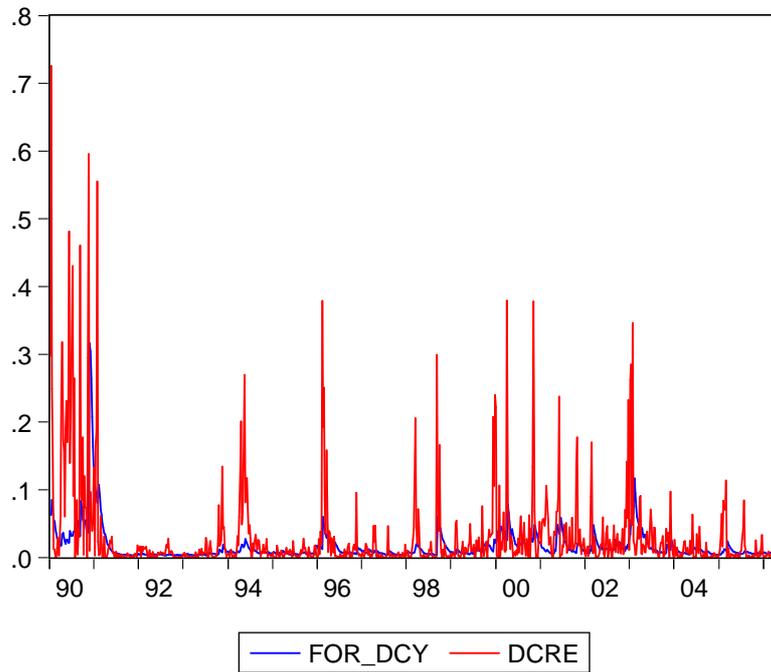


Figure 3.7: Realised and GARCH Volatility of Convenience Yield

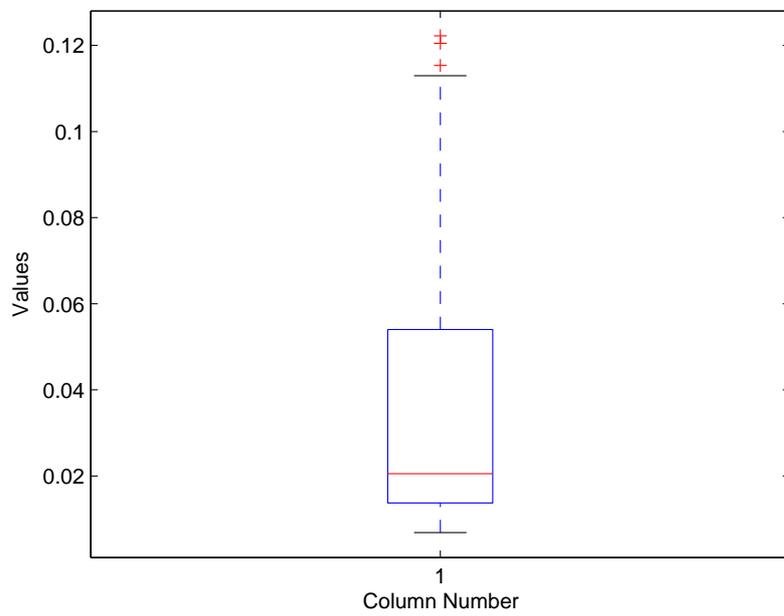


Figure 3.8: RMSE for GARCH model of Convenience Yield

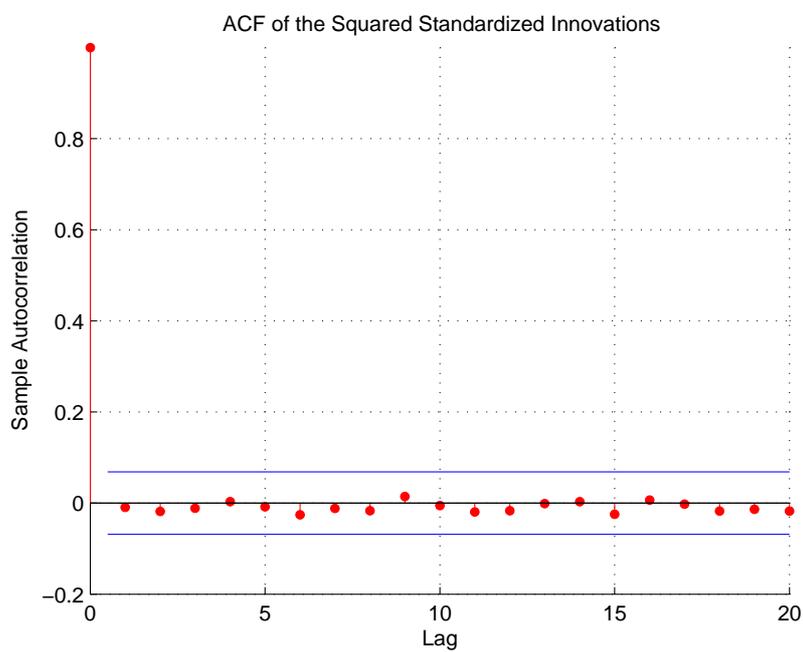


Figure 3.9: Autocorrelation Function of Standardized Residuals of GARCH model for convenience yield. Comparison of this ACF to the ACF estimated for squared detrended series of convenience yield shows that fitted model sufficiently explains the heteroscedasticity in the raw series.

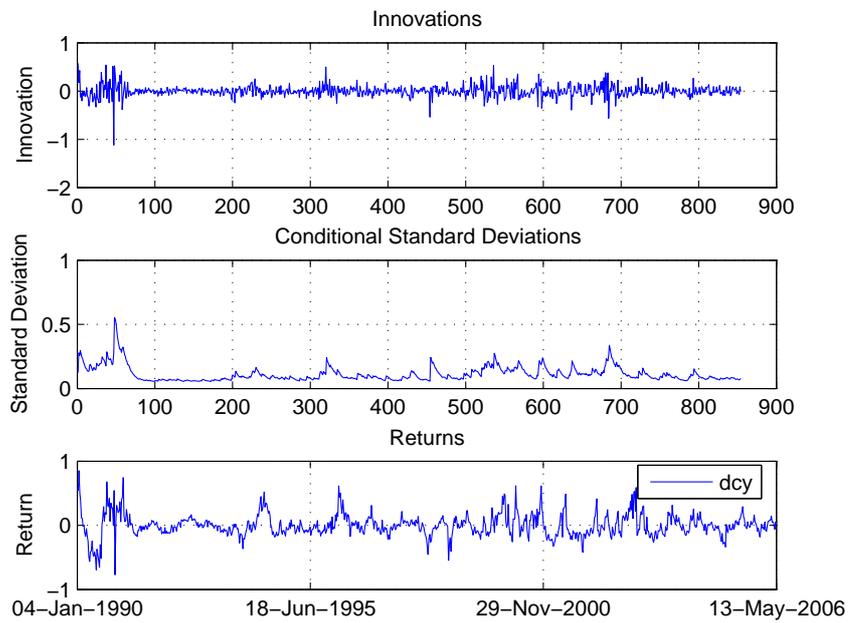


Figure 3.10: Relationship between the innovations (i.e., residuals) derived from the fitted model, the corresponding conditional standard deviations, and the observed detrended series of convenience yield.

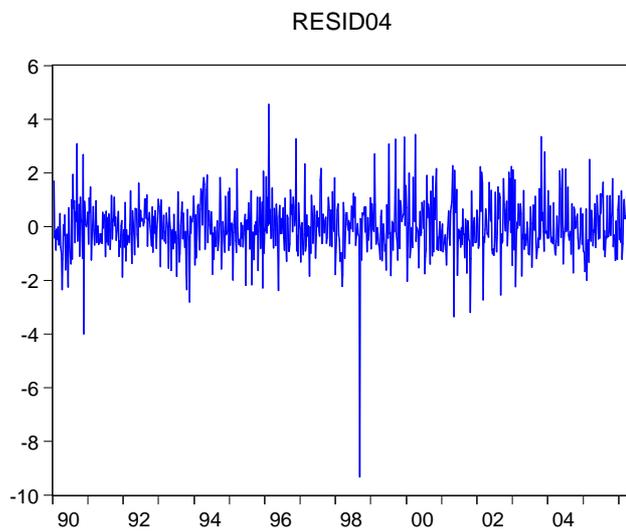


Figure 3.11: Standardized Residuals of GARCH model for detrended Convenience Yield. We can see that residuals distribution is close to normal, therefore our model provides good explanation of the data.

## Chapter 4

# A Multivariate GARCH Approach for Hedging in Commodity Markets

In recent years, oil markets have experienced sustained periods of extreme conditional volatility. Episodes of persistent market uncertainty are not unprecedented. Historically, prolonged sharp increases in energy prices have led to inflation and adverse economic performance for oil importing countries. Similarly, sharp decreases create serious budgetary problems for oil exporting countries. A number of studies have appeared that have addressed the efficiency of the oil futures market Crowder and Hamed (1993) and Peroni and McNown (1998). However, the literature to date does not provide any clear consensus. Moreover, no studies have appeared that examine market behavior during periods of sustained extreme conditional volatility or of the effectiveness of alternative hedging models during such periods. Lorne N. Switzer and Mario El-Khoury in their paper "Extreme Volatility, Speculative Efficiency, and the Hedge Effectiveness of the Oil Futures Markets" (2006) test the efficiency of the oil futures during periods of extreme conditional volatility. Using Fama (1984) regression approach with monthly data, a daily regression tests, as well as the Johansen (1988) co-integration techniques that are robust to varying error structures, they find that futures prices are unbiased predictors of future spot prices, consistent with the speculative efficiency hypothesis during the recent episodes of extreme volatility from the onset of the Iraqi war until the formation of the new Iraqi government. Most applications of time-varying models of hedging have imposed symmetry in the responses of volatility to positive or negative shocks. In a recent study, Brooks, Henry, and Persaud (2002) demonstrate that there are benefits to accounting for asymmetry in volatility in deriving optimal hedge ratios for equities with futures. The asymmetry that is typically found for equities associates negative price shocks with greater volatility increases

than positive price shocks, due to leverage effects (e.g., (Glosten, Jagannathan, and Runkle 1993). For oil, however, the asymmetry that they observe involves positive price shocks leading to greater volatility increases than negative shocks. The traditional approach to estimating the hedge ratio examines the unconditional covariance between cash and future prices relative to the unconditional variance of the future price. This ratio gives an estimate of the optimal proportion of the cash position that should be hedged using an opposite transaction in the futures market. The traditional method of estimating the hedge ration can be criticized on a number of grounds. (Myers and Thomson 1989) demonstrate that this approach provides an estimate of the optimal hedge ratio only under very restrictive assumptions. In particular, the expected return to holding a futures contract must be zero, and the processes generating the covariance matrix of cash and futures prices must be constant over time. If the expected return to the futures position is not zero, the traditional method of calculating the hedge ratio provides an estimate of the risk minimizing rather than the optimal hedge ratio.

This is clear if one considers traders who attempt to maximize the utility of a portfolio of cash and futures, where utility is given as a linear function of expected return( $R_t$ ) and variance  $V(R_t)$  according to (4.1), with  $\phi$  the degree of absolute risk aversion.

$$\text{Utility} = (R_t) - \phi V(R_t). \quad (4.1)$$

If  $R_t^c$  and  $R_t^f$  denote the expected returns on the cash and futures positions between two periods, respectively, and the hedge ratio  $h_{t-1}$  is defined as the proportion of the cash position to be offset in the futures market, the expected return on the portfolio is given by  $R_t = R_t^c - h_{t-1}R_t^f$ . The variance of this portfolio is  $V = \sigma_c^2 + h^2\sigma_f^2 - 2h\sigma_{cf}$ , where  $\sigma_c^2$ ,  $\sigma_f^2$ ,  $\sigma_{cf}$  are the variances of the cash position, future position, and their covariance respectively. Maximizing (4.1) with respect to  $h_{t-1}$  leads to an optimal value of  $h_{t-1}$  given by (4.2), whence it is clear that the risk-minimizing hedge ratio  $\phi$  coincides with the optimal hedge ratio only if the expected return to the futures contract is zero. Otherwise, the simple ratio of covariance to variance does not yield an estimate of the optimal proportion of the cash position to hedge using futures.

$$h_{t-1} = (\sigma_{cf}/\sigma_f^2) - (R_t^f/2\phi\sigma_f^2). \quad (4.2)$$

A vast volume of empirical work has concluded that the covariance matrices of many high frequency asset prices (including commodity prices) are conditionally heteroscedastic (Bollerslev and Kroner (1992)). This suggest that the variance and covariance terms in equation (4.2) should be indexed over time. Indeed, (Myers and Thomson 1989), report that many commodity prices are characterized by time-varying second moments, and that variants of the

GARCH model due to (Bollerslev 1986) adequately describe many covariance processes. This suggests that the optimal hedge ratio should be determined as a function of the conditional covariance matrix of cash and futures prices, since, as new information is revealed, it alters the optimal proportion of the cash position to be covered. Baillie and Myers (1991) found that the hedge ratios for six commodities were non stationary, and that GARCH based hedge ratios out-performed those obtained by the traditional regression approach. Similar results were reported by Myers(1991). The purpose of this section of my work is to extend Baillie and Myers(1991) methodology to crude oil. The object is to determine whether the GARCH modeling approach to estimating the optimal hedge ratio provides good fit to WTI data. The next section provides survey of multivariate GARCH models and choose the best model on the basis of estimation of our data. Following section will demonstrate that the series of spot and future prices are integrated of an identical order and that they are co-integrated, using maximum likelihood tests due to Johansen and Juselius (1990). Afterwards we will approach hedging problem with aid of bivariate Vector Autoregressive Model with GARCH variance structure. Cointegration between spot and future prices will allow us to reduce our estimation to one equation error correction model with GARCH structure and to derive optimal hedging ratios which will provide us with superior hedging strategies in comparison to other described here methods. Previous literature, for example Alexander and Dimitriu (2002), apply co-integration technique to portfolio managing and asset allocation. Lien (1996), Lien (2004) apply co-integration to deriving of optimal hedge ratio. (Schwartz 2006) applies co-integration to fund management, Sephon (1993).

## 4.1 Multivariate GARCH Models

It is now widely accepted that financial volatility move together over time across assets and markets. Recognizing this feature through a multivariate modeling framework leads to more relevant empirical models than working with separate univariate models. From a financial point of view, it opens the door to better decision tools in various areas, such as asset pricing portfolio selection, option pricing, hedging, and risk management. Indeed, unlike at the beginning of 1990s, several institutions have now developed the necessary skills to use the econometric theory in a financial perspective. In this section we would like to bring an overview of multivariate GARCH models as it was presented by Bauwens, Laurent, and Rombouts (2006). Since the seminal paper of Engle (1982), traditional time series tools such as autoregressive moving average(ARMA) models (Box (1970)) for mean have been extended to essentially analogous models for variance. The most obvious application of MGARCH (multivariate GARCH)

models is the study of the relations between the volatilities and co-volatilities of several markets. Does volatility of one market influence the volatility of other markets? Is the volatility of an asset transmitted to another asset directly (through its conditional variance) or indirectly (through its conditional covariances)? Does a shock in a market increase the volatility on another market and by how much? Is the impact the same for positive and negative shocks of the same amplitude? A related issue is whether the correlations between asset returns change over time (correlation between oil and gold, for example). Are they higher during periods of higher volatility (sometimes associated with periods of financial crises)? Are they increasing in the long run, perhaps because of globalisation of financial markets? Another application of MGARCH models is the computation of time-varying hedge ratio. Since a bivariate GARCH model for the spot and futures returns directly specifies their conditional variance-covariance matrix, the hedge ratio can be computed as a by-product of estimation and may be updated by using new observation as they become available. See Lien and Tse (2000) for a survey on hedging and additional references.

### Overview of models

Consider a vector stochastic process  $\{y_t\}$  of dimension  $N \times 1$ . As usual, we condition on the sigma field, denoted by  $I_{t-1}$ , generated by the past information (here the  $y_t$ 's) until time  $t - 1$ . Lets denote by  $\theta$  a finite vector of parameters and we write:

$$y_t = \mu_t(\theta) + \epsilon_t \quad (4.3)$$

where  $\mu_t(\theta)$  is the conditional mean vector and,

$$\epsilon_t = H_t^{1/2}(\theta)z_t \quad (4.4)$$

where  $H_t^{1/2}(\theta)$  is a  $N \times N$  positive definite matrix. Furthermore, assume the  $N \times 1$  random vector  $z_t$  to have the following first two moments:

$$E(z_t) = 0 \quad (4.5)$$

$$\text{Var}(z_t) = I_N \quad (4.6)$$

where  $I_N$  is the identity matrix of order  $N$ . We still have to explain what  $H_t^{1/2}$  is (for convenience we leave out  $\theta$  in the notation). To make this clear we calculate the conditional variance matrix of  $y_t$ :

$$\text{Var}(y_t | I_{t-1}) = \text{Var}_{t-1}(y_t) = \text{Var}_{t-1}(\epsilon_t) = H_t^{1/2}\text{Var}_{t-1}(z_t)(H_t^{1/2})' = H_t \quad (4.7)$$

Hence  $H_t^{1/2}$  is any  $N \times N$  positive definite matrix such that  $H_t$  is the conditional variance matrix of  $y_t$  (e.g.  $H_t^{1/2}$  may be obtained by the Cholesky factorization of  $H_t$ ). Both  $H_t$  and  $\mu_t$  depend on the unknown parameter vector  $\theta$ , which can in most cases be split into two disjoint parts, one for  $\mu_t$  and one for  $H_t$ . A case where this not true is that of GARCH-in-mean models, where  $\mu_t$  is functionally dependent on  $H_t$ . In this section, we make abstraction of the conditional mean vector for notational ease. It is usually specified as a function of the past, through a vectorial autoregressive moving average (VARMA) representation for the level of  $y_t$ . In the next subsections we review different specifications of  $H_t$ . They differ in various aspects. We identify three non mutually exclusive approaches for constructing multivariate GARCH models: (i) direct generalizations of the univariate GARCH model of Bollerslev (1986), (ii) linear combinations of univariate GARCH models, (iii) nonlinear combinations of univariate GARCH models. In the first category we have VEC, BEKK and factor models. Related models like the flexible MGARCH, Riskmetrics, Cholesky and full factor GARCH models are also in this category. In the second category we have (generalized) orthogonal models and latent factor models. The last category contains constant and dynamic conditional correlation models and the general dynamic covariance model. To keep the notational burden low, we present the models in their "g(1,1)" form rather than in their general "(p,q)" form.

