ORBIT - Online Repository of Birkbeck Institutional Theses

Enabling Open Access to Birkbeck's Research Degree output

A study in the financial valuation of a topping oil refinery

https://eprints.bbk.ac.uk/id/eprint/40232/

Version: Full Version

Citation: O'Driscoll, Patrick J. (2016) A study in the financial valuation of a topping oil refinery. [Thesis] (Unpublished)

© 2020 The Author(s)

All material available through ORBIT is protected by intellectual property law, including copyright law.
Any use made of the contents should comply with the relevant law.

Deposit Guide
Contact: email

A study in the financial valuation of a topping oil refinery

Thesis submitted for the degree of PhD

Birkbeck College, University of London

May 2016

by ***Patrick J O'Driscoll***

Declaration

I declare that the work in thesis is my own only, Patrick J. O'Driscoll.

Signed...

Date: ....May 2016.

Abstract

Oil refineries underpin modern day economics, finance and engineering – without their refined products the world would stand still, as vehicles would not have petrol, planes grounded without kerosene and homes not heated, without heating oil. In this thesis I study the refinery as a financial asset; it is not too dissimilar to a chemical plant, in this respect. There are a number of reasons for this research; over recent years there have been legal disputes based on a refiner's value, investors and entrepreneurs are interested in purchasing refineries, and finally the research in this arena is sparse. In this thesis I utilise knowledge and techniques within finance, optimisation, stochastic mathematics and commodities to build programs that obtain a financial value for an oil refinery. In chapter one I introduce the background of crude oil and the significance of the refinery in the oil value chain. In chapter two I construct a traditional discounted cash flow valuation often applied within practical finance. In chapter three I program an extensive piecewise non linear optimisation solution on the entire state space, leveraging off a simulation of the refined products using a set of single factor Schwartz (1997) stochastic equations often applied to commodities. In chapter four I program an optimisation using an approximation on crack spread option data with the aim of lowering the duration of solution found in chapter three; this is achieved by utilising a two-factor Hull & White sub-trinomial tree based numerical scheme; see Hull & White (1994) articles I & II for a thorough description. I obtain realistic and accurate numbers for a topping oil refinery using financial market contracts and other real data for the Vadinar refinery based in Gujurat India.

Declaration 2

Acknowledgements 13

Introduction 15

Background and main contribution 15

Assumptions 18

Current state of research 20

Chapter 1 23

1.0 Introduction to crude oil and its refined products 24

(Figure 1.1: A hundred years history of crude oil prices) 26

1.1 Trading crude and its products 27

(Figure 1.2: US Refinery capacity and locations in Jan 2012 - EIA) 28

(Figure 1.3: A quadratic fit to US oil extraction projection from 2015 onwards) 30

1.2 OPEC 31

(Figure 1.4: OPEC production's relationship with the WTI crude price over ten years) 32

1.3 The IEA 33

1.4 The refinery's role 34

(Figure 1.5: Typical historical GRM for a Coking Refinery over ten years) 35

(Figure 1.6: Dependency of the US on crude imports) 36

(Figure 1.7: The leading role of Western Europe in petroleum exports to the US) 37

(Figure 1.8: The US problem of growing consumption and reduced production) 38

(Figure 1.9: Count on active US refineries) 39

(Figure 1.10: Amount of product yield from a barrel of crude in the US (Retrieved 08/01/2013)) 40

(Figure 1.11: Sales volumes of gasoline over time (Retrieved 08/01/2013)) 41

(Table 1.1: Summary of UK Refining Capacity, (2010)) 42

1.5 A brief history of petroleum 43

1.5.1 The modern oil industry 44

(Figure 1.12: The oil and gas value chain) 45

(Table 1.2: Basic Features of Benchmark Crudes, Q1 2010 averages by Argus) 46

1.6. Statistics of crude oil and petroleum products over the last 10 years 47
1.6.1 WTI Crude Oil 48
(Figure1.13: NYMEX Futures price on the front month contract over 28 years) 48
(Table 1.3.1: Descriptive Statistics, WTI Oklahoma Contract and the European Brent Spot Price FOB, ($/bbl) over 20 years) 49
1.6.2 Brent Crude Oil 50
(Figure 1.14: Spot FOB European Brent over 25 years, ($/bbl)) 50
1.6.3 Gasoline 51
(Figure 1.15: NYMEX Futures price of front month Gasoline contract, ($/gallon)) 51
(Table 1.3.2: New York Harbour Conventional Gasoline Regular Spot Price FOB ($/Gallon) 52
1.6.4 Heating Oil 53
(Figure 1.16: NYMEX Futures price of front-month Heating Oil contract, ($/gallon)) 53
(Table 1.3.3: New York “Harbour” No. 2 Heating Oil Spot Price FOB ($/Gallon)) 54
1.6.5 Kerosene 55
(Figure 1.17: Spot prices of Kerosene FOB, ($/Gallon), Gulf Coast) 55
(Table 1.3.4: U.S. Gulf Coast Kerosene-Type Jet Fuel Spot Price FOB ($/Gallon)) 56
1.6.6 Diesel Fuel 57
(Figure 1.18: Spot prices of Diesel, ($/gallon), New York) 58
(Table 1.3.5: Los Angeles, CA Ultra-Low Sulphur Diesel Spot Price ($/Gallon)) 58
1.6.7 Naphtha 59
(Table 1.3.6: OSN Naphtha CFR Japan Front month Contract ($/tonne)) 60
1.6.7 Fuel Oil 61
(Figure 1.19: Refiner prices of fuel oil, ($/gallon)) 62
(Table 1.3.7: US residual fuel oil ($/gallon)) 62
1.7 Refinery Optimisation 63
(Figure 1.20: Futures curve expectations) 67
(Figure 1.21: WTI/Brent spread over the last five years) 68
1.4.1 Refinery Risks 69
Chapter 2 71

2.1 A traditional valuation of an oil refinery 72

2.2 Refinery cash margin, RCM........ 77

2.3. Vadinar gross refining margin calculation 79

(Table 2.1: Reconciling Revenue from the refining part of the business to GRM, for the period 1 May 2008 – 31 March 2009 and the 9 month period to 31 December 2009) 79

(Table 2.2: Operating Costs per barrel at the Vadinar refinery) 80

(Table 2.3: Production of the refinery broken down by refined product) 80

(Table 2.4: Revenue generated by each refined product during periods shown) 81

2.4 Valuation risks to the Oil Refinery 83

2.5 Discounted cash flow value (DCF) 84

2.6 Weighted Average Cost of Capital (WACC) 87

2.7. Free Cash Flow Method (FCFF) 90

Step 1 - Calculate the free cash flow 90

Step 2 - Forecast FCF and the terminal value 90

Step 3 - Calculate the weighted average cost of capital (WACC) 91

Step 4 - Discount the free cash flows at WACC and aggregate to obtain value of the firm 91

Step 5 - Calculate the equity value 91

2.8. Oil refinery FCF calculation 91

(Table 2.5: FCF Calculation) 92

(Table 2.6: Indian refined product demand, -'000 bbl/d) 94

2.9 Real Option Approach 96

2.10. Problems with simple real option approach 99

Chapter 3 101

3.1 INTRODUCING COMMODITY LINKED ASSET VALUATION 102

3.1. The Refiner's Optionality 104

3.1.1 Real data from an Indian Oil Refinery 105

3.1.2. The refinery's risk 110

3.1.3. Introducing refinery planning 112

3.1.4. Multi-period refinery planning 113

3.2. STOCHASTIC PROGRAMMING FOR OIL REFINERIES 114

(Figure 3.1: Crude oil distillation and fractionation) 115

3.2.1. Refinery physical flow constraints 115

(Figure 3.2: Fractionation of crude oil) 116

3.2.2. Refined oil products market data 118

(Figure 3.3: WTI crude front month contract historical price series) 119

(Figure 3.4: Diesel fuel low sulphur historical price series) 120

(Figure 3.5: NY RBOB gasoline front month contract historical price series) 121

(Figure 3.6: Daily jet fuel retail historical price series) 122

(Figure 3.7: New York Harbor No. 2 daily heating oil historical price future contract series) .. 123

(Figure 3.8: Propane front month future historical price series) 124

3.2.3. The producer midterm planning framework 125

3.2.4. Decision Variables 125

3.2.4.1. Fixed Yields 128

3.2.4.2. Fixed Blends 130

3.2.4.3. Unrestricted balances 130

3.2.4.4. Raw material availability constraints 131

3.2.5. Deterministic objective function 132

3.2.6. Solving the deterministic objective function 133

3.3 MULTI-PERIOD OPTIMISATION 135

3.3.0.3. N-period (N+1 dates) optimisation 136

3.3.2. Foundation of the oil refinery intrinsic value 139

(Figure 3.12: N period optimisation) 140

3.3.2.1. Optimisation data flow chart 142

(Figure 3.13: Flow chart process behind the refinery optimisation) 143

3.3.3. Stochastic commodity behaviour 144

(Figure 3.14: NYMEX future price data on refinery commodities) 145

3.3.3.1. The Ornstein Uhlenbeck process 146

3.3.3.2. The Samuelson effect, mean reversion and positivity 147

3.3.4. Two period - three dates, refinery optimisation 151

3.3.4.1. Discrete process for crude oil 153

3.3.4.2. Discrete process for naphtha 153

3.3.4.3. Discrete process for kerosene 154

3.3.5. Transitional Probabilities 155

3.3.5.1. Crude price probability conditions 156

3.3.5.2. Naphtha price probability conditions 156

3.3.5.3. Kerosene price probability conditions 157

(Figure 3.15 Node transitioning on the trinomial tree) 158

3.3.6. Steps for a two period dynamic optimisation 158

3.3.7. Refinery valuation using the backward recursion methodology 160

3.4. A CONTINUOUS TIME MODEL OF COMMODITY SPOT PRICE PROCESS EVOLUTION 162

3.4.0.1. The calibration of the spot price process 162

3.4.1. Generating correlated commodity forward price series 163

3.4.1.1. Approximation for generating a symmetric positive definite matrix 165

*(Table 3.1: Correlations of commodity price return series – 10 years of futures prices NYMEX)*166

(Table 3.2: Correlations of commodity price return series – positive definite) 167

3.4.2. The discrete-time spot price process 169

(Figure 3.16 Node transitioning on the trinomial tree with probability inputs) 171

3.4.2.1. Trinomial tree building procedure: Stage one 172

(Figure 3.17: Branching for the stochastic process along the trinomial tree) 173

3.4.2.2. Trinomial tree building procedure: Stage two 176

3.4.3. Forward curves calibration procedure using Levenberg-Marquardt (1944) 177

3.4.3.1. Results of forward curves calibration 180

(Table 3.3: The calibration results of the stochastic one factor model over ten years using crude future contract prices from NYMEX.*)* 180

(Figure 3.18: Crude oil model prices versus futures market prices, NYMEX) 181

*(Figure 3.19: Four period (months) trinomial tree of spot crude oil market prices.) *72 periods were calculated in practice.* 182

3.4.3.2. Delta of the oil refinery 183

3.5. NUMERICAL PROCEDURE FOR THE VALUATION 185

3.5.1. Numerical Approximation using stages to obtain a valuation 190

3.5.2. Methods available to solve the dynamic program 190

3.5.3. Mathematical issues with the valuation 193

3.5.4. Pseudo Code for the static refinery valuation 193

3.5.5. Pseudo Code for the dynamic refinery valuation 195

3.6. DYNAMMIC PROGRAMMING USING NODES 199

3.6.1.0. Set notation for optimisation 199

3.6.1.1. Decision Variables 199

3.6.1.2. Sets and Indices 200

3.6.1.3. Parameters 200

3.6.2. Expected profit maximisation 201

3.6.3. Maximisation of dynamic objective function 201

3.6.4. Solving the multi-period refinery portfolio model 202

3.6.5. Maximising of final wealth 204

3.7. Numerical Results 206

*(Table 3.6: Summary of stochastic trinomial tree optimisation of a Topping refinery with risk values, including the number of scenarios, and the CPU solution times) *All values were calculated 1000 times and then averaged.* 208

3.8. Graphical results 208

(Figure 3.20: As the number of stages input into the program is increased the solution time increases in a non linear way) 208

(Figure 3.21: As the number of years increases towards six the value begins to stabilise) 209

(Figure 3.22: The refinery owner chooses an amount of crude to refine at each monthly period over the lifetime of the refinery) 210

(Figure 3.23: At different periods the valuation increases more than others as the refinery manager optimises the decisions along the trinomial tree) 211

3.9. Out of sample stability analysis 212

*(Table 3.7: Stability associated with scenario generation procedure decision variables. *Trees were generated and optimised 1000 times)* 212

3.10 Analysis of the Valuation Model 212

3.11. CONCLUSION 213

Chapter 4 216

A CRACK SPREAD OPTION REFINERY VALUATION 216

INTRODUCTION 217

4.1.1 Crack spread effects 222

(Table 4.1: Factors affecting the crack spread value, assuming the other refined products remain the same value) 222

4.1.2. Calculating the crack spread value with traded contracts 224

4.1.3. Descriptive statistics for commodity future price returns 226

(Table 4.2: Descriptive statistics for future price returns) 227

4.1.4. Assumptions underlying the crack spread valuation 228

*(Figure 4.2: Daily calculated crack spread prices, Gulf Coast 3-2-1 underlying in $/barrel)*Prices for the crack spread are calculated as standard with the 3-2-1 contract; 1 heating oil barrel spot price + 2 gasoline barrels - 3 crude barrels, all divided by 3.(remembering to multiply gallon prices by 42).* 230

(Figure 4.3: Daily log return crack spread prices, Gulf Coast 3-2-1 underlying) 231

4.1.5. Crack spread modelling 231

4.1.6. Types of crack spread options listed on the NYMEX exchange 233

*(Table 4.3: Crack spread options listed on the NYMEX exchange*There are days when open interest and volumes on all of these products is extremely high)(Data accessed on 05/03/2013)* 234

Valuing the refinery 235

4.2.1. Data 235

(Table 4.4: Sample crack spread option prices) 236

(Table 4.5: Gasoline Crack Spread Option prices) 236

4.2.2. A strip of European crack spread options 237

4.2.3. The Linear program valuation construction 239

4.2.4. Crack spread option pricing methods 242

4.2.4.1 The Bachelier method applied to the crack spread option 242

4.2.4.2. Kirk's Approximation applied to the crack spread option 244

(Figure 4.4: Calculated crack spread prices, versus Kirk approx. of the option value $/barrel, strike at $2.44 per barrel) 245

(Figure 4.5: Calculated option price difference between Bachelier and Kirk $/barrel) 246

4.2.4.3. Alexander and Venkatramanan's approximation.. 246

4.2.4.4. Schwartz and Smith(2000) estimation applied to the crack spread option 248

4.2.5. Hull and White Two factor model 249

(Figure 4.6: Crude oil trinomial price series $/barrel) 251

(Figure 4.7: Gasoline trinomial price series $/Gallon) 251

(Figure 4.8: Heating oil trinomial price series $/Gallon) 252

(Table 4.6: European call crack spread option price errors) 252

4.3. The refinery valuation results.. 253

(Table 4.7: Oil refinery valuation methods with 100,000 simulations, values in million $. Standard error in valuation in brackets.* NYMEX call crack spread option prices over ten years) 253

4.3.1. The Updated linear program valuation results applied to the refinery 254

(Table 4.8: All methods with an updated oil refinery valuation, values in million $ Standard errors of MC in brackets. *NYMEX call option crack spread prices) 255

(Figure 4.9: Daily crack spread value ($ per barrel) versus the crude oil purchase decisions (tons); (the left y axis is $ per barrel, the right y axis is tons/100).. 256

4.3.2. Analysis 256

4.3.3. Comparison of all oil refinery valuation methods 258

(Table 4.9: Alternative valuation approaches to the oil refinery: five methods from chapter 4, three from chapter 3, one from chapter I and one from chapter II) 258

4.3.4. Refinery valuation research extensions 258

4.4. Conclusion 262

Thesis Conclusion 263

APPENDIX... 269

A.1.1. Calibration of gasoline 269

A.1.2. Calibration of naphtha 269

A.1.3. Calibration of heating oil.. 270

A.1.4. Calibration of kerosene 270

A.1.5. Calibration of cracker feed 271

A.1.6. Calibration of diesel 271

(Figure A.2 Naphtha model prices versus market residential prices, EIA) 273

(Figure A.3 Heating oil model prices versus market futures prices, NYMEX) 273

(Figure A.4. Diesel model prices versus futures market prices, NYMEX) 274

A.2.0. Calibration of Alexander and Venkatramanan (2011) 274

(Figure: A.5. m as a function of strike K on the 12/02/2014, for option data from NYMEX) 275

A.2.1. Calibration of Hull and White two factor stochastic process 275

A.2.2. Trinomial tree construction for Hull and White two factor process 276

A.2.3. Calibration results for Hull and White two factor process 278

Table A.2.3.1: The results of the calibration to crack-spread futures volatility on 12/02/2014 278

*Table A.2.3.2: At-the-money European crack spread options-volatility quotes on 12/02/2014*279

Table A.2.3.3: At-the-money Hull & White two factor implied volatilities. 279

A.2.4. Sub-tree calibration for Schwartz and Smith (2000) two-factor process 280

(Table A.2.4.1: Resultant Calibration results for Schwartz and Smith (2000) on WTI crude oil for Dec 2000 - Dec 2010) 281

(Table A.2.4.2: Resultant Calibration results for Schwartz and Smith (2000) on No.2 heating oil for Dec 2000 - Dec 2010) 282

(Table A.2.4.3: Resultant Calibration results for Schwartz and Smith (2000) on RBOB NYH Gasoline for Dec 2000 - Dec 2010) 282

(Figure A.2.4.1 WTI Crude Oil calibration results using Schwartz and Smith (2000)): 284

(Figure A.2.4.2 RBOB Gasoline calibration results using Schwartz and Smith (2000)): 285

*(Figure A.2.4.3 No.2 Heating Oil NYMEX calibration results using Schwartz and Smith (2000)):):*286

A 3.0 Comparison of forward curves for the one factor and two factor models 287

Published Paper 288

REFERENCES 315

Acknowledgements

I would firstly like to thank my supervisor, Professor Helyette Geman, for providing me with the original idea and the environment in which to work undistracted and relentlessly. Secondly, I would like to thank my second supervisor, Professor Raymond Brummehlius, who has provided the support of my chapters to ensure that they have passed at each stage of the PhD process. I would also like to thank Professor Ron Smith, for his general advice. I hope that the practical approach to real option pricing in this thesis will be a contribution to the literature, that at least, sets some threshold for others in this very exciting and practical area to overcome or leverage from.

I would like to thank Reflection, a small statistics company based in Soho, for providing me with the statistical forecasting practical paid work; using R for regressions and other relevant statistical exercises. This work enabled me to continue the PhD well past the 3rd year after the government funding ran out; I wish them well in their endeavors. I would also like to thank the financial accountancy team at Hermes Investment Management for their contribution and help with accountancy techniques for chapter two. Thanks also to Spiral software in Cambridge for their suggestions and input on the optimisation coding carried out in chapters three and four.

Originally this thesis was supposed to be unfinished business from 2003, whence I quit a PhD in astrophysics to join the great befallen Merrill Lynch – although I never rue decisions, this path was certainly the harder route; there have been many lessons, from extreme patience, managing the nuances of an environment that should be rewarding meritocracy, to knowing what is in ones control and what is not. As the original thesis was to be dedicated to my mother Patricia Johnson O'Driscoll, who I owe it all to – a very eloquent and sophisticated woman who knew that life is all about being happy; it is not too far that I have to look to find the next person to thank, my father Patrick O'Driscoll, without his parenting and disciplinary lessons I would not have succeeded at

UCL or in the dojo – a true practical person who epitomises work ethic and practicality; and under whatever strenuous circumstance one is under he always taught: “one cannot keep a winning man down, in the end he will get there”.

I would also like to thank my fellow PhD students for providing support and the necessary adversity when required; without the competition and the great discussions with Georg Dettmann, Tobias Grasl, Alexander Karalis and Marco Pellicia, the interest for this thesis would have left me a long time ago.

Finally, I would like to thank my wife Ying Wu-O'Driscoll, for her constant supervision, advice and focus, she has truly been inspirational. This thesis is dedicated to her; without her love I would also have been nowhere, to quote a very famous individual: “we are lost in the water, we must close the door, we are done for” - luckily I was never close to feeling this way, but was looking for fun with this thesis and creating something that others may be able to use somewhere, somehow.

Introduction

Background and main contribution

Without the flow and refining of petroleum, planes would be grounded, plastic containers ridiculously expensive, butane for lighters non existent, the heating of homes all over the US would be impossible to maintain. Without this multi-billion dollar industry, many would be unemployed. It is without doubt one of the most significant and interesting sectors of the global economy and refineries themselves are the foundation of the flow from raw crude to the final products that the market purchases.

In recent years many refineries have gone under due to the difficult climate; in the seven years up to January 2015, the four main EU countries in refining lost a total of 2M barrels per day, b/d, of operable refining capacity, that is approximately 23%, *IEA (2009-2015)*. The UK has closed two refineries in the last five years, with France closing four. The WTI and Brent crude oils have hit huge lows, with Brent going below \$30 per barrel in Jan 2016 and WTI at \$31 per barrel in the same period. This benefit to the consumer has been contrasted by a volatile environment and oil organisations altering their business and investment strategies. However, refineries will be required as long as there is a demand for kerosene, butane, gasoline, and others, but the processes need to be more efficient, more environmentally friendly, and the complexes themselves, will not survive unless their complexity increases inline with the most advanced refiners. The decision to alter output product prices needs to be done in a way that means, volumetrically, profits are maximised within applicable constraints.

High tech refineries can refine very heavy raw crude; regarded as lower quality due to the waste inherent in this type of oil but it is in a businesses ability to adapt that ensures its survival. A measurement named the Nelson Complexity factor measures the technology level of a refinery; a number of 9 is regarded as a high level; yet the highest currently is the former BP Texas refinery, acquired by Marathon Petroleum with an index of 15.3. To expand the complex up to this level can be expensive, but in the right local or market region it can be extremely profitable – enabling the refinery to serve overseas customers with inexpensive crude input. There are many types of software packages in use at a refinery complex, from health and safety reasons to optimising the quality within the crude distillation unit, or CDU, and the catalytic cracker; in this thesis we investigate how optimisation of the volumetric decisions affects the profit for the owner, who can buy crude oil on the market, and choose to produce as much as physically possible of the best performing petroleum products. This, to our knowledge, is yet to be achieved in the literature. It is important to note that we are in a setting of incomplete markets as we cannot replicated the value of the refinery with available assets in the market - the refinery is an illiquid asset in a part of the world without relevant available financial contracts. We attempt to produce a credible range of real asset values; in that they are realistic in the sense that the approach to represent value is logical and representative of the oil refinery's business incorporating the underlying commodities. Whilst using a risk neutral assumption on the price return series of the commodities the embedded optionality of choosing an alternative refined product is captured by the numerical method presented. There is in fact no available method to remove the uncertainty in the refinery cash flows; with this in mind we include a value at risk figure on the final wealth of the refinery applicable to the horizon represented. Due to this real option valuation considering all seven refined products it means that the owner has the embedded optionality of many more products than usually considered in refinery approaches. If the correlated commodity price trees are constructed correctly it means that we have a numerical method that can value the refinery based upon this choice of refining more or less of a particular product at a certain point in time.

The spine of the thesis is the topping refinery itself - one of four standard types of refinery and considered the most basic; this category of refiner has been chosen despite the cracking refinery, a very advanced refinery, being ubiquitous, as the linear programs used in industry are very difficult to obtain for investigative purposes, whereas for the topping programs there are many open source code bases which can be accessed. Throughout the thesis a standard linear program for a topping oil refinery, which is detailed in the appendix is used as the foundation for each valuation. In addition to providing a valuation, the optimisations provide a number of other useful outputs not explored in the literature. It is also important to note that wherever in the thesis there is a calibration present it is the returns of the prices that have been input into any program and not the absolute prices - despite this as frequently evident in the older papers like Schwartz (1997) we graph the price series and not the returns to display the movements of the commodities in reality.

In this thesis we firstly introduce the industry and describe the characteristics of the products in a statistical analysis; petroleum product moments and events in the time series are analysed along with an introduction to the mathematical programming literature available in the refining space.

In chapter two, we describe a standardised discounted cash flow valuation, which is representative of the financial market's way to value a real asset. Many auditors and finance professionals will use an income based approach to value an asset. In this chapter we discuss the various approaches and the lag to the more advanced methods within the industry, like the real option approach.

In chapter three we go beyond standard calculations and introduce a real option valuation; that is more complex and computationally more demanding, but proves to be a worthwhile investigation, and in terms of the uncertainty within the prices, continuous stochastic equations are chosen. It is trivial to discretise them onto a relevant simulation tree to enable a dynamic program to be optimised at each node. The choice of equations is based on models required that would enable a number of aspects to be included for a realistic and dynamic commodity price series - the valuation is higher than DCF reveals and despite being an incomplete market settings gives a range of values that represent the optionality for the refinery owner.

Finally, in chapter four we find ways to value the complex, using a dimensional shortcut based on using the crack-spread rather than seven prices, which can be compared to other real option valuations within an oil refinery linear program. Due to the simplification introduced in chapter four we can elaborate on the complexity of the underlying stochastic processes. To our knowledge, the valuation procedure introduced in chapter four is the only real option valuation of an oil refinery in the current literature including the option pricing analyses. The method we use is also a unique combination of stochastic mathematics to represent the commodity prices over time and numerical methods to obtain a value. The program created is able to manage the extensive state space due to the use of dynamic set assignment; to our knowledge not applied in the refinery literature. Too many studies, theoretical or numerical, avoid implementing a model if solving over a colossal dimensional space.

We provide various alternative but realistic valuations in chapters two to four and the paper published in the appendix. These procedures open the door to further areas of research worth studying. The main contribution of this thesis is to provide a valuation for a refinery complex. An additional contribution is the combination of the stochastic simulation, with the discretisation onto a trinomial tree, into the eventual dynamic program implemented in GAMS, which introduces a number of computational shortcuts to find the valuation - including the dynamic set construction.