#### 4.1.1 Generalizations of the Univariate Standard GARCH model

The models in this category are multivariate extensions of the univariate GARCH model. When we consider VARMA models for the conditional mean of several time series the number of parameters increases rapidly. Likewise, the same happens for multivariate GARCH models as straightforward extensions of the univariate GARCH model.

Furthermore, since  $H_t$  is a variance matrix, positive definiteness has to be ensured. To make the model tractable for applied purposes, additional structure may be imposed, for example, in the form of factors or diagonal parameter matrices. This class of models lends itself to relatively easy theoretical derivations of stationarity and ergodicity conditions, and unconditional moments (see e.g. He and Terasvirta, 2002a).

**Vec and BEKK models** A general formulation of  $H_t$  has been proposed by Bollerslev, Engle, and Wooldbrige(1988). In the general VEC model, each element of  $H_t$  is a linear function of the lagged squared errors and cross products of errors and lagged values of the elements of  $H_t$ .

**Definition 1:** *Vec model is defined as:*

$$h_t = c + A\eta_{t-1} + Gh_{t-1} \quad (4.8)$$

where

$$h_t = \text{vech}(H_t)\eta_t = \text{vech}(\epsilon_t\epsilon_t') \quad (4.9)$$

$A$  and  $G$  are square parameter matrices of order  $((N + 1)N/2)$  and  $c$  is a  $((N + 1)N/2) \times 1$  parameter vector.

The number of parameters is  $(N(N + 1)N(N + 1) + 1)/2$  (e.g. for  $N = 3$  it is equal to 78) which implies that in practice this model is used in bivariate case. To overcome the problem some simplifying assumptions have to be imposed. Bollerslev, Engle, and Wooldridge(1988) suggest the diagonal (DVEC) model in which the  $A$  and  $G$  matrices are assumed to be diagonal: each element  $h_t$  depending only on its own lag and on the previous value of  $\epsilon_{i,t}\epsilon_{j,t}$ . This restriction reduces the number of parameters to  $[N(N + 5)]/2$  (for  $N = 3$  it is equal to 12). But even under this diagonality assumption, large scale systems are still highly parameterized and difficult to work with in practice. The flexible MGARCH model of Ledoit, Santa-Clara, and Wolf(2003) circumvents the problem of estimating  $c$ ,  $A$  and  $G$  jointly for the DVEC model by estimating each variance and covariance equation separately. The resulting estimates do not necessarily guarantee positive semi-definite  $H_t's$ . Therefore, in a second step, the estimates are transformed in order to achieve the requirement, keeping the disruptive effects as small as possible. The transformed estimates are still consistent with respect to the parameters of the DVEC model. J.P. Morgan(1996) uses the exponentially weighted moving average model (EWMA) to forecast variance and covariances. Practitioners who study volatility processes often observe that their model is very close to the unit root case. To take this into account, Risk metrics defines the variances and covariances as IGARCH type models Engle (1982) and Bollerslev (1986):

$$h_{ij,t} = (1 - \lambda)\epsilon_{i,t-1}\epsilon_{j,t-1} + \lambda h_{ij,t-1} \quad (4.10)$$

In terms of the VEC model we have:

$$h_t = (1 - \lambda)\eta_{t-1} + \lambda h_{t-1} \quad (4.11)$$

which is a scalar VEC model. The decay factor  $\lambda$  proposed by Risk metrics is equal to 0.94 for daily data and 0.97 for monthly data. The decay factor is not estimated but suggested by Risk metrics. In this respect, this model is easy to work with in practice. However, imposing the same dynamics on every component in a multivariate GARCH model, no matter which data are used, is difficult to motivate. Because it is difficult to guarantee the positivity of

$H_t$  in the VEC representation without imposing strong restrictions on parameters, Engle and Kroner (1995) propose a new parametrization for  $H_t$  that easily imposes its positivity, i.e. the BEKK model (the acronym comes from synthesized work on multivariate models by Baba, Engle, Kraft and Kroner).

**Definition 2** *The BEKK(1,1,K) model is defined as:*

$$H_t = C^{*'}C^* + \sum_{k=1}^K A_k^{*'}\epsilon_{t-1}\epsilon_{t-1}'A_k^* + \sum_{k=1}^K G_k^{*'}H_{t-1}G_k^* \quad (4.12)$$

where  $C^*$ ,  $A_k^*$  and  $G_k^*$  are  $(N \times N)$  matrices but  $C^*$  is upper triangular.

The summation limit  $K$  determines the generality of the process. The parameters of the BEKK model do not represent directly the impact of the different lagged terms on the elements of  $H_t$ , like in the VEC model. The BEKK model is a special case of the VEC model. We refer to Engle and Kroner(1995) for propositions and proofs about VEC and BEKK models. The number of parameters in the BEKK(1,1,1) model is  $\{(N(5n + 1))/2\}$ . To reduce this number, and consequently to reduce the generality, one can impose a diagonal BEKK model, i.e.  $A_k^*$  and  $G_k^*$  are diagonal matrices.

Besides the BEKK model, another option to guarantee the positivity of  $H_t$  in the VEC representation is given in Kawakatsu(2003) who proposes the Cholesky factor GARCH model. Instead of specifying a functional form for  $H_t$ , he specifies a model on  $L_t$  where  $H_t = L_tL_t'$ . The advantage of this specification is that  $H_t$  is always positive definite without any restrictions on the parameters. The disadvantage is that identification restriction are needed which implies that the order of the series  $y_t$  is relevant and that the interpretation of the parameters is difficult. A similar model based on Cholesky decomposition can be found in Tsay (2002).

The difficulty when estimating a VEC or even a BEKK model is the high number of unknown parameters, even after imposing several restriction. It is thus not surprising that these models are rarely used when the number of series is larger than 3 or 4 Kearney and Patton (2000). Factor and orthogonal models circumvent this difficulty by imposing a common dynamic structure on all the elements of  $H_t$ , which results in less parameterized models.

**Factor Models** Engle (1990) proposes parametrization of  $H_t$  using the idea that co-movements of the stock returns are driven by a small number of common underlying variables, which are called factors. Bollerslev (1993) use this parametrization to model common persistence in conditional variances. The factor model can be seen as a particular BEKK model. We take the definition of (Lin 1992).

**Definition 3** The BEKK(1,1,K) model in definition 2 is a factor GARCH Model, denoted by F-GARCH(1,1,K), if for each  $k = 1, \dots, K$ ,  $A_k^*$  and  $G_k^*$  have rank one and have the same left and right elements,  $\lambda_k$  and  $w_k$ , i.e.

$$A_k^* = \alpha_k w_k \lambda_k', \text{ i.e. and } G_k^* = \beta_k w_k \lambda_k', \quad (4.13)$$

where  $\alpha_k$  and  $\beta_k$  are scalars, and  $\lambda_k$  and  $w_k$  (for  $k = 1, \dots, K$ ) are  $N \times 1$  vectors satisfying

$$w_k' \lambda_i = \begin{cases} 0 & \text{for } k \neq i \\ 1 & \text{for } k = i \end{cases} \quad (4.14)$$

$$\sum_{n=1}^N w_k w_n = 1 \quad (4.15)$$

If we substitute (4.14) and (4.15) into (4.13) we get

$$H_t = \Omega + \sum \lambda_k \lambda_k' (\alpha_k^2 w_k' \epsilon_{t-1} w_k + \beta_k^2 w_k' H_{t-1} w_k) \quad (4.16)$$

The restriction (4.16) is an identification restriction. Therefore, the K-factor GARCH model implies that the time varying part of  $H_t$  has reduced rank K, but  $H_t$  remains of full rank because of  $\Omega$ . Notice that  $\lambda_k$  and  $w_k'$  are also called the k-th factor loading and the k-th factor respectively. The k-th factor (denoted as the scalar  $f_{kt}$  hereafter) summarizes the information in the vector  $\epsilon_t$  using the vector of weights  $w_k$ . Note further that in (4.16) the first summation is over K rank one matrices and that the expression between brackets can be extended to handle more complex GARCH specifications. The number of parameters in the F-GARCH(1,1,1) is  $[N(n+5)]/2$ .

For example the conditional variance matrix of F-GARCH(1,1,2) model is:

$$H_t = \Omega + \lambda_1 \lambda_1' [\alpha_1^2 \omega_1' \epsilon_{t-1} \epsilon_{t-1}' \omega_1 + \beta_1^2 \omega_1' H_{t-1} \omega_1] + \lambda_2 \lambda_2' [\alpha_2^2 \omega_2' \epsilon_{t-1} \epsilon_{t-1}' \omega_2 + \beta_2^2 \omega_2' H_{t-1} \omega_2] \quad (4.17)$$

where the parameter vectors  $\lambda_k = (\lambda_{k1}, \lambda_{k2}, \dots, \lambda_{kN})'$  and  $\omega_k$  are of dimensions  $N \times 1$  while  $\alpha_k^2$  and  $\beta_k^2$  are scalar parameters. Denoting  $h_{kt} = \omega_k' H_t \omega_k$  we can write (4.17) in a more familiar way as:

$$h_{ijt} = \tau_{ij} + \lambda_{1i} \lambda_{1j} h_{1t} + \lambda_{2i} \lambda_{2j} h_{2t}, \text{ for } i, j = 1, \dots, N \quad (4.18)$$

$$h_{kt} = \omega_k + \alpha_k^2 f_{k,t-1}^2 + \beta_k^2 h_{k,t-1}; \quad k = 1, 2 \quad (4.19)$$

where  $\tau_{ij} = \omega_{ij} - \lambda_{1i}\lambda_{1j}\omega_1 - \lambda_{2i}\lambda_{2j}\omega_2$  and  $\omega_k = \omega'_k\Omega\omega_k$ . Hence  $h_{kt}$  is defined as a univariate GARCH(1,1) model. The persistence of the conditional variance in (4.17) is measured by  $\sum_{k=1}^2(\alpha_k^2 + \beta_k^2)$  and can also be interpreted as common persistence. In other words, the dynamics of the elements of  $H_t$  is the same. We can write  $H_t$  as:

$$H_t = \Omega^* + \lambda_1\lambda'_1 h_{1t} + \lambda_2\lambda'_2 h_{2t}, \quad (4.20)$$

where  $\Omega^* = \Omega - \lambda_1\lambda'_1\omega_1 - \lambda_2\lambda'_2\omega_2$ . Note that:  $E_{t-1}(f_{1t}f_{2t}) = \omega'_1\Omega\omega_2$  because  $\omega'_k\lambda_l = 0$  for  $k$  not equal, see (4.18). This implies that in the case of more than one factor we have the result that any pair of factors has a time-invariant conditional covariance. Note that alternatively, the two factor model described in (4.20) can be obtained from

$$\varepsilon_t = \lambda_1 f_{1t} + \lambda_2 f_{2t} + \epsilon_t \quad (4.21)$$

where  $\epsilon_t$  represents an idiosyncratic shock with constant variance matrix and uncorrelated with the two factors. Each factor  $f_{kt}$  has zero conditional mean and conditional variance like a GARCH(1,1) process, see (4.19). The K-factor model can be written as

$$\varepsilon_t = \lambda F_t + \epsilon_t \quad (4.22)$$

A factor is observable if it is specified as a function of  $\varepsilon_t$ , like in (4.17). Section (4.2.2) describes latent factor models. Several variants of the factor model are proposed in the literature. For example, Vrontos, Dellaportas, and Politis (2003) introduce the full-factor multivariate GARCH model.

**Definition 4.** *The FF-GARCH model is defined as*

$$H_t = W\Sigma_t W', \quad (4.23)$$

where  $W$  is a  $N \times N$  triangular parameter matrix with ones on the diagonal and

$$\Sigma_t = \text{diag}(\sigma_{1t}^2, \dots, \sigma_{Nt}^2)$$

where  $\sigma_{it}^2$  is the conditional variance of the  $i$ th factor, i.e the  $i$ th element of  $W^{-1}\epsilon_t$ , which can be separately defined as any univariate GARCH model. From the construction of the model,  $H_t$  is always positive definite. This model can also be interpreted as a dynamic conditional correlation model, see section (4.2.3). Note that  $H_t$  has a structure that depends on the ordering of the time series in  $y_t$ . Rigobon and Sack (2003) start from a structural model where

the conditional variances of the innovations are jointly specified. By deriving the reduced form model one obtains innovations with a conditional variance matrix that can be compared with other unrestricted reduced form MGARCH models. The structural model imposes a number of restrictions on the functional form of the conditional variance of the reduced form innovations resulting in less parameters than in a VEC model for example.

#### 4.1.2 Linear combinations of univariate GARCH models

In this category, we consider models that are linear combinations of several univariate models, each of which is not necessarily a standard GARCH models.

##### Orthogonal models

In the orthogonal GARCH model, the observed data are an orthogonal transformation of  $N$  (or a smaller number) of univariate GARCH processes. The matrix of the linear transformation is the orthogonal matrix (or a selection) of eigenvectors of the population unconditional covariance matrix of the standardized returns. In the generalized version, this matrix must only be invertible. The orthogonal models can also be considered as a factor models, where the factors are univariate GARCH-types processes with zero mean. In the orthogonal GARCH model of Kariya (1998), Alexander and Chibumba (1997) and Alexander (2001b), the  $N \times N$  time varying matrix  $H_t$  is generated by  $m$  univariate GARCH models.

**Definition 5.** The O-GARCH(1,1,m) model is defined as

$$V^{-1/2}\epsilon_t = u_t = \Lambda_m f_t, \quad (4.24)$$

where  $[-V] = \text{diag}(v_1, V_2, \dots, V_n)$ , with  $v_i$  the population variance of  $\epsilon_{it}$ , and  $[-\Lambda_m]$  is a matrix of dimension  $N \times m$  given by:

$$\Lambda_m = P_m \text{diag}(l_1^{1/2}, \dots, l_m^{1/2}), \quad (4.25)$$

$l_1 \geq \dots, l_m > 0$  being the  $m$  largest eigenvalues of population correlation matrix of  $u_t$ , and  $P_m$  the  $N \times m$  matrix of associated (mutually orthogonal) eigenvectors;  $-f_t = (f_{1t}, \dots, f_{mt})'$  is a random process such that:

$$E_{t-1}(f_t) = 0 \quad (4.26)$$

$$\text{Var}_{t-1}(f_t) = Q_t = \text{diag}(\sigma^2 f_{1t}, \dots, \sigma^2 f_{mt}), \quad (4.27)$$

$$\sigma^2(f_{it}) = (1 - \alpha_i - \beta_i) + \alpha_i f_{i,t-1}^2 + \beta_i f_{i,t-1}^2, \quad i = 1, \dots, m \quad (4.28)$$

Consequently,

$$H_t = \text{Var}_t - 1(\epsilon_t) = V^{1/2} V_t V^{1/2} \text{ where } V_t = \text{Var}_t - 1(u_t) = \Lambda_m Q_t \Lambda_m' \quad (4.29)$$

In practice, population parameters (such as  $V$  and  $\Lambda_m$ ) are replaced by their sample counterparts, and  $m$  is chosen by principal component analysis applied to the standardized residuals  $\hat{u}_t$ . Alexander(2001a) illustrates the use of the O-GARCH model. She emphasizes that using a small number of principal components compared to the number of assets is the strength of the approach (in one example she fixes  $m$  at 2 for 12 assets)

**Definition 6. The GO-GARCH(1,1) model** is defined as in definition 5, where  $m = N$  and  $\Lambda$  is a non singular matrix of parameters. The implied conditional correlation matrix of  $\epsilon_t$ , can be expressed as:

$$R_t = J_t^{-1} V_t J_t^{-1}, \text{ where } J_t = (V_t \odot I_m)^{1/2} \text{ and } V_t = \Lambda Q_t \Lambda' \quad (4.30)$$

In Van der Weide (2002), the singular value decomposition of the matrix  $\Lambda$  is used as a parametrization, i.e.  $\Lambda = PL1/2U$ , where the matrix  $U$  is orthogonal, and  $P$  and  $L$  are defined as above (from the eigenvectors and eigenvalues). The O-GARCH model (when  $m = N$ ) corresponds then to the particular choice  $U = IN$ . He expresses  $U$  as the product of  $N(N - 1)/2$  rotation matrices:

$$U = \prod G_{ij}(\delta_{ij}), \quad -\pi \leq \delta_{ij} \leq \pi \quad (4.31)$$

$i, j = 1, 2, \dots, n$ . where  $G_{ij}(\delta_{ij})$  performs a rotation in the plane spanned by the  $i$ -th and the  $j$ -th vectors of the canonical basis of  $\mathbb{R}^N$  over an angle  $\delta_{ij}$ .