Assumptions

The main assumptions in obtaining a valuation in this thesis are: investors are rational and we are making calculations in a risk neutral setting; despite being in incomplete markets with no available financial contracts for replication. Many refinery papers in this area assume the same, as it allows a depth of equations and techniques for option pricing that can be manipulated without additional extraneous tampering, and they can be justified in a real market.

One of the common issues within pricing is finding a market price of risk, which is representative in various scenarios. The market price of risk is the return expected above the risk free-rate that the market demands as compensation for taking the specified risk. In classic economics no rational investor would commit their own resources without expecting to earn above the risk free rate. The market price of risk is usually measured as a ratio of the expected return for the asset over the risk-free rate divided by the standard deviation of returns. In option theory we often attempt to model the state variables as random. This stochasticity leads to risk, how much should we expect to earn for taking this risk? Is this quantity directly tradable in the market?

If the quantity is traded directly in the market then we can hedge away the risk in an option by dynamically buying and selling a particular quantity of the underlying. The refinery does not have this luxury due to it being an illiquid real asset in a remote and deregulated market. Despite the issues we aim to construct a risk neutral valuation that takes into these practical considerations and provides a statistically correct value. Throughout we assume no financial trading transactions costs, as is commonly applied in option pricing as it is not the focus of our problem. We do not consider stochastic interest rates however, but apply a deterministic risk free curve whenever discounting is required, built using standard products for treasuries and swaps using appropriate data from Bloomberg.

Where we use specific techniques, we explicitly state the assumptions made, for example in chapters three and four we use a Hull and White trinomial tree (1994). There are very standard assumptions for this tree, like the size of the step relative to the volatility, along with others described in detail later in the thesis.

The major assumption inherent in the final two chapters, which carry out unique valuations and comparisons respectively is that the oil refinery asset value can be represented as a daily strip of monthly European call options on the crack spread. This is described in depth in chapters three and four, but simplistically means the owner has optionality on each day to defer or execute production at the refinery – the advantages and disadvantages of our approach is discussed in depth. In terms of the refinery itself we ignore taxes and transaction costs, along with preventative maintenance and assume start up and shutdown costs are considered insignificant.

Current state of research

As mentioned there is no current oil refinery valuation methodology within either the engineering or finance literature, whereas in practice, a very basic and flawed method, DCF, is commonly applied to real assets; see Brandao (2002 & 2005) for a detailed discussion on replacing DCF with real option approaches. There is an extensive literature on refinery planning and on balancing the fluid flows in chemical engineering; Neiro and Pinto (2003), Pongsadki (2005) and Nan and Marc (2013) all provide refinery planning optimisations at different timescales. There is also a literature on the financial valuation of real assets utilising real option approaches, but the relevant techniques are yet to be applied to oil refineries. This is the first such study.

Benyoucef (2010) considers a mixed integer non linear program applied to a network of refineries, but not an individual complex and without a financial valuation being the objective. Pindyck and Dixit (1994) develop the theory, especially in the dynamic programming space for real option valuations, but do not implement the methods to actual plants or any real assets. Hull and White (1994-1995) provide many numerical and mathematical tools applied to real options and the stochastic simulation of the relevant equations for forecasting purposes, yet no valuation for a refinery is provided. There are a number of engineering papers: Lin et al (2009) and Cao et al (2009), and others, that consider the oil refinery owner's decisions over one period (commonly known as two stage stochastic programming), although very similar in terms of the objective function, and also being extremely helpful, the formulations are basic when compared to a multi-period solution, hence used as a foundation for this thesis. Ribas et al (2012), examine the detailed fluid movements within a Brazilian refinery; the authors optimise at a granularity that is extremely detailed and useful in terms of the mathematical program constructed; yet we are interested in the financial ramifications. Leiras et al (2011) collect refinery planning papers and provide a thorough comparison of methods and results; very few authors in these studies attempt to capture the uncertainty of petroleum prices, especially using continuous stochastic equations, which are adept at capturing the relevant bespoke commodity behaviour. In this study there are many two stage programs, discrete normal simulations but rarely are there financial objectives.

We leverage off these previous studies, that use linear programming for financial goals; Aldaihani and Al-Deehani (2010) examine the efficiency of the Kuwaiti stock exchange using an optimisation for an equally weighted basket of stocks – despite its simplicity in terms of the dimension of the problem, the authors find a way to lower risk for investors for a threshold of specific return. Elkamel and Al-Qahatani (2011) use discrete scenario generation, where an assumption of the probability of each scenario is applied to a PVC complex on top of a refiner – a sample average approximation is applied; despite being non-deterministic, the method leaves space for stochastically fluctuating uncertainty. The type of problem we are developing comes under the bracket of tactical midterm refining planning; this is very complex and can be optimising different units in the refinery for example, efficient catalytic cracking – in our case we concentrate on the net present value over multiple periods.

Decision based systems at refiners aim to aid the process; PIMS (AspenTech) is a complex system designed to perfect the chemical flows within the parts of the refinery to create the best product possible. It utilises mathematical programming and this is the most common system implemented at refineries in the world (Kelly and Mann, 2003). Wang (2013) manages the optimisation by using a finite planning horizon with periodic preventative maintenance, and random failure; however, his optimisation is discrete and cannot manage the dimensional problem. An extensive literature review is provided at the start of each chapter relevant to that chapter's objective. Before moving onto the first attempt at valuing the refinery we provide some context by introducing the refining industry and lay the foundations for the mathematical models introduced in later chapters.

Chapter 1

An introduction to the oil industry and its refineries

“It is wise to apply the oil of refined politeness to the mechanism of friendship.” Sidonie Gabrielle Colette.

In this first chapter we begin with a discussion of the petroleum industry and its most important products, i.e. the refined outputs from a typical cracking or topping oil refinery. In recent months prices on the WTI and Brent products have dropped dramatically – the effects in the outer economy are widespread, but within the oil industry itself it is unprecedented. We leave the fundamental drivers of this market within the macro economy to others but emphasise the detailed refinery processes and their importance. Next, we give a grounding in the basics of the technical indicators of crude and its by products; along with some historical context and applications of modelling within the oil industry – highlighting statistical moments and graphing typical behaviour of the spot prices. The refining sector is a multi billion dollar industry with huge impacts on the global environment, global economy and the global energy industry, hence the significance of the prices of the products derived from crude and its refining. Oil refineries are studied in detail within chemical engineering, but analysis of the optimisation of decisions of a refinery owner and its knock on effects to the market products is sparse – the financial impacts are even more rare; leaving space for investigation and discovery in linking financial and oil refinery analytics. Here we discuss the individual refined products, their behaviour and their significance in the refiner's *decision process. We conclude with a discussion of the risks associated with refineries, and the main products utilised to managed these risks.*

The goal in this thesis is to investigate the details of an optimisation model that will maximise the profit of an oil refinery, subject to a standard set of *Topping* refinery constraints. The final model is required to be optimal from a statistical perspective in that probability distributions are obtained for the commodity prices, and their financial risk can then be included if required. There are a number of standard models derived in the 1950 and 1960s that form the foundation of the current mathematical programming models developed for refinery optimisation software. These ideas underpin the optimisation model throughout the thesis.

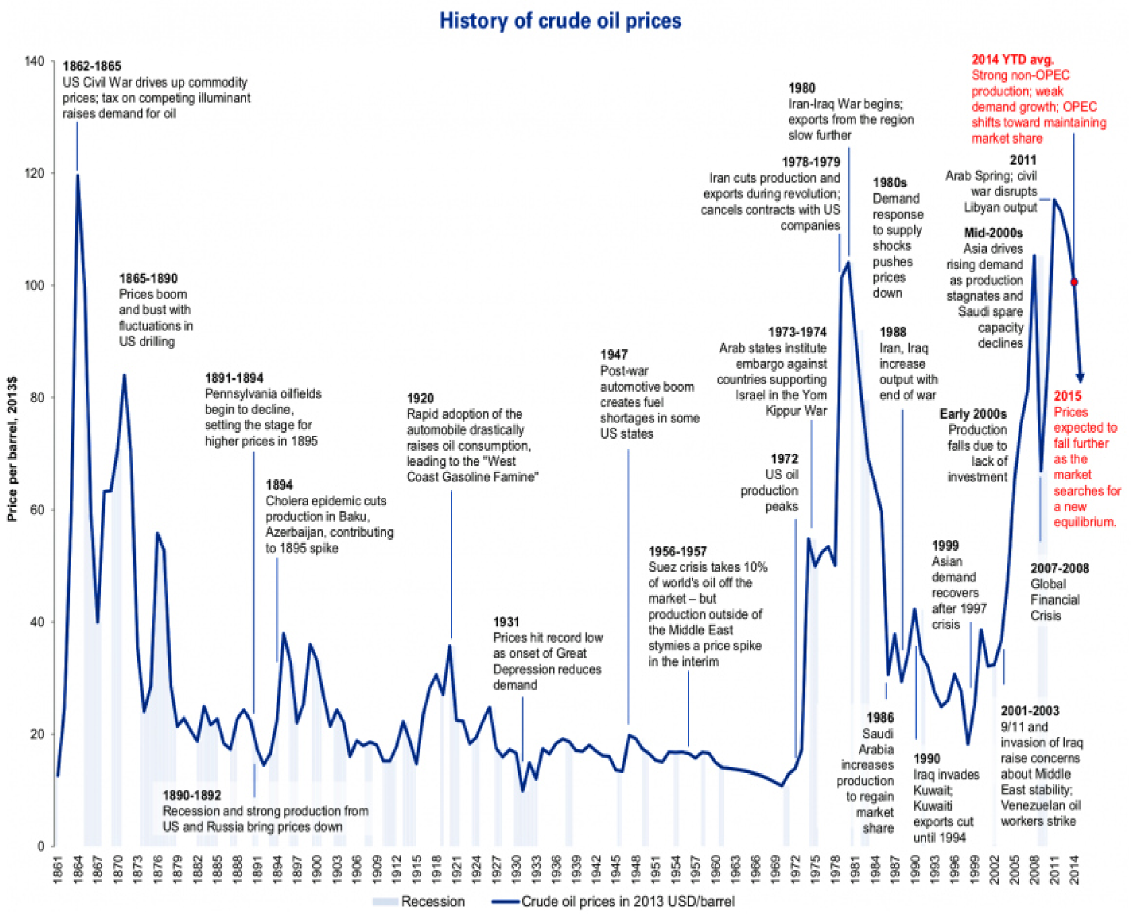
1.0 Introduction to crude oil and its refined products

The subject of crude oil and its products is so rich and varied that it cannot be fully covered; hence a number of issues will not be discussed in this work, for instance the technicalities of oil production and exploration or the details of crude transportation across the globe and its oceans. However, ideas based upon why the world has depended on oil for centuries, and will continue to do so for a long time yet, are described in this chapter. Secondly, what constitutes crude oil and its refined products, and what practical uses they have is exemplified. Finally, the refined product price series characteristics are analysed through statistics applied to the data over the last ten years.

The price of crude oil affects us all; it has direct and indirect effects on inflation, cost of production in corporations, mining activities and agricultural farming. Its price volatility moves the value of some of the largest funds heavily invested in the petroleum family, and it is often imputed as one of the causes for Governments to go to war. There are many variations of raw crude oil and its refinement upholds the entire transport industry due to the gasoline product. The diverse products refined from petroleum are classified by a number of properties, the most important ones being: American Petroleum Institute (API) gravity (a measure of how heavy or light a petroleum fuel is relative to water, most values fall between 10 and 70 API gravity degrees); the "cetane" number (a measure of light distillate diesel oil's ignition delay); octane number (a measure of resistance to detonation of the fuel) and sulphur content. Crude oil is considered light if it has a low density, and sweet if it contains little sulphur. There are a profuse number of factors that influence the type of products that a refinery will create, the primary ones being the modernity of the refinery and the distilled products' prices in particular. Most refineries are at ports due to the expensive costs of the raw crude transportation[1]. The most abundantly refined output product is petrol (gasoline in the US) for use in the transportation industry. A higher yield of petrol comes from light crude and there are fewer environmental problems with the sweet type; hence, light sweet types of crude are more expensive. The benchmark in the US and all over the world until 2009 was the West Texas Intermediate (WTI) crude. It is held, after extraction, in huge facilities before travelling across America's pipelines. It is a high quality, sweet, light oil delivered at Cushing, Oklahoma as far as the NYMEX Futures contracts are concerned (this financial contract trades in units of 1,000 barrels). The U.S. Governments' decontrol of oil prices, starting in January 1981, influenced the volumes and transparency of the spot markets in oil. The decontrolling of oil prices created the circumstances for the development of the WTI contract in 1983 priced on crude at Cushing. More recently, the Brent contract has replaced WTI, in that it is used as a benchmark for over two thirds of the world's crude oil. The differential between WTI and the Brent oil contract is a large area of research, and many experts believe this shift is only temporary until WTI is not landlocked. The Brent blend in comparison, entails 15 oils from different fields in the East Shetland Basin.

[1] The Port of Antwerp provides an example of a refinery cluster: http://www.portofantwerp.com/en/chemical-cluster

Benchmarks for crudes were fashioned as there are globally many different varieties of the raw fuel demanded. The Brent crude is regarded as the benchmark crude in Europe; Dubai-Oman is used as the benchmark for the Middle East. In this recent decade the Platts Dubai benchmark has become the reference for deliveries of crude out of the East Siberian port of refineries: Kozmino; its price on Sep 28th 2015 was $44.5 per barrel. The physical is underpinned by several million barrels per day derivative spread trades on the Brent/Dubai difference. In practise, the physical trade is miniscule in comparison to the paper market; paper contracts have provided investors with the instrument to buy and sell this versatile commodity in a liquid, hence reliable way. Figure 1.1 captures a history of crude oil and some significant events.



(Figure 1.1: A hundred years history of crude oil prices)

1.1 Trading crude and its products

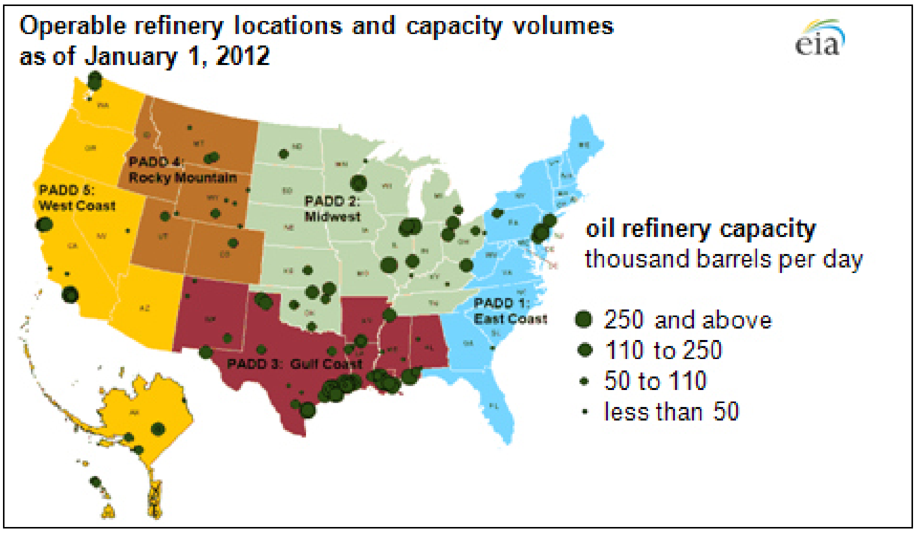
The first Futures contracts can be traced to European trade fairs in the 12th century. In this period, travelling with large quantities of goods was perilous. Vendors resorted to taking the road with simple samples and sold Futures for quantities to be delivered at future dates. These contracts have become the primary mechanism for trading crude oil and other fossil fuels in the modern financial arena. Crude oils, as discussed, can be of various compositions, enabling a medley of products to be processed. Trading of these various oil and oil related products, is carried out all over the world at exchanges with specific contracts. The exchanges transacting the largest volumes are the New York Mercantile Exchange (NYMEX) and the Intercontinental Exchange (ICE), which absorbed the International Petroleum Exchange (IPE) in June 2001. Dubai, Mumbai and Tokyo have become the other large trading platforms for crude oil. These financial contracts enable oil companies, refiners and investors to buy or sell crude oil for future dates. Trading of these contracts on Exchanges directly determines the market prices for crude and its products. Refineries are interested in buying crude in physical form, and selling the distilled products back to the market. However, due to the growth of the financial markets in the oil industry, derivative contracts related to an asset "on paper" have seen a meteoric rise in transaction volume. Theretofore, paper contracts are frequently utilised to hedge the refinery's exposure to the volatility of crude prices.

Refineries create a variety of products using processes known as distillation, cracking and purification. Refinery products demanded by the market are liquid petroleum gas (LPG), gasoline, fuel oil, heating oil, naphtha, residues and many others. Therefore, crude oil needs to be first refined to be valuable, hence useful to the end customer. For example, gasoline is refined in the largest cut of the slate due to its use as a fuel for vehicles; retail prices at the "pump" in the UK in mid 2012 were around 149 pence per litre[2]. There are different grades and combinations of products that can be distilled to produce over 4,000 petrochemical products for the market. Once

---

[2] BP ultimate 28/06/2012

refined, the uses of crude throughout industries all over the world are phenomenal. In the US and Canada, the transportation industry leads all others in volumes of crude processed by a large margin. Elsewhere, it is exploited for energy production and is globally a key ingredient to the agricultural industry that feeds the world's population.



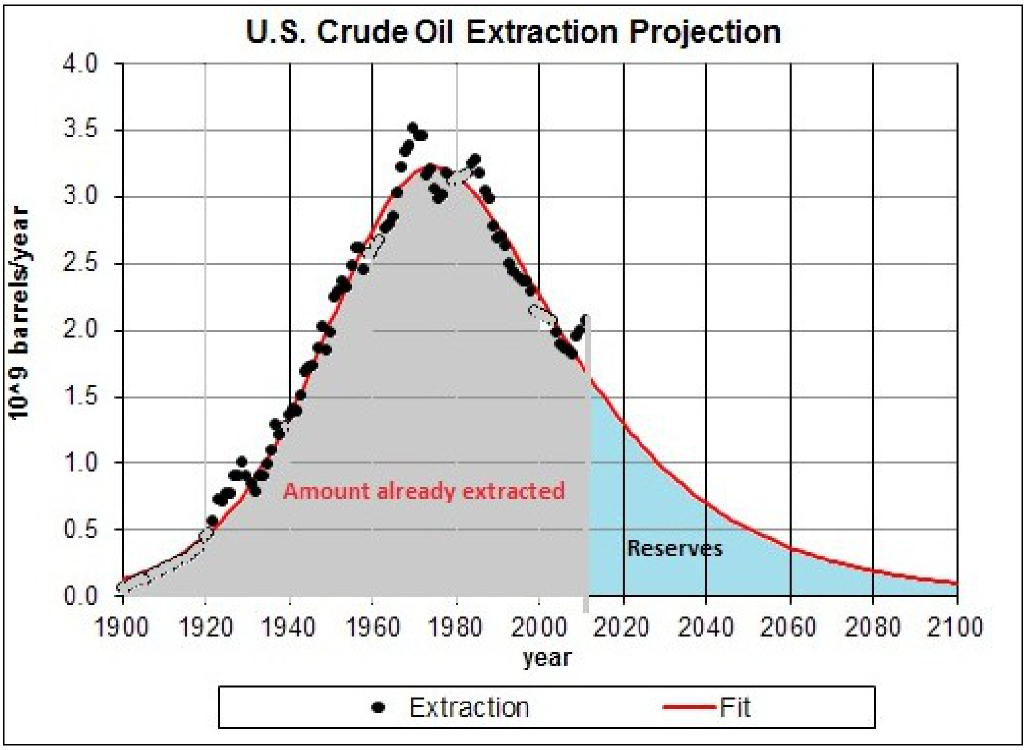
*(**Figure 1.2:** US Refinery capacity and locations in Jan 2012 - EIA)*

The survival of refineries is dependent upon the type of crude they can buy and the technology available at that particular refinery - the raw crude and its quality can make or break the local region's refinery; see figure 1.2 for density of refineries on the Gulf Coast of the US. Market data on crude oil has an indispensable impact: prices, reserve levels and volumes traded on the financial markets all act as significant indicators of the worldwide economy and global trade levels. Current estimates from the Energy Information Administration (EIA) state that 30% of the world's energy needs are being met by crude oil. On average, a substantial proportion of a country's GDP is directly accounted for by the energy sector - for example, the UK's energy sector was 4.4% of GDP in 2011[3]. Products from petroleum account for huge proportions of a number of countries

3 DECC UK Government research, 2011

export's value. For example, the UK exported 38 million tonnes of petroleum-related products in 2011, roughly equal in value to $26 billion.

Environmentalists are concerned about the crude oil sulphur content, and therefore try to ensure governments increase regulatory specifications. The mechanism that Governments usually employ to meet these specifications is to introduce a specific taxation based on the amount of sulphur contained in the crude - for example, in 2012 the UK government was attempting to introduce increases in taxes on refined products[4], but not without fierce opposition from the Grangemouth refinery in Scotland. As this wave of opposition has grown, the oil industry has adapted. Many companies now have entire budgets devoted to green technology and corporate responsibility. For instance, advanced technology based car companies are attempting to reduce transportation reliance upon crude oils due to its diminishing reserve numbers. The Gas Journal (Dec 21st 2009) states that Saudi Arabia holds the world's largest reserves at 262.4 billion barrels, but these are rapidly declining.



[4] http://www.falkirkherald.co.uk/news/local-headlines/grangemouth-meeting-gets-the-message-across-1-2537090

(**<u>Figure 1.3:</u>** *A quadratic fit to US oil extraction projection from 2015 onwards)*

A number of bespoke companies now invest in research that attempts to obtain crude oil reserve numbers to predict depletion rates. Crude oil reserve forecasting is usually calculated using Hubbert curves. Recent research predicts that at the current pace of exploration and development, the world has less than 50 years left of traditional oil[5]; see figure 1.3 for a standard quadratic fit to EIA crude extraction figures. In 1956, Hubbert formulated a model for estimating future production for fossil fuels. His main assumption was that after reserves of coal, oil and other fossil fuels are discovered, the production curve follows a distinct shape. Production at first increases exponentially; at some point a peak is reached, and finally production brings an exponential decline. Hubbert curves on crude take into account current oil reserves; forecasted usage; forecasted discoveries and copious other factors. In response to the data on depletion rates, unconventional oil extraction techniques have sprung up in the last ten years. For example, oil sands – an unconventional extra heavy petroleum deposit - represent 97% of Canada's oil reserves. Despite the fact that it will be many years before reserves have noticeably diminished, investment from a proportion of companies has noticeably shifted towards those product areas that could replace or substitute refined products. In the short term however, a sudden supply blockage, unusual weather or other logistical issues will dramatically impact the pricing of refined products. This statement holds true only if an entire region important to the refining market is impacted, for example the Gulf Coast of the United States - in 2005 when hurricane Katrina hit the Gulf Coast, Florida and Texas, its effects had a major impact on gasoline prices, amongst the other disastrous environmental effects. WTI oil Futures reached record highs of $70 per barrel and gasoline prices more than tripled in areas such as Louisiana. On August 31st, eight of the Gulf of Mexico refineries remained shut down. By September the 7th, Gulf oil production had returned to 42% of normal levels. Out of

---

5 Gas Journal (Dec 21st 2009)

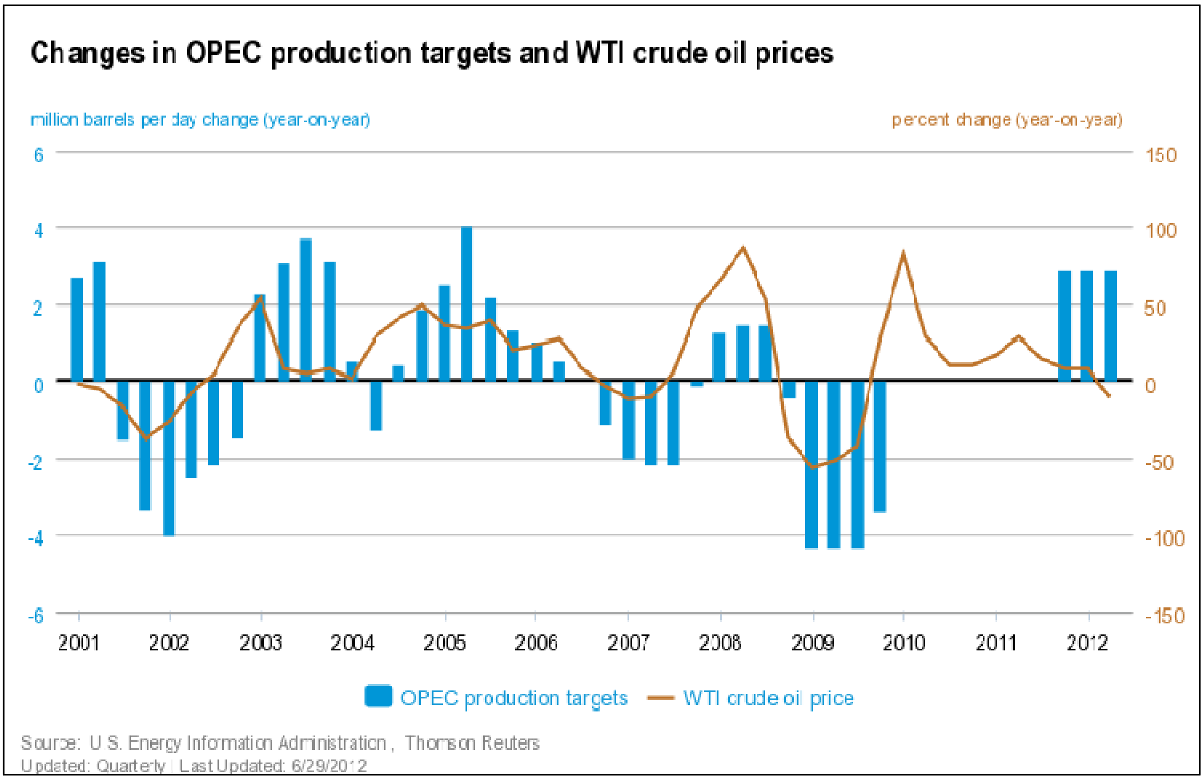
ten refineries that were closed, four were back at full capacity within a week, whereas the other refineries were out of commission for months.