For estimation, van der Weide (2002) replaces in a first step  $P$  and  $L$  by their sample counterparts and the remaining parameters (those of  $U$ ) are estimated together with the parameters of the GARCH factors in a second step. Note that such a two step estimation method is not applicable if a MGARCH-in-mean effect is included. More generally, as pointed out by a referee, the matrix  $\Lambda$  as such could be estimated together with the GARCH parameters of the factors, in a single step. The orthogonal models are nested in the F-GARCH model and thus in the BEKK model. As a consequence, the results about stationarity, unconditional moments and estimation available for the BEKK model can be applied to the factor models. In particular, it is obvious that the (G)O-GARCH model is covariance-stationary if the  $m$

univariate GARCH processes are themselves stationary. Note that (G)O-GARCH models can be extended easily by considering other univariate GARCH models than the usual specification in (4.28).

### Latent factor models

The factor model in (4.22) becomes a latent model if  $F_t$  is latent which means that it is not included in  $I_t$  implying that the conditional variance matrix, see for example (4.20), is not measurable anymore. This is in contrast with orthogonal models where the conditional variance of the factors are specified as a function of the past data ( $\varepsilon_t$ ). Therefore, latent factor models can be classified as stochastic volatility models as mentioned in Shephard (1996). The elements of  $F_t$  typically follow dynamic heteroscedastic processes, Diebold and Nerlove (1989) for example use ARCH models. The fact that the factor is considered as non-observable complicates inference considerably since the likelihood function must be marginalized with respect to it (see Gouriéroux, 1997, Section 6.3). The conditional covariance between the factors is usually assumed to be equal to zero. See Sentana and Fiorentini (2001) for more details on identification and estimation of factor models. Sentana (1998) shows that the observed factor model is observationally equivalent (up to conditional second moments) to a class of conditionally heteroscedastic factor models including latent factor models. Doz and Renault (2003) elaborate on this result and draw the conclusions in terms of model specification and identification, and in terms of inference methodologies. The multivariate latent factor model is used in several applications. Diebold and Nerlove (1989) study the dynamics of exchange rate volatility patterns in seven nominal dollar spot exchange rates by the use of a multivariate latent one-factor model. King, Sentana, and Wadhvani (1994), generalizing the work of Diebold and Nerlove (1989), develop a multivariate factor model with observable and latent factors to assess the extent of capital market integration on sixteen national stock markets. Dungey, Martin, and Pagan (2000) analyze bond yield spreads between five countries by decomposing international interest rate spreads into national and global latent factors.

#### 4.1.3 Nonlinear combinations of Univariate GARCH Models

This section collects models that may be viewed as nonlinear combinations of univariate GARCH models. This allows for models where one can specify separately, on the one hand, the individual conditional variances, and on the other hand, the conditional correlation matrix or another measure of dependence between the individual series (the copula of the conditional joint density). For models of this category theoretical results on stationarity, ergodicity and

moments may not be so straightforward to obtain as for models presented in the preceding sections. Nevertheless, they are less greedy in parameters than the models of the first category, and therefore they are more easily estimable.

### Conditional correlation models

The conditional variance matrix for this class of models is specified in a hierarchical way. First, one chooses a model for each conditional variance. For example, some conditional variances may follow a conventional GARCH model while others may be described as an EGARCH model (Nelson, 1991) or an APARCH model (Ding, Granger, and Engle, 1993). Second, based on the conditional variances one models the conditional correlation matrix (imposing its positive definiteness  $\forall t$ ). Bollerslev (1990) proposes a class of MGARCH models in which the conditional correlations are constant and thus the conditional covariances are proportional to the product of the corresponding conditional standard deviations. This restriction highly reduces the number of unknown parameters and thus simplifies estimation.

**Definition 7** *The CCC model* is defined as:

$$H_t = D_t R D_t = (\rho_{ij} \sqrt{h_{ijt} h_{ijt}}) \quad (4.32)$$

where

$$D_t = \text{diag}(h_{11t}^{1/2} \dots h_{NNt}^{1/2}) \quad (4.33)$$

$h_{iit}$  can be defined as any univariate GARCH model, and  $R = (\rho_{ij})$  is a symmetric positive definite matrix with  $\rho_{ii} = 1, \forall i$ .

$R$  is the matrix containing the constant conditional correlations  $\rho_{ij}$ . The original DCC model has a GARCH(1,1) equation for each conditional variance in  $D_t$ :

$$H_{iit} = w_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i h_{iit,t-1}, \quad i = 1, \dots, N \quad (4.34)$$

This CCC model contains  $[N(N + 5)]/2$  parameters.  $H_t$  is positive definite if and only if all the  $N$  conditional variances are positive and  $R$  is positive definite. The unconditional variances are easily obtained, as in the univariate case, but the unconditional covariances are difficult to calculate because of the nonlinearity in (4.32). He and Terasvirta (2002b) use a VEC-type formulation for  $(h_{11t}, h_{22t}, \dots, h_{NNt})'$ , to allow for interactions between the conditional variances. They call this the extended CCC model. The assumption that the conditional correlations are constant may seem unrealistic in many empirical applications. Christodoulakis and Satchell (2002), Engle (2002) and Tse and Tsui (2002) propose a general-

ization of the CCC model by making the conditional correlation matrix time dependent. The model is then called a dynamic conditional correlation (DCC) model. An additional difficulty is that the time dependent conditional correlation matrix has to be positive definite for all  $t$ . The DCC models guarantee this under simple conditions on the parameters. The DCC model of Christodoulakis and Satchell (2002) uses the Fisher transformation of the correlation coefficient. The specification of the correlation coefficient is  $\rho_{2r_t} = (e^{2r_t} - 1)/(e^{2r_t} + 1)$  where  $r_t$  can be defined as any GARCH model using  $\epsilon_{1t}\epsilon_{2t}/\sqrt{h_{11t}h_{22t}}$  as innovation. This model is easy to implement because the positive definiteness of the conditional correlation matrix is guaranteed by the transformation. However, it is only a bivariate model. The DCC models of Tse and Tsui (2002) and Engle (2002) are genuinely multivariate and are useful when modeling high dimensional data sets.

**Definition 8** *The DCC model of Tse and Tsui (2002) or DCCT* (M) is defined as:

$$H_t = D_t R_t D_t, \quad (4.35)$$

where  $D_t$  is defined in (4.33),  $h_{iit}$  can be defined as any univariate GARCH model, and

$$R_t = (1 - \theta_1 - \theta_2)R + \theta_1\psi_{t-1} + \theta_2R_{t-1} \quad (4.36)$$

$\theta_1$  and  $\theta_2$  are non-negative parameters satisfying  $\theta_1 + \theta_2 < 1$ ,  $R$  is a symmetric  $N \times N$  positive definite parameter matrix with  $\rho_{ii} = 1$ , and  $\psi_{t-1}$  is the  $N \times N$  correlation matrix of  $\epsilon_\tau$  for  $\tau = t - M, t - M + 1, \dots, t - 1$ . Its  $i, j$ -th element is given by:

where  $u_{it} = \epsilon_{it}/\sqrt{h_{it}}$ . The matrix  $\psi_{t-1}$  can be expressed as:

$$\psi_{t-1} = B_{t-1}^{-1} L_{t-1} L'_{t-1} B_{t-1}^{-1}, \quad (4.37)$$

with  $B_{t-1}$  a  $N \times N$  diagonal matrix with  $i$ -th diagonal element being  $(\sum_{h=1}^M u_{i,t-h}^2)^{1/2}$  and  $L_{t-1} = (u_{t-1}, \dots, u_{t-M})$ , an  $N \times M$  matrix.

**Definition 9** *The DCC model of Engle(2002) or DCC<sub>E</sub>(1, 1)* is defined as:

$$H_t = D_t R_t D_t \quad (4.38)$$

where  $D_t$  is defined as before,  $h_{iit}$  can be defined as any univariate GARCH model,

$$R_t = (\text{diag}Q_t)^{-1/2} Q_t (\text{diag}Q_t)^{-1/2}. \quad (4.39)$$

and where the  $N \times N$  symmetric positive definite matrix  $Q_t$  is given by:

$$Q_t = (1 - \alpha - \beta)\bar{Q} + \alpha u_{t-1} u'_{t-1} + \beta Q_{t-1}, \quad (4.40)$$

with  $u_{it} = \epsilon_{it}/\sqrt{H_{it}}$ ,  $Q$  is the  $N \times N$  unconditional variance matrix of  $u_t$ , and  $\alpha (\geq 0)$  and  $\beta (\geq 0)$  are scalar parameters satisfying  $\alpha + \beta < 1$ . The elements of  $Q$  can be estimated or alternatively set to their empirical counterpart to render the estimation even simpler.

Unlike in the  $DCC_T$  model, the  $DCC_E$  model does not model the conditional correlation as a weighted sum of past correlations. Indeed, the matrix  $Q_t$  is written like a GARCH equation, and then transformed to a correlation matrix. A drawback of the DCC models is that  $\theta_1, \theta_2$  in  $DCC_T$  and  $\alpha, \beta$  in  $DCC_E$  are scalars, or in other words, all the conditional correlations obey the same dynamics. This is necessary to ensure that  $R_t$  is positive definite for all  $t$  through sufficient conditions on the parameters. If the conditional variances are specified as GARCH(1,1) models then the  $DCC_T$  and the  $DCC_E$  models contain  $[(N + 1)(N + 4)]/2$  parameters.

**General dynamic covariance model** A model somewhat different from the previous ones but that nests several of them is the general dynamic covariance (GDC) model proposed by Kroner and Ng (1998). They illustrate that the choice of a multivariate volatility model can lead to substantially different conclusions in an application that involves forecasting dynamic variance matrices. The definition of Kroner and Ng (1998) is extended here to cover the models of dynamic conditional correlations.

**Definition 10** *The GDC model is defined as:*

$$H_t = D_t R_t D_t + \Phi \odot \Theta_t. \quad (4.41)$$

where  $D_t = (d_{ijt})$ ,  $d_{iit} = \sqrt{\theta_{iit}} \forall i, d_{ijt} = 0 \forall i \neq j$ ,  $\Theta_t = (\theta_{ijt})$   $R_t$  is specified as  $DCC_T(M)$ , or as  $DCC_E(1, 1)$ ,

$\Phi = (\phi_{ijt}), \phi_{ii} = 0 \forall i, \phi_{ij} = \phi_{ji}$   $\theta_{ijt} = \omega_{ij} + \alpha'_i \epsilon_{t-1} \epsilon'_{t-1} \alpha_j + g'_i H_{t-1} g_j \forall i, j$   $\alpha_i, \beta_i, i = 1, \dots, N$  are  $(N \times 1)$  vector of parameters, and  $\Omega = (\omega_{ij})$  is positive definite and symmetric.

Element wise we have:

$$h_{iit} = \theta_{iit} \quad (4.42)$$

$$h_{ijt} = \rho_{ijt} \sqrt{\theta_{iit} \theta_{jjt}} + \phi_{ij} \theta_{ijt}, \forall i \neq j \quad (4.43)$$

where  $\theta_{ijt}$  are given by the BEKK formulation. The GDC model contains several MGARCH models as special case. To show this we adapt a proposition from Kroner and Ng (1998).

Consider the following set of conditions:

1. **a.**  $\theta_1 = \theta_2 = 0(DCC_T)$  or  $\alpha = \beta = 0(DCC_E)$ .  
**b.**  $R = I_N(DCC_T)$  or  $\bar{Q} = I_N(DCC_E)$ .
2.  $\alpha_i = \alpha_i l_i$  and  $g_i = \beta_i l_i \quad \forall i$ , where  $l_i$  is the  $i$ th column of an  $(N \times N)$  identity matrix, and  $\alpha_i$  and  $\beta_i, i = 1, \dots, N$  are scalars.
3.  $\phi_{ij} = 0 \quad \forall i \neq j$ .
4.  $\phi_{ij} = 1 \quad \forall i \neq j$ .
5.  $A = \alpha(\omega\lambda')$  and  $G = \beta(\omega\lambda')$  where  $A = [a_1, \dots, a_N]$  and  $G = [g_1, \dots, g_N]$  are  $N \times N$  matrices,  $\omega$  and  $\lambda$  are  $N \times 1$  vectors, and  $\alpha$  and  $\beta$  are scalars.

The GDC model reduces to different multivariate GARCH models under different combinations of these conditions. Specifically, the GDC model becomes:

- the DCCT or the DCCE(1, 1) model with GARCH(1,1) conditional variances under conditions (ii) and (iii)
- the CCC model with GARCH(1,1) conditional variances under conditions (1a), (2) and (3),
- a restricted DVEC(1,1) model under conditions (1) and (2),
- the BEKK(1,1,1) model under conditions (1) and (4),
- the F-GARCH(1,1,1) model under conditions (1), (4) and (5).

Condition (1b) serves as an identification restriction for the VEC, BEKK and F-GARCH models. As we can see, the GDC model is an encompassing model. This requires a large number of parameters (i.e.  $[N(7N - 1) + 4]/2$ ). For example in the bivariate case there are 11 parameters in  $\Theta_t$ , 3 in  $R_t$  and 1 in  $\Phi$  which makes a total of 15. This is less than for an unrestricted *VEC* model (21 parameters), but more than for the BEKK model (11 parameters).

### Copula MGARCH models

Another approach for modeling the conditional dependencies is known as the copula-GARCH model. This approach makes use of the theorem due to Sklar (1959) stating that any  $N$  dimensional joint distribution function may be decomposed into its  $N$  marginal distributions, and a

copula function that completely describes the dependence between the  $N$  variables. See Nelsen (1999) for a comprehensive introduction to copulas. (2000) and Jondeau and Rockinger (2001) have proposed copula-GARCH models. These models are specified by GARCH equations for the conditional variances (possibly with each variance depending on the lag of the other variance and of the other shock), marginal distributions for each series (e.g.  $t$  distributions), and a conditional copula function. Both papers highlight the need to allow for time-variation in the conditional copula, extending in some sense the DCC models to other specifications of the conditional dependence. The copula function is rendered time-varying through its parameters, which can be functions of past data. In this respect, like the DCC model of Engle (2002), copula GARCH models are estimated using a two-step maximum likelihood approach which solves the dimensionality problem. An interesting feature of copula-GARCH models is the ease with which very flexible joint distributions may be obtained in a bivariate framework. Their extension to a higher dimension is however somewhat tricky and a challenge for the future.

### Leverage effects in MGARCH models

For stock returns, negative shocks may have a larger impact on their volatility than positive shocks of the same absolute value (this is most often interpreted as the leverage effect unveiled by Black, 1976). In other words, the news impact curve, which traces the relation between volatility and the previous shock, is asymmetric. Univariate models that allow for this effect are the EGARCH model of Nelson (1991), the GJR model of Glosten, Jagannathan, and Runkle (1993), and the threshold ARCH model of Zakoian (1994), among others. For multivariate series the same argument applies: the variances and the covariances may react differently to a positive than to a negative shock. In the multivariate case, a shock can be defined in terms of  $\epsilon_t$  or of  $z_t$ . Note that the signs of  $\epsilon_{it}$  and  $z_{it}$  do not necessarily coincide, see (2).

The MGARCH models reviewed in the previous subsections define the conditional variance matrix as a function of lagged values of  $\epsilon_t \epsilon_t'$ . For example, each conditional variance in the VEC model is a function of its own squared error but it is also a function of the squared errors of the other series as well as the cross-products of errors. A model that takes explicitly the sign of the errors into account is the asymmetric dynamic covariance (ADC) model of Kroner and Ng (1998). The only difference with Definition 10 is an extra term based on the vector  $v_t = \max[0, -\epsilon_t]$  in  $\theta_{ijt}$  to take into account the sign of  $\epsilon_{it}$ :

$$\theta_{ijt} = \omega_{ij} + \alpha_i' \epsilon_{t-1} \epsilon_{t-1}' \alpha_j + g_i' H_{t-1} g_j + b_i' H_{t-1} g_j + b_i' v_{t-1} b_j \quad (4.44)$$

The ADC model nests some natural extensions of MGARCH models that incorporate the leverage effect. Kroner and Ng (1998) apply the model to large and small firm returns. They find that bad news about large firms can cause additional volatility in both small-firm and large-firm returns. Furthermore, this bad news increases the conditional covariance. Small firm news has only minimal effects. Hansson and Hordahl (1998) add the term  $D \odot v_{t-1} v'_{t-1}$ , in a DVEC model like (8), where  $D$  is a diagonal matrix of parameters. To incorporate the leverage effect in the (bivariate) BEKK model, Hafner and Herwartz (1998) add the terms  $D'_1 \epsilon_{t-1} \epsilon'_{t-1} D_1 1_{\epsilon_{1,t-1} < 0}$ , where  $D_1$  and  $D_2$  are  $2 \times 2$  matrices of parameters and  $1_{\{\dots\}}$  is the indicator function. This generalizes the univariate GJR specification.

### Transformations of MGARCH models

Not all MGARCH models are invariant with respect to linear transformations. By invariance of the model we mean that it stays in the same class if a linear transformation is applied to  $y_t$ , say  $\tilde{y}_t = F y_t$ , where  $F$  is a matrix of constants (for simplicity we assume  $F$  is square). If  $y_t$  is a vector of returns, a linear transformation corresponds to new assets (portfolios combining the original assets). It seems sensible that a model should be invariant, otherwise the question arises which basic assets should be modeled. In some cases (stocks), these are naturally defined, in other cases, like exchange rates, they are not, since a reference currency must be chosen (see Gouriéroux and Jasiak, 2001, p 140). Lack of invariance of a model does not imply that the model is not suitable at all for use in empirical work. Implications of invariance, or lack of invariance, are an open issue. For example, if the model is invariant, one can estimate it with some number of basic assets, as well as with a smaller number of portfolios of the basic assets. Estimates of the larger model imply estimates of the smaller models, which could be compared to the direct estimates of the latter. Very different estimates may lead to question the specification. Lack of invariance occurs whenever a diagonal matrix in the equation defining  $H_t$  is pre-multiplied by the matrix  $F$  defined above. The general VEC and BEKK models are invariant, but their diagonal versions are not. Conditional correlation models are not invariant, since  $F D_t$  is not diagonal when  $D_t$  is diagonal (4.33). Another question is that of temporal aggregation of MGARCH processes. Hafner (2004) shows that, like Drost and Nijman (1993) in the univariate case, the class of weak multivariate GARCH processes is closed under temporal aggregation. Weak multivariate GARCH models are characterized by a weak VARMA structure of  $\eta_t$  in (4.11). Fourth moment characteristics turn out to be crucial for deriving the low frequency dynamics. The issue of estimation of the parameters of the low frequency model is difficult because the probability law of the innovation vector is unknown, since it is only assumed to be a weak white noise. See Hafner and Rombouts (2003) for more

details.

## 4.2 Choosing the Best Model

In this part of my work I would like to analyze performance of different GARCH multivariate models for hedging in crude oil market. Our data set consists of daily spot and near to expiration future prices of crude oil for the period of 01/1990-03/2006. We switch every 15-th of each month between next to expire and second to expire future prices in order to avoid Samuelson effect. We also build time series of constant 30 days until expiration future prices by interpolating between prices of 4 daily available prices of futures contract with different maturities. In this way we will be able to compare between two sets of results. We obtain 4260 observations.

Since ARCH family of models requires stationary series, tests for the presence of unit roots in the logarithms of the data were undertaken. Philips(1987) and Philips and Perron(1988) unit root tests indicate that at the 5 per cent level of significance, all series are first difference stationary. Tests based on different lag length lead to the same conclusions. Further testing for the presence of a second unit root did not identify evidence that series were  $I(2)$  at a 5 per cent level of significance. For the sake of brevity these results will not be reported.

Engle and Granger(1987) argued that multivariate VAR models in first-difference form are valid only if the series under study are integrated of identical order and they are not co-integrated. Since the logarithms of the oil prices are  $I(1)$ , their rates of change will be modeled in the construction of the hedge ratios. A multivariate model in this form can be interpreted as simple VAR model. We use Johansen and Juselius (1990) test for co-integration and find evidence of a Co-integrating relationship between the logarithms of cash and future price series at a 5 per cent level of significance.

### 4.2.1 Estimation of Univariate Models

From an examination of the time paths of the first differences of the logarithms of the price data, there seem to be periods during which there is above average volatility in the approximate rates of inflation, GARCH models can be used to explain this time dependent heteroscedasticity. Bollerslev (1986) proposed a generalized autoregressive conditional heteroscedastic model in which the error variance, conditional on available information as of time  $t-1$  (denoted by  $\Omega_{t-1}$ ) is allowed to evolve according to its own history and the history of the residuals. Under the assumption of  $t$ -distribution, the GARCH(1,1) model contained in equation was estimated for variable  $X$ .

$$100\Delta(X_t) = \mu + \epsilon_t \epsilon_t / \Omega_{t-1} N(0, h_t) h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (4.45)$$

Parameter	Value	SE	T statistic
$\mu$	0.044981	0.028063	1.6029
$\alpha_0$	0.063262	0.015730	4.0217
$\alpha_1$	0.056842	0.007069	8.0416
$\beta_1$	0.932455	0.007902	118.0011
DoF	5.762335	0.444218	12.9719
	LLF	AIC	BIC
	-9275.0086	18560.0172	18591.8011

Table 4.1: Univariate GARCH(1,1) estimates for cash log-returns.

The next table contains estimates for univariate GARCH(1,1) for interpolated future prices.

Parameter	Value	SE	T statistic
$\mu$	0.030367	0.025538	1.1891
$\alpha_0$	0.029965	0.009286	3.2270
$\alpha_1$	0.051502	0.006607	7.7951
$\beta_1$	0.942396	0.007206	130.7829
DoF	7.250003	0.670882	10.8067
	LLF	AIC	BIC
	-8745.2676	17500.5351	17532.3191

Table 4.2: Univariate GARCH(1,1) estimates for (interpolated) future log-returns.

We also estimate GARCH(1,1) model for switched future series

The GARCH-regression model can be generalized to include the cases where the conditional mean  $\mu$  is a function of  $h_t$ . This is the so-called GARCH-in-mean model (see Engle, Lilien, and Robins (1987) and Domowitz and Hakkio (1985)). For example, for an GARCH(1,1)-in-mean, regression model is generalized to

$$\Delta X_t = \mu + \delta f(h_t) + \epsilon_t \quad (4.46)$$

where  $h_t$  is given by as before and  $f$  is usually  $\sqrt{h}$  or  $h$ . (Other functions like  $\ln(h)$  are also possible.) The GARCH-in-mean model's estimates with  $f = \sqrt{h_t}$  are represented by the

Parameter	Value	SE	T statistic
$\mu$	0.034332	0.026465	1.2972
$\alpha_0$	0.034057	0.010200	3.3390
$\alpha_1$	0.054102	0.006716	8.0556
$\beta_1$	0.939745	0.007247	129.6774
DoF	7.080893	0.683175	10.3647
	LLF	AIC	BIC
	-8932.4255	17874.8510	17906.6349

Table 4.3: Univariate GARCH(1,1) estimates for (switched) future log-returns.

following tables:

Parameter	Value	SE	
$\mu$	-0.023595	0.001246	
$\delta$	0.012486	0.000391	
$\alpha_0$	0.058323	0.000260	
$\alpha_1$	0.085552	0.000103	
$\beta_1$	0.910602	0.000106	
	LLF	AIC	BIC
	-9442.2627	18894.5254	18926.3094

Table 4.4: GARCH-in-mean estimates for cash log-returns.

Now we will compare this univariate models and choose the optimal one. Our decision will be based on the values of AIC, BIC and Log likelihood function.

We see that GARCH(1,1) modeling with t-distribution gives the best result for cash prices of crude oil.

For future prices of crude oil we receive the same conclusion.

## 4.2.2 Estimation of Multivariate Models

We will report results for switched future series only due to lack of space.

### Diagonal BEKK model

$$\epsilon_t = H_t^{1/2} Z_t \quad (4.47)$$

$$H_t = C'C + A'\epsilon_{t-1}\epsilon'_{t-1}A + B'H_{t-1}B \quad (4.48)$$

We assume that C is lower triangular and A,B are diagonal matrices.

Parameter	Value	SE	
$\mu$	-0.039863	0.001996	
$\delta$	0.033120	0.000629	
$\alpha_0$	0.035794	0.000101	
$\alpha_1$	0.070021	0.000055	
$\beta_1$	0.925647	0.000058	
	LLF	AIC	BIC
	-9025.2488	18060.4977	18092.2816

Table 4.5: GARCH-in-mean estimates for (switched) future log-returns.

$$\hat{C} = \begin{pmatrix} 0.445552 & 0 \\ 0.285369 & 0.030678 \end{pmatrix} \quad (4.49)$$

$$\hat{A} = \begin{pmatrix} 0.351021 & 0 \\ 0 & 0.274306 \end{pmatrix} \quad (4.50)$$

$$\hat{B} = \begin{pmatrix} 0.924395 & 0 \\ 0 & 0.956248 \end{pmatrix} \quad (4.51)$$

Next, we will estimate the Full BEKK Model

### Full BEKK Model

C is a lower triangular matrices

$$Y_t = H_t^{1/2} Z_t = \epsilon_t \quad (4.52)$$

$$H_t = C'C + A'\epsilon_{t-1}\epsilon'_{t-1}A + B'H_{t-1}B \quad (4.53)$$

$$\hat{C} = \begin{pmatrix} 0.155291 & 0 \\ 0.054235 & 0.114556 \end{pmatrix}$$

$$\hat{B} = \begin{pmatrix} 0.901550 & 0.090789 \\ -0.022000 & 1.004274 \end{pmatrix}$$

$$\hat{A} = \begin{pmatrix} 0.456811 & -0.356468 \\ 0.075904 & 0.096831 \end{pmatrix}$$

### Constant Conditional Correlation Model (CCC)

$$Y_t = H_t^{1/2} Z_t = \epsilon_t \quad (4.54)$$

$$H_t = D_t R D_t \quad (4.55)$$

$$R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} D_t = \begin{pmatrix} \sigma_{1t} & 0 \\ 0 & \sigma_{2t} \end{pmatrix} \sigma_{it}^2 = \alpha_{i0} + \alpha_i Y_{t-1}^2 + \beta_{i1} \epsilon_{t-1}^2 \quad (4.56)$$

### Dynamic Conditional Correlation

$$Y_t = H_t^{1/2} Z_t = \epsilon_t \quad (4.57)$$

$$H_t = D_t R D_t \quad (4.58)$$

$$Q_t = (1 - dc\alpha - dc\beta)\bar{Q} + dc\alpha\epsilon_{t-1}\epsilon'_{t-1} + dc\beta Q_{t-1} \quad (4.59)$$

$$\bar{Q} = Cov(Z_t) \quad (4.60)$$

$$D_t = \begin{pmatrix} \sigma_{1t} & 0 \\ 0 & \sigma_{2t} \end{pmatrix} \quad (4.61)$$

$$\sigma_{it}^2 = \alpha_{i0} + \alpha_i Y_{t-1}^2 + \beta_{i1} \epsilon_{t-1}^2 \quad (4.62)$$

Constant Conditional Correlation  $D_t$  is constructed with standard GARCH(1,1) component.

$\bar{Q} = Cov(Z_t)$  is estimated directly via sample Variance-Covariance matrix

$$\bar{Q} = Cov(Z_t) = \begin{pmatrix} 1.000931 & 0.902914 \\ 0.902914 & 1.105938 \end{pmatrix}$$

We can notice that estimation of Dynamic Conditional Correlation and Constant Conditional Correlation Models produce very similar results.

In order to choose model with best performance we compare maximum likelihood value and AIC and BIC criteria. All these criteria suggest using of Full BEKK model.

## 4.3 Bivariate Full Bekk GARCH Model of Cash and Future Prices

Our dataset consists of daily spot and near to expiration future prices of crude oil for the period of 01/1990-03/2006. We switch every 15-th of each month between next to expire and second to expire future prices in order to avoid Samuelson effect.

Diagonal BEKK			Full BEKK			CCC			DCC		
	Value	SE		Value	SE		Value	SE		Value	SE
$c_{11}$	0.445552	0.016040	$c_{11}$	0.155291	0.002488	$\alpha_{10}$	0.059549	0.000806	$\alpha_{10}$	0.059549	0.000819
$c_{21}$	0.285369	0.006083	$c_{21}$	0.054235	0.000726	$\alpha_{11}$	0.085961	0.000634	$\alpha_{11}$	0.085961	0.000607
$c_{22}$	0.030678	0.004150	$c_{22}$	0.114556	0.000633	$\beta_{11}$	0.910049	0.000544	$\beta_{11}$	0.910049	0.000534
$a_{11}$	0.351021	0.001837	$a_{11}$	0.456811	0.002256	$\alpha_{20}$	0.000000	0.000009	$\alpha_{20}$	0.000000	0.000009
$a_{22}$	0.274306	0.000673	$a_{21}$	0.075904	0.001178	$\alpha_{21}$	0.067262	0.000107	$\alpha_{21}$	0.067262	0.000107
$b_{11}$	0.924395	0.000524	$a_{12}$	-0.356468	0.006689	$\beta_{21}$	0.932736	0.000104	$\beta_{21}$	0.932736	0.000103
$b_{22}$	0.956248	0.000092	$a_{22}$	0.096831	0.002488	$\rho$	0.858181	0.000080	$dca$	0.059158	0.000114
			$b_{11}$	0.901550	0.000401				$dc\beta$	0.905170	0.000345
			$b_{21}$	-0.022000	0.000130						
			$b_{12}$	0.090789	0.000547						
			$b_{22}$	1.004274	0.000168						
LLF	-15307.7722		LLF	-15257.7376		LLF	-15667.4861		LLF	-15389.9831	
AIC	30629.5444		AIC	30537.4752		AIC	31348.9721		AIC	30795.9663	
BIC	30674.0419		BIC	30607.3999		BIC	31393.4696		BIC	30846.8206	

Table 4.6: Estimation of Multivariate GARCH Models

Afterwards, we build time series of logreturns of cash and futures prices.  $X_t$  is a bivariate vector of log returns and  $Y_t = \Delta X_t$ . We perform test of cointegration on the vector  $Y_t$  and find that the hypothesis of no co-integration between first differences of log returns can not be rejected. Due to the fact that first differences of log returns are not cointegrated we can build bivariate GARCH model.

$$Y_t = \mu + \epsilon_t \quad (4.63)$$

$$\epsilon_t/\Omega \quad t-1 N(0, H_t) \quad (4.64)$$

$$\text{Vech}(H_t) = C + A\text{Vech}(\epsilon_t\epsilon'_{t-1}) + B\text{Vech}(H_{t-1}) \quad (4.65)$$

$H_t$  is a (2×2)conditional covariance matrix; C is a (3×1) parameter vector; A and B are (3×3) parameter matrices; and *vech* is the column stacking operator that stacks the lower triangular portion of a symmetric matrix. Following Bollerslev, Engle, and Wooldbrige(1988), a more parsimonious representation can be obtained by assuming that A and B matrices are diagonal. This is a natural simplification because it implies that each variance and covariance depends only on its own past values and prediction errors. Since both cash and future prices changes have no autocorrelation, the dynamics of the conditional mean are trivial to specify. Any extensions can be tested by means of a sequence of LM tests. The vector of disturbances  $\epsilon_t$  is assumed to have a conditional normal distribution with time dependent covariance matrix  $H_t$ . There are several possible parameterizations of multivariate GARCH processes (Engle and Bollerslev(1986)). An alternative to the diagonal Vec parameterization is the positive definite parameterization.

$$\Delta y_t = \mu + \epsilon_t \quad (4.66)$$

$$\epsilon_t/\Omega_{t-1}\tilde{N}(0, H_t) \quad (4.67)$$

$$H_t = C'C + A'\epsilon_{t-1}\epsilon'_{t-1}A + B'H_{t-1}B \quad (4.68)$$

$$(4.69)$$

where C is a symmetric 2×2 parameter matrix and A and B are unrestricted 2×2 parameter matrices. In a bivariate GARCH(1,1) systems, the diagonal VECM parametrization involves nine conditional variance parameters while the positive definite We estimate the following model using MATLAB software. OHR is the ratio of conditional covariance of future and spot prices to its conditional variance.

### 4.3.1 Empirical Results

We obtain the following coefficients:

$$\begin{aligned}\Omega &= \begin{pmatrix} 0.0522 & 0.043 \\ 0.0430 & 0.0354 \end{pmatrix}, \\ ARCH_1 &= \begin{pmatrix} 0.6867 & -0.5606 \\ -0.05587 & -0.0847 \end{pmatrix} \\ ARCH_2 &= \begin{pmatrix} 0.3494 & -0.491 \\ -0.0298 & -0.2996 \end{pmatrix} \\ GARCH_1 &= \begin{pmatrix} 0.5120 & 0.2819 \\ 0.1314 & 0.7845 \end{pmatrix} \\ GARCH_2 &= \begin{pmatrix} 0.3173 & 0.2662 \\ 0.1430 & 0.1510 \end{pmatrix}\end{aligned}$$

The following table presents hedging performance of six strategies in sample based on OHR derived from bivariate GARCH model.

	Unhedged Mean	Unhedged Volatility	Hedged Mean	Hedged Volatility
1	8.38	48.08	1.13	25.58
2	8.38	48.08	0.44	25.39
3	8.38	48.08	0.38	25.40
4	8.38	48.08	1.88	25.53
5	8.38	48.08	1.04	25.31
6	8.38	48.08	0.88	25.33

Table 4.7: Performance in Sample of Bivariate GARCH Model

We see that hedging significantly reduces mean of the return but also volatility decreases and this is valuable result for oil industry. However, later we will analyse more performance based criteria like Sharpe Ratio for each strategy.

## 4.4 Cointegrated VAR-GARCH Model

On the other hand we find that logreturns of cash and future prices are cointegrated. To test for cointegration we use Johansen Juselius methodology, Johansen and Juselius (1990), and we reject the hypothesis of no cointegration of 5 per cent level.