In the midterm, under normal market circumstances, it is the cost of refining that has the largest impact on the market price of oil products. In what is regarded as the standard oil market model, spare capacity at refineries used to be common; these days, it has reduced dramatically. In 2003 spare capacity had dropped from 3 million barrels per day (bpd) to 1 million barrels per day. During the gasoline season (April to August), the price of gasoline can reach highs of $3.406 per Gallon to consumers[6]. In general, demand increase occurs every year in the summer period. There are innumerable other factors that can impact oil prices, including wars, political interventions, government policy announcements and news from organisations such as OPEC.

1.2 OPEC

The Organisation of the Oil Exporting Countries, OPEC, was founded in Baghdad, Iraq, in 1960. It has its' headquarters in Vienna and is an inter-governmental group of 12 oil producing states, ranging from Algeria to the United Arab Emirates. It was known as the Organisation of Arab Petroleum Exporting Countries or OAPEC (consisting of the Arab members of OPEC, plus Egypt, Syria and Tunisia) in the early 1970s. One of its principal goals is the safeguarding of the governments involved in the organisation. Another is stabilisation of the oil price within international markets; ensuring a predictable income for those nations involved, and predictable costs for importing countries. Despite these published goals, OPEC was blamed for the 1973 oil crisis by implementing oil embargoes in response to the U.S. decision to re-supply the Israeli military during the Kippur war.

[6] EIA website April 2012



*(**Figure 1.4**: OPEC production's relationship with the WTI crude price over ten years)*

OPEC was the key provider to the West for oil and in the past, maintained control of the prices. Since the discoveries in Alaska and the Gulf of Mexico, the control of prices has loosened somewhat, nevertheless the members of OPEC still own somewhere near 80% of proven world crude oil reserves[7].

In 1974, after the end of the Bretton Woods agreement, OPEC announced that they would price a barrel of oil against gold[8]. Following these events, the volatility and price of oil both increased and have remained in and around this new threshold since. The overall effect of the Arab oil embargo in economic terms was to increase the real price of crude oil to the refineries. In 1972, the price of a barrel (bbl) of crude oil was $3; by the end of 1974 it had quadrupled to $12 per barrel. In recent times, the price of WTI crude has hovered around $85 a barrel (after a peak to

[7] OPEC's annual statistical bulletin (2012).

[8] Cohen, Benjamin. "Bretton Woods System", 28/06/2001.

$140 in July 2008 and a collapse to $38 in December 2008). In 1988 OPEC decided to use Brent, underlying the IPE/ICE Futures contracts, as the reference price of crude oil in its trading activities. Note that only 25% of the physical market at anytime constitutes the commercial crude oil market. The price variations depend highly on the choices made at the heart of the value chain, i.e. the refinery. As figure 1.4 shows, oil price changes are correlated with OPEC production levels. Oil related businesses, investors and the general public are the parties that can suffer from high or volatile oil prices.

1.3 The IEA

The International Energy Association is an autonomous Paris based intergovernmental organisation established in 1974 in response to the 1973 oil crisis[9]. It acts as a policy advisor to 28 member states. Its aims are to ensure that oil policies are focused on energy security, economic development and environmental protection. It comes under the umbrella of the Organisation for Economic Co-Operation and Development (OECD) – it originated in 1961 to stimulate economic progress and world trade. The OECD was first established to help administer the Marshall Plan for reconstruction of Europe after World War II. It is also famous for promoting climate change related policies and alternate energy sources. In 2011, the IEA stated that the $409 billion of fossil fuels subsidies were the cause of wasted energy and a viable policy to address this problem would be to cut subsidies.

The influence of organisations like OPEC is certainly significant for the price of crude; the IEA is less influential in constructing environmental and economic policies, but can act on preponderant short-term issues: for instance, it intervened on the oil markets three times by releasing oil stocks in 1991 during the Gulf War; in 2005 for a month after Hurricane Katrina, and

[9] http://www.iea.ma/fr/

recently in 2011 to offset the Libyan civil war. IEA member countries are currently required to maintain stock levels equivalent to 90 days of the previous year's net imports.

1.4 The refinery's role

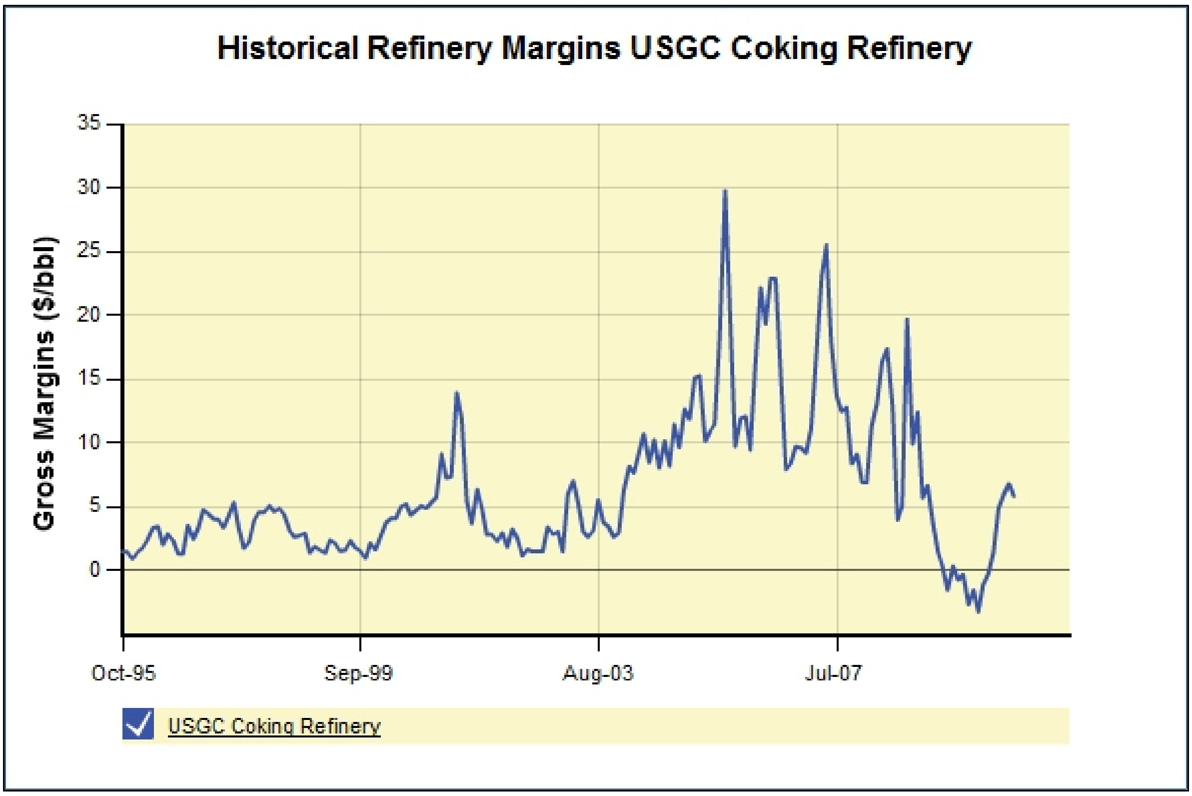
There are currently around 655 oil refineries in the world, located in 116 countries, with a daily total capacity of 88 million barrels per day[10]. The annual throughput is roughly 75 million barrels a day with an average utilisation rate of 85%. The refinery is built around the process of acquiring crude oil and ensuring that converting it into the market products is as smooth and profitable a process as possible. The equipment is very expensive and the technology and complexity of most refineries lag behind the top performers within the industry. There is a notable correlation between greater profits and refineries with a high Nelson complexity index[11]. This index was created by Wilbur Nelson in 1960 to quantify the relative costs of the components of a refinery. It is associated with equipment that can refine even the heavier crudes effectively and extract the products that are in high demand. The increased regulation surrounding the sulphur content of the raw crude has meant that only the most complex refineries, capable of refining both the heaviest and the lightest oils, are able to manage a volatile Gross Refining Margin (GRM): equal to the total value of the petroleum products from one barrel of crude after cracking, minus the cost of one barrel of crude oil. This measure is often reported within the industry as an indicator over time of refinery profit. In a regulated business environment, the GRM is most likely set by an agency, but in open markets like the US and Europe it is determined openly without regulatory interference. In tough economic conditions, the GRM can become negative, i.e., -$1.6/bbl[12]. In times of high demand with very little spare capacity a GRM of $20/bbl is not uncommon. Shown

[10] *Source: BP Statistical Review of World Energy, June 2011*

[11] Reliance Industries: http://www.ril.com/downloads/pdf/business_petroleum_refiningmktg_lc_ncf.pdf

[12] ftp://ftp.eia.doe.gov/pub/oil_gas/petroleum/analysis_publications/petroleum_issues_trends_1996/CHAPTER7.PDF

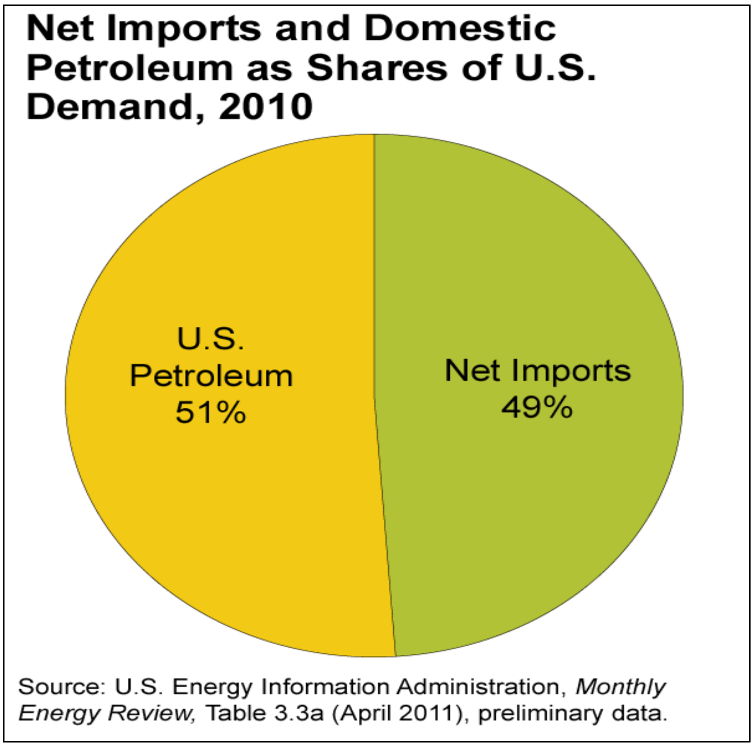
below in figure 1.5 is the evolution of the GRM of a typical coking refinery in the US over ten years. A topping refinery is a basic form of refinery, usually unable to produce gasoline. A hydro-skimming refinery is slightly more complex, and produces gasoline and excess fuel that is often in low demand. A cracking refinery is even more complex and allows the cracking of residual fuel into the middle distillates. The middle distillates comprise a wide range of products stemming from, jet fuel and kerosene to diesel. Finally, a coking refinery has an additional coking unit enabling the secondary unit to turn residues into high quality products. The Topping category is a small share of global refineries as the cracking type is the most prevalent - yet we focus in this thesis on a Topping type that can produce gasoline due to the ease of access to Topping linear programs.



(***Figure 1.5:*** *Typical historical GRM for a Coking Refinery over ten years)*

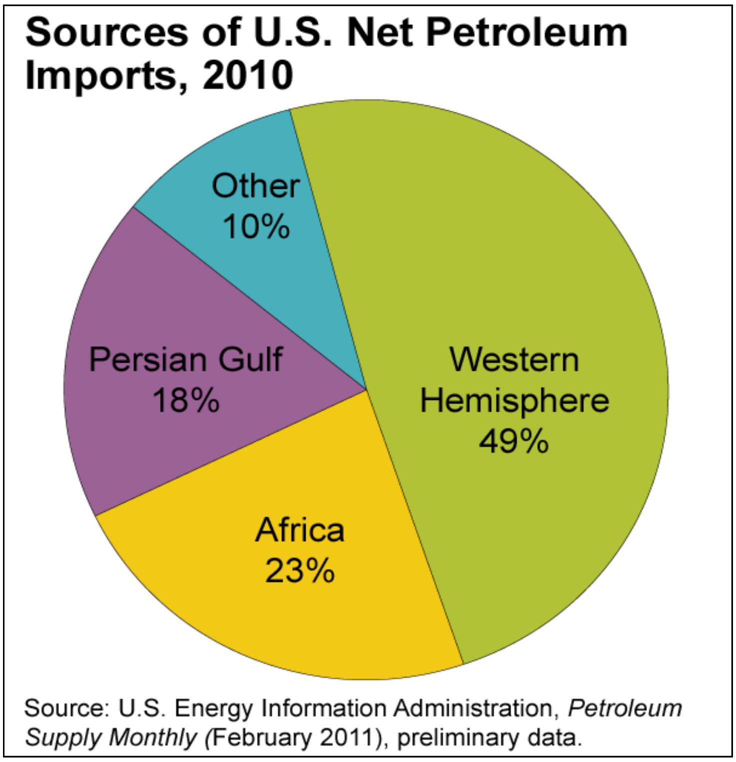
If the refinery does not have the inherent flexibility to meet the demand in its location, then it may import the required products. In recent times, conversion refineries run at full utilisation, but if the market prices of the refined products dip dramatically, and the refining margin goes below some pre-specified threshold, the operator can switch the refining equipment off. Utilisation rates of refineries highlight the market conditions, which have seen much lower spare capacity over the last ten years. To recognise the state of the oil market, one can observe the market prices of crude versus the cracked products.

The location and complexity of a refinery determine which strategy it will choose to operate on the crude assay, as the markets across the globe demand alternate variations on the typical crude slate. In the US, the main domestic demand in summer is for gasoline, and switches to heating oil in winter. Both of these product's demand is not met by domestic refiners and large amounts are therefore imported.



*(**Figure 1.6:** Dependency of the US on crude imports)*

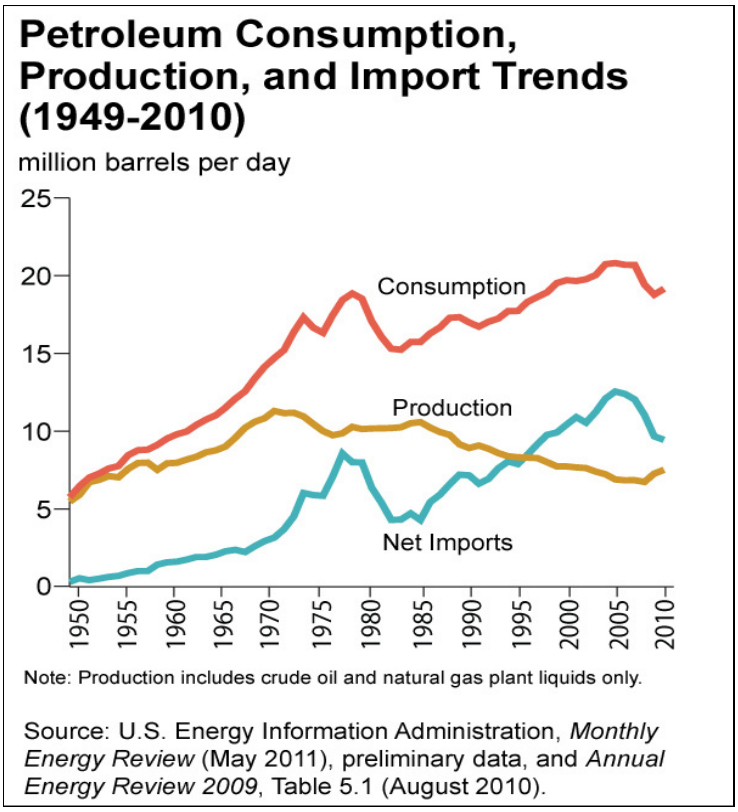
After the sophisticated refineries in the Middle East were destroyed in 1990, the US invested in light and middle distillate refineries to supply Europe and the Far East. Consequently, the US became a net exporter of distillates whilst remaining a huge importer of gasoline. Shipping a product reduces profits dramatically; hence if a refiner can meet demand in its local market, it will focus its efforts there. The refineries in Northern Europe, especially in Amsterdam and parts of Italy, supply the majority of their gasoline to the US. Shell's Pernis Refinery in the Netherlands is one of the company's largest, with pipelines serving the Schiphol airport in Amsterdam – it has a crude oil processing capacity of 416,000 barrels per day (bpd).



*(**Figure 1.7:** The leading role of Western Europe in petroleum exports to the US)*

Refined products are traded all over the globe; hence the possibilities for *geographical arbitrage* by trading houses can be huge. In 2006, 30 million tonnes of gasoil/diesel were traded from the CIS (Commonwealth of Independent States) region to Europe; 28 million tonnes of naphtha were exported from the Middle East to Asia, and 26 million tonnes of gasoline from Europe to the US. In Western Europe, there are a large number of sophisticated refineries able to import low quality feedstock from Russia and the Middle East and turn them into high quality gasoline and kerosene for exporting mostly to the US.

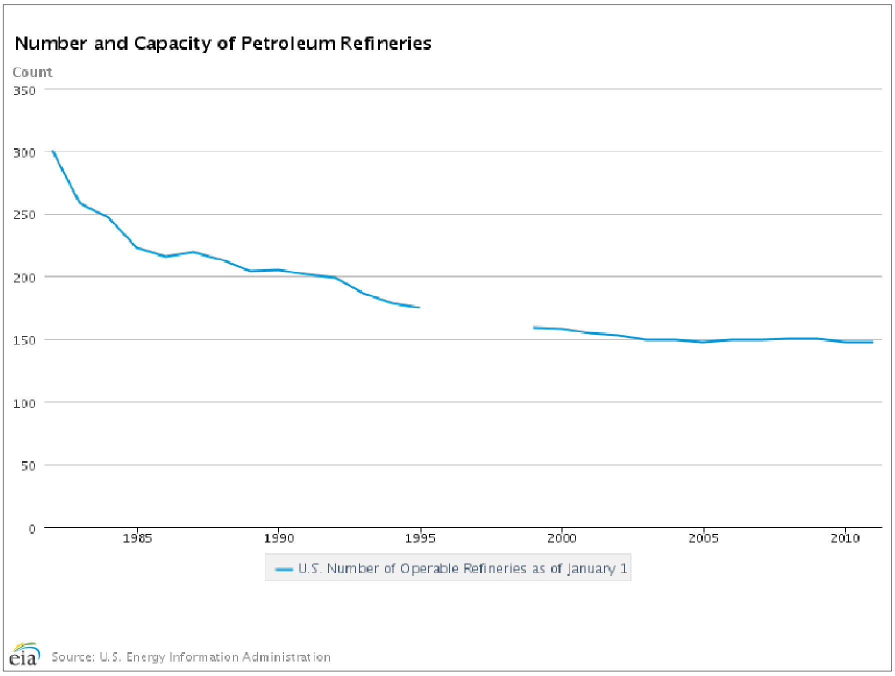
The US is likely to continue to be a net importer of gasoline for some time, although the investment in more complex refineries is gathering traction as shown by figure 1.7 below. The US is increasingly able to manage oil supply disruptions after the government in 1980 implemented a policy that ensured at least 1 billion barrels of crude were stored to hedge against volatile imports[13]. This was a reaction to the Yom Kippur War and the Arab Oil Embargo's effects. Consumption in the US has not slowed and more refineries are closing due to production having fallen. Recently, the US government has acted to change this situation with more refinery upgrades and lower taxation on gasoline production requirements. This is in direct contrast to the increase in taxation on the European refineries.



*(**Figure 1.8:** The US problem of growing consumption and reduced production)*

[13] Carter energy: http://www.pbs.org/wgbh/americanexperience/features/primary-resources/carter-energy/

Despite the recent trend in general production, gasoline production seems to be rising, the number of refineries has been falling for a long time due to increased strains economically and operationally. In 1980, not long after the oil embargo, there were around 250 oil refineries in the United States capable of supplying petroleum products demanded by the market. Now there are only 150, as shown in figure 1.8. Industry experts speculate that the reason why this number is so low is that this is a rigorous effort by the oil majors to squeeze the market[14]. Others state that the planning restrictions in the US have increased and not allowed the growth demanded by energy politicians. The report released in March 2011 from the US department of energy, stated that federal regulations were a significant factor in the closing of 66 US refineries[15].



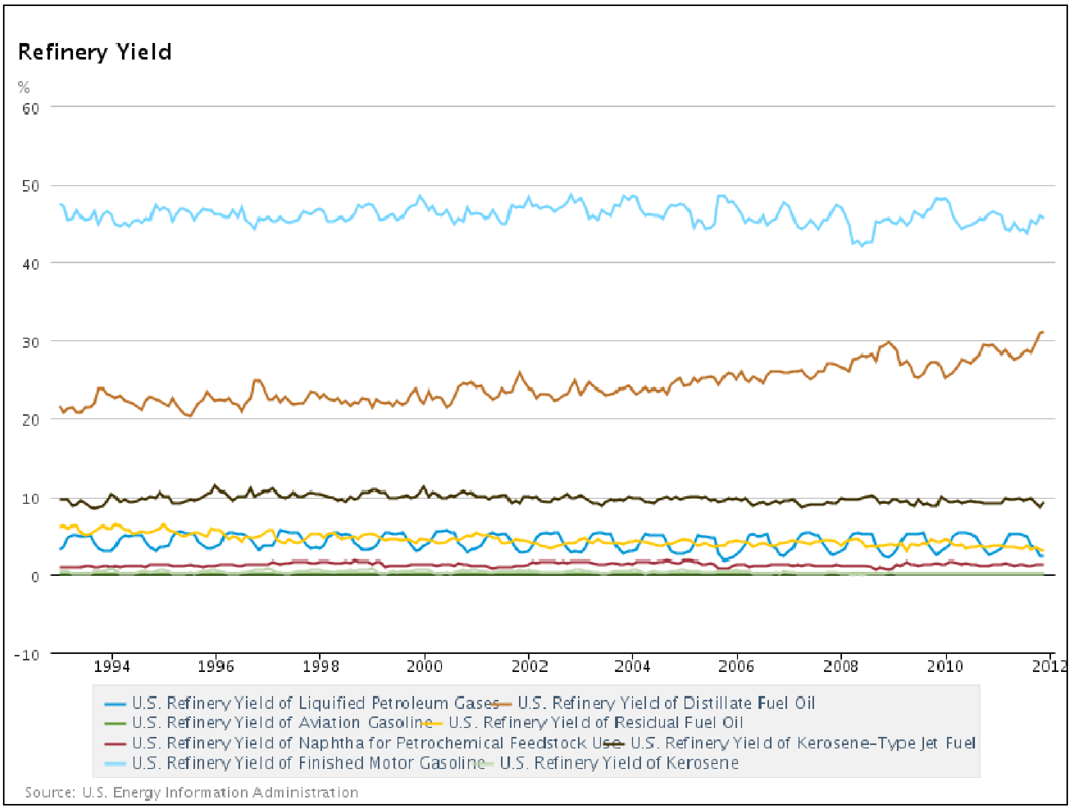
*(**Figure 1.9:** Count on active US refineries)*

Various economic and industrial changes have occurred over the past 30 years, yet, the refining product yield has stayed more or less constant as displayed in figure 1.10. Gasoline is usually the

[14] Wall Street Journal: http://online.wsj.com/news/articles/SB10001424052702304591604579291432462690714

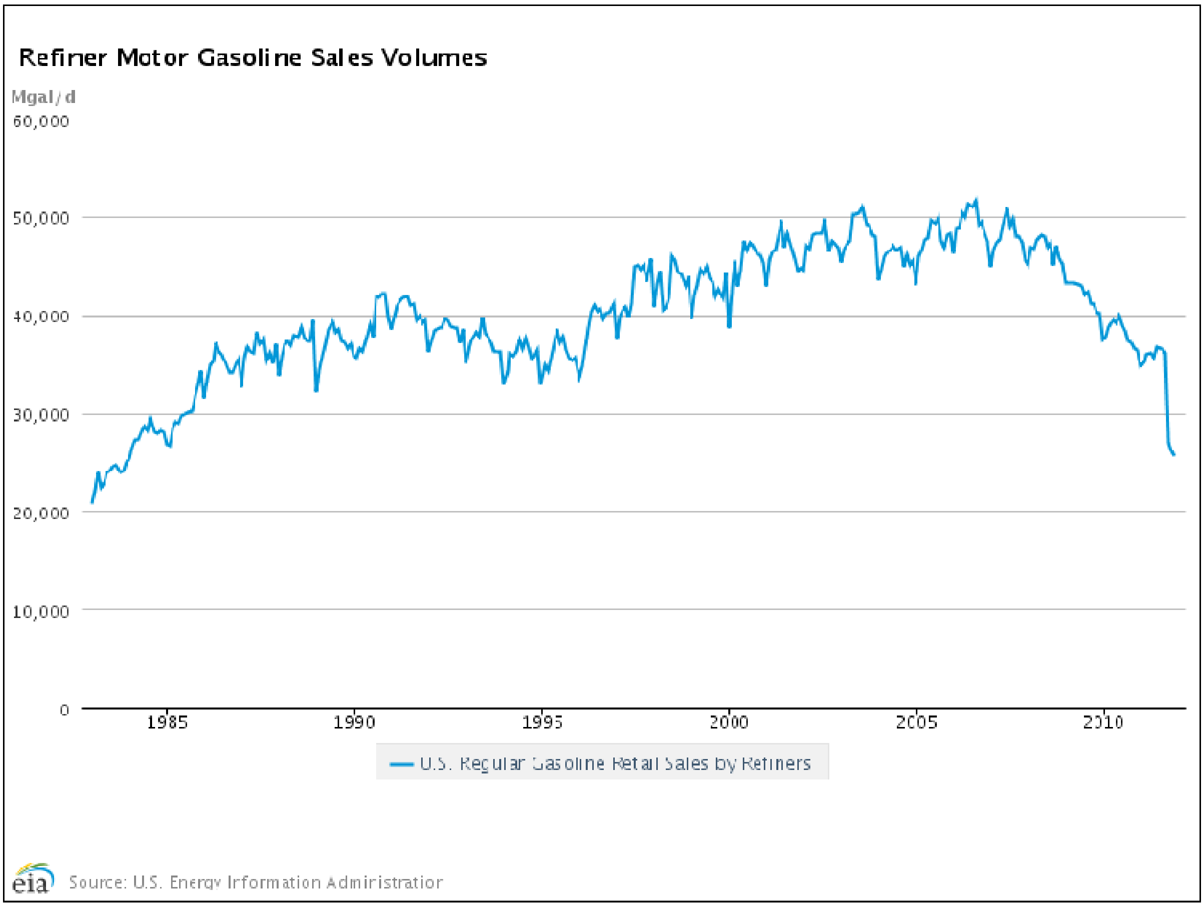
[15] http://www.instituteforenergyresearch.org/2012/05/03/over-regulation-of-the-nations-refineries/

highest yield from a barrel of crude oil and is on average 47% of the barrel. In the US, Liquid Petroleum Gas (LPG) is at a lower yield per barrel whereas in the Middle East, due to petrochemicals, LPG is in higher demand. Since the demand for middle distillates has grown, and the capacity for its refining in the US has recently increased, there has been a growing trend in fewer imports.