Null Hypotheses	Trace Statistic	Crit 0.95	Crit 0.99
$r \leq 0$	453.867	15.494	19.935
$r \leq 1$	0.395	2.705	6.635
one cointegration relation			

Table 4.8: Johansen Trace Statistic Test

Null Hypotheses	EV Statist	Crit 0.95	Crit 0.99
$r \leq 0$	453.472	12.297	18.520
$r \leq 1$	0.395	2.705	6.635
one cointegration relation			

Table 4.9: Johansen Eigenvalue Statistic Test

The presence of cointegration means that it can be shown that the quasi-spread between logarithm of cash prices  $C_t$  and logarithm of future prices  $F_t$ :

$$S_t = F_t - \beta C_t \quad (4.70)$$

is stationary Campbell and Mankiw (1987); Phillips and Perron (1988). Therefore  $F_t$  and  $C_t$  are co-integrated with co-integration vector  $(1 - \beta)'$ . Campbell and Shiller (1987) derive and test the parameter restrictions implied by the present value relation on a VAR model for  $\Delta(r_t)$  (the first difference of  $r_t$ -interest rate) and  $S_t$ -asset price. We use a VAR model for the futures and cash prices of crude oil and we test if cointegration holds. Since a look at the difference time series of crude oil prices reveals that the volatility of the series is not constant, we enlarge the model with a multivariate GARCH effect on the innovations. We will follow the paper of Bauwens (1997). They use a bivariate VAR model to model and predict the joint evolution of short term and long term interest rates. They introduce GARCH effect on the innovations of the model in order to account for the changing volatility of the series. We parameterize the VAR model in the error correction mechanism form:

$$\Delta y_t = \gamma + \Pi y_{t-1} + \sum_{i=1}^m \Pi_i \Delta y_{t-i} + \epsilon_t \quad (4.71)$$

where  $\Pi_i (i = 1, \dots, m)$  and  $\Pi$  are square matrices of order 2 and  $\gamma$  is the vector of intercepts. With the co-integration relation (4.70) between the prices, the matrix  $\Pi$  is of rank one, and we include the constant in the co-integrating relation, so that:

$$\Pi = AB', \quad \gamma = A\mu, \quad A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \quad B = \begin{pmatrix} 1 \\ -\beta \end{pmatrix} \quad (4.72)$$

The error vector  $\epsilon_t$  follows a GARCH(p,q) process, defined by

$$E(\epsilon_t/I_{t-1}) = 0 \quad (4.73)$$

$$E(\epsilon_t \epsilon_{t-1}'/I_{t-1}) = H_t \quad (4.74)$$

$$(4.75)$$

where  $I_{t-1}$  denotes the past information set and  $H_t$  has the BEKK formulation of Engle and Kroner (1995):

$$H_t = \Omega + \sum_{i=1}^p (A_i)' \epsilon_t (\epsilon_{t-i})' A_i + \sum_{j=1}^q (G_j)' H_{t-j} G_j \quad (4.76)$$

$\Omega$ ,  $A_i (i=1\dots p)$  and  $G_j (j=1\dots q)$  are square matrices of order 2. Estimation is done by the quasi-maximum likelihood method assuming that  $\epsilon_t$  is Gaussian with moments given by (4.73) and (4.74). Tuncer(1994) and Bauwens (1995) have shown that the QML estimator for this type of model is consistent under some "high-level" conditions. We assume that asymptotic normality holds, but it has not been proved, as far as we know.

**Specification strategy** The VAR-GARCH model may involve a lot of parameters. With  $p = q = 2$ , there are 19 parameters in the GARCH part. Since we use monthly data, a reasonable value for the lag order (m) in the VAR system may be 5 or 6, in which case there are 24 or 28 parameters in the VAR part. The total number of parameters is thus of the order of 40 to 50. Estimation is then typically demanding in computing time. A reasonable specification strategy is to start with a "general" starting model, meaning that it is preferable to start with an over parameterized model rather than an underparameterized one. Moreover the starting point of a simplification strategy should be checked for misspecification, in particular for residual autocorrelation. Autocorrelation tests can be performed as follows: i) one estimates the starting model and saves its residuals; ii) one adds lags of the two residuals (up to the order of autocorrelation to be tested) in both equations; and iii) one tests that all the added variables have zero coefficients. If the restriction is accepted, one can assume that the lag order of the starting model is sufficient. This method involves estimating more parameters at step ii) than in the starting model. It turned out to be practically unfeasible. Another test

of autocorrelation is the multivariate version of the Portmanteau test (see e.g. Luktepohl, 1991, p. 152) which uses only the residuals of the model estimated under the null hypothesis. We decided to introduce five lags of the two variables  $F_t$  and  $C_t$  in both equations and to apply Portmanteau tests to the residuals of the VAR model estimated without a GARCH part. As the test was passed (with a few exceptions), we proceeded with the estimation of the specified VAR model together with the GARCH effect. For practical purposes, we have considered at most GARCH(2,2) specifications. The residuals of the VAR-GARCH model were then submitted to the Portmanteau test. Obviously, before the specification search and the estimation of the VAR, we tested for the presence of a unit root in the two price series, and we tested for cointegration between them. Cointegration can be tested by using the Engle-Granger test or the likelihood ratio test (i.e. Johansen's trace test). Likelihood ratio tests that take into account the GARCH effects can be computed after the estimations of the VAR model (4.71) with the GARCH part (4.73)-(4.76) under the three hypotheses that  $\Pi$  has a rank equal to two, to one - as in (4.72), or to zero. We decided to use the Johansen's trace test critical values even in the presence of GARCH effects, since we conjecture that they should remain valid asymptotically (given that Johansen's results rely on a functional central limit theorem). Moreover, Cheung and Lai (1993) studied the size of the trace test in finite samples under non-normality and conclude that it shows little bias in the presence of skewness or of excess kurtosis. As they did not specify the non-normal distribution of their Monte Carlo experiments as a GARCH process, other Monte-Carlo experiments should be performed to check if their conclusion holds in this case. As our estimation results seem to imply that the covariance matrix of the errors of the VAR does not exist, we also used critical values computed by Caner (1996), which seem to be relevant in such a case. Caner has computed several tables as the asymptotic distribution of the trace statistic depends in his setup on an nuisance parameter (the index of the stable law of the errors). As this parameter is unknown, we used the table corresponding to the smallest index, which gives critical values which are the most distant from Johansen's ones.

#### 4.4.1 Empirical Results

We use the same data set as before: daily spot and near to expiration future prices of crude oil for the period of 01/1990-03/2006. We switch every 15-th of each month between next to expire and second to expire future prices in order to avoid Samuelson effect. We also build time series of constant 30 days until expiration future prices by interpolating between prices of 4 daily available prices of futures contract with different maturities. We obtain 4260 observations.

However it is much more convenient to estimate our model with 100 times logreturns. This transformation doesn't affect the variance.

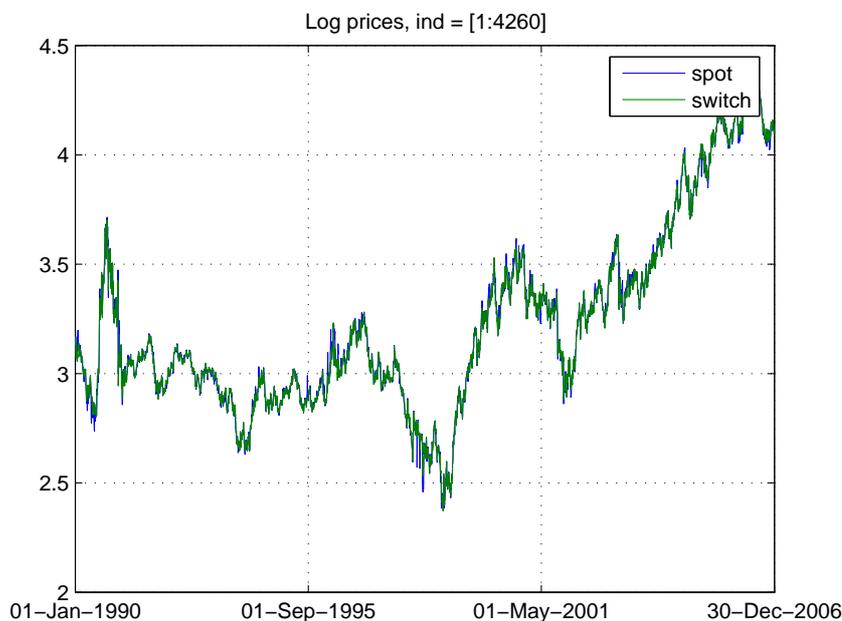


Figure 4.1: Log Returns

A look at the time series plots reveals that the individual series may well be  $I(1)$ , that there may be a common movement in the spot and the future prices of crude oil and that the differenced series exhibit the typical ARCH behavior.

### Misspecification Tests

Multivariate Portmanteau tests reveal that there remains some significant autocorrelation in the models. Compared to the model without GARCH effect, this is due to the change of the estimates of the coefficients of the VAR, induced by the estimation of the GARCH parameters. It is therefore not clear that this problem can be solved by adding more lags in the VAR part. Moreover, the residual autocorrelogram show that there are some large autocorrelations at some rather long lags. It is practically unfeasible to extend the VAR specification to include more lags in the VAR part given the number of parameters to be estimated. A possible approach to tackle this issue is to consider the inclusion of moving average terms on the errors. Another issue is the normality of the errors. Skewness and kurtosis coefficients (not reported) of the standardized residuals reveal that they are far from being normal. Since we use QML estimation, non-normality is not crucial, since standard errors are adjusted to take

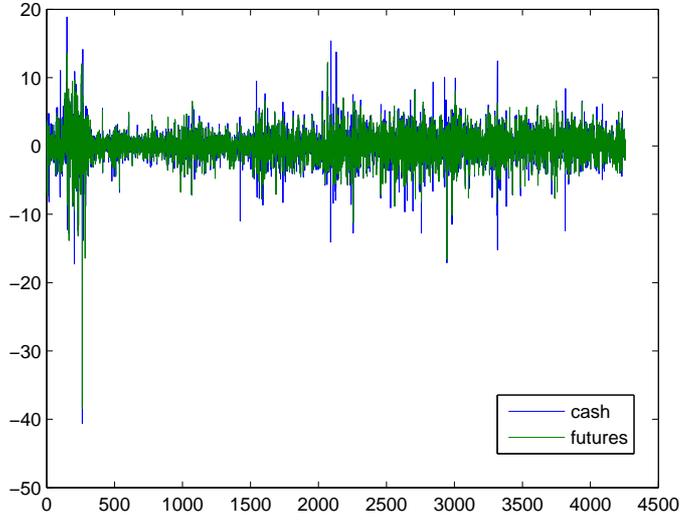


Figure 4.2: Differenced Time Series

account of the possible non-normality.

#### Cointegration and Error-Correction Parameters

We estimate bivariate VAR model and impose on its residuals Full-Bekk GARCH structure. Assume:  $X_t$  is a bivariate vector of logreturns of futures and cash prices and  $Y_t = \Delta X_t$ . We reduce the number of lags in VAR due to estimation heaviness and high number of parameters in case of Full Bekk model.

The model takes the following form:

$$Y_t = \gamma + \Pi Y_{t-1} + \sum_{i=1..m} \Pi \delta Y_{t-i} + \epsilon_t$$

Where  $\gamma$  is a  $2 \times 1$  vector of constant;  $\Pi = AB'$ ,  $A = [\alpha_{1i}, \alpha_{2i}, \dots, \alpha_{2j} \dots]$  is a adjustment coefficients matrix and  $B' = [1 - \beta_{12} - \beta_{13} \dots - \beta_{1d}; 1 - \beta_{22} - \beta_{23} \dots - \beta_{2d}]$ ; is a matrix of co-integrating relationships.

The variance has the following structure:

$$\begin{aligned}
E[\epsilon_t|I_{t-1}] &= 0 \\
Var(\epsilon_t|I_{t-1}) &= E[\epsilon_t\epsilon_t'|I_{t-1}] = H_t \\
\epsilon_t|I_{t-1} &\sim N(0, H_t)
\end{aligned}$$

$$\begin{aligned}
H_t &= \Omega + \sum_{i=1..p} \Pi A_i' \epsilon_{t-i} \epsilon_{t-i}' A_i \dots + \sum_{j=1..q} \Pi G_j' * H_{t-j} * G_j \\
\Omega &= C' * C
\end{aligned}$$

Where  $C$  is upper triangular matrix,  $A_i$ ,  $G_j$  are full matrices. The cointegrating relation between the cash and future prices should have a slope equal to one if the present value relation holds, so that the spread  $F_t - C_t$  is stationary around the constant  $\mu$ . We also notice that adjustment coefficient of second equation is not significantly different from zero. We estimate restricted model imposing joint hypotheses. In order to test if we can accept this hypotheses we use Likelihood Ratio Test. We estimate original Cointegrated VAR model with GARCH error structure. From this model we have Likelihood Value,  $LLH_u$ . Afterwards, we estimate restricted model on the same log returns, imposing that adjustment coefficient of the second equation,  $A(2, 1)$  is equal to zero and the co-integrating relation is  $B = (1, -1)$  and from this model we receive Likelihood Value,  $LLH_r$ . In the next step we build the following statistic:

$$LR = 2[LLH_u - LLH_r] \sim \chi^2(m).$$

$m$  is a number of restrictions, in our case  $m = 2$ .  $LR = 3.5294$  in our case. From the  $\chi^2(m)$  distribution table the probability that a  $\chi^2(2)$  variable exceeds 5.99 is 5%. Since  $3.5294 < 5.99$  we accept the null hypotheses of  $A(2, 1) = 0$  and  $B = (1, -1)$ . Therefore, we conclude that logarithm of future prices is weakly exogenous for the system. Meaning of this is that future prices do not adjust to long run cointegration relationship and depends on the short run dynamic only. This means that  $\sum_{i=1}^t \epsilon_{2i}$  is a common trend in the sense that the errors in the equations for  $X_{2t}$  (future prices) cumulate in the system and give rise to the non stationarity. Because  $X_{1t}$  is univariate our model can be represented in single equation form which is an Error Correction Model for the changes in  $X_{1t}$  as explained by simultaneous values of  $\Delta X_{2t}$ , the lags of  $\Delta X_t$  and the error correction term  $\beta' X_{t-1}$ . Johansen (1995) proves a

theorem of weak exogeneity according to which if  $\alpha_2 = 0$ , then  $X_{2t}$  is weak exogenous for the parameter  $(\beta, \alpha_1)$  and the maximum likelihood estimator of  $\beta$  and  $\alpha_1$  can be calculated from the conditional model. It means that for inference concerning  $\beta$  and  $\alpha_1$  we only need to do a single equation analysis, which is much easier.

After imposing restrictions and verifying their significance our system takes the following form for futures prices time series composed with switches after each 15-th of the month

$$\begin{pmatrix} \Delta LC_t \\ \Delta LF_t \end{pmatrix} = \begin{pmatrix} 0.0201 \\ 0.0186 \end{pmatrix} + \begin{pmatrix} -0.4092 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} LC_{t-1} \\ LF_{t-1} \end{pmatrix} + \begin{pmatrix} 0.1295 & -0.1511 \\ 0.1317 & -0.1506 \end{pmatrix} \begin{pmatrix} \Delta LC_{t-1} \\ \Delta LF_{t-1} \end{pmatrix} \\ + \begin{pmatrix} 0.0760 & -0.0981 \\ 0.0749 & -0.0885 \end{pmatrix} \begin{pmatrix} \Delta LC_{t-2} \\ \Delta LF_{t-2} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

Variance matrix has the following form:

$$H_t = \Omega + A'_1 \epsilon_{t-1} \epsilon'_{t-1} A_1 + A'_2 \epsilon_{t-2} \epsilon'_{t-2} A_2 + G'_1 H_{t-1} G_1 + G'_2 H_{t-2} G_2$$

The estimates are as follows:

$$\begin{aligned} \Omega &= \begin{pmatrix} 0.3051 & 0.2712 \\ 0.2712 & 0.2410 \end{pmatrix}, \\ ARCH_1 &= \begin{pmatrix} 1.0625 & 0.2795 \\ -0.8572 & -0.0845 \end{pmatrix} \\ ARCH_2 &= \begin{pmatrix} 0.2206 & 0.0318 \\ -0.2336 & -0.2049 \end{pmatrix} \\ GARCH_1 &= \begin{pmatrix} 0.4586 & -0.0124 \\ -0.4559 & 0.2388 \end{pmatrix} \\ GARCH_2 &= \begin{pmatrix} -0.6575 & -0.6285 \\ 1.4767 & 1.4520 \end{pmatrix} \end{aligned}$$

We can notice that co-integrating relationship is quite stable but with notable number of outliers

OHR can be expressed as a ratio of covariances obtained from the VAR model with GARCH, taking into account cointegration relationship between cash and future prices of