(__Figure 1.10:__ Amount of product yield from a barrel of crude in the US (Retrieved 08/01/2013))

Capacity over the last 30 years in the US has stayed approximately constant and the trends in sales remain coupled with the state of the economy. After the financial recession hit the US in 2008/2009, the sales of gasoline by refineries to retail outlets fell dramatically to 25 billion gallons per day as depicted in figure 1.11. This was close to levels seen in the early 1980s and a clear strain on the revenues of the oil majors. The retailing of petrol and gasoline is the lion's share of the profit for oil majors but also a significant contributor to the world economy.



*(**Figure 1.11:** Sales volumes of gasoline over time (Retrieved 08/01/2013))*

The entrance of retail stores into the domestic oil market has meant the closure of many smaller filling stations. In spite of this, sales had been increasing until the global recession of 2008. The large demand of unleaded and leaded petrol in the UK at the pump is the main driver for UK refineries. In mid 2012 the price of unleaded petrol was approximately 149 pence per litre; with 75-90% being taxes, VAT and costs - leaving well under 40 pence for the refiner per litre[16]. Globally, as constraints on refiners has grown the number of refiners has decreased.

[16] Source: Wood / Mackenzie, http://www.ukpia.com/files/pdf/understanding-pump-prices-april-2012.pdf

*(**Table 1.1:** Summary of UK Refining Capacity, (2010))*

Refinery	**Owner**	**Primary Distillation Capacity (Million tonnes per annum, MTA)**	**Primary Distillation Capacity (Thousands of barrels per day, KB/D)**	**Nelson Complexity Factor**
Fawley	ExxonMobil	15.9	326	9.1
Stanlow	Shell	14.4	296	7.4
South Killingholme	ConocoPhillips	10.8	221	11.3
Lindsey	Total	10.8	221	5.9
Pembroke	Chevron	10.2	210	8.6
Grangemouth	Ineos	10.0	205	7.9
Coryton	Petroplus	8.4	172	8.3
Milford Haven	Murphy	5.2	106	8.0
Eastham	Shell/Nynas	1.3	27	3.5
Dundee	Nunas	0.6	12	3.5
Total		87	1796	

Table 1.1 illustrates that the UK market is much smaller than that of the US. The US has a refining capacity of approximately 15 million bbl/day, whereas the UK processes 1.796 million bbl/day. However, the UK capacity has grown in the last 30 years despite increased regulation and refinery closures. The most complex refinery in the UK is the one at South Killingholme owned by ConocoPhillips. The largest capacity is at the Fawley refinery owned by ExxonMobil, with a primary distillation capacity of 326,000 barrels per day.

Crude and its derivatives are of prime importance to many areas of local and global economies, yet alternative energy sources are continuously being researched and examined. The transition has begun, with many groups of people interested in the emergence of climate friendly fuel sources and a removal of polluting fuels. It will take many years for these changes to have a practical impact on the energy industry, but it is inevitable according to senior scientists and petroleum practitioners[17]. The economy's dependency on crude oil is known ubiquitously and this reliance has stemmed from a number of significant events.

1.5 A brief history of petroleum

Crude oil has been in use for thousands of years. It is mentioned in the Bible for building purposes in Babylon: oil pits near Ardericca are described as having great quantities on the banks of the river Issus. The "black sludge" is mentioned in the literature all over the Middle East. The oil wells drilled in China in 347AD were some of the earliest known. In Japan, petroleum was known as the *burning water* in the 7th century. In the 9th century the first oil fields were exploited in Azerbaijan, mainly used for lighting purposes. In 1745 the first oil well and refinery was built in *Ukhta*[18]. In 1745 Russia, most households still relied upon candles for lighting. The Russians realised that through distillation of "rock oil" or petroleum, a type of kerosene could be obtained. This technique was then used by Russian churches in oil lamps. It was only after Lukasiewicz had improved Gesner's[19] method to develop a means of refining kerosene from "rock oil" that the first oil extraction well was built in central Poland. This knowledge diffused around the world and by 1861, the first modern Russian refinery was built in Baku. Momentum gained and at some point Russia was producing 90% of the world's oil. From data records in 1842 from the Caspian Chamber in the Department of State Property Ministry, it was found that around 23,000 barrels per year were extracted and most frequently sent to Persia. The first recorded economic oil boom was in 1871, when an Armenian entrepreneur in Baku built the first wooden oil derrick. Consequently, the

[17] Raymond Kopp: http://www.rff.org/Publications/Resources/Pages/Replacing-Oil.aspx

[18] Roman Ambramovitche's home town: http://www.themoscowtimes.com/beyond_moscow/ukhta.html

[19] Inventor of Kerosene: http://en.wikipedia.org/wiki/Abraham_Pineo_Gesner

industry exploded and in 1901 Baku's share of production fell to only half of the world's oil at 212,000 barrels per day.

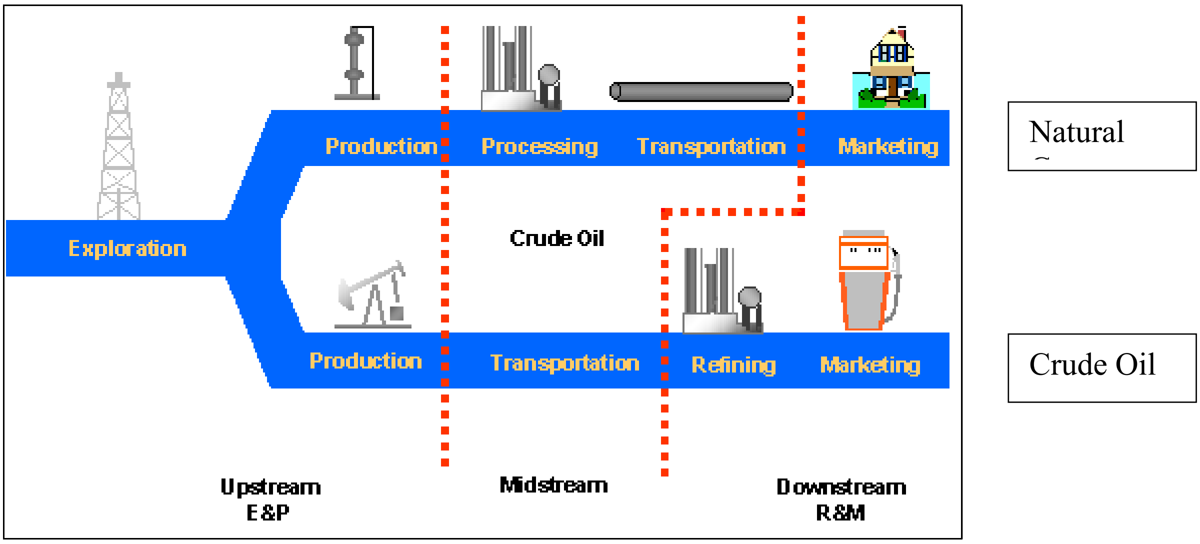
The modern uses of crude oil really began in the era when the internal combustion engine was invented in 1889 by Gottlieb Daimler[20]. Before this, engines ran on steam. With the turn of the century came Ford's implementation of the assembly line for car production; hence, Texas and Oklahoma became the epicentres of US crude production. Even with the American and British "Seven Sisters" (the seven great oil corporations that dominated the international oil industry from the 1920s to the 1970s), the Middle East began to gain oil concessions. After World War II, the Middle East became a major supplier of oil and joined the list of already profitable fuel producing countries.

1.5.1 The modern oil industry

Within the modern oil industry, there are three levels in the supply chain: upstream, midstream and downstream. Upstream is the drilling, exploration and production of the crude oil. The midstream is the transportation and trading of the crude oil to refineries. Refining the raw crude into its marketable products is the downstream business; this includes storage and marketing of the refined products. The collection of these segments constitutes the oil value chain.

---

[20] The main precursor to today's modern cars: http://inventors.about.com/od/dstartinventors/a/Gottlieb_Daimler.htm



(***Figure 1.12:*** *The oil and gas value chain)*

From procurement of the oil out of the fields, the crude is then delivered through products into the engines or hands of the consumers. Shown in figure 1.12 are the stages of the oil value chain: development, production, processing, transportation and marketing of the products.

The level of production is driven by the demand of customers. The major companies in the oil and gas business are: ExxonMobil, Shell, BP, Chevron, Total, ConocoPhillips, Petrobras and many others. These companies are often structured to be vertically integrated. This means having economies of scale: owning the infrastructure within the entire value chain. This can greatly reduce risks that are not easily avoidable by companies only involved in one part of the chain. Most of the companies mentioned above dominate the oil market globally, and are either international oil companies (IOCs) or national oil companies (NOCs). Currently, the industry is in a transition, moving from investment in its refineries to focusing on the upstream parts of the business. The expansion of the oil reserves remains at the core of the business: the more crude found in the underground reserves, the safer the revenues in the future will be. Although many refineries have closed in the last decade, the ones that are left are utilising capacities at much higher rates to meet market obligations. Nationally owned companies are regularly seen in the top rankings of oil companies by revenues and this is primarily because they have access to expansive oil reserves.

Table 1.2 shows key financial trading figures published in the press of the oil industry's benchmark crudes.

*(**Table 1.2**: Basic Features of Benchmark Crudes, Q1 2010 averages by Argus)*

	ASCI (Argus Sour Crude Index)	**WTI CMA + WTI P-Plus (Calendar Month Average & Postings Plus)**	**Forties (First substantial oil field discovered in British Territory)**	**BFOE (Average of 21 days in Brent, Forties, Oseberg and Ekofisk)**	**Dubai**	**Oman**
Production *(MBPD)	736	300-400	562	1,220	70-80	710
Volume Spot Traded (MBPD)	579	939	514	635	86	246
Number of Spot Trades per Cal Month	260	330	18	98	3.5	10
Number of Spot Trades Per Day	13	16	<1	5	<1	<1
Number of Different Spot Buyers per Cal Month	26	27	7	10	3	5
Number of Different Spot Sellers per Cal Month	24	36	6	9	3	6
Largest 3 Buyers % of Total Spot Volume	43%	38%	63%	72%	100%	50%
Largest 3 Sellers % of Total Spot Volume	38%	51%	76%	56%	100%	80.00%
Source: Argus **Million Barrels per day*						

As is evident in table 2.1, the crude produced in the largest volume is Brent at 1.22 billion bpd and WTI is the most spot traded crude at 939 million bpd.

The current financial trading of petroleum products is a trillion dollar business predominantly carried out in electronic form. It is a complex network of buyers and sellers with differing positions, agendas, algorithms and complex software to reach their profit targets. Black box trading[21], carried out at large investment banks, aims to buy and sell different commodities without the market observing the trade. Products like calendar swaps allow trading of one commodity for another at a future date, with the trade being protected by the exchange upon which it occurs. The market has grown more dynamic and complicated but the petroleum economic fundamentals still underpin and constitute the petroleum business. Capturing the underlying behaviour using sophisticated models and statistics is the cutting edge trading mechanism.

1.6. Statistics of crude oil and petroleum products over the last 10 years

Prices quoted on oil related markets are usually stated as either FOB (free on board, which means that the buyer pays for shipment and loading costs) or CIF (Cost, Insurance and Freight, a term expressing that the seller arranges for insurance and transportation of goods to a port and provides the buyer with documentation) for imports of cargoes. These costs are important for refiners that need to export outside their domestic targets. As Indian and other Asian refiners come on line at an ever increasing pace, the European gasoline export business is expected to experience a stiffer competitive climate. The Middle East will find it increasingly more efficient to import gasoline from Asian refineries. According to the EIA, in Europe there is limited growth expected until 2020 for gasoline - 120 million tonnes, compared to 330 million tonnes for diesel. Demand for diesel has been growing within Europe due to the number of diesel engine cars sold in the last decade. This can also be explained by taxation: while gasoline is being taxed as a consumer "luxury" good, diesel is being taxed at a lower rate, due to its use in the construction and transportation industries.

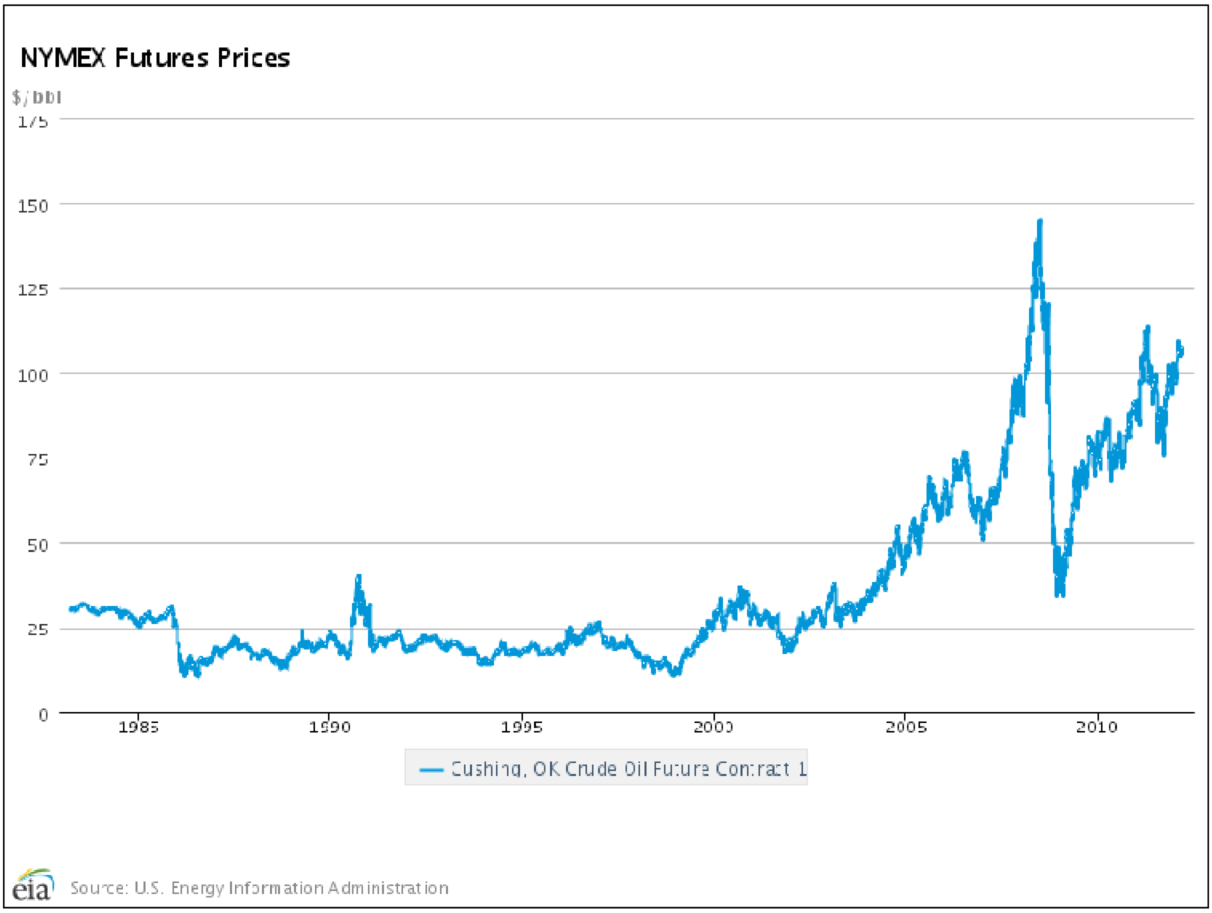
Experts in the analysis of time series data, rigorously search historical prices for a grasp on at least four key characteristics: the trend, level, noise and seasonality components of the data.

[21] Definition and example: https://www.quantopian.com/posts/black-box-trading-sample-algorithm-explanation

Prior to the construction of a trading model, an analysis of each commodity refined and processed is required.

1.6.1 WTI Crude Oil

In the US, one barrel of oil is equivalent in units to 42 gallons. For the WTI future contract the underlying oil, if settled physically, is for 1,000 barrels. Contracts on the Chicago Mercantile Exchange (CME) are the widest listed of options and futures of any exchange on commodities in the world. Crude oil futures are generally listed nine years forward using a particular schedule.



(***Figure1.13:*** *NYMEX Futures price on the front month contract over 28 years)*

Shown here are descriptive statistics of the price series for the two most popular crude contracts in the world. The statistics are analysed within three different time periods and show the large swings in price between the two contracts.

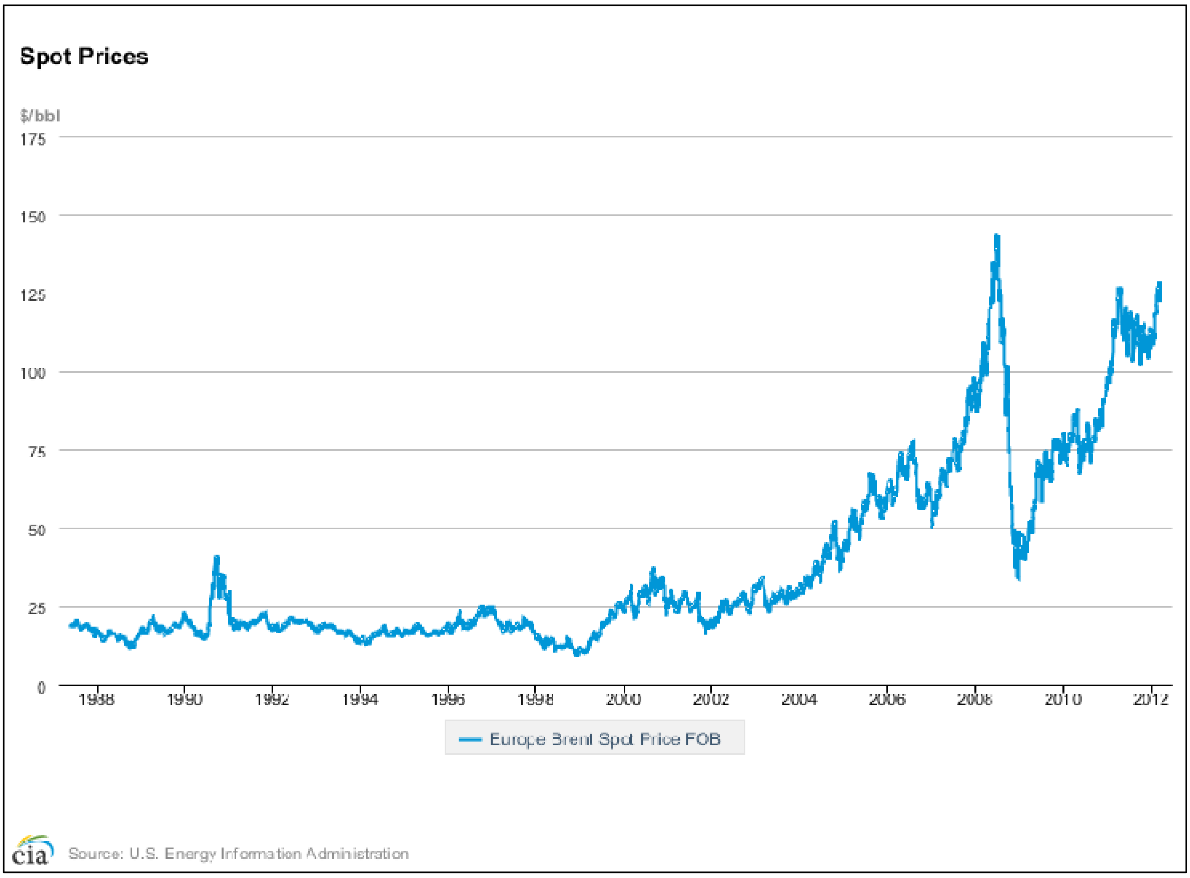
(Table 1.3.1: *Descriptive Statistics, WTI Oklahoma Contract* [22] *and the European Brent Spot Price FOB, ($/bbl) over 20 years)*

Jan 1983 – Dec 2001:	**WTI**	**Brent**
Mean	21.50015	19.24116
Standard Error	0.080589	0.081241
Median	20.1	18.43
Mode	18.67	18.48
Standard Deviation	5.525491	4.947685
Sample Variance	30.53105	24.47959
Kurtosis	-0.48625	1.858936
Skewness	0.552815	1.172789
Range	30	32.35
Minimum value	10.42	9.1
Maximum value	40.42	41.45
Number of observations	4701	3709
Jan 2002 – Dec 2008:	**WTI**	**Brent**
Mean	54.74521871	52.86125
Standard Error	0.645720219	0.641014
Median	53.555	50.86
Mode	26.86	25.51
Standard Deviation	26.19746695	26.28938
Sample Variance	686.3072745	691.1314
Kurtosis	1.084630188	0.959895
Skewness	1.084774994	1.064137
Range	127.32	125.78
Minimum value	17.97	18.17
Maximum value	145.29	143.95
Number of observations	1646	1682
Jan 2009 – Dec 2012:	**WTI**	**Brent**
Mean	80.5739154	85.25954
Standard Error	0.620011611	0.798837
Median	81.245	80.115
Mode	87.81	93.52
Standard Deviation	18.82631338	24.20358
Sample Variance	354.4300753	585.8135

[22] Data are split into three sub periods: up to end of 2001, up to August 2008 and after August 2008

Kurtosis	-0.217778406	-1.00624
Skewness	-0.37987559	-0.0736
Range	91.23	94.41
Minimum value	33.87	33.73
Maximum value	125.1	128.14
Number of observations	922	918

1.6.2 Brent Crude Oil

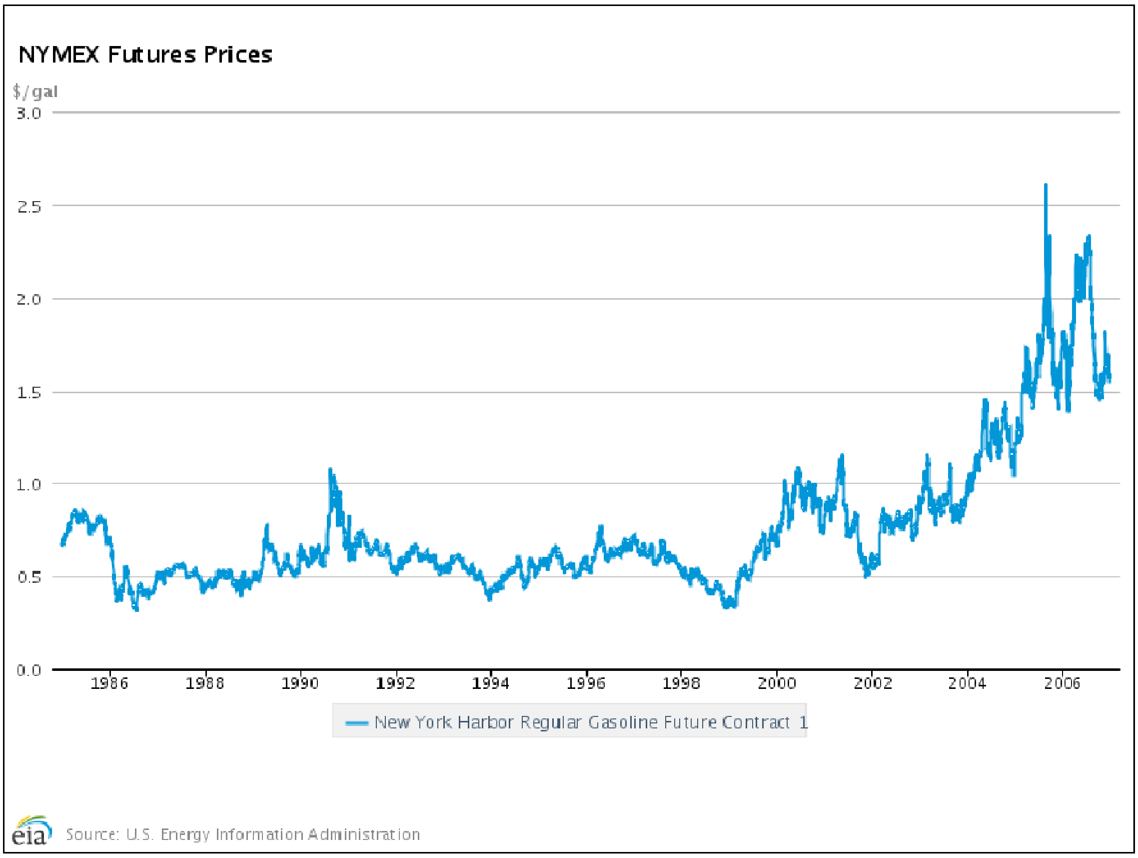


***(Figure 1.14**: Spot FOB European Brent over 25 years, ($/bbl))*

Dated Brent, the 15 day contract, collapsed a number of energy trading firms in 1986; the resultant replacement Brent crude futures contract (IPE 1988) transformed the market from a physical one to a purely financial market; this is a cash settled contract, figure 1.14 is for the underlying's price.