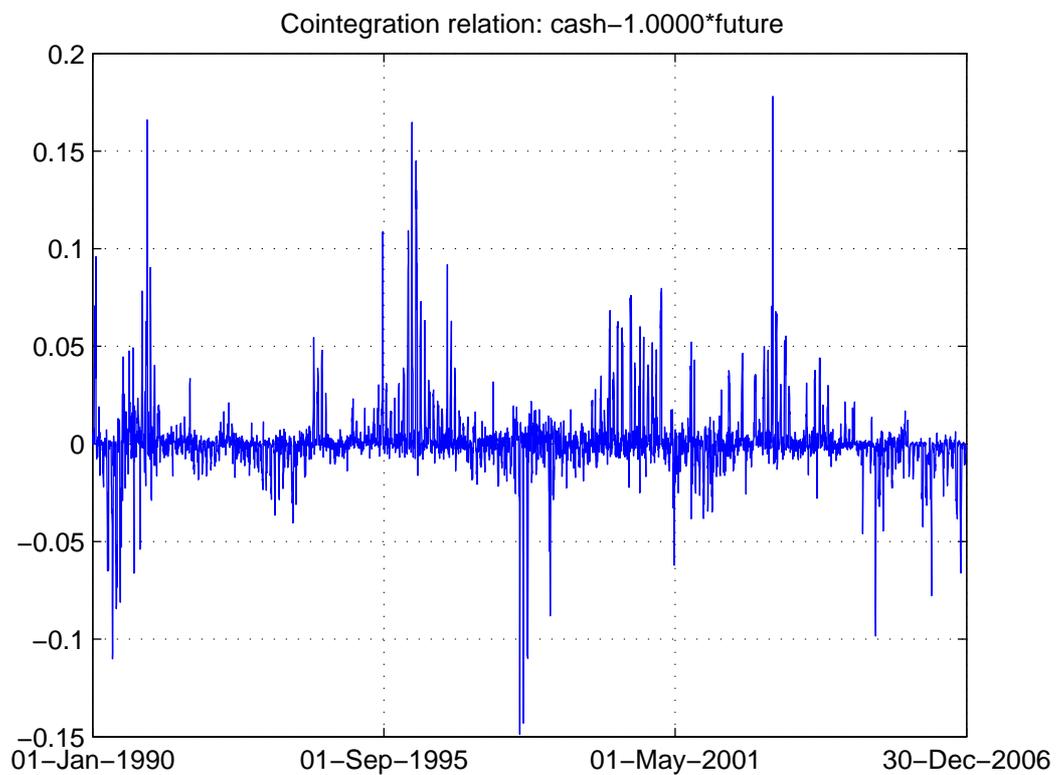
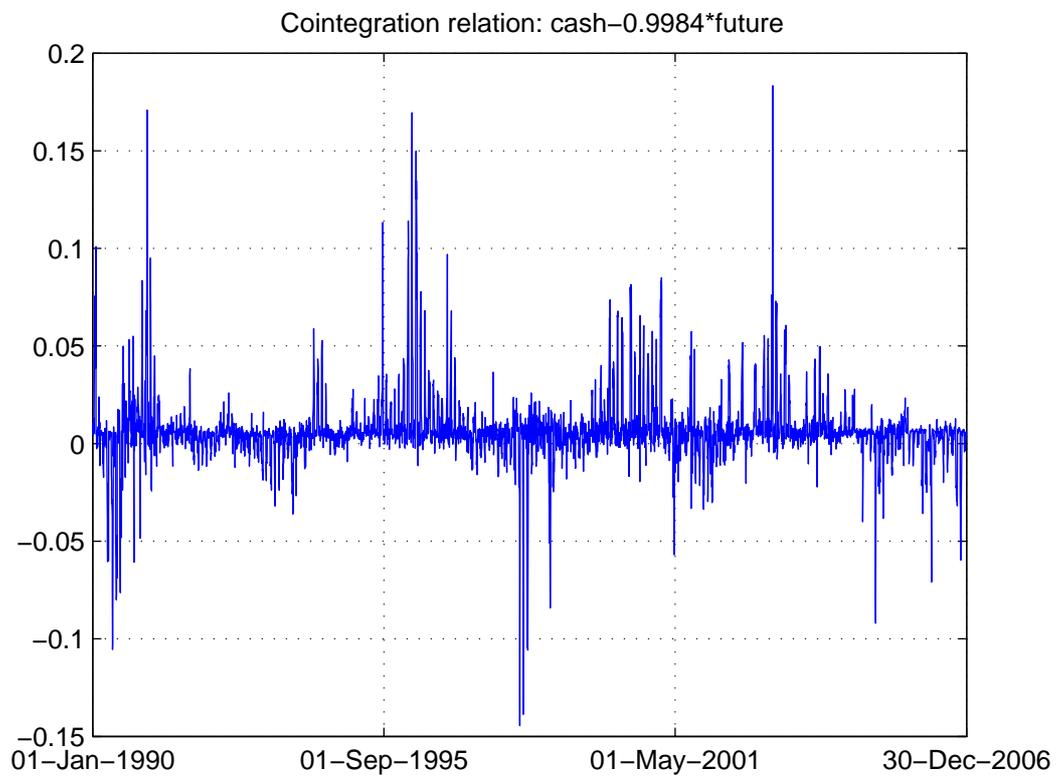


Figure 4.3: Co-integrating Relationship. Upper panel: the unrestricted estimate. Lower panel: restricted estimate.

crude oil

$$b_{t-1} = \sigma_{21,t}/\sigma_{22,t} \quad (4.77)$$

where  $\sigma_{i,j}$  is the element in the  $i$ th row and  $j$ th column of the conditional covariance matrix  $H_t$ .

#### 4.4.2 Hedging Performance In Sample

Our model produces good results in sample. Let us look on its goodness of fit:

A quantile-quantile plot shows relationship between empirical quantiles of the data and theoretical quantiles of a reference distribution, with a lack of linearity showing evidence against the hypothesized reference distribution. We estimate our model with Quasi Maximum Likelihood thus we assume multivariate normal distribution of errors. The graph of residuals shows us that distribution has heavier tails than normal distribution. This is quite typical for financial data.

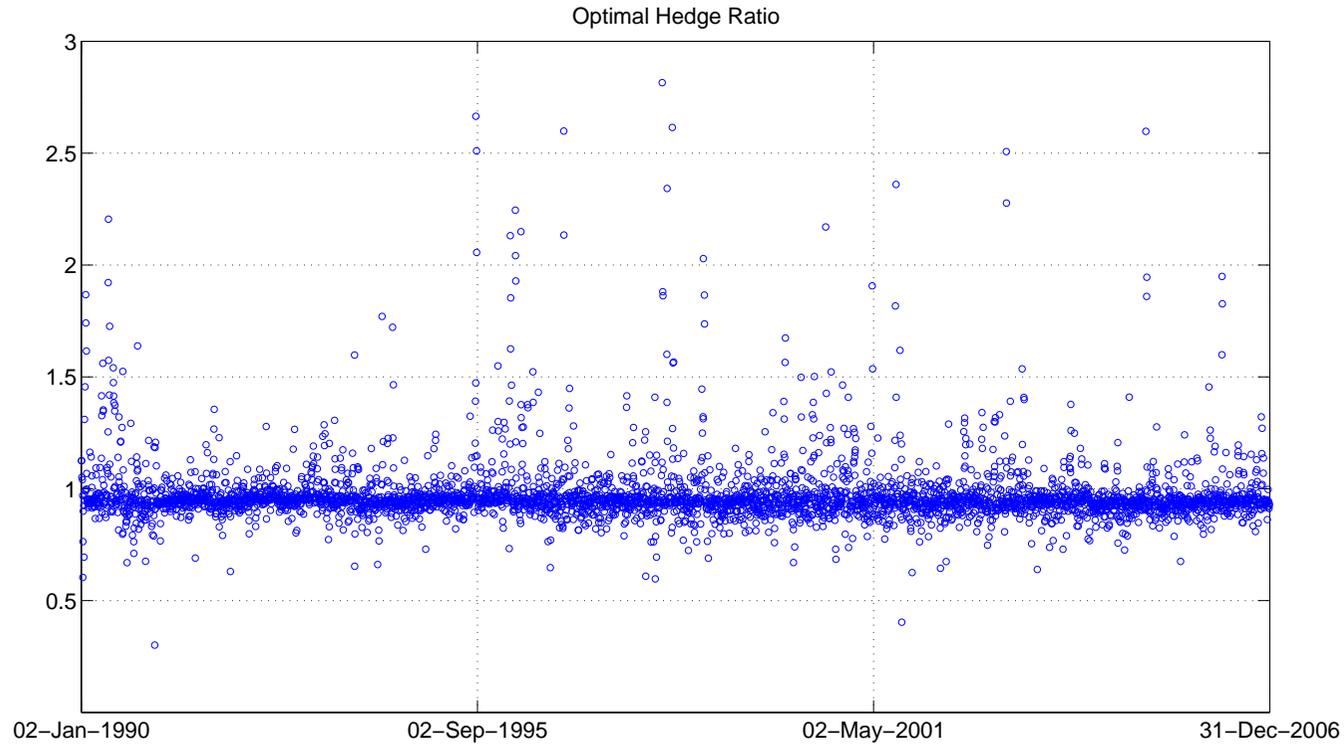
As we can see most of the time hedging ratio is 0.9724. The next graphs show annualized values of variances and covariances and optimal hedge ratio in sample.

Now we will propose different hedging strategies based on hedge ratios calculated from cointegrated GARCH model. We will evaluate their performance on the basis of portfolio return and variance first in sample and afterwards out of sample. As in previous case, OHR can be expressed as a ratio of covariances obtained from the VAR model with GARCH, taking into account cointegration relationship between cash and future prices of crude oil

$$b_{t-1} = \sigma_{21,t}/\sigma_{22,t} \quad (4.78)$$

where  $\sigma_{i,j}$  is the element in the  $i$ th row and  $j$ th column of the conditional covariance matrix  $H_t$ . We will evaluate the following strategies of hedging for portfolio of spot and futures of crude oil.

1. hedged returns=cash return -time varying hedge ratio  $\times$  future return
2. hedged return=cash return-mean of time varying HR of cointegrated GARCH  $\times$  future return
3. hedged return= cash return -median of time varying HR of cointegrated GARCH  $\times$  futures return
4. hedged return= cash return-HR (lies in the interval  $[-1, 1]$ )  $\times$  futures return



Q0	Q1	Q2	Q3	Q4	IQR	LW	UW	OL	mean	std	var	skew	kurt
0.3004	0.9216	0.9464	0.9764	2.8148	0.0548	0.8396	1.0584	594	0.9724	0.1489	0.0222	5.2532	45.5466

Figure 4.4: Optimal hedge ratio (OHR) time series plot and descriptive statistics. Q0-Q4 stand for minimum, 25% quantile, median, 75% quantile and maximum. IQR is the interquartile range. LW and UW are lower and upper whisker of a standard boxplot, OL is the number of observations outside lower and upper whisker (outliers). Mean is the average of all observations, *std* the standard deviation, *var* the variance, *skew* stands for the skewness and *kurt* for the kurtosis.

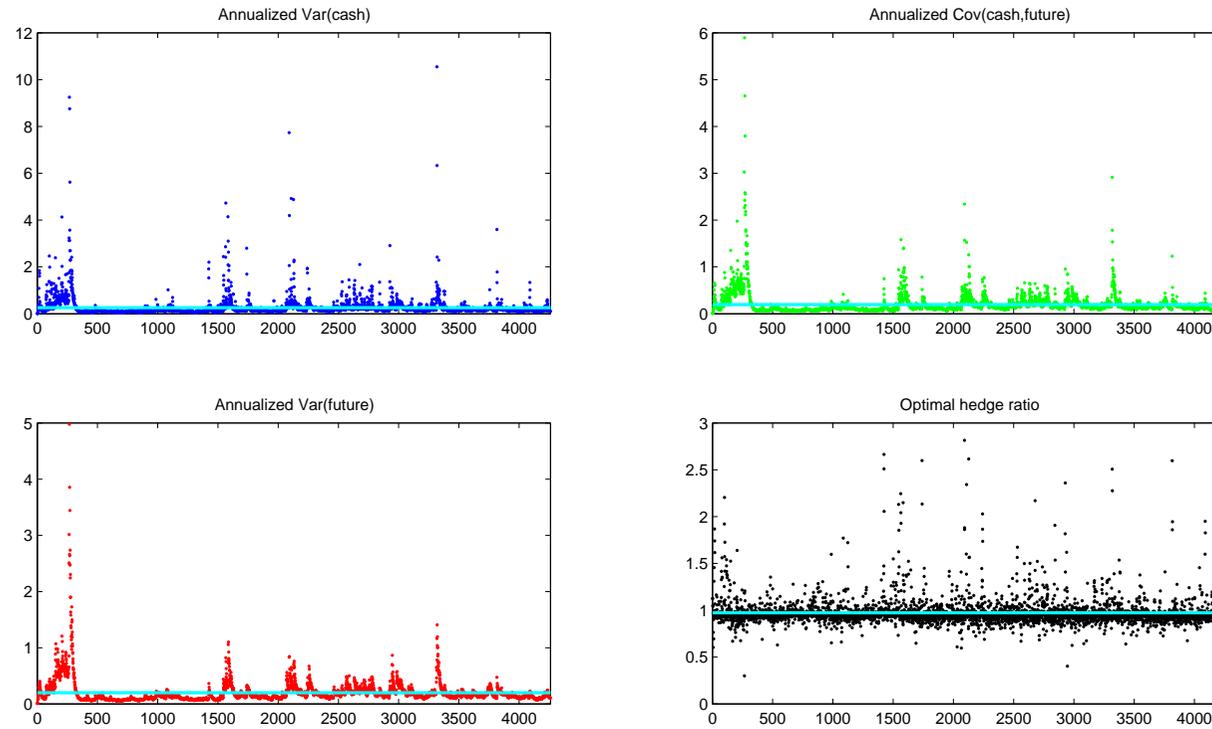


Figure 4.5: Full BEKK estimates for annualized conditional variances and covariances of cash and future. The optimal hedge ratio is the ratio of  $\text{Cov}(\text{cash}, \text{future})$  and  $\text{Var}(\text{future})$ . The horizontal bright blue line marks the average over time in each case. The mean for the annualized  $\text{Var}(\text{cash})$  is 0.2620, for the annualized  $\text{Cov}(\text{cash}, \text{future})$  0.1967, for the annualized  $\text{Var}(\text{future})$  0.1986 and for the optimal hedge ratio 0.9724.

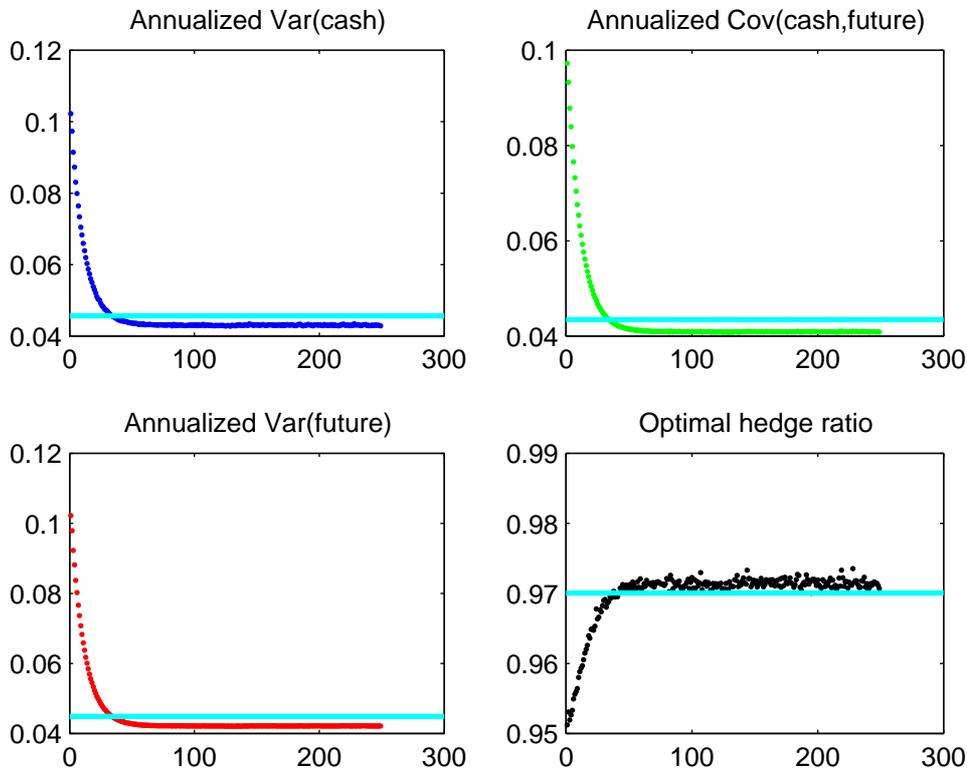


Figure 4.6: Annualized Variance, Covariance and Hedge Ratio in sample

5. hedged return= cash return-HR (lies in the interval [lower whisker, upper whisker])  $\times$  futures return
6. hedged return= cash return-HR (lies in the interval [Q1=25percentile, Q3=75 percentile])  $\times$  futures return

Hedging causes loss in terms of the mean of the return but gain in reducing volatility of the return.

#### 4.4.3 Performance out of sample

In order to test forecasting ability of our model we will analyze its performance out of sample. In order to produce a forecast we will use Filtered Historical Simulations method ,(FHS), developed by Barone-Adesi, Bourgoin, and Giannopoulos (1998), and Barone-Adesi, Giannopoulos, and Vosper (1999).