1.6.3 Gasoline

Gasoline exists as premium leaded and regular unleaded for the automotive industry, and the US is globally the largest importer. Gasoline is different from most other products as its price is posted outside gas stations and there are very few substitutes. The retail price of gasoline is usually two to three times higher in Europe than in the US due to various factors. For instance, on Mar 13th 2012 the cost of filling the same 39-gallon tank on a Chevrolet Suburban in the US was $2.26, in Norway it was $9.97. Across the globe, the $/gallon price of gasoline is generally closely inline before taxes; following taxation the oil prices widely vary from region to region. In Italy for example, the taxes are approximately three times higher than those in the US.



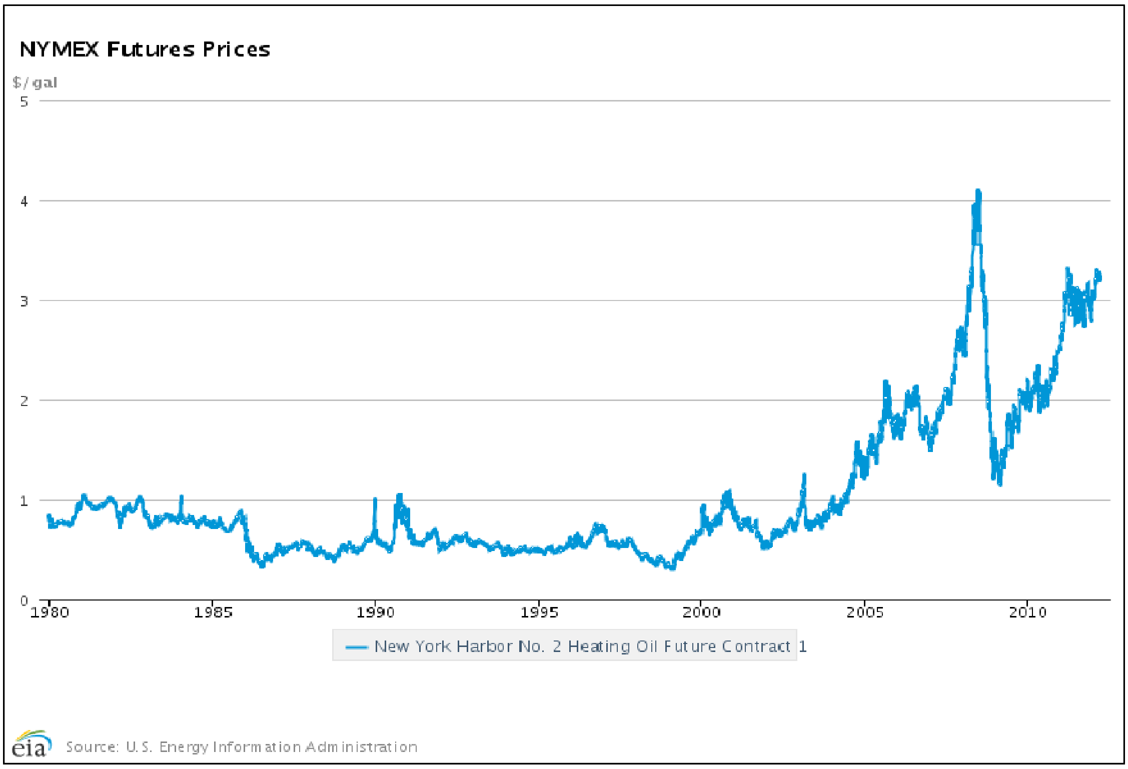
*(**Figure 1.15**: NYMEX Futures price of front month Gasoline contract, ($/gallon))*

(Table 1.3.2: New York Harbour Conventional Gasoline Regular Spot Price FOB ($/Gallon)

Jan 1983 – Dec 2001:	**New York Harbour Conventional Gasoline**
Mean	0.587061
Standard Error	0.002235
Median	0.559
Mode	0.516
Standard Deviation	0.140056
Sample Variance	0.019616
Kurtosis	1.045828
Skewness	0.975132
Range	0.813
Minimum value	0.29
Maximum value	1.103
Number of observations	3927
Jan 2002 – Dec 2008:	
Mean	1.498597
Standard Error	0.016161
Median	1.407
Mode	1.235
Standard Deviation	0.655859
Sample Variance	0.430152
Kurtosis	-0.13778
Skewness	0.691086
Range	2.909
Minimum value	0.507
Maximum value	3.416
Number of observations	1647
Jan 2009 – Dec 2012:	
Mean	2.214921
Standard Error	0.019357
Median	2.1235
Mode	2.862
Standard Deviation	0.587775
Sample Variance	0.345479
Kurtosis	-0.74927
Skewness	-0.17852
Range	2.542
Minimum value	0.788
Maximum value	3.33
Number of observations	922

1.6.4 Heating Oil

Heating oil commonly refers to No.2 Fuel Oil, often used in the US as distillate home heating oil. Its old names included: range oil, stove oil and coal oil. It is the fraction from distillation regarded as a residue. Heating oil has a seasonal pricing series[23], with demand growing in the winter. In the UK it is the office of fair trading who deal with adverse price reactions for customers. Weather and logistical issues often hit the price of heating oil, which some customers rely upon, along with LPG, for hot water and heating. Futures contracts are usually employed to trade heating oil on the markets.



*(**Figure 1.16**: NYMEX Futures price of front-month Heating Oil contract, ($/gallon))*

---

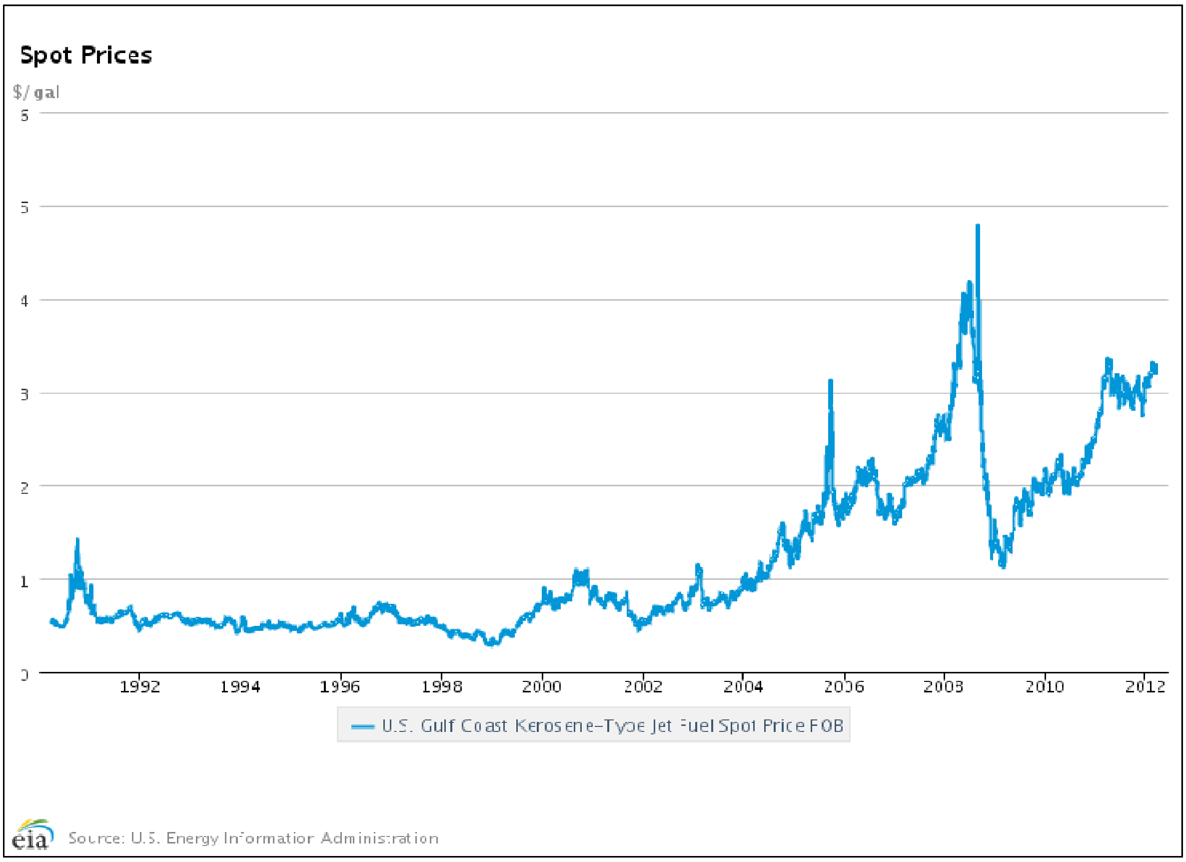
[23] http://commodities.about.com/b/2010/09/06/is-there-a-seasonal-trade-in-heating-oil-this-year.htm

*(**Table 1.3.3:** New York "Harbour" No. 2 Heating Oil Spot Price FOB ($/Gallon))*

Jan 1983 – Dec 2001:	**No. 2 Heating Oil Spot Price**
Mean	0.566397
Standard Error	0.002291
Median	0.538
Mode	0.551
Standard Deviation	0.143515
Sample Variance	0.020597
Kurtosis	3.591524
Skewness	1.408758
Range	1.481
Minimum value	0.284
Maximum value	1.765
Number of observations	3925
Jan 2002 – Dec 2008:	
Mean	1.518494
Standard Error	0.018653
Median	1.506
Mode	0.747
Standard Deviation	0.75698
Sample Variance	0.573018
Kurtosis	1.110434
Skewness	1.066993
Range	3.576
Minimum value	0.507
Maximum value	4.083
Number of observations	1647
Jan 2009 – Dec 2012:	
Mean	2.308702
Standard Error	0.020148
Median	2.155
Mode	3.048
Standard Deviation	0.611789
Sample Variance	0.374286
Kurtosis	-1.25683
Skewness	0.012403
Range	2.314
Minimum value	1.121
Maximum value	3.435
Number of observations	922

1.6.5 Kerosene

The word kerosene is derived from the Greek word for "wax". A Persian scholar named Razi (Rhazes)[24] was the first to distil this fuel in the 9th Century. Its major use in industry began in 1854 in New York. Throughout Europe, it was firstly used in lamps and then became a fuel for engines after they were invented. In the UK there are 2 popular basic grades: premium kerosene class C1 which is used for lanterns and in some combustion engines; and class 2, a heavier distillate used in domestic heating oil. Its main role on the market and real value is for its use as a fuel for jet engines and rockets. In the UK in 2006, jet fuel demand grew heavily due to low cost airlines. Post the 2008 financial recession this demand dipped and declined 5.2% in 2009 to 11.5 million tonnes. Demand slightly increased in 2011 and growth is expected in the near future because of plans for two major new airports in the UK.



(Figure 1.17: Spot prices of Kerosene FOB, ($/Gallon), Gulf Coast)

[24] http://elementsunearthed.com/tag/kerosene/

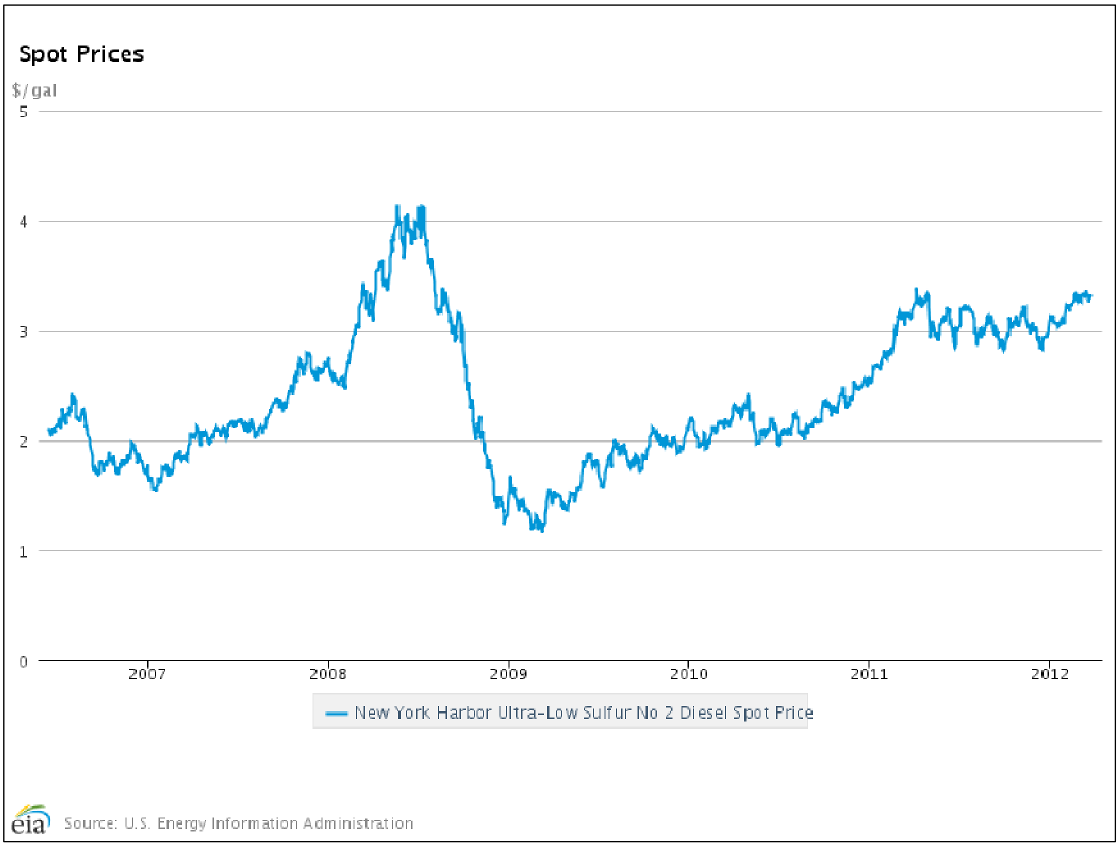
(*Table 1.3.4*: ***U.S. Gulf Coast Kerosene-Type Jet Fuel Spot Price FOB ($/Gallon))***

Jan 1983 – Dec 2001:	**Kerosene-Type Jet Fuel**
Mean	0.588439
Standard Error	0.002863
Median	0.551
Mode	0.496
Standard Deviation	0.155811
Sample Variance	0.024277
Kurtosis	2.840876
Skewness	1.359925
Range	1.149
Minimum value	0.282
Maximum value	1.431
Number of observations	2961
Jan 2002 – Dec 2008:	
Mean	1.575769
Standard Error	0.019791
Median	1.557
Mode	0.68
Standard Deviation	0.803197
Sample Variance	0.645126
Kurtosis	0.689352
Skewness	0.949616
Range	3.702
Minimum value	0.505
Maximum value	4.207
Number of observations	1647
Jan 2009 – Dec 2012:	
Mean	2.350438
Standard Error	0.021219
Median	2.184
Mode	3.079
Standard Deviation	0.643952
Sample Variance	0.414674
Kurtosis	-0.93374
Skewness	0.114147
Range	3.703
Minimum value	1.111
Maximum value	4.814
Number of observations	921

1.6.6 Diesel Fuel

Diesel is used as fuel for diesel engines; the word is derived from the name of the German inventor, Rudolf Diesel, who invented the diesel engine in 1892 and unveiled it to the public at an exhibition in Paris in 1900[25]. Diesel engines are lean burn engines, burning the fuel in more air than is necessary for the sparking reaction. Due to the high compression of the engine and it having no throttle, diesel engines are more efficient than standard spark-ignited (SI) engines. In many parts of the world, diesel will be priced higher than petrol. This is due to demand where diesel cars are more popular. Diesel has a slightly higher density than that of ethanol free petrol (gasoline), and it offers the same net heating value. The polluting emissions of diesel are lower than those of petrol, with most SI engines producing ten times the amount of CO and slightly more $C0_2$. The demand for this fuel usually rises in the colder months, corresponding with rises in heating oil. The reason for less production in recent years on the slate compared to petrol is increases in the sulphur controls introduced in the US. A consequence of these controls, is that the diesel in the US has a lower "cetane number" (a measure of ignition quality and the primary measure of diesel quality) than European diesel; resulting in worse cold weather performance and some increase in emissions. The taxation levied on each refined product varies and is a component of the refinery's crude slate decision process, in that it is incorporated into the market price of the product. The taxes applied to diesel fuel are generally lower than those applied to gasoline but are higher on the diesel vehicles themselves.

[25] http://www.tested.com/tech/454861-inventions-debuted-worlds-fair/item/diesel-engine-1900/



*(**Figure 1.18:** Spot prices of Diesel, ($/gallon), New York)*

---

*(**Table 1.3.5**: Los Angeles, CA Ultra-Low Sulphur Diesel Spot Price ($/Gallon))*

Jan 1983 – Dec 2001:	**Diesel Spot Price**
Mean	0.725304
Standard Error	0.00507
Median	0.72
Mode	0.875
Standard Deviation	0.191914
Sample Variance	0.036831
Kurtosis	0.029947
Skewness	0.435905
Range	0.905
Minimum value	0.375
Maximum value	1.28
Number of observations	1433
Jan 2002 – Dec 2008:	

Mean	1.679812
Standard Error	0.019436
Median	1.683
Mode	0.83
Standard Deviation	0.788793
Sample Variance	0.622194
Kurtosis	0.41353
Skewness	0.78469
Range	3.62
Minimum value	0.513
Maximum value	4.133
Number of observations	1647
Jan 2009 – Dec 2012:	
Mean	2.378716
Standard Error	0.02106
Median	2.23
Mode	2.129
Standard Deviation	0.639467
Sample Variance	0.408918
Kurtosis	-1.16353
Skewness	0.01105
Range	2.428
Minimum value	1.099
Maximum value	3.527
Number of observations	922

1.6.7 Naphtha

Naphtha is a product resulting from the distillation process applied to petroleum, coal tar or peat. It is an all encompassing term, as it refers to the lightest and most volatile fractions of the hydrocarbons in petroleum. In the US, refined Naphtha is primarily used as feedstock for high octane gasoline. It can also be used as a solvent. It boils between 30 and 200 degrees Celsius; has a specific gravity of 0.7 and is volatile and flammable. There are a number of categories for Naphtha, these can generally be split into less dense *lighter* Naphtha and more dense *heavier* Naphtha. The lighter Naphtha is referred to as paraffinic Naphtha[26]. These are mainly used as feedstock of olefins; sometimes these are also called “Light Virgin Naphtha”, LVN, or “straight run gasoline”,

[26] http://www.petroleumhpv.org/docs/gasoline/052003_gasoline_robustsummary_pnaphthas_revisedfinal.pdf

SRG. The heavier or denser types are richer in aromatics and naphthenes, known as “N&A s” or “straight run benzene”, SRB. N&As are used in the reformate part of the refining process, where the heavy Naphtha is cracked into butane.

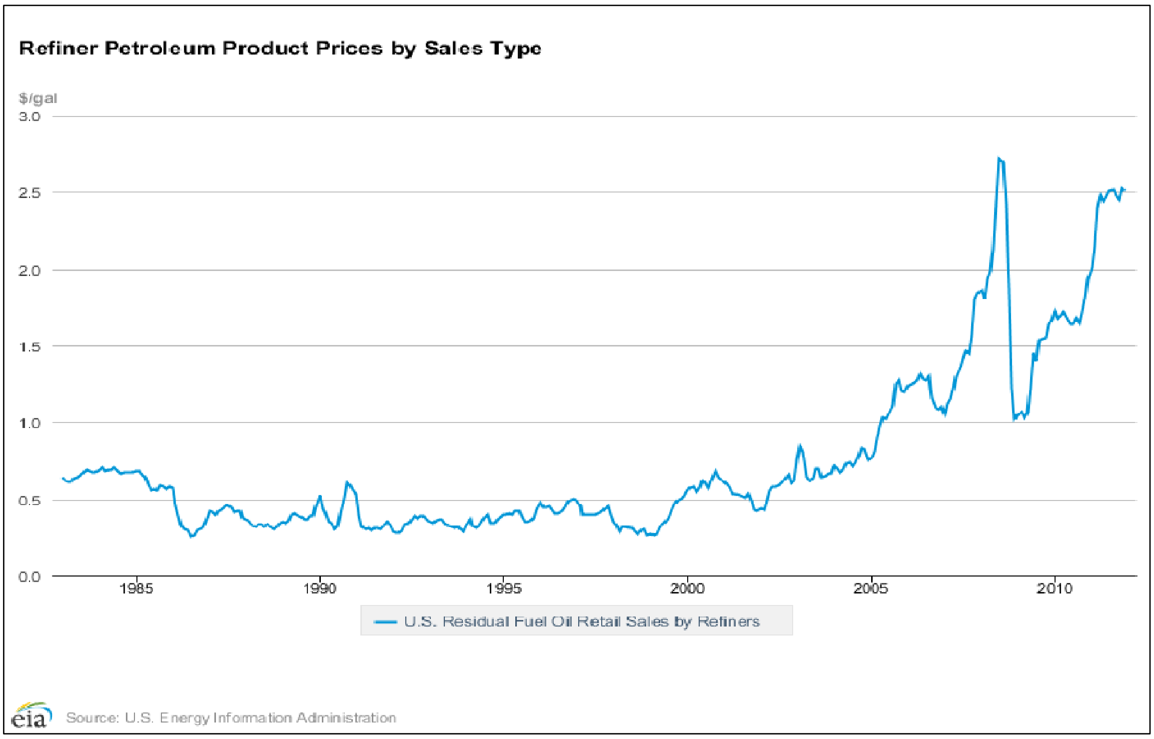
***(Table 1.3.6:** OSN Naphtha CFR Japan Front month Contract ($/tonne))*

Jan 1983 – Dec 2001:	**Naphtha CFR Japan Front month Contract**
Mean	188.3587698
Standard Error	0.743143425
Median	180
Mode	177
Standard Deviation	46.49251614
Sample Variance	2161.554057
Kurtosis	2.052506241
Skewness	1.167322624
Range	345
Minimum value	87.5
Maximum value	432.5
Number of observations	3914
Jan 2002 – Dec 2008:	
Mean	493.3758978
Standard Error	5.65536246
Median	460.5
Mode	245.5
Standard Deviation	226.49709
Sample Variance	51300.93179
Kurtosis	0.614234028
Skewness	0.934071984
Range	1087.5
Minimum value	167
Maximum value	1254.5
Number of observations	1604
Jan 2009 – Dec 2012:	
Mean	739.5837099
Standard Error	7.038602064
Median	738
Mode	735
Standard Deviation	204.1194598
Sample Variance	41664.75389
Kurtosis	-0.622241093
Skewness	-0.393942553
Range	862.5
Minimum value	244
Maximum value	1106.5

Number of observations	841

1.6.7 Fuel Oil

This is a term used to refer to only the heaviest commercial fuel that can be obtained from petroleum distillation: heavier than that of gasoline and Naphtha. In the US there are six grades of fuel oil and its uses vary from stove oil to heating oil for the home. Due to the recent widespread penetration of natural gas, the heating oil has had less use in homes and sales over the last decade or so have decreased dramatically. There are many areas in the US where it is however still copious. Residual fuel oil is less useful due to its viscosity, it must be heated with a special system before use and it contains a high level of sulphur. However, it is very cheap and it can be used as fuel on boats or small ships and as fuel for power plants. Fuel oil is transported worldwide by super tankers from ports all over the world including: Houston, Singapore and Rotterdam. In Europe, the Rhine is used for the transportation of fuel oil. Throughout the thesis Fuel oil will refer to No.4 Fuel Oil alias Residual Fuel Oil, it is frequently utilised in the US as commercial heating oil for burner installations.



(__Figure 1.19__: Refiner prices of fuel oil, ($/gallon))

(__Table 1.3.7:__ US residual fuel oil ($/gallon))

Jan 1983 – Dec 2001:	**US residual fuel oil**
Mean	0.438219
Standard Error	0.008312
Median	0.402
Mode	0.42
Standard Deviation	0.125513
Sample Variance	0.015754
Kurtosis	-0.73377
Skewness	0.69265
Range	0.451
Minimum value	0.259
Maximum value	0.71
Number of observations	228
Jan 2002 – Dec 2008:	
Mean	1.049215
Standard Error	0.052855
Median	0.978
Mode	1.252
Standard Deviation	0.469787
Sample Variance	0.2207
Kurtosis	1.72599
Skewness	1.279405

Range	2.291
Minimum value	0.433
Maximum value	2.724
Number of observations	79
Jan 2009 – Dec 2012:	
Mean	1.827756
Standard Error	0.078414
Median	1.692
Mode	2.473
Standard Deviation	0.502092
Sample Variance	0.252096
Kurtosis	-1.09027
Skewness	0.122145
Range	1.673
Minimum value	1.021
Maximum value	2.694
Number of observations	41

1.7 Refinery Optimisation

Crude oil by itself is not of any practical use; to use the hydrocarbon energy the chains must be broken. Hydrocarbons are practical for two reasons; the first is that the chains contain a significant amount of energy, and the second is that they take on many different structures. Why does a refiner need an optimisation model? The scheduling of crude oil unloading, inventories, blending and feed to oil refineries is a complicated decision process requiring optimising. There are a number of exogenous and endogenous factors that affect the choices being made by the refinery owner. All these separate areas can and are in fact modelled with linear programs in practise and within academia. In this thesis we have a financial objective for a refinery owner - with the objective of maximising the profit: we construct a mathematical model for the optimisation of chemical processes under uncertainty in the medium term. In spite of the fact that many refineries have closed down due to decreasing gross refining margins (GRMs), there are those recording bumper profits. This is sometimes due to location, where a refinery has closed and another is left servicing the same demand. In the Midwest US where the imports from Canada are easily transported, the refineries are very profitable. However, there are certainly cases where demand in a location is at its lowest in recent times yet the refinery is running at full capacity and recording strong profits[27]. This is not due to new capacity or increasing demand. Western Europe and the USA have not seen refinery capacity increases in the last 20 years due to the environmental difficulties. India is however a different case, growth in petroleum products consumption has been growing for a long time and refineries are constantly expanding to meet the new demand. We seek a model that works for a refinery within India but the results are globally applicable. Despite the increase in demand, a number of refineries closed in the early 2000s due to the intense competition. A set of competitive optimisation decisions allows a refiner to survive: they need alternative strategies and more efficient operations than their competitors. Within the industry a GRM of $4/barrel is historically profitable: many refinery owners argue it should be closer to $9/barrel. To reduce the effect of operating costs eating into the fluctuating margins the only choices the refinery has is either to refine more efficiently, or expand using newer technology, which emanates in a higher GRM. If a refiner concentrates on optimising one unit, for example the Coker in a blending optimisation, it does not guarantee a complete set of optimal decisions. This is due to there being many integrated units and a web of nested decisions that need to be made to produce an optimal result. Consequently, refiners often concentrate on two key areas when optimising:

1 – Raw input crude selection – if the choices or decisions made early on are the correct ones over time, it saves huge effort changing them later. Once the crude slate is chosen, the opportunity for optimisation is much more limited.

2 – Feedstock quality management – ensuring that the right crude mix, i.e. specification of the oil, is accurately produced by the refinery along the processes reaching yield levels within the individual units.

---

[27] http://uk.reuters.com/article/2013/07/19/usa-fuel-exports-idUKL1N0FO1HG20130719

Linear Programming techniques are relatively mature and many commercial software systems are implemented onsite at most refineries including: Aspen PIMS (AspenTech), GRTMPS (Haverely Systems), TRIOS (UOP Limited) and many others. In the short term, the problem for a refiner is the mechanical day to day decisions: operational optimisation, i.e. analysing units for elementary composition mixes have great affects on the physical properties and processing required for producing saleable products. This is a large area of research in chemical engineering. In practise, there is a lack of systematic integration between *intertemporal* planning and operational blending optimisation - this is the gap we will exploit in this thesis. The software listed above can be utilised to focus on the medium term decision sets. The midterm planning arena is starting to be examined using ideas from mathematical finance as the price series dynamics of the crude and the refined products are vital to the mathematical programming problem at this scale. This is still a very open area, and the model formulation in this thesis seeks to address this problem.

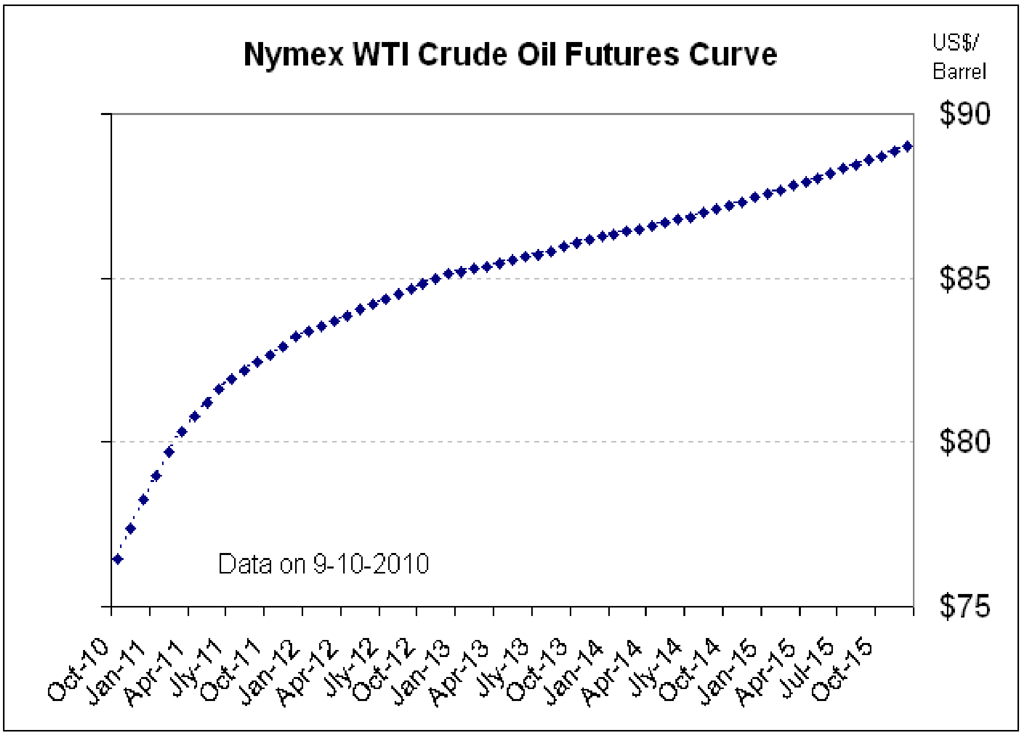
Dependent on the market that the particular refinery is serving, it may be more profitable to invest in cutting edge refining equipment. This could enable the refining of heavier crude oil and allow products to be sold that can only originate from the cheaper, less pure and more common crudes. An example, would be the Vadinar refinery owned by Essar Energy, in Gujarat, India - we use data for this refinery in our linear program optimisation. Recent investment for an additional CDU and Cracker has introduced an extra 100,000 barrels per day throughput. The denouncement from an expansion usually more than offsets the investment in the equipment that allows heavier crudes to be processed within the first year of operation.

In the UK, ever since the 1990s, the refining margin has been squeezed. This is due to the increased burden of regulation on the refinery's operations and the increase in competition from emerging markets. The UK government has increased mandates looking for cleaner processing and customers have demanded higher quality refined products, laterally the supply has become *heavier*. This had made life for the refiners strenuous, and demand for huge improvements and adaptations to the optimisation software has increased. This software has to cope with the stricter regulation on the removal of sulphur and nitrogen; the choice of crude oil to purchase and the decision yields to maximise in the profit function in the face of all the various constraints present at the refinery.

The refiner must make the correct set of decisions at each refining period based on the internal and external factors that affect the price of petroleum and its constituent products. There are restrictions on, for instance, the physical constraints present at the facilities in the refinery. The capacity constraints state the maximum volumes permitted to be stored by the refinery, typically the storage containers split this 50:50 between crude oil and its products. The mass balance constraints are applied at each unit used to process the crude within the refinery, for example the crude distillation unit (CDU). Internally, the mass balance equations have to ensure that the amounts of fuel entering and exiting any refinery unit are balanced. Externally, the local and global market prices of the refined products impact the decision maker's choices of amounts to produce and consume. It is generally assumed within optimisation models, that these decisions do not however have an impact on the exogenous prices. This can easily been seen if one imagines the impact of one refinery out of the world's 700 - it can only be a price taker. The cost of the raw crude, is influenced by: product quality, shipping costs, global petroleum prices, local taxes and logistics.

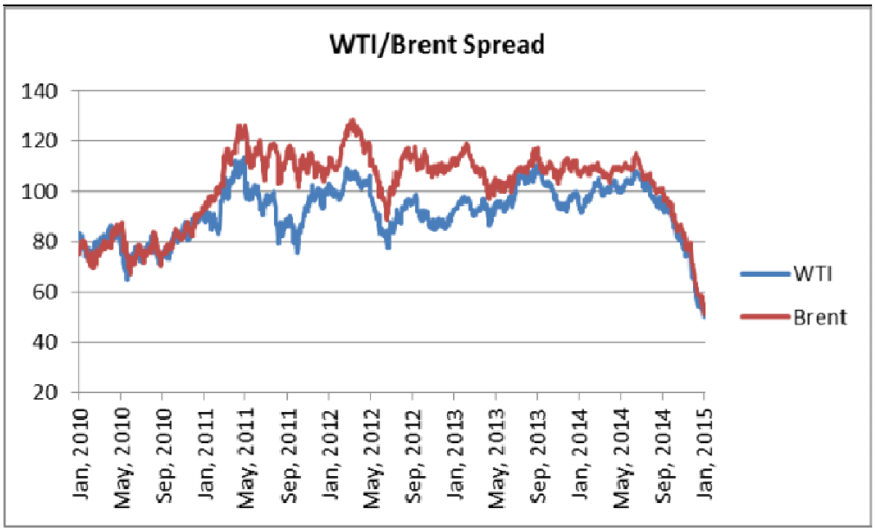
The market prices of refined products can have a huge effect on the cost of raw crude. For example, in 2009 there was a major blockage in getting jet fuel to destinations across the US from bottlenecked Cushing, Oklahoma. Consequently, not only did the price of kerosene go up but the cost of crude to the refineries dropped. Optimisation models can use market data as input, hence if calibrated correctly, will price in these external fundamental factors for accurate and realistic modelling.

Refining crude, not only means acquiring a raw fuel and processing it, but decision makers must attempt to manage the uncertainty in the prices of the end products to lock in value. The volatility of gross refining margins is a sign of the complications within the industry. Where is this volatility coming from? One of the driving factors is the crude oil price. The crude futures curve, which depicts the expectations of the spot price, dramatically affects the future prices of the refined products.



(Figure 1.20: Futures curve expectations)

In the US refiners are holding less spare capacity; therefore, any negative event on the refining cycle has a larger impact than has previously been the case. The fundamentals that affect this price volatility have shifted, and recently, the spread between sweet and sour crude oil was at an historical high. The spread between Brent and WTI has gone from an average of $4 to $25 per barrel in Jan 2013, to only ~ $1 per barrel in September 2014.



***(Figure 1.21:** WTI/Brent spread over the last five years)*

In terms of managing this spread, the refiner usually purchases the cheapest crude, obviously dependent upon its refining complexity, and hedges using a relevant contract on an exchange.

Market prices of both the crude and the refined products change continuously every day, the more volatile these prices relative to each other become, the more uncertainty introduced into the decision process. The refinery revenues over a year must exceed its operating costs, asset depreciation and corporate taxes. To maximise this return on investment, the refinery manager, intends to pay the lowest price for crude oil, and sell the yield of the products at the highest market value, whilst controlling operating and regulatory costs. During the 1980s and 1990s, there was huge refining capacity over the globe and GRMs were low. After the demise of the majority of these refineries, the GRMs began to increase again up to the year 2000. Knowing that the GRM can drop so low, the largest US oil majors slowed investment and this has by and large remained the case ever since. However, outside the OECD regions there have been many refinery expansions despite for instance, the high capital costs required, the increase in the costs of steel and the restrictions placed on the refined products quality. The solicitation of optimisation models has thus boomed – the risks needing representation in these models, in practise, is vast.

1.4.1 Refinery Risks

There are a number of key risks that impact the refinery owner's profits. *Price Risk* relates to the fluctuation of crude oil prices and refined petroleum prices affecting margins. Operating efficiency and access to crude oil of the required quantity, quality and price has a significant impact on the refinery's performance. While refined products normally track the changes in the feedstock prices, there is a lag which can impact short-term working capital requirements.

Foreign Exchange Risk relates to the foreign exchange fluctuations that impact the refinery due to its imports and exports as a part of its general operations. This is an area where refiners' risk management teams can learn from professional foreign exchange analysts, as often huge amounts of capital can be lost due to this hedging risk exposure[28].

Reputational Risk relates to the potential commercial and reputational damage that could result from health, safety or environmental incident or conflicts with local communities, terrorism, or the geo-political location of the refinery[29].

"Crack risk" refers to the oil refiner's crack spread. This is a term used in the industry for the difference between the price of crude oil and the products extracted from it; or in other words, the profit the refinery can expect to make by "cracking" the crude. The oil majors are vertically integrated, ergo they own the supply chain from exploration, to production, to retail; in this case there is already a natural hedge in place. For independent refiners however, like those renting the refinery's equipment, adverse price movements provide the bulk of the risk, whereas oil majors already own the crude. A refiner can utilise the futures market and purchase relevant contracts to lock in fixed prices. To hedge the crack spread exposure a DM could remain neutral by longing crude futures and shorting the refined products contracts separately. However, instead of purchasing so many contracts, the basis risk is usually managed by using a "3-2-1 crack spread option"; or some constructed alternative based on the specific products refined in the refiners'

---

28 http://people.stern.nyu.edu/igiddy/fxrisk.htm

29 http://www.ft.com/cms/s/0/84d6a0cc-9977-11e3-91cd-00144feab7de.html#axzz308BTG7mI

region. The "Chicago 3-2-1" is the most popular quoted crack spread contract in terms of volumes traded on the NYSE. The 3-2-1 refers to the volume ratio of the commodities being hedged; from a petroleum refinery one would expect two barrels of gasoline and one barrel of heating oil from three barrels of crude. For a premium, a refiner can then execute the option to alleviate the uncertain crack spread. Hedging strategies require the refiner to tie up funds in a margin account to trade options or futures. Alternatively, swaps and more complicated OTC trades can also be executed by the refiner to manage its price risk.

Liquidity is usually in excess in the futures market for crude contracts on the near month contract: the issue for refiners is the availability of contracts for each market product at future dates. This is can be implied by the prices for the contracts on the financial exchanges. Some refiners will tie in a futures price by entering into an OTC transaction, or use a future contract for example, to sell gasoline up to 1 year ahead. If the price increases above the agreed contracted level, then the trader has lost profit, managing the expectations of these future prices is the foundation of the hedging business. Hurricanes and weather events are risks that remain very significant. This is certainly more significant for refineries in the US on the Gulf Coast.

The following chapters detail a refining optimisation model including relevant constraints and financial risks. The primary goal of such a model is to aid the refinery decision maker in maximising profitable choices, however, with a number of alterations the model can be implemented to obtain a financial valuation of the refinery as an asset, whilst still outputting the decision set for the entire crude assay.

Chapter 2

Valuing an oil refinery with discounted cash flows

"*Thus the art and science of valuation has seen a constant debate between what something is worth versus what the market thinks it's worth and versus what a strategic or motivated buyer thinks it's worth."*[30]

Financial valuation relies upon one of three pillars: comparable company analysis, discounted cash flow calculation or transaction analysis. Auditors generally apply traditional accountancy calculations rather than out of the box financial option related approaches as provided by professional valuers of businesses and assets. In this chapter we provide a standard calculation to set the foundation for more complex and accurate approaches in the following chapters; this valuation is done as standard within industry, it is practically useful for many reasons from tax and litigation to mergers and financial reporting. In traditional financial theory it is assumed that investors are rational and risk averse; hence requiring some additional return to accept risk – the time value of money, means that we can discount expected future cash flows (DCF) to value the refinery asset today, using an applicable rate of interest for the investor in question. Finally, we collect standard financial values from an oil refinery in Gujurat India, and using these ratios

[30] **Keith Spence, Hons. B.Sc., MBA**
Export Development Corporation
Co-Chair, CIM Valuation Committee (CIMVal)

produce a DCF calculation that gives a present value as a standard auditor would within industry – this calculation has been verified by a fully qualified ACA accountant of the ICAEW.

2.1 A traditional valuation of an oil refinery

Valuing a project or a company is a vital process for investors and the company itself; the three most common applied approaches are: the asset-based approach, the market approach and the income method. The asset-based method, values the business on the sum of the parts of the business, the market approach compares the company to be valued on companies within the same industry or a comparable entity, whereas the income based method, attempts to calculate the benefit stream of the business. Combinations of the above do exist, the traditional Discounted Cash Flow ("DCF") analysis, capitalisation of earnings, and the excess earnings method are the most common utilised to value a financial entity where the economic principle of expectation is suitable: valuing a company relying upon the expected economic benefit subject to a relevant level of risk. Surveys have been carried out on the FTSE 100 companies in the UK which consistently show that for a project's NPV, Net Present Value is the most commonly applied business valuation method by far[31]. There are many valuation methods and some more relevant depending on the nature of the company or project. The refinery is an illiquid asset, hence prices are not observable; the DCF calculation needs to address this. The two most acceptable valuation methods in the UK, dependent upon the type of company, are asset based methods and income based methods. Most often the valuation is calculated using DCF analysis, an income based method. Corporations generate profits in various ways; yet generally projects are undertaken after some amount of analysis has indicated that they are worthwhile, i.e. there needs to be evidence of profitability. This is a simple method which takes into account the time value of money and the cost of capital to the company in question; the cost of capital is a open and difficult to calculate. The

[31] http://www.rics.org/site/download_feed.aspx?fileID=2369&fileExtension=PDF, *Royal Institute of Chartered Surveyors*

capital asset pricing model (CAPM) is one way to determine the discount rate to apply in a business valuation. NPV itself, is a tool that dates back at least to the 19^{th} century, Karl Marx wrote:

"The forming of a fictitious capital is called capitalising. Every periodic repeated income is capitalised by calculating it on the average rate of interest, as an income which would be realised by a capital at this rate of interest. "

Why is it important to be able to value a company, for example an oil refinery? Here are two examples from the financial news:

A) A private investor Gary Klesch was close to buying Total's (TOTF.PA) British Lindsey Refinery for as much as £2 billion, as reported by the Sunday Telegraph on April 2010. This refinery produces 223,000 barrels per day and employs 500 people.

B) In the US on October 16^{th} 2009 there was an important dispute about the method to value a Dutch Oil refinery *Lyondell Basell*, which claimed its refinery was worth $3.43 billion where as an alternative value by auditors gave it a value of $950 million. This refinery produces 105,000 barrels a day.

Clearly the methodology to value is vital for many reasons whether it be as a fair price for investors or a dispute in the courts.

The financial statements that apply to a number of years of operation of an oil refinery based in India will be examined to value it with the traditional method of DCF analysis. This will give a fair value, according to UK accountancy rules, and allow a comparison to be made of other methods used to value the refinery in the future chapters using statistics and more extensive mathematical models.

For an oil refinery there are a number of figures on the financial statements that will be different to other businesses. Officially under International Financial Reporting Standards (IFRS) rules, a valuation needs to be carried out using an official method:

- Absolute Value models determine the present value of an asset's expected future cash flows, these are either multi-period like DCF or single period like the Gordon Growth model.
- Relative Value models obtain a price by using market prices of similar assets.
- Option pricing models are used less frequently but are used for certain type of assets, most often derivatives. Black-Scholes (1973) and lattice models are the most common.

The choice between these methods is made when there is no active market in which to value the asset. The IFRS has a key role to play within the world of valuation, in IAS 39 fair value accounting is described. IFRS 13 was adopted in May 12, 2011 and provides guidance on performing fair value measurement under IFRS but it is not defined in which situations these measurements should be done. Generally they state that a financial statement should reflect the true and fair view of the business affairs of an organisation. The financial reports are seen by various constituents of society and need to reflect the true value and financial position of an organisation at a point in time. FAS 157, issued on September 2006, states that the fair value on an asset is the value that it could be sold at to a willing party. This conflicts with the historical cost method that says for example, if a piece of land was bought in 1980 for $1 million, it would still be recorded on the balance sheet as a historical cost basis. The fair value tries to capture the current price, in the Efficient Market Hypothesis, the market incorporates new information quickly and so the market price is assumed the fair price - yet there are strong and weak forms of this hypothesis, and it assumes that prices are observed. Within behavioural finance there are anomalies, as agents are influenced by irrationality and so prices in the market are believed to diverge from fair value. The IFRS authorise three underlying assumptions, which define how professionals should incorporate the concepts behind valuation:

1 – Going concern: An organisation will continue for the foreseeable future under the historical cost paradigm and using units of constant purchasing power paradigm.

2 – Stable measuring unit assumption: Accountants should consider the changes in purchasing power of the functional currency up to 26% per annum for 3 years in a row as immaterial.

3 – Units of constant purchasing power: considering remedies for an indefinite time period the erosion caused by the historical cost accounting of the real values of constant real value non-monetary items.

Under these three assumptions the framework under FAS 157 has three levels for inputs to a calculation when valuing:

1. The first level prefers inputs to valuation that are "quoted prices in active markets for identical assets or liabilities", for example a stock traded on the New York Stock Exchange. This is based on direct observations of transactions.
2. The second level is based on market observables, acknowledging that active markets for identical assets and liabilities are relatively uncommon and even if they exists there may be little liquidity. An example would be the price of an illiquid option based on the Black-Scholes model (1973) using implied volatility techniques. Here fair value is estimated using a valuation technique, the use of inputs on the markets is sometimes required.
3. This level is regarded as "unobservable", if one and two inputs are not available a valuation technique is required, however, significant assumptions which are not observable on the market are required, this is known as *mark to management*.

IFRS 13 gives guidance on carrying out the measurement of valuation, these show where new methodologies could be of benefit:

[IFRS 13:11] An entity should consider the properties of the asset or liability being measured that an investor would consider when pricing at a particular date.

[IFRS 13:15] Fair value measurement conjectures that a transaction is carried out orderly between buyer and seller under the current market conditions.

[IFRS 13:24] Fair value measurement conjectures that the transaction takes place in the most advantageous market at that period in time.

[IFRS 13:27] A fair value measurement of a non-financial asset considers the highest and best application.

[IFRS 13:34] A transaction of an entity's own equity products is given to an investor without settlement, extinguishment or cancellation at the valuation date.

[IFRS 13:42] The fair value of a liability represents the credit and non-performance risk before and after the transaction.

[IFRS 13:48] An exception is applied for financial assets and financial liabilities with positions or counterparty credit risk that offsets the original transaction, with disclosure.

Annual reporting periods beginning after the 1st January 2013 can apply IFRS 13. A previous accounting time period can make use of these valuation definitions but it must be stated on the financial reports.

Most of the value of a business is in its ability to generate profit. Within business the term Gross Margin is often used as an input to EBITDA which gives a guideline to a company's value when inserted into a DCF analysis. EBITDA is earnings before interest, tax, depreciation and amortisation. A significant reporting figure calculated in the refining industry is the *gross refining margin ("GRM")*, this represents the difference in the cost of crude oil and the selling prices of the refined products sold. Another is the Gross Marketing Margin (*GMM),* this is a figure quoted heavily by managers within the industry; the GMM represents the price from the refinery to selling to the retail customers, for example petrol stations. The GRM can be calculated differently dependent upon the company and that is why the IFRS does not recognise it as an official measure[32], but management and investors within the industry consider it an important figure representative of the profitability of an oil refinery. The issue with the GRM is it does not take into account the operating expenses. Another significant financial figure is the refinery cash margin.

[32] FAS 130

2.2 Refinery cash margin, RCM

RCM ($/bbl) = product price ($/bbl) – cost of crude ($/bbl) – operating expenses ($/bbl) (1)

Here we assume that the *RCM* multiplied by crude throughput (the capacity for refining crude oil over a given period of time) is roughly equal to EBITDA. It is a measure of cash earnings or profit / (loss) in accrual accounting. This is another figure that the IFRS does not recognise but it is utilised within industry by investors comparing the performance of companies in the same sectors as credit periods, which will be broadly similar. EBITDA may not be representative of the Company's historical operating results and it is not meant to be forecasted. However, it is a measure commonly sought by investors to get a handle on a company's performance. The *RCM* is then a convenient calculation to represent the performance of the refinery in annual terms. We also assume that if the throughout annually is multiplied by the RCM for a reporting period it will represent a valuation figure. This is primarily due to the fact that an oil refinery generates profits on the difference between the prices of the refined product slate and the cost of crude oil, therefore, the RCM will encapsulate most of its business value. This will not include all capital expenditures and some other important financial figures, like the associated debt, but we argue that it is a relevant estimate of how much profit is being intrinsically generated. There are however many types of refinery in different geographical complexes. The aim was to set up a standard valuation for an average oil refinery in terms of size and complexity; laying the groundwork for others. A useful measure within the industry to represent a refinery's complexity is the Nelson Complexity Index.

The Nelson Complexity Index is a useful figure to gage the level of refining complexity being carried out at a particular refinery. It assigns a factor to each piece of refining machinery that stems from its complexity and cost compared to the basic process of crude distillation (a complexity factor of 1.0). Multiplying this factor by the throughput ratio as a fraction of the distillation capacity gives the complexity; aggregating these final values together for each part of the refinery gives the final Nelson Complexity Value. It was developed by Wilbur L. Nelson as described in the Oil and Gas journal (1960). The greater this index number, the greater the cost of the refinery, and the higher the value of its products. US refineries on average rank the highest in complexity index with a value of 9.5, Europe at 6.5; the Jamnagar refinery owned by Reliance Industries Ltd is one of the most complex in the world at 14; the most being the Texas City Refinery at 15.3.

We sought a refinery that has a Nelson Complexity Index above 5 in order for an optimisation to be worthwhile; the less complex refineries, on the whole, do not generate enough revenue for a comparison of alternative decisions within the program to be valid. A Nelson Complexity Index lower than this, means that only a few products are refined and the selection choices for optimisation are simplified. Those refineries that exist with a lower complexity index are finding the economic environment much more strenuous, due to their smaller range of products, giving them less flexibility and a fraction of the revenue of the major refineries.

It is important to obtain realistic financial data to test the methods of valuation that are being researched and calculated. In this chapter the following refinery's financial statements are examined with the aim of creating an additional optimisation analysis based on available company information from Companies House UK or on company web pages:

- Grangemouth, Scotland
- Coryton (Owned by Petroplus), England
- Humber (Owned by ConcoPhillips), England
- Lindsey Oil (Owned by Total), England
- Fawley (ExxonMobil), England
- Pembroke (was sold to Valero for £750 million in 2009), England
- Milford Haven, England
- Stanlow (Owned by Essar energy), England
- Vadinar, India

We decide that the ***Vadinar*** refinery in Gujurat is the most appropriate to analyse and value, due to its huge expansion ambitions and its slightly greater than mid level complexity.