	Unhedged Mean	Unhedged Volatility	Hedged Mean	Hedged Volatility
1	5.14	41.3	0.27	23.89
2	5.14	41.3	-0.11	23.5
3	5.14	41.3	-0.09	23.49
4	5.14	41.3	0.29	23.6
5	5.14	41.3	0.26	23.53
6	5.14	41.3	0.04	23.49

Table 4.10: Performance in Sample of Cointegrated GARCH Model

Our methodology is non-parametric in the sense that simulations do not rely on any theoretical distribution on the data as we start from the historical distribution of the return series. We use 2500 observations of earlier data to calibrate our GARCH models, for asset returns and to build the data bases necessary to our simulation. By calibrating GARCH models to the historical data we form residual returns from the returns series. Residual returns are then filtered to become identically and independently distributed, removing serial correlation and volatility clusters. As the computation of the i.i.d. residual logreturns involves the calibration of the appropriate GARCH model, the overall approach can be described as semi-parametric. For example, assuming a GARCH(1,1) process with both moving average  $\theta$  and autoregressive  $\mu$  terms, our estimates of the residuals  $\epsilon_t$  and the variance  $h_t$  are :

$$x_t = \mu x_{t-1} + \theta \epsilon_{t-1} + \epsilon_t \quad (4.79)$$

$$h_t = \omega + \alpha(\epsilon_{t-1} - \gamma)^2 + \beta h_{t-1} \quad (4.80)$$

To bring residuals close to a stationary i.i.d. distribution, so that they are suitable for historical simulation, we divide the residual  $\epsilon_t$  by the corresponding daily volatility estimate:

$$e_t = \frac{\epsilon_t}{\sqrt{h_t}} \quad (4.81)$$

We now randomly draw standardised residual returns from the dataset and use them to form a pathway of variances to be used in our backtest. To do this the first-drawn standardised residual is scaled by the deterministic volatility forecast one day ahead:

$$z_{t+1} = e_1 \times \sqrt{h_{t+1}} \quad (4.82)$$

This forecast is used to form the one-day ahead forecast of the logarithm of asset price:

$$p_{t+1} = p_t + p_t(\mu x_t + \theta z_t + z_{t+1}) \quad (4.83)$$

Forecasts of volatility for subsequent days ahead are simulated by the recursive substitution of scaled residuals into the variance equation (4.80). Thus the first-drawn standardised residual from (4.81) with which we form the price forecast one-day ahead in (??) also allows for the simulation of the volatility forecast two days ahead. This volatility depends on the return simulated on the first day. Therefore it is stochastic. We then scale the second-drawn standardised residual, which is used to simulate the price two days ahead. Similarly the volatility three days ahead is formed from the previous second-drawn scaled residual and allows for the scaling of the third-drawn residual and so on up to 50 days ahead. Generally we have for the volatility pathway:

$$\sqrt{h_{t+i}} = \sqrt{\omega + \alpha(z_{t+i-1})^2 + \beta h_{t+i-1}}, \quad i \geq 2 \quad (4.84)$$

Our procedure enables successive scaling of drawn residuals and for the price pathway to be constructed for the forecasted period. Repetition of the approach enables the formation of many pathways of the logreturns  $Y = \Delta X_t$ .

With those forecasted values of log returns and epsilons using definitions of our model(Full Bekk GARCH) we construct conditional Variance-Covariance matrix-forecasted  $H_t$ . From this matrix we calculate the forecasted value of dynamic hedge ratio.

As in previous case, Optimal Hedge Ratio can be expressed as a ratio of covariances obtained from the forecasted  $H_t$ .

$$b_{t-1} = \sigma_{21,t} / \sigma_{22,t} \quad (4.85)$$

where  $\sigma_{i,j}$  is the element in the  $i$ th row and  $j$ th column of the forecasted conditional covariance matrix  $H_t$ .

We reselect 35 subsamples and we evaluate hedging performance for each subsample and calculate Sharpe Ratio and Root Mean Square Error as well.

The following graphs indicate that our model's forecast provides good fit to the real data and that relative error is small.

The quantile plot shows that our data has fat tails which is typical for financial time series. However, we can see that errors distribution is close to normal. Out of sample forecasting results confirms our in sample results: hedging based on ECM with GARCH variance structure significantly reduces volatility of crude oil returns.

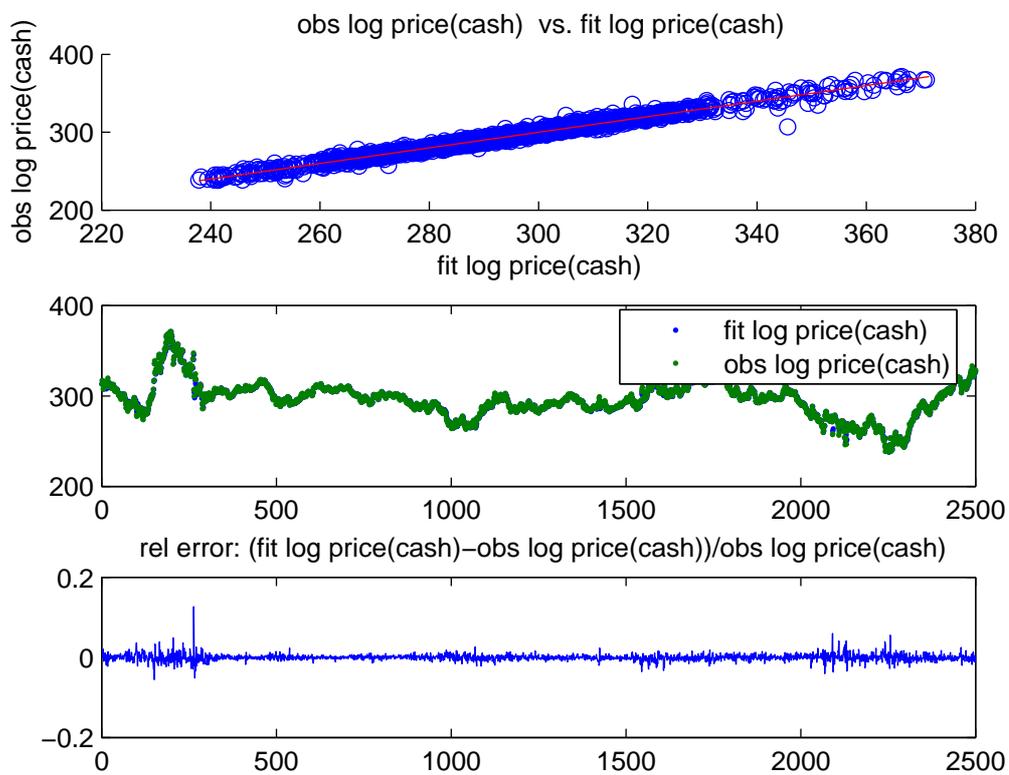


Figure 4.7: Observed versus Fitted Logarithms of Cash Prices for ECM model

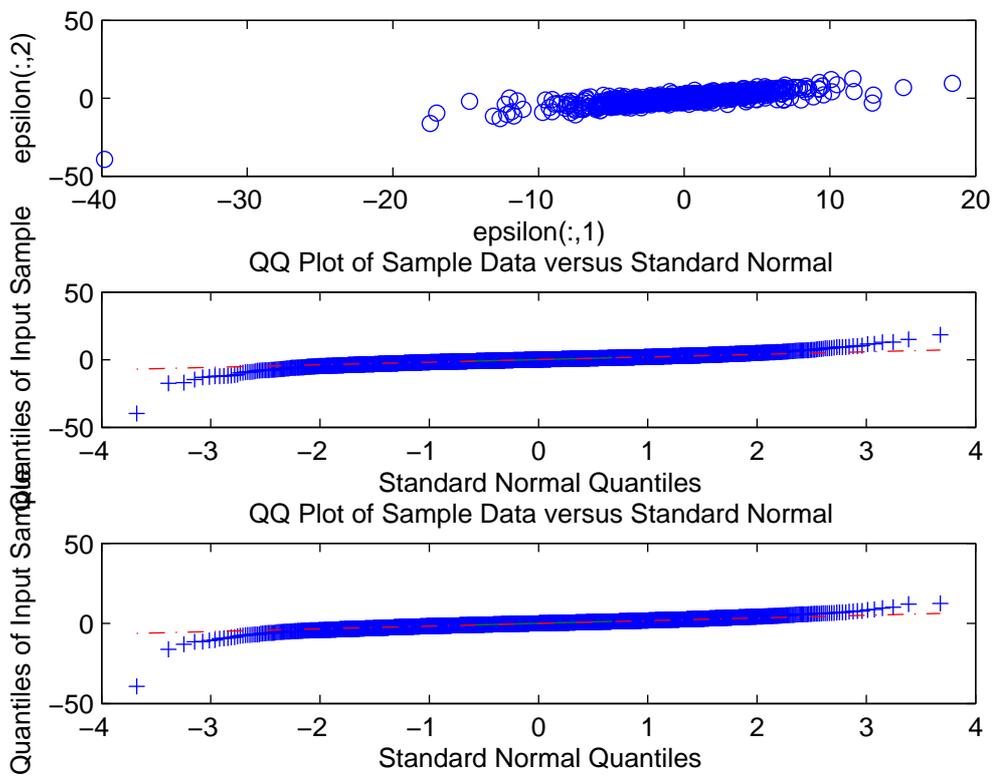


Figure 4.8: Quantiles of sample Data

## 4.5 Comparison of Results

In this section we would like to compare RMSE for various GARCH models we have estimated in the previous section. In the first step we compute RMSE for bivariate GARCH model, with and without cointegration, for spot and futures time series. In the second step we compute the relative difference between them. Mean RMSE for spot prices is equal to 1.212 % for bivariate GARCH and 1.225 % for cointegrated GARCH. For future prices mean RMSE is 0.89 % for bivariate GARCH and 0.9 % for cointegrated GARCH. We obtain that mean RMSE is 1.6% higher in cointegrated GARCH model than in simple bivariate GARCH model for future prices, but is 1.14 % lower for spot prices. One possible explanation of this kind of results might be the fact that quality of forecast is increasing when longer maturities periods are used for futures contracts

We also obtain that for all 35 subsamples the log likelihood is higher for cointegrated GARCH model than for simple bivariate GARCH model.

In order to evaluate performance of hedging based on OHR derived from different models we compute annual Sharpe Ratio for each strategy for each subsample and compare between Sharpe Ratios from bivariate GARCH and cointegrated GARCH using mean and standard deviation criteria.

$$\text{Sharpe Ratio} = \frac{\text{Hedged Return} - \text{Risk Free Return}}{\text{Hedged Volatility}}$$

For the dynamic hedge strategy Sharp Ratio is higher for cointegrated GARCH than for bivariate GARCH 27 out of 35 times.

The following table reports comparison between mean Sharpe Ratios calculated for various hedging strategies based on OHR derived from bivariate GARCH model and cointegrated GARCH model for 35 subsamples.

This table demonstrates that Sharpe Ratio calculated for hedging strategies based on OHR from cointegrated GARCH model is higher than Sharpe Ratio based on OHR from bivariate GARCH model in all cases except the strategy 3 (median OHR) for out of sample estimation. On the basis of these results we can conclude that OHR calculated from cointegrated GARCH model allow us to hedge more effectively. Therefore, this is the model we will recommend for practitioners to use.

Cointegrated GARCH Model Sharpe Ratios													
In Sample							Out of Sample						
<i>Spot</i>	Str 1	Str 2	Str 3	Str 4	Str 5	Str 6	<i>Spot</i>	Str 1	Str 2	Str 3	Str 4	Str 5	Str 6
0.27	0.094	0.028	0.019	0.097	0.07	0.04	0.951	0.454	0.143	0.123	0.414	0.248	0.179
Bivariate GARCH Model Sharpe Ratios													
In Sample							Out of Sample						
<i>Spot</i>	Str 1	Str 2	Str 3	Str 4	Str 5	Str 6	<i>Spot</i>	Str 1	Str 2	Str 3	Str 4	Str 5	Str 6
0.18	0.02	0.022	0.018	0.041	0.020	0.021	0.505	0.204	0.13	0.125	0.202	0.178	0.149
Relative Difference in Sharpe Ratios between the models													
<i>Spot</i>	Str 1	Str 2	Str 3	Str 4	Str 5	Str 6	<i>Spot</i>	Str 1	Str 2	Str 3	Str 4	Str 5	Str 6
0.09	0.074	0.005	0.0005	0.056	0.049	0.0199	0.446	0.249	0.0136	-0.0012	0.212	0.07	0.029

Table 4.11: Overview of Sharpe Ratios

## Chapter 5

# Recent Developments and Conclusions

### 5.1 Recent Developments

The commodities market is being affected by a lot more than just events in the Middle East. Prices of raw materials are seeing widespread gains as talk of a "super cycle" is in the air. Copper, lead, soybeans, wheat, cotton, coffee, cocoa and feeder cattle have all registered double-digit percentage gains this year.

Some of these gains can be traced back to the rise in oil prices. The planned substitution of ethanol for petrol encouraged corn planting, which took acreage away from crops like soya beans and led to a surge in prices. Higher livestock prices can, in turn, be explained by higher grain costs.

But the broad strength of commodity prices may also reflect the appeal of the sector as an "alternative asset," along with hedge funds and private equity. Ever since the dotcom bubble burst, investors have been keen to diversify away from shares and government bonds. That has led to the launch of a whole series of exchange-traded funds based on commodities, which have made the asset class accessible for a much wider range of investors. The latest example, from Barclays Global Investors, a big asset manager, is a fund based on timber prices. Wall Street has been gearing up to meet demand: a survey by Options Group, a recruitment consultant, found that the hiring rate of commodity traders is up by 33 percent on last year.

The recent credit crunch may have given commodities a further lift. Speculative money that had been flowing into high-yield bonds and structured credit is now looking for a new home. Some commodities, particularly gold, are also seen as a hedge against a declining dollar.

Individual commodity prices are still highly volatile thanks to speculative demand. A sharp rise tends to attract "momentum" investors, who push prices up even further until end users start looking for alternatives. At that point, the momentum buyers retreat. But oil's attractions to investors have increased recently because the market has moved into "backwardation" where futures prices are lower than the current price. Investors can thus earn a "roll yield" by buying the future and waiting for the price to rise to the spot level.

Oil prices are hovering near historic highs, but consuming nations shouldn't expect quick relief from OPEC, the world's only source for big, quick supplies. OPEC officials insist that geopolitical jitters and speculative cash are driving the price surge, not a crimp in supply. Any step to boost output would come on top of the cartel's decision in October 2007 to add about 500,000 barrels a day to world supplies as of Nov. 1, a move that did little to calm the market.

Saudi Arabia's powerful oil minister, Ali Naimi, has boasted that OPEC is now immune to political infighting. But member countries such as Iran and Venezuela, faced with slumping production and widespread economic travails, have an interest in keeping supplies tight and prices high. Saudi Arabia, on the other hand, sees itself more as a steward of the world economy.

Over the longer term, high oil prices could have the unusual effect of boosting consumption in one increasingly important growth spot: the Middle East. Saudi Arabia and its neighbors are using more oil, natural gas and gasoline to fuel their own surging energy needs. Higher oil prices could fuel even headier economic growth, which will in turn deepen the region's thirst for energy and limit exports.

Saudi Arabia has little to fear from the world's other major producers, such as Russia, which in decades past have ramped up supplies in an effort to capture a greater market share. But at the moment, the world's major producers for the most part are already pumping flat-out.

The world economy has managed, with some indigestion, to swallow the rise of oil prices past 80 USD a barrel. How well could it survive 100 USD a barrel?

The answer is quite well- so long as several conditions still hold true. The price rise would probably have to be gradual. Inflation couldn't get so bad as to force big interest-rate hikes. Oil-rich nations would need to pump their profits back into U.S. and European economies. All of this has happened so far. The happy confluence may continue, though fears remain strong that high energy prices will tip the U.S. into recession. High oil prices could lead to ugly consequences if they hit consumers' pocketbooks - especially in the U.S., where the housing slump is already hurting the economy. Consumer spending has been the primary engine of

growth in the U.S. in recent years.

For all the concern, the world today is better equipped to swallow expensive oil than it was when Jimmy Carter was installing solar panels and a wood-burning stove in the White House.

The main reason has to do with what some call the Wall-Mart effect. For every extra dollar taken from drivers' pockets at the pump in the form of higher prices in recent years, low-cost exporters from China and elsewhere have put roughly 1.50 USD back in the form of cheaper retail goods. Even at today's near-record prices, U.S. households today spend less than 4 percent of their disposable income at the pump, vs. over 6 percent in 1980. Historically, oil prices have doubled or trebled in a matter of weeks because of sudden and sharp supply disruptions, such as those in 1980 following the Iranian revolution and the outbreak of the Iran-Iraq war. That prompted the Fed to raise interest rates sharply in an effort to head off a spiral of inflation.

Current Fed chairman Ben Bernanke has spent a lot of time trying to understand such shocks. The majority of the impact of an oil price shock on the real economy is attributable to the central bank's response to the inflationary pressures engendered by the shock," he wrote. Today, that view is fairly mainstream among central bankers.

Mr. Bernanke's Fed recently responded to the sub prime mortgage crisis by cutting benchmark interest rates for the first time in four years. By implication, the Fed was saying it was more worried about the fallout from credit-market gloom than about the risk of inflation. At a time of record energy prices, that's a risky but educated bet.

Growing fuel efficiency could also blunt the blow of higher prices. James Barnes, a Union Pacific Corp. spokesman, says the railroad has bought more fuel-efficient locomotives and trained engineers to operate trains in ways that conserve fuel. "From a macro level, we would anticipate that rising oil costs will make us more competitive [with trucks] and potentially drive more business our way," Mr. Barnes says.

In China, the engine of growth on which many are counting, other energy sources can make up for oil. China uses oil for only 21percent of its energy needs, with most of the rest coming from coal. Unlike in the U.S., where imported oil goes to fill people's gasoline tanks, China mainly uses oil in industrial settings, where coal may be an alternative. Greater coal use, however, would also exacerbate China's already serious pollution problem and speed up emissions of gases that contribute to global warming.