The Vadinar Refinery owned by Essar energy is a relatively new and medium sized refinery in India, yet it has plans to be one of the biggest in the next couple of years. Inserting numbers from the Vadinar oil refinery in India into a formula, the Nelson Complexity Index of the refinery is equal to 6.1 and is supposed to increase to 11.8 following completion of the Phase I Refinery Project and to 12.8 after Phase II of the refinery capacity expansion project. The following is a calculation of the Vadinar refinery's GRM:

2.3. Vadinar gross refining margin calculation

For example, the financial statements for the Vadinar refinery for the year ended 31 March 2009 enable a calculation of the GRM:

***(Table 2.1**: Reconciling Revenue from the refining part of the business to GRM, for the period 1 May 2008 – 31 March 2009 and the 9 month period to 31 December 2009)*

	Period Ended 31 Mar 2009	Period Ended 31 Dec 2009
Revenue – Refined Petroleum Products ……….	**7,689.8**	**4,965.3**
Cost of crude oil ……………………………….	**(7,331.2)**	**(4,771.9)**
Sales tax incentives ……………………………….	**256.3**	**168.4**
Commodity hedging gains/losses ……………….	**78.9**	**(40.0)**
GRM (including sales tax incentives) ...………...	**693.8**	**321.8**
GRM (excluding sales tax incentives) ...………...	**437.5**	**153.4**
Number of barrels (in millions)…………………	**87**	**72**
GRM per barrel (including sales tax incentives)…	**7.97**	**4.46**
GRM per barrel (excluding sales tax incentives)…	**5.02**	**2.12**

[The IEA benchmark was US$2.33 per barrel in the same period.]
[All values are in US$ millions]

Revenue above is defined as the cash inflows in the period for producing petroleum products. The cost of crude oil is the cash outflow for the barrels of the raw crude needed to produce the refined products. Sales tax incentives are allowances made by the state of Gujurat that may not allow tax breaks to continue and may not allow a direct comparison of GRMs with other refineries.

The Vadinar refinery processed ***11.95*** million metric tonnes (mmt) of crude oil between 1 May 2008 and 31st March 2009 and ***9.90*** mmt of crude oil between 1 April 2009 and 31st December 2009. Hence in this period, the throughput was approximately ***75.24*** million barrels, see above, approximately ***72*** million barrels (see calculation below).

In this period the refinery sold ***9.26*** mmt of refined petroleum products.

The on average conversion factor used is:

1 metric tonne = 7.6 barrels of crude oil.

The Vadinar refinery's current refining operating costs are $1 lower than the average across the industry, below in table 2.2 are the operating costs per barrel, based on data from their financial statements made public, for the 9 months ended 31 December 2009.

(Table 2.2: Operating Costs per barrel at the Vadinar refinery)

Cost Item	**US$/barrel**
Asset Management Costs[1] …………………………………………………….	0.30
Manpower Costs …………………………………………………………..	0.15
Purchased energy ………………………………………………………….	0.31
Other overhead Costs ……………………………………………………...	0.39
Operational variable Costs ………………………………………………...	0.19
Total Costs ………………………………………………………………...	**1.34**

(1) Includes costs of shutdowns, turnarounds, maintenance and repair and other routine maintenance

The company in this period also had corporate and marketing expenses of US$ ***0.96*** per barrel.

In total, operating costs were $US ***2.29*** per barrel. Therefore an estimate of the refinery's profitability for one year is:

RCM = [($4.46 – $2.29)*(72e6 barrels)]*(12/9) = **US$ 208.32 million (With tax incentives)**

Below table 2.3 shows the value each refined product adds to the refinery complex.

(Table 2.3: *Production of the refinery broken down by refined product)*

	1 May 2008 to 31 March 2009 Percentage	**1 April 2009 to 31 December 2009 Percentage**

	of total mmt production		of total mmt production	
Production:				
Liquefied petroleum gas	0.46	3.8%	0.40	4.1%
Naphtha	0.03	0.3%	0.11	1.1%
Motor spirit	2.10	17.6%	1.67	16.8%
Kerosene/Aviation turbine fuel	0.57	4.8%	0.66	6.7%
High speed diesel	5.10	42.7%	4.02	40.6%
Fuel oil	2.72	22.7%	1.97	19.9%
Sulphur	0.07	0.6%	0.06	0.6%
Bitumen	0.13	1.1%	0.40	4.1%
Fuel loss	0.77	6.4%	0.61	6.1%
Total	**11.95**	**100%**	**9.90**	**100%**

[Motor Spirits is Gasoline and Petrol]

The above Vadinar Refinery figures meant that the revenue could be broken down into the value generated per refined product. Even though the price of each refined product will have been fluctuating over this period, using an *average price* we can depict the value per product sold during this period.

(Table 2.4: *Revenue generated by each refined product during periods shown)*

	1 May 2008 to 31 March 2009 US$ millions	Percentage of total revenue	1 April 2009 to 31 Dec 2009 US$ millions	Percentage of total revenue
Production:				
Liquefied petroleum gas	$292.21	3.8%	$203.56	4.1%
Naphtha	$23.07	0.3%	$54.62	1.1%
Motor spirit	$1353.45	17.6%	$834.17	16.8%
Kerosene/Aviation turbine fuel	$369.11	4.8%	$332.68	6.7%
High speed diesel	$3283.55	42.7%	$2015.92	0.6%
Fuel oil	$1745.58	22.7%	$988.09	9.9%
Sulphur	$46.14	0.6%	$29.79	0.6%
Bitumen	$84.59	1.1%	$203.58	4.1%
Fuel loss	$492.15	6.4%	$302.88	6.1%
Total	**$7689.8**	**100%**	**$4965.3**	**100%**

Growing a company through project value is considered under investment decisions; typically this is analysed with the hegemonic discounted cash flow analysis. Net Present Value (NPV), the internal rate of return (IRR), and payback period, are key methods to evaluate the attractiveness of investments within the finance industry. This is true whether it is an acquisition of a company or an asset that is to be purchased.

The goal of discounted cash flow analysis is to determine:

- The net present value of a stream of expected future cash revenues and expenditures
- The rate of return which the expected future cash flows will yield on a given level of initial investment

The oil refinery presented here has fixed operational costs, hence the analysis of costs in the refinery will not be as involved as the valuation of the commodity prices of crude bought and refined products sold - this is the greatest source of financial risk the refinery has. Applying income based methods is industry consistent; the calculation here is in 2 parts:

- Calculation of each cash flow for some future number of years
- The aggregation of the present value worth of each cash flow

A refinery is generating its profits from the net difference between the price of a barrel of crude oil and the value of the refined products. On top of this, the cost of operating the refinery *(fuel and staff costs)* and the costs of freight to the port of the refinery must be injected into the calculation. Example figures are given below:

The Vadinar Gross refining margin in US$ for December 2009 for 11 months is: US$2.12 per barrel. On average it takes the Vadinar refinery 20-30 days to process crude oil from the day it receives the crude oil, hence producing an estimate for the refining process sold timeline. Each month's price series can be forecasted, in terms of optimising the profits of the process with a mathematical model. The higher the *GRM* the more profitable the refinery will be. The way to increase the *GRM* is to refine as efficiently as possible the particular crude assay into the highest priced market products for each time period. In general transportation fuels are the most valuable, however, one particular barrel of crude yields approximate volumes of each product, barring yield uncertainty in the CDU and an amount due to CDU reactions that can increase the production of particular products. After we obtain the DCF valuation of the refinery with traditional methods, the goal for the next chapter is to create a model to optimise the refinery's processes and maximise its *GRM* each month over a number of years.

Any valuation should take into account the risk of an investment, and although the refinery's greatest risks are the prices of the commodities involved, other valuation risks are investigated.

2.4 Valuation risks to the Oil Refinery

Risks for a typical oil refinery would be as follows:

- Fluctuation of crude oil and refined petroleum product prices and refining margins
- Inability to enter into term contracts for mid to long term crude oil purchases
- Contract dependence on more favourable pricing of its refined petroleum product sales to the markets and companies the refinery is located in
- Ability to find, develop and commercially exploit resources and reserves
- Foreign exchange currency fluctuation risk

Most oil refineries purchase third party software to determine how to optimise the refinery's product slate due to the many factors that impact the chemical engineers decision. In the following chapter the uncertainty from a financial aspect based upon choices of which products to market (planning) are considered in a model, but there is also uncertainty present in crude delivery time (scheduling), yield uncertainty processed in the CDU (blending), quality specifications and these decisions adhere to all the physical, mass balance and financial constraints on the refinery. This decision process is far too complex for a human to completely manage no matter how experienced the refiner. Some companies creating optimisation industry software that deal with some or all of these complexities are: *IBM, Deliotte and Touche*, *Aspens Process Industry Modelling Systems, PetroSim (KBC, kinetic based models used in over 100 refineries worldwide) and Spiral software optimisation (claimed in person that they can save an average complexity refinery $8.5 million per year through blend optimisation).* Optimising the buying and selling of the products is vital to the performance of the refinery in a very competitive business that has recently seen many smaller refineries close down and pushed the majors into shrinking their refining interests. It is vitally important that the risks of the refinery are taken into account in its financial valuation.

A more accurate valuation methodology implemented less often due to its slight increase in complexity is the Real option valuation approach. In practise, however, DCF and NPV are applied ubiquitously.

2.5 Discounted cash flow value (DCF)

The real option approach utilises an option pricing equation in the investment decision. A real option is one in which the user has the right, but not the obligation, to commit to a decision or not, to purchase a real asset at a particular fixed price, named the strike price. For example, the option to invest in a gold mine or not, would be a real call option approach to valuing the mine, where the gold value is the state variable and the strike price would be the investment cost. It is a logical approach with which to value an investment, whilst taking into account the uncertainty often present in financial decisions, that are missed by static income based methods. Industry usually deliberates between using the standard methods for an investment decision, NPV, DCF or the more advanced Real option approach. The DCF approach has to be calculated for at least 6 to 10 years into the future, dependent upon the type of company, to estimate a fair value of the company or asset; for instance, all UK/US courts in a financial prosecution regard it as an official method with which to value a company33. It is not sensible for certain types of businesses and so the real options approach can be adopted instead. There are arguments within the literature that claim option pricing is over valued. Trigeorgis and Mason, 1987 argue that the existence of a 'twin security' is implicitly assumed in NPV for purposes of estimating the required rate of return. The assumptions underlying the application of option pricing to real assets are no stronger than the assumptions underlying the application of NPV to real assets. The accuracy of using NPV relies however on the assumption of market completeness; the underlying asset itself for the refinery here is not traded. The DCF analysis is an approximation of the value of the refinery if

[33] http://www.nybusinessdivorce.com/2012/08/articles/delaware/recent-fair-value-cases-in-the-delaware-chancery-court/

it were traded. The random behaviour of the underlying assets that must be applied with real options is an approximation of the future behaviour that the refinery would have if it were traded. Consequently, the real option value is an estimate of the value the option would have if the underlying assets were traded and behaved in the manner that the random prices are depicted. The majority of real asset option-pricing applications are where the option value depends on the market price of a commodity such as crude oil. Simulating the random prices and solving with a lattice approach, however, enables the valuer to solve a much wider set of valuations than ever before, in a manner that ensure managerial support due to its simplicity.

The cash flows generated in the future for the DCF are estimated from industry wide figures like forecasted macro-economic figures or relevant demand figures published by respected organisations (KBC regularly carry out in depth reports on these exact figures). The issue with the DCF is that it is a static measure due to the growth rate of projected cash flows is already ignoring the instability or behaviour of asset prices in financial markets and neglecting management's right to alternate strategies. The real option approach deals with these issues.

In the DCF approach, only the most likely outcomes are considered, and the flexibility of the decisions available to management is ignored, a statistical approach to the random variables in question are not considered. The future cash flows are discounted back to present value, but the rate at which this is done leaves much room for error. The company that is being valued usually cannot borrow and lend at the risk free rate and has risk attached to its operations. Trying to capture this risk by a number called the cost of capital is an intricate and often inaccurate calculation. The NPV approach assumes a capital investment once committed is left to play out and so is fixed. With a varying cost of capital calculations do not properly account for the changes in risk over the lifecycle of the business. Whereas, the real option approach assumes that the investment can be modified contingent upon the various scenarios that can happen, this is its advantage. Many methods of valuation require an estimation of uncertain cash flows; an example would be the *Gordon Growth Dividend* model.

"If future outcomes are uncertain, then any estimation of growth and risk may not be sufficiently well-informed that the assumptions underlying the simple dividend discount model can be rejected."[34]

In the real option approach the underlying concept is that of risk adjusted valuation, or applying the risk neutral measure. Typically the real options approach gives a greater value than the NPV; attaching an amount of value to alternative strategic decisions. The initial assumptions inherited for a DCF or a NPV approach is the same for a real options approach. Complete markets and absence of arbitrage are formally required for the calculation to hold. The aim here is to calculate a DCF value for the Essar Oil refinery and then show that the real option approach could be more representative. The real option approach is a difficult calculation in comparison to DCF, hence a method with a one period stochastic program is developed to value the refinery whilst accepting the inherent uncertainty. In the process of understanding NPV analysis it is assumed that all future cash flows are certain for the next six to ten years, based on historical figures, which removes the uncertainty inherent in the generated cash flows of a business. The exact assumptions behind NPV analysis are listed below:

Assumption 1) any new cash flow stream can be exactly replicated by a combination of securities that already exist in financial markets. *(Complete markets)*

Assumption 2) the financial markets do not allow arbitrage.

Any NPV valuation that is positive is considered to be adding value, meaning it should be undertaken. The advantage of DCF over other traditional methods are numerous, for instance often P/E ratios and multiples are forecasted and implemented in a valuation. These multiples are not of much use if an entire sector is over or undervalued, whereas DCF will give a fair value in comparison to this approach, but it can over or under value to an extent if the economic environment becomes different to the historical values injected as inputs. For a NPV calculation, the vital component is the discount rate; an input to the valuation formula, and it is generally calculated with the method named Weighted Average Cost of Capital (WACC) or CAPM.

[34] The idea is that the stock is valued by using predicted dividends and discounting them back. It has no relevance for companies that do not pay dividends.

If the asset being valued is illiquid then the discount rates must be adjusted; there are three typical ways in which this can be done:

- Estimate the value of the refinery as if it were a liquid asset and then discount that value for illiquidity
- Adjust the discount rates and use a higher discount rate for the refinery
- Estimate the illiquidity discount by looking at comparable refineries and recording how much their values are impacted by illiquidity

As carried out by KBC we value the Vadinar refinery using a cost of capital using WACC, which is adjusted for the industry we are in; the petroleum sector.

2.6 Weighted Average Cost of Capital (WACC)

Weighted average cost of capital, WACC, is the rate of return a company pays on average to all its shareholders. The company must earn this rate on its' investments and services in order to satisfy all shareholders and interested parties. The WACC takes into account the capital structure of the company and averages all the relative weights of value generated by the different sources of capital within it. What should this discount rate be for the oil refinery? The challenges that arise in applying CAPM or WACC to illiquid real assets, such as a chemical plant or a refinery raise a fundamental issue. How do we assess risk in illiquid asset classes where market prices cannot be observed (as is the case for an Indian refinery) and hence subject to substantial uncertainty?

A simple version of this discount rate calculation is shown below:

$$\textbf{WACC} = k_e*(E/(E+D)) + k_d*(D/(E+D)) \qquad (2)$$

Where: D = the value of debt

E = the value of equity

(Note that k_d is the cost of debt after tax)

Evaluating the WACC for a company is different than evaluating the WACC for an individual energy project. When WACC for a company is evaluated, we are often trying to determine (under imperfect information) what a company's costs of capital are. In this case we would utilise as much financial data as possible to estimate the various terms in the WACC equation. For an individual energy project, the various terms in the WACC equation are determined in large part by the type of investment being made, the type of market (regulated versus unregulated) in which the investment is occurring and the individual making the investments. Historically, (looking at KBC reports)[35] the petroleum industry has financed around 15-20% of its activities through debt (this can be seen on the "cost of capital by sector" web page [36], there are sector numbers available for India and KBR reports support the numbers within 2 decimal places; see the second to last column in the petroleum rows - showing the ratio of debt to equity). Hence, we can estimate that Essar energy (the owners of the Vadinar refinery) would finance 15% of its activities with debt. Oil company bonds have historically high ratings, so we assume that Essar energy's bonds are AAA; we see that the 10 year AAA corporate bond would have a yield of 3.08% (at the time of this writing) - the cost of debt financing. Turning now to the cost of equity financing, we need the return on the safe asset; the market risk premium; and the Beta for the petroleum industry. The yield on the 30 year treasury bond was 3.78%. We can assume a risk premium of 5% and from the *cost of capital* web page we see beta for the petroleum industry is between 1.17 and 1.45. Thus, the cost of equity for Essar energy would be 0.0378 + (1.45 x 0.05) = 11%. The WACC represents the discount rate that a company should use in conducting a discounted cash flow analysis of a given energy project. We use the cost of equity of 3.08% and a cost of debt as 11%, which is very close to KBC's calculations for the same period.

In a state of the economy where a payoff is less likely, investors are willing to pay more for it. This is where the confusion of the discount rate can arise. Each future state of nature has its own unique discount

[35] http://www.essar.com/upload/pdf/EOL_AR_2012-13.pdf

[36] http://people.stern.nyu.edu/adamodar/New_Home_Page/datafile/wacc.htm

rate, but the DCF approach does not take this into account. As an example consider a power plant; this can be switched on and off based on demand - a DCF calculation ignores this embedded optionality in this real asset, there is value attached to being able to switch strategy. An important real option approach was created by Cox, Ross and Rubinstein (CRR) (1979), they apply the binomial method. This is due to Harrison and Kreps (1979*),* where the risk neutral probabilities are actually state prices for securities scaled by the risk free factor. If markets are complete and free of arbitrage opportunities, then there exists a unique set of 'risk-neutral probabilities' that can be found. Usually option pricing techniques can only be applied if markets are complete and free of arbitrage opportunities.

This is enough to claim that if a manager is willing to use DCF analysis on an illiquid new investment, a gold mine, factory or structured product, then the assumptions behind real option analysis have already been assumed and the benefits should not be ignored. Valuation of the refinery is similar to asking the following: if the refinery was bought and then immediately sold to the financial markets, what would it be worth? The present value of the cash flows generated by the refinery are those that the capital markets would pay for now. The NPV is the difference between the price actually paid for the new real asset (the PV of the outflows) and the price that would be received for the cash flows on the refinery in the financial market (the PV of the inflows). Therefore the capital budgeting method of valuation is figuring out what the cash flows of the refinery are worth to the financial markets. A static approach to valuation is applied first to gage the result and consequently uncertainty is introduced. Surveys completed by corporate finance practitioners have continuously found that real options are only used at most in 25% of cases[37].

A DCF method generally relied upon in the FTSE 100 and S&P 200, is the Free Cash Flow approach, FCF - see Graham and Campbell (2001) for an example and analysis of various valuations. Investment analysts compute the free cash flow: Free Cash Flow is equal to Operating cash flow, less expenditures to maintain assets (not including increases in working capital) less the interest charges. It is the actual amount of cash that the company has left from its operations that could be invested to pursue profitable projects to enhance shareholder value. The method itself has less vagueness in its calculation than the net income approach, hence manipulation is trickier. For companies with stable capital expenditures,

[37] Graham, J. and H. Campbell, "Theory and Practice of Corporate Finance: Evidence from the Field," *Journal of Financial Economics*, 60 (2001): 187-243.

free cash flow will be approximately equal to earnings, see Jensen C, (1986). The steps for the Vadinar refinery in Gujurat are shown below. (This is the free cash flow method to the firm, FCFF).

2.7. Free Cash Flow Method (FCFF)

Step 1 - Calculate the free cash flow

The free cash flow is defined as:

FCF = Earnings before interest and taxes (EBIT)

Less: Tax on EBIT

Add back: Non-cash charges

Less: Capital Expenditure

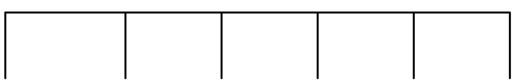
Less: Net Working Capital Increases

Plus: Net Working Capital Decreases

Plus: Salvage Values received

Step 2 - Forecast FCF and the terminal value

For a time horizon of six years the following values are estimated: FCF1 → FCF6, the terminal value. Six years is chosen as Vadinar refinery's owner is a large company with expansion plans for the next five to ten years, values beyond this are generally discounted into much smaller values, hence considered irrelevant for our purposes. When forecasting these cash flows many factors should to be taken into account as they are estimated by obtaining a growth rate. The economic environment globally, locally, the industry the company is in, the company’s current status amongst its plans and its leverage level are all relevant in estimates of future cash flows.



FCF1 FCF2 FCF3 FCF4 FCF5 FCF6

Step 3 - Calculate the weighted average cost of capital (WACC)

The WACC is calculated using the cost of equity (k_e) and the cost of debt (k_d)

(See equation (2) above)

Step 4 - Discount the free cash flows at WACC and aggregate to obtain value of the firm

Step 5 - Calculate the equity value

Firm value = Equity + Debt (3)

2.8. Oil refinery FCF calculation

Step 1

*FCF = NI + NCC + [Int * (1- tax rate)] – FCInv – WCInv* (4)

Where:

NI = net income

NCC = non-cash charges

Int = interest expense

FCInv = fixed capital investment (capital expenditures) = Change in Gross PP&E

WCInv = working capital investment = Change in working Capital Accounts = Change in (AcctsRec + Inventory – AcctsPay)

We can also start with earnings before interest, taxes, depreciation, and amortization (EBITDA) to arrive at FCF:

FCF = [EBITDA * (1- tax rate)] + (Dep * tax rate)- FCInv – WCInv (5)

Where:

EBITDA = earnings before interest, taxes, depreciation, and amortization

FCF can also be estimated by starting with cash flow from operations (CFO) from the statement of cash flows:

$$FCF = CFO + [Int * (1 - \text{tax rate})] - FCInv \quad (6)$$

Where: CFO = cash flow from operations.

Indian corporation tax rate as December 2009 = 33.99%

Depreciation for 31st March 2009 to 31st December 2009 = US$ 89.9 million

WCInv = - US$ 80.10 million

In table 2.5 the calculation for the free cash flows for the company Essar Energy are obtained and by making an assumption on how much value is generated by refining, the value of the refinery follows:

***(Table 2.5:** FCF Calculation[38])*

1) Calculate the free cash flow

All figures in US$millions			**$m**	
FCF = Earnings before interest and taxes (EBIT)			363.00	
EBITDA Dec 09		433.10		
Power	120.80			
Oil	312.30			
	433.10			
Less tax on EBIT			- (27.60)	
Add back non-cash charges			89.90	
Less cap expenditure			- (405.50)	Cash flow statement

38 Obtained by using official material from the ICAEW, The Institute of Accountants for England and Wales

Less net working cap increases	- (656.60)	Cash flow statement
Plus net working cap decreases	736.7	Cash flow statement
FCF	**99.9**	
Per (12/9months)	133.2	
(WACC = Ke*E/(E+D) + Kd*D/(E+D))	*10.77%*	

Main Assumption for valuation: 72% of profits and costs are due to the oil refinery the rest are due to the company's power business; therefore as an estimate, numbers generated above representing the free cash flow value are multiplied by 72%:

FCF Oil Refinery: 0.72*133.2 = US$ 95.9 million

Step 2

There are 2 methods generally applicable to forecasting the cash flows, one is to calculate the historical free cash flow and apply a growth rate, under the assumptions that the growth will be constant and fundamental factors will be maintained. The second method, is to forecast the underlying components of free cash flows: costs, revenues, tax changes for each year in the future and calculate each year separately. The simpler first method is utilised here as a first draft calculation. How does one estimate growth? This is entirely down to discretion and a full in depth economic report of predicted oil demand and prices within the Indian region by a company called KBC[39] is referenced for growth figures in this region.

Gross Profits for the refinery in US$ millions for the following years were:

2007	2008	2009
40	**74.9**	**681.7**

[39] KBC's (2011) *Outlook for the World Refining Industry*

The last figure is a big jump and is down to the refinery profits being realised and some derivative contracts maturing.

If in 2011, the phase I project is realised for the refinery, the capacity will double meaning twice the products can be refined. These figures were difficult to estimate due to the current economic climate.

Instead we apply KBC forecasts for the oil demand for the next 6 years going from *3107kbpd* to *4262kbpd*, due to the growth in the region, a growth rate per year is applicable to the refinery of:

$3107 *(1+g)^{7} = 4262$, therefore **g = 4.62%**

(KBC are a company that carry out extensive research in the commodities space and have been hired by Essar Energy).

Global oil demand in Europe and the US will most likely decline and the majority of demand in the next decade will be from the Asia-Pacific region as shown below in table 2.6:

*(**Table 2.6:** Indian refined product demand, -'000 bbl/d)*

	2008	**2009**	**2010**	**2011**	**2012**	**2015**
LPG	406	416	432	450	468	526
Naphtha	336	352	352	366	380	426
Gasoline	260	299	332	363	394	478
Kerosene	303	301	316	329	344	395
Gas/Diesel	1050	1112	1184	1246	1314	1553
Fuel Oil	390	404	402	401	400	398
Others	362	379	395	412	429	487
Total	3107	3264	3414	3568	3729	4262

Source: KBC

Step 3

WACC was calculated as approximately equal to ***10.77%***

Step 4

FCF1 = 96 *(1.0462) = 100.33

FCF2 = 96 *$(1.0462)^2$ = 105

FCF3 = 96 *$(1.0462)^3$ = 110

FCF4 = 96 *$(1.0462)^4$ = 115

FCF5 = 96 *$(1.0462)^5$ = 120.20

FCF6 = 96 *$(1.0462)^6$ = 126

Step 5

Fair Value of the Vadinar Refinery = (\$100.33/ 1.1077) + (\$105/ 1.1077^2) + (\$110/ 1.1077^3) + (\$115/ 1.1077^4) + (\$120.20/ 1.1077^5) + (\$126/ 1.1077^6) = <u>US\$ 569.75 million</u>

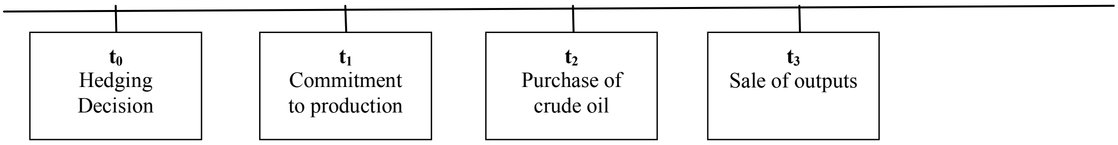
With the refining capacity of Vadinar currently at around 267 kbpd this is on the lower side of estimates, but looking at the latest share price for Essar Energy it looks like a reasonable estimate of terminal value. *(Share Price of Essar on June 2011 is [ESSR on LSE] 405 pence).*

The problem with the above value is that it does not take into account the variation in the profits of the oil refinery; due to the estimates of the economic variables being static. If a weather event for example hits India, this could seriously affect profits for a number of business quarters. This method states that growth will be consistent with estimates of oil demand over the next decade. Another effect ignored is that of management changing strategic decisions, they could for example decide to change output products due to various economic effects that there is no control over. This would drastically alter the nature of the cash flows over time. Another issue is the calculation of the cost of capital for the company, which is always an equivocal calculation. An alternative would be to use a cost of capital that is industry wide accepted, sometimes printed by various financial organisations, to take account of the capital structure of the company. Again although DCF is a reliable and tractable method, it has many pitfalls. The FCF method gave an approximate value of $570 million for the refinery complex; in contrast a financial market data calculation which is applied in the next chapter.