Higher oil prices could hit the beleaguered auto and airline industries. Detroit is still digging out from the fall in demand for sport-utility vehicles caused by the climb in gasoline prices. Paul Ballew, General Motors Corp.'s top sales analyst, explained sluggish industry

sales earlier this month by citing in part high fuel prices, which he called "effectively a tax on U.S. households."

For now, most economists expect oil prices will stay high through next year. An unexpected hurricane in the Gulf or a sudden disruption to oil flows from a big producer like Iran or Mexico could push oil to 100-150 USD, they say. Demand is chugging along. The Paris-based International Energy Agency sees world oil demand in the fourth quarter rising by 2.8 percent, or 2.3 million barrels a day from a year ago, to nearly 88 million barrels a day.

Of course, those forecasts could go awry if the U.S. economy tanks and brings Europe and Japan along with it. Then demand would likely ease, and oil prices could fall, perhaps significantly. And then, the world would have something else to worry about.

## 5.2 Conclusion

This thesis was devoted to empirical investigation of crude oil market. Commodities markets are especially popular recently because various economies, such as Russia, China, India, Brazil etc., growing very fast and global consumption of commodities has risen accordingly. Moreover, recent weakness and turbulence of stock markets made commodities markets especially attractive for investors. In my thesis I choose to concentrate on the market for crude oil. This commodity has an important part in our life and this importance receives growing place in global agenda during last years because its prices has surged from 52USD a barrel in May 2005 to 140 USD a barrel in June 2008. The rise was continuous and its impact on global economy is difficult to ignore. In the first part of my thesis I was interested in estimating the dynamics of convenience yield of crude oil. Physical ownership of the commodity carries an associated flow of services. On the one hand, the agent has the option of flexibility with regards to consumption (no risk of commodity shortage). On the other hand, the decision to postpone consumption implies storage expenses. The net flow of these services per unit of time is termed the convenience yield. Since convenience yield appears as a factor which cannot be directly observed, GARCH process was used for its modeling. We find that GARCH model provides good fit for convenience yield process. It paves the way for new term structure models of commodities prices and its empirical testing. Improving the understanding of term structure of oil can be of value for forecasting and planning within oil industry.

For the risk management purpose convenience yield values should be observed at weekly and higher frequencies and in this case they are not independent. I choose to model convenience yield series as a GARCH process. While observations in these series are uncorrelated or nearly uncorrelated, the series contain higher-order dependence. Models of Autoregressive

Conditional Heteroscedasticity (ARCH) form the most popular way of parameterizing this dependence.

Even a cursory look at financial data suggests that some time periods are riskier than others; that is, the expected value of the magnitude of error terms at some times is greater than at others. Moreover, these risky times are not scattered randomly across quarterly or annual data. Instead, there is a degree of autocorrelation in the risk of financial returns. Financial analysts, looking at plots of daily returns of an asset, notice that the amplitude of the returns varies over time and describe this as "volatility clustering". The ARCH and GARCH models, which stand for autoregressive conditional heteroscedasticity and generalized autoregressive conditional heteroscedasticity, are designed to deal with just this set of issues. They have become widespread tools for dealing with time series heteroscedastic models. The goal of such models is to provide a volatility measure-like a standard deviation-that can be used in financial decisions concerning risk analysis, portfolio selection and derivative pricing.

ARCH/GARCH

The first ARCH model assumed that the variance of tomorrow's return is an equally weighted average of the squared residuals from the last days. The assumption of equal weights seems unattractive, as one would think that the more recent events would be more relevant and therefore should have higher weights. Furthermore the assumption of zero weights for observations more than one month old is also unattractive. A useful generalization of this model is the GARCH parameterizations introduced by Bollerslev (1986). This model is also a weighted average of past square residuals, but it has declining weights that never go completely to zero. It gives parsimonious models that are easy to estimate and, even in its simplest form, has proven surprisingly successful in predicting conditional variances. The most widely used GARCH specification asserts that the best predictor of the variance in the next period is a weighted average of the long-run average variance, the variance predicted for this period, and the new information in this period that is captured by the most recent squared residual. Such an updating rule is a simple description of adaptive or learning behavior and can be thought of as Bayesian updating.

Chapter 2 describes models of commodities markets, concentrating particularly on convenience yield and current strand of literature in this field. Chapter 3 first provides an overview of univariate GARCH models and compares between them. Afterwards it applies them to convenience yield time series. It turns out that convenience yield time series has a cyclical component and should be detrended before GARCH model can be estimated. Detrended convenience yield's dynamics are best captured by GARCH(1,1) model for the variance and ARMA(2,0) model for its mean. GARCH model provides a good fit for convenience yield time series in and out of sample. Chapter 4 addresses the issue of hedging in

oil industry. For this purpose we consider portfolio of cash and future position in oil market. Our data set consists of daily spot and near to expiration future prices of crude oil for the period of 01/1990-03/2006. In order to avoid Samuelson effect (increased volatility close to the expiration date of future contract) we adjust the futures prices in the following way: we switch every 15-th of each month between next to expire and second to expire future prices in order to avoid Samuelson effect. We also build time series of constant 30 days until expiration future prices by interpolating between prices of 4 daily available prices of futures contract with different maturities. In order to address this issue we propose Multivariate GARCH model. At the beginning we provide a survey of Multivariate GARCH models highlighting each one's strengths and weaknesses. Afterwards we chose best model's performance on the basis of our dataset. This is a FULL BEKK GARCH(1,1) model. We build bivariate GARCH model for spot and future prices and derive optimal hedge ratio (OHR) from this. Afterwards we build six hedging strategies and compare its performance to hedging on the basis of regular OHR. Consequently we propose and test the hypothesis of co-integration between spot and future prices. We find presence of co-integration and it allows us to build error correction model with GARCH structure for residuals. Building hedging strategies on basis of this model improves portfolio's performance in and out of sample even further. We see that hedging strategy based on Optimal Hedging Ratio from co-integrated GARCH model over beat hedging performance based on OHR of GARCH model without taking into account co-integration relationship between spot and future prices of crude oil. Between hedging strategies of OHR from co-integrated GARCH we can see that strategy 1, dynamic hedge, gives the best performance of portfolio in terms of return and volatility.

Further research is needed to develop the idea of implementing GARCH model in describing convenience yield dynamics in term structure models. We believe that this can improve risk management within the oil industry. An interesting extension of the second part of our work would be to simulate portfolio performance on the basis of hedging with GARCH and co-integration effects and to compare it to the performance on the basis of the model which does not take into account co-integration between spot and future prices of crude oil.

# References

- Alexander, C. and A. N. C. A. Dimitriu (2002, June). The cointegration alpha: Enhanced index tracking and long-short equity market neutral strategies. *Social Science Research Network Working Paper Series*.
- Andersen, T. G. and T. Bollerslev (1998). *Encyclopedia of Statistical Sciences*, Volume 2, Chapter ARCH and GARCH Models. New York: John Wiley and Sons Inc. II.
- Anderson, H. M., K. Nam, and Vahid (1999). *Nonlinear time series analysis of economic and financial data*, Chapter Asymmetric nonlinear smooth transition GARCH models, pp. 191–207. Boston: Kluwer.
- Anderson, R. W. (1985). Some determinants of the volatility of futures prices. *Journal of futures markets* 5(3), 169–204.
- Audrino, F. and P. Bühlmann (2001). Tree-structured GARCH models. *Journal of the Royal Statistical Society, Series B* 63(4), 727–744.
- Audrino, F. and E. De Giorgi (2007). Beta regimes for the yield curve. *Journal of Financial Econometrics* 5, 456–490.
- Baillie, R. T. and R. Myers (1991). Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics* 6, 109–124.
- Barone-Adesi, G., F. Bourgoin, and K. Giannopoulos (1998, August). Don't look back. *Risk* 11, 100–104.
- Barone-Adesi, G., K. Giannopoulos, and L. Vosper (1999). Var without correlations for non-linear portfolios. *Journal of Futures Markets* 19, 583–602.
- Bauwens, L. (1997, october). Modelling interest rates with a cointegrated var-garch model. CORE discussion paper 9780.
- Bauwens, L., S. Laurent, and J. Rombouts (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21, 79–109.

- Black, F. (1976). The pricing of commodity contracts, *Journal of Financial Economics*, 3, 167-179. *Journal of Financial Economics* 3, 167–179.
- Bollerslev, T., R. C. and K. Kroner (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*. 52, 5–59.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of econometrics* 31, 307–27.
- Box, G.E. P. & Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden -Day.
- Brennan, Michael, J. (1958). The supply of storage. *American Economic Review XLVII*, 50–72.
- Brennan, M. (1991). *Stochastic Models and Option Values*, Chapter The Price of Convenience and the Valuation of Commodity Contingent Claims, pp. 33–71. Amsterdam: North Holland.
- Brennan, M. and Schwartz (1985). Evaluating natural resource investments. *Journal of Business* 58(2), 135–157.
- Brooks, C., O. Henry, and G. Persaud (2002). The effects of asymmetries on optimal hedge ratios. *Journal of Business* 75, 333–352.
- Campbell, J. and G. Mankiw (1987). Permanent and transitory component in macro- economic fluctuations. *American Economic Review* 77, 111–117.
- Chambers, M. J. and R. Bailey (1996). A theory of commodity price fluctuations. *Journal of Political Economy* 104(5), 924–57.
- Cortazar, G. and E. S. Schwartz (1994). The valuation of commodity contingent claim. *Journal of derivatives* 1:4, 27–39.
- Crowder, W. J. and A. Hamed (1993). A cointegration test for oil futures market unbiasedness. *Journal of Futures Market* 13, 933–941.
- Deaton, A. S. (1992). On the behaviour of commodity prices. *Review of economic studies* 59(1), 1–23.
- Deaton, A. S. G. L. (1996). Competitive storage and commodity price dynamics. *Journal of Political Economy* 104, 896–923.
- Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of UNITED KINGDOM inflation. *Econometrica* 50, 987–1007.

- Engle, R. and K. Kroner (1995). Multivariate simultaneous generalized ARCH model. *Econometric Theory* 11, 122–150.
- Engle, R. and V. Ng (1993). Measuring and testing the impact of news on volatility. *Journal of Finance* 48, 1749–1778.
- Fama, E. F. (1984). Forward and spot exchange rates. *Journal of Monetary Economics* 14, 319–38.
- Fama E. F., F. K. (1987). Commodity futures prices: Some evidence on forecast power, premiums and the theory of storage. *The Journal of Business* 60(1), 55–73.
- Fong, W. and K. See (2003). Basis variations and regime shifts in the oil futures market. *The European Journal of Finance* 9, 499–513.
- Friedman, M. (1977). Inflation and unemployment. *Journal of Economic Dynamics and Control* 12, 231–254. Nobel lecture.
- Gabillon, J. (1992). The term structure of oil futures prices. Oxford Institute for Energy Studies. Working Paper.
- Gabillon, J. (1995). *Managing Energy Price Risk*, Chapter Analysing the forward curve, pp. 29–38. Risk Publications.
- Gai, J. (1994). Markov model of unconditional variance in ARCH. *Journal of Business and Economics Statistics* 12, 309–316.
- Gibson, R. and E. S. Schwartz (1990, July). Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance* 45(3), 959–976.
- Glosten, L., R. Jagannathan, and D. Runkle (1993). On the relation between expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance* 48, 1779–1801.
- Gonzalez-Rivera, G. (1998.). Smooth-transition GARCH models. *Studies in Nonlinear Dynamics and Econometrics*, 3(2): 61-78, 1998. 3(2), 61–78.
- Gray, S. (1996). Modelling the conditional distribution of interest rate as a regime switching process. *Journal of Financial Economics* 42, 27–62.
- Hagerud, G. E. (1997). *A New Non-Linear GARCH Model*. Ph. D. thesis, IFE, Stockholm School of Economics.
- Hamilton, J. D. and R. Susmel (1994). Autoregressive conditional heteroscedasticity and changes in regime. *Journal of Econometrics* 64, 307–333.

- Heaney, R. (2002). Approximation for convenience yield in commodity futures pricing. *The Journal of Futures Markets* 22, 1005–1017.
- Hilliard, J. R. J. (1998). Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions in the spot. *The Journal of Financial and Quantitative Analysis* 33, 61–86.
- Hotelling, H. (1931). The economics of exhaustible resources. *Journal of Political Economy* 39(2), 137–175.
- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12, 231–254.
- Johansen, S. (1995). *Likelihood Based Inference on Cointegration in the Vector Autoregressive Model* (2 nd ed.). Oxford University Press, Oxford,.
- Johansen, S. and K. Juselius (1990). Maximum likelihood estimation and inference on cointegration - with applications to the demand for money. *Oxford Bulletin of Economics and Statistics* 52, 169–210.
- Kaldor, N. (1939). Speculation and economic stability. *Review of economic studies* 7, 1–27.
- Kearney, C. and A. Patton (2000, February). Multivariate garch modelling of exchange rate volatility transmission in european monetary system. *The Financial Review* 35(1), 29–48.
- Keynes, J. (1930). *A treatise on Money: The applied Theory of Money.*, Volume 2. London: Risk Publications.
- Lautier (2003). The informational value of crude oil futures prices. Working Paper 2003-6, Cereg, University Paris IX.
- Lautier (2005). Term structure models of commodity prices:a review. *The Journal of Alternative Investments* 8, 42–64.
- Lautier, D. and A. Galli (2001). A term structure model of commodity prices with asymmetrical behaviour of the convenience yield. *Finéco* 11, 73–95.
- Li, C. W. and W. K. Li (1996). On a double-threshold autoregressive heteroscedastic time series model. *Journal of Applied Econometrics* 11, 253–274.
- Lien, D. (2004). Cointegration and the optimal hedge ratio: the general case. *The Quarterly Review of Economics and Finance* 44, 654–658.
- Lien, D. and Y. Tse (2000). Hedging downside risk with futures contracts. *Applied Financial Econometrics* 10, 163–70.

- Lien, D.-H. D. (1996, October). The effect of the cointegration relationship on futures hedging: a note. *The Journal of Futures Markets* 16(7), 1986–1998.
- Lin, W. (1992). Alternative estimators for factor GARCH models - a monte carlo comparison. *Journal of Applied Econometrics* 7, 259–279.
- Litzenberger, R. and N. Rabinowitz (1995). Backwardation on oil futures markets: theory and empirical evidence. *Journal of Finance* L(5), 1517–1545.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business* 36, 394–419.
- Mandelbrot, B. and H. Taylor (1967). On the distribution of stock price differences. *Operational Research* 15, 1057–1052.
- Milonas, N. (1986). Price variability and the maturity effect in futures markets. *Journal of futures markets* 6, 443–460.
- Modigliani, F. and R. Sutch (1966). Innovations in interest rate policy. *American Economic review* 56(1/2), 178–197.
- Myers, R. and S. R. Thomson (1989). Generalized optimal hedge ratio estimation. *American Journal of Agricultural Economics* 71, 858–867.
- Peroni, E. and R. McNown (1998). Noninformative and informative tests of efficiency in three energy futures markets. *Journal of Futures Markets* 18, 939–964.
- Phillips, P. B. C. and P. Perron (1988). Testing for a unit root in time series regression. *Biometrika* 65, 335–346.
- Rabemananjara, R. and J. M. Zakoian (1993). Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics* 8, 31–49.
- Rachev, S. T. and S. Mittnik (2000). *Stable Paretian Models in Finance*. John Wiley and Sons, Financial Economics and Quantitative Analysis Series.
- Richter, M. and C. Sorensen (2002). Stochastic volatility and seasonality in commodity futures and options: the case of soybeans. Working paper. University of Copenhagen Business School.
- Routledge, B. R., D. J. Seppi, and C. Spatt (2000). Equilibrium forward curves for commodities. *Journal of Finance* 55, 1297–1338.
- Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial management review* 6, 41–63.

- Scheinkman, J. and J. Schechtman (1983). A simple competitive model with production and storage. *Review of Economic Studies* 50, 427–441.
- Schwartz, E. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance* 52, 923–973.
- Schwartz, E. (2006). Cointegration and the application to fund management. University of Witwatersrand.
- Sephon, P. S. (1993, februar). Optimal hedge ratios at the winnipeg commodity exchange. *The Canadian Journal of Economics* 26(1), 175–193.
- Smith, J. and E. Schwartz (2000). Short term variations and long term dynamics in commodity prices. *Management Science* 47, 893–911.
- Taylor, H. (1986). *Modelling Financial Time Series*. New York: Wiley.
- Telser, L. G. (1958). Futures trading and the storage of cotton and wheat. *Journal of Political Economy* 66(3), 233–255.
- Teräsvirta, T. (2006, December). An introduction to univariate GARCH models. SSE/EFI Working Papers in Economics and Finance, No.646.
- Tolmasky, C., D. (2002). Principal components analysis for correlated curves and seasonal commodities: The case of the petroleum market. *The Journal of Futures Markets* 22, 1019–1035.
- Working, H. (1948). The theory of the price of storage. *American Econometric review* 39, 1254–1262.
- Yan, X. (2002). Valuation of commodity derivatives in a new multi-factor model. *Review of Derivatives Research* 5, 251–271.
- Zakoian, J. M. (1994). Threshold heteroscedastic models. *Journal of Economic Dynamics and Control* 18(Zakoian, J. M.), 934–955.