2.9 Real Option Approach

Most refinery related contracts are traded on NYMEX; where there are numerous contracts that expose quantities and values related to the refinery market. These exchanges trade what is referred to as "light-sweet" crude oil and a single contract, or "lot", refers to the purchase or sale of 1,000 barrels of oil. A large fraction of the activity on oil Futures markets is concentrated on delivery for the next three months. In our mathematical model we will specify the delivery to be one month ahead, or 30 days. The quoted Futures price is the price at which the owner of the contract will buy/sell crude or the refined products during this delivery period per 1000 barrels. For each buying or selling of the contract, we do not consider transaction fees or a bid-ask spread. "Fair value" of an asset or liability should be based on the current market price, this is known as marking-to-market. This type of accounting can change values on the balance sheet as the financial markets evolve. It supersedes the historical cost account based on past revenues. A book value should always be compared to a fair value. In substitution of the above method, companies often apply marking-to model, where a market for the asset or product is not available. Models can give an artificial illusion of liquidity but are absolutely necessary for illiquid or complex assets. Any advanced model dealing with valuation should attempt to capture market stress whence bid-ask spreads widen. In 2003 Warren Buffet was famous for saying "mark-to-myth" when referring to the methods applied by financial institutions to value their derivative positions, a statement that emphasises the difficulty in making fair a valuation.

Distillation of a liquid can occur almost instantaneously, in the refinery however the production is not taking place instantaneously; imagine the process time line present for the oil refiner:



(***Figure 2.1:*** Timeline for refinery decision maker)

Contracts listed on an exchange that manage the differential risk between crude and its products, can be purchased by the refiner to lock in a refining margin and ensure that either long term contractual obligations are met or the crack risk is hedged accordingly. Chicago board of trade (CBOT) offers a number of contracts on refinery spreads that are relevant to refineries throughout the US. Initially, before production and sales are decided, the DM can assess the market dynamics and consider trading financial contracts to manage the inherent risk. The table below defines factors that are significant to the decision maker considering hedging or speculating on the crack spread's value.

This data shows that the financial contracts listed can be indicative of the profit that can be extracted from the crack spread, as they enable traders to manage the factors impacting the refiner's profit margin.

A standard real option comparison was the next stage in assessing the financial worth of the refinery. An oil refinery is similar to a chemical plant in many ways and can be considered a producer with exposure to a differential spread. The real difficulty in an option approach is the pricing of the choices embedded within it. The choices available to real assets usually fall into one of the following categories: the option to expand, the option to abandon and the option to defer. Specific to the oil refinery is the choice available to the decision maker to vary purchasing and production decisions of the crude and refined products. The risk-neutral valuation approach can be extended to allow real option valuation of land, plant, equipment and assets. This approach does not require risk adjusted discount rates but it does need market price of risk parameters for all stochastic variables. Many investments are underpinned by uncertainty connected to future prices. These future prices can be used explicitly to estimate the risk-neutral stochastic processes involved in the calculation and avoid the requirement to obtain a market price of risk for the random variables exemplar set.

Based on the decision maker's (DM) initial set of decisions the other choices available would have been altered due to the knock on effects of the limitations on the particular crude assay being refined. For example, producing more gasoline means less heating oil can be produced. The CDU must maintain the mass balances at all times, therefore some choices are not possible, e.g. producing all kerosene or no Naphtha at all. If for instance the refiner knew that residual fuel oil would be the highest valued product in one month time, then as much of this fuel oil as could be physically stored could be processed. The refiner has an embedded option that is strictly constrained due to the chemical processes that are viable.

This is investigated as a cacophony of commodity future prices and correlations in the following chapter. Trivially this value can be captured by the 3-2-1 crack spread. There are a number of spreads available on financial exchanges, an example of a few follow:

- 1:1 one crude oil versus one gasoline or one heating oil
- 2:1:1 two crude oil versus one gasoline and one heating oil
- 3:2:1 three crude oil versus two gasoline and one heating oil
- 5:2:1 five crude oil versus two gasoline and one heating oil

The 3-2-1 contract is the most frequently traded spread contract and represents the amount of value in a barrel of WTI crude refined into its two most valuable products, gasoline and heating oil. Utilising data from CBOT on the 3-2-1 crack spread an option approach of valuation is built. If the refining cash margin is positive the DM should refine, if negative, the refinery should be switched off.

This idiosyncrasy of pricing would be of interest in the current commodity climate, where independent traders, private equity firms and even airlines are invested heavily in the refining industry[40]. This is due to companies either yearning to be vertically integrated or investing to capture diversification opportunities. Delta airlines for instance bought the Trainer refinery, 185,000 bpd, in June 2012 and reconfigured it to produce more jet fuel.

At any period in time, without purchasing additional financial contracts, the refiner is short crude oil, as it needs buying, and is long the refined products, as the DM has the ability to produce them. Therefore the refiner is long the crack spread. In practise refinery traders will construct trades on the constituents of the crack spread based on their estimation of the spread being under or overvalued. If for instance a dispersion trader feels the crack spread is overvalued then it will be sold short, and vice versa. The hedging replication argument allows the refinery to be valued using a strip of crack spread options; it is then compared to the DCF valuation previously calculated.

"The discounted cash flow (DCF) criterion induces investment too early (i.e., when prices are too low), but the real options approach induces investment too late (i.e. when prices are too high) when it neglects mean reversion in prices".

2.10. Problems with simple real option approach

In practice, the valuation of the refinery will involve a mixture of positions in each commodity in forward contracts to secure future revenues whilst allowing for maximisation in the next time period. These contracts are however correlated and this is one of the difficulties with a commodity pricing model. This valuation is calculated whilst always taking into account the physical constraints; if for example there is a surplus of crude oil from one period to the next, the model

[40] http://www.ft.com/cms/s/0/e1c96862-9516-11e1-ad72-00144feab49a.html#axzz2V9JYyX17

needs to address this as there is less available to then refine in the following period. The valuation is more difficult than the strip of options as the choice on one option affects the next, and the option is not as simple as a call; it is in fact an option with choices on the weights of each refined product minus the crude cost. There are two underlying stochastic processes that represent the commodities that constitute a spread, $S_{i,j}$.

The S_i and S_1 are correlated, and the above spread is subject to the physical constraints of the refinery. There are several methods that can be envisioned to solve the corresponding optimisation problem described.

- The most elementary one consists in using the current forward curve to find the optimal portfolio of long and short forward contracts attached to the refining period of one month up to 10 years for a legitimate financial valuation. The corresponding values are the intrinsic value of the refinery facility.
- Stochastic optimisation on a tree or through Monte Carlo simulations: the quantity of cumulative production denoted Q_t and one can work back in time to solve the optimisation problem, this is a dynamic programming approach.

This former method is precisely the approach that is considered in the following chapter that to investigate the possibilities of obtaining a much more accurate representation of the uncertainty facing the refinery DM. In the next chapter an accurate way to value the refinery, the prevalence of the financial market risks and the uncertain behaviour of the commodity price series are analysed extensively.

Chapter 3

A multi-period dynamically consistent valuation of a topping oil refinery, with mass balance and capacity constraints

"The only thing that makes life possible is permanent, intolerable uncertainty; not knowing what comes next." Ursula K. Le Guin.

Valuing an oil refinery is similar but not the same as obtaining a market price for a chemical plant or a hydropower station; it is a real asset that generates a commodity product/s sold on the market that enable it to continue to operate. A real option approach can capture the optionality of the owner's decision process, which can be deferred or executed, and if the future returns of the commodities are simulated accurately, enable a realistic and robust valuation. We present the issues surrounding the modelling of the underlying commodity prices and design a calculation for optimising the volumes of crude to purchase and refined outputs to be sold on the market – all whilst adhering to the capacity and fluid flow constraints present at a topping oil refinery. The valuation relies upon the assumptions in dynamic programming and maximising profit as a rational *and risk adverse agent; a trinomial tree is utilised along with mean reverting stochastic equations to simulate the prices and produce the volumetric decisions over the lifetime of the oil refinery. To our knowledge this is the first application of dynamic set assignment to the valuation of a refinery in the literature. The valuation program is created within GAMS and the optimisation is convex, non-linear and therefore it can be shown that the value is unique locally and globally.*

3.1 INTRODUCING COMMODITY LINKED ASSET VALUATION

Commodity linked investment provides investors with a way to diversify beyond traditional stocks and bonds, whether they are related to food, energy or precious metals. Historically, investors wishing to purchase commodities had many barriers to entry; significant amounts of money, expertise and time were required. There are now a number of avenues enabling investors to participate in this dynamic asset class, from specific products to investing in the real assets themselves. Consequently, there has been increased activity in commodity related investment, risk management and asset valuation. The uncertainty associated with commodities has a particular impact on commodity linked asset valuation. The aim is to investigate how a stochastic one factor model impacts the investment decisions and valuation associated with an oil refinery. In constructing improved models for commodity prices we have a two sided problem; increased realism through extra factors, such as jumps, and the difficulty of solving for the contingent claims that rely upon them. This chapter searches for a tractable model that captures the main features of the refined commodity processes and can be incorporated into the real option valuation of the refinery. This will contribute to the understanding of the valuation of commodity linked investments applying optimal decisions where the commodity prices have active futures markets available.

The real option valuation technique has been used extensively over the last decade to improve upon the errors apparent in discounted cash flow analysis (DCF). As mentioned in chapter two, the problem with DCF valuation is that it fails to account for the value of managerial flexibility inherent in many types of projects and assets. The options derived from managerial flexibility are commonly called "real options" to reflect their association with real assets rather than financial assets see Brandao, Dyer and Hahn (2008) for details. In Eydeland and Geman (1999), the authors construct a model to value electricity derivatives; a particularly onerous calculation due to the specific properties of electricity prices. It is noted however, that operating a merchant power plant is financially equivalent to owning a portfolio of daily options between electricity and fuel (spark spread options). Therefore, obtaining a model to price these options enables the power plant itself to be valued as a real asset. We will use a similar approach in chapter four, where we use a strip of crack spread options to value the refinery.

Valuation of options on real assets has been limited by the mathematical complexity of the approach, the restrictive theoretical assumptions required, and by the lack of its intuitive appeal. This complexity stems from the fact that the problem requires a probabilistic solution to a firm's optimal investment decision policy, not only at present but also at all instances in time up to maturity of its options. To solve this problem of dynamic optimisation, the evolution of uncertainty in the value of the real asset over time is first modelled as a stochastic process. Then the value of the firm's optimal policy is represented, if possible, by a partial differential equation - obtained as a solution to a value function represented by Bellman's principle of optimality, see Bellman (1957), where appropriate boundary conditions reflect the initial conditions and terminal payoff characteristics. When closed-form mathematical solutions are not possible, which is usually the case for more complex problems, where the asset may be subject to several sources of uncertainty and more than one type of option, numerical methods and discrete dynamic programming must be applied to obtain a solution. Generally, a discrete approximation to the underlying stochastic process can be achieved enabling a computationally efficient model of the valuation problem. The first example of this was a binomial tree that converges weakly to a lognormal diffusion of stock prices as shown in Cox et al. (1979). The binomial tree can be implemented to accurately approximate solutions from the Black-Scholes-Merton continuous-time valuation model for financial options – it provides solutions for the early-exercise American options and enables the pricing of exotic options in a tractable manner. However, the use of traditional option-pricing methods and the replicating portfolio approach is complicated by the fact that, for most real assets, no such replicating portfolio of securities exists, so markets are incomplete. With this in mind, Pindyck and Dixit (1994) suggest the use of dynamic programming, using a subjectively defined discount rate. An oil refinery is impacted by a number of uncertain market prices, and although it can be approached as an option on these sources of volatility, it gives the owner a complicated set of knock on decisions to make over the refinery's lifetime.

3.1. The Refiner's Optionality

Fundamentally, the refining plant's aim is to produce a high quality product sold in local and global markets while taking into account efficiency and regulatory requirements. The decision makers recognise that they possess vast "optionality" but do not correctly extract it. At each point in time the refiner can decide to refine or not based on whether it is profitable to do so. If the price of the raw crude spikes producing a gross refining margin that is negative then choosing not to refine is the most sensible decision. In fact, this is just one of the decisions available to the refinery manager. In standard market conditions it is the volumes of each refined product that constitute the majority of the decision making. If the price of jet fuel increases enough, the shift in the refinery processes should be towards producing more of it, whilst adhering to the strict physical, mass balance and blending constraints inherent in the refinery's processing. This is a complex and dynamic set of decisions that can be augmented by the optimisation calculations. In this work we construct a model to aid the refiner in deciphering the optimum amounts of petroleum products to purchase and sell. To our knowledge, this modelling and quantifying of the optionality is yet to have been achieved in the literature. The goal of this work is to quantify and better understand the optionality of oil refineries without access to liquid crack spread products, under a model of optimal refinery management. We have three major findings:

(a) the stochastic model for the asset prices fits historical prices of crude and refined petroleum product's futures well

(b) the actual optionality that could have been extracted allows a valuation much higher than DCF reveals

(c) the model can be run in a reasonable time

We show that the mean reversion present in the commodity price series, which is exploited by the embedded refinery optionality, appears to increase the value of the refinery to a much higher price than DCF indicates. One reason for this is the opportunity for the owner to switch to another product if one is not performing well. The model is calibrated on real financial and refinery data, and results obtained after using a large enough simulation indicate that a statistically consistent valuation is accomplished - despite the strong assumptions present.

3.1.1 Real data from an Indian Oil Refinery

India currently has over 22 refineries and has accomplished much in a short period through expanding its capacity and complexity to compete on a global scale. It is emerging as a refinery hub and with capacity clearly exceeding demand; the country's refining capacity has grown from 62 million metric tonnes per annum (MMTPA) in 1998 to 215 MMTPA in 2012 – 4.4% of global capacity. Out of the 22 refineries – 17 are public sector, three are private sector and two are joint venture. Since August 2009, India is the largest exporter of petroleum products in Asia (Platts assessment). The Vadinar refinery in Gujurat, has an installed capacity of 10.5 MMTPA, equivalent to 210,000 barrels per stream day (BPSD) of crude along with secondary processing units like: fluidised catalytic cracking, naphtha and diesel hydrotreater, vis breaker and product treating units. Capacity was increased to 20 MMTPA on June 5th 2012. We use this refinery in our analysis to analyse the model's capabilities.

Owners of refineries buy crude and have the right to refine for an indefinite period of time – assuming financial solvency. There are of course daily minimum and maximum refining rates as the different products induce various choices to the owner. As the owner refines, two types of transaction costs are present: the financial transaction cost if the financial market is used to sell the products, a percentage charge, and the commodity charge, a dollar amount associated with the operational costs of refining a barrel of crude. We choose to focus on the uncertainty in the gross refining margin (GRM) costs, simply the sale price minus the overall cost of purchasing and refining the crude; rather than the detailed operational costs of refining. This is achieved by modelling the financial decision set.

As posited in Carmona and Ludkovski (2010), the storage of a commodity that has optionality when being sent to market, derives from the ability to exploit the mean reverting trends in the crude and refined products markets. The authors value a gas dome storage facility and a hydroelectric pump using optimal switching and stochastic optimal control. The authors use a numerical method to deal with the path dependency, which is based on an old quasi-variational inequality approach. The difference with a refining complex in contrast to a natural gas dome storage complex, is the ability to "shift" to an alternative product if the price is more favourable on the market. Oil majors, for example, have financial traders working within their refineries doing just this: capturing the optionality with hedging and speculating transactions. In the short term, current forward/future prices on crude and the refined products provide information regarding the expected spot price mean reversion. This mean reversion derives from demand and supply characteristics on the market; refined petroleum products are a major part of the developed world's GDP. In the next section, we introduce the one period program and show how the model would capture value in this simplified setting, laying the foundation for the multi-period optimisation.

No such models, to our knowledge, have been developed to capture this refining value; this is mainly due to the huge dynamic programming "curse-of-dimensionality" that exists. A global auditing firm, for example KPMG, value oil refineries as financial assets, like the one in Gujurat – US GAAP of UK ICAEW rules are commonly applied due to their extensive application. To obtain a valuation under UK accountancy standards: a typical DCF is employed over six years, with seven random variables representing the decision variables, taking years in monthly blocks, and assuming a deterministic optimisation, a state space of 7 x 72 exists. However, commodity prices are in constant flux on the market and a representation of this uncertainty is required for the model to be considered realistic. Even a small increase in uncertainty can alter a valuation "materially"; see Pindyck and Dixit (1994), this is because it increases the upside potential payoff of the refiner's option, leaving the downside potential unchanged at zero – assuming a downside decision to switch off the refinery and incur no operational costs. Stochastic modelling is a minefield, and this addition to the model increases the state space further. This is a huge computation for any optimiser – here we choose to discretise the stochastic prices onto a trinomial tree to alleviate the complexity; such as those used by Hull and White (1997). By including the uncertainty in decision trees the investment problem is much more effective, as there are now alternative investment strategies that can be evaluated. The value of the investment option must be calculated for each possible rule and is a function of the number of paths on the tree. The difficulties associated with generating accurate discretisation processes is discussed in Zenios (1995). The author applies a discrete space binomial process for generating interest rate scenarios for asset liability management of fixed income securities. The generation of return scenarios concentrates on fixed income securities with contingent claims, and it is emphasized that the future prices are Markov. The focus in this part of the model is to generate a scenario tree; decision analysis and stochastic programming rely heavily upon its accuracy. Another method commonly applied to capture uncertainty is via a Monte Carlo simulation as shown in Carmon and Ludkovski (2010). For reasons explained later, discrete trees are more realistic within this framework. Alternatively, stochastic optimal control as seen in Chen and Forsyth (2007) can be applied. The authors use a regime switching price model with storage characteristics - these models capture the uncertainty by finding a control function that represents the refinery manager's set of choices over time. At the time of writing, no model utilising optimal refining management, optimal control theory applied to the refinery decision problem, has been constructed or back tested on historical data to quantify its optionality - ultimately providing a financial valuation.

For our valuation model we examine a one-factor tree model of optimal refining management that can be back tested, meaning we focussed on: (a) developing a realistic price model (mean-reverting) of spot and forward price processes, and (b) developed a methodology for capturing the embedded optionality in a refining complex, given all the physical and mass balance constraints and the uncertain price processes. Ensuring that the price model is calibrated to the market and constructing a tree that enables each commodity to be optimised is difficult due to the state space. Strong assumptions must be made to implement this model calibration due to being in an unregulated market without access to liquid futures contracts and having no certain cost of capital. As described in Schwartz (1997) the mean reversion behaviour of the commodities is an important feature to capture with a representative model - the coefficients of calibration are shown in the results. Risk-neutrality is an important assumption when using dynamic programming in real options valuation; where it is common to use a constant discount rate.

In back testing the model over the same period that our DCF value was calculated with, we found it was very successful at consistently capturing vast amounts of optionality. In chapter four we tested the model using crack spread data over the same period of ten years using US contracts available on the Chicago board of trade (CBOT). The sensitivity of the model to investor risk aversion, mean reversion level of the commodity prices and the granularity of decision making were investigated. Inventory of petroleum products on site at the refinery is significant. Having zero inventory means we cannot choose to refine immediately; only crude can be bought, whereas being partly filled means refining can happen or be delayed if the price is not trending high enough. A fully stocked refinery gives us optionality to refine, but not to buy crude, hence if the products move much lower and crude much higher, this crude trade cannot physically be made. This is indicative of the embedded option available to the owner. An interesting computational extension would be to consider how the optionality is affected by the refinery inventory level at the beginning and end of the horizon – as these values clearly impact dramatically the value obtained for this asset.

The optionality in this problem is similar to an American option valuation. An American call option gives the investor the right to exercise the contract at any point in time up to maturity. The underlying will be bought at a specified strike, stated in the contract, enabling the owner of the option to observe market prices and execute if the momentum moves into favour, see Longstaff and Schwartz (2001) for a least squares Monte Carlo valuation approach. The owner can decide to remain dormant allowing the contract to expire if conditions are not profitable; one difference in comparison to the refinery is that there is a continuous batch of contracts, each day the decision to refine and in what proportion can be executed. Further, there is not just one underlying as is the standard case with an American option, where a stock or a bond represents the intrinsic value – the refinery has seven underlyings and each one is correlated with one another. Finally, the underlyings are not stocks, where features such as geometric Brownian motion and dividends are significant, they are commodities, which exhibit particular properties and idiosyncratic qualities that differentiate them from other asset classes. For instance, the futures on commodities display a fundamental difference to those on stocks as they approach maturity – their at the money volatility decreases in comparison with shorter dated contracts rather than increasing at any given date, this is known as the Samuelson effect; hence each futures volatility itself increases as it approaches its own maturity date - see Geman (2005) for a detailed discussion. To construct a realistic model, we investigate these peculiarities and the specifics of the refinery that underpin our construction. In the next section, we describe the risks facing the owner and introduce the refinery planning literature.

3.1.2. The refinery's risk

Price, safety, operational, regulatory and economic risks are just a few of the concerns facing the owner of the refinery. Recently, as mentioned in chapter one, the refining environment has become particularly competitive. Hence risk management becomes a more pressing issue for refiner consumers and producers, and knowledge of the product prices can help reduce this risk. The refiner's portfolio includes his own production and a set of sales and purchases over periods of time.

The goal of using such an approach is to reduce the economic risk and ultimately obtain a value for the asset. This is connected to the fact that the oil spot price may be highly volatile due to the various different unpredictable reasons; geopolitical factors, OPEC decisions, wars and to the possibility of unknown market factors. When the refinery buys or sells the commodities, it is exposed to the risk that the spot price and the futures price do not converge on expiration of the future contract; this is known as basis risk. The basis risk factors here are: the wholesale spot crude price and the refined products spot prices. The refiner is a price taker with respect to the crude and refined products' prices. Since we want to concentrate on the financial risks; the detailed operational refinery flows (*known as system planning*), the operational uncertainties, like yield uncertainty in the crude distillation unit (CDU), are ignored. The data generated by the mid-term production planning problem, however, can be used as input data to the scheduling problem, which determines detailed crude oil processing schedule, process unit schedule, blending and shipping schedules.

There are two mathematical terms that explain the decision that confronts the refinery owner. At the start of the first period, the decision maker chooses the production variables - the volumes of crude to buy and refined products to sell, whereas, the second term considers the risk that the profit induces. Due to the nature of the Vadinar refinery, the hedging decisions will be volumes of the refined products to buy or sell when the uncertain prices have unfolded at the end of the period. The literature on this subject differs in the way the expectation and the risk is considered when optimising over one time period. Rarely considered in the literature, is to approach the problem as a multi-period sequence of decisions over some fixed time horizon. This approach enables a financial valuation to be obtained, as the cash flows can be modelled over time and discounted back to the present day. This area is considered significant in finance as it impacts valuation risk; the risk that the asset is overvalued and is worth less than expected when it matures or is sold, as investigated in Mun (2006). There are two main contributors to this valuation risk:

(a) The risk that valuations are incorrect due to control, operational and management errors.

(b) The valuation uncertainty due to model risk and assumptions underpinning a valuation calculation.

Dependent upon the asset being valued, factors such as tail risk, credit risk and many others can have differing and substantial effects on a value. The ubiquitous DCF approach in the case of an oil refinery gives erroneous values and is often the cause of financial value disputes. In the next section we discuss the current planning literature.

3.1.3. Introducing refinery planning

There are four common types of refinery: coking, cracking, topping and hydro-skimming. Due to the detailed analysis in the literature available for the topping type, its granular construction was considered parsimonious and representative enough for a real refinery model. Ravi and Reddy (1996) present a one period model to capture the set up of a modern topping refinery, which we utilise here but extend it over all time periods until maturity. They present a fuzzy (a form of many-valued logic; approximate rather exact sets) fractional programming approach to the optimisation of operations at an oil refinery. This is optimising using a ratio within the objective function, in this case, very similar to the Sharpe ratio for the refinery. The authors set as an objective function the mean profit divided by a function representing risk – it is named fuzzy as it is based on simple fuzzy logic. This term applies when the objectives cannot be expressed as "true" or "false" but rather as "partially true". The authors have 22 constraints to represent mass balances, and two fuzzy goals, which maximise profit relative to capacity; here the objective goals are between zero and one. They elicit a very small increase in profit over previous deterministic studies. The authors avoid the inclusion of uncertainty and relevant risk functions in their optimisation.

A great introduction to stochastic programming is discussed in Sen and Higle (1999). Historically, optimisation models of chemical plants were mostly linear due to the computational issues: Gao et al (2008) use a mixed integer linear programming model for the midterm planning problem of a fuel-oil plant in China; Micheletto et al (2007) use a MILP model for a refinery over multiple periods of time where demand satisfaction is the focus. Neiro and Pinto (2003) present the problem of managing multiple refinery operations using a mixed integer non linear program. Their objective function maximises the net present value under raw material and product inventory constraints. These include mass balance and operating constraints for each refinery in their network and demand scenarios are used to include uncertainty. Although the authors claim advantages over single refinery optimisation by considering the supply chain, the risk preferences of the decision maker are not explored.

A number of studies in the chemical engineering literature approach the refinery as a *blendshop* problem (considering the results from blending of the fractionated products on a very short time scale), see Grossman and Zamora (1998), which is clearly necessary in short term planning, but exogenous effects or the financial side of the problem are usually not integrated at this time horizon. Pongsadki (2005) presents a mathematical program that maximises the profit whilst minimising the risk of the refinery. Their model uses data from the Bangchak Petroleum refinery to decide on the production level for given forecasts of demand. The scenario generation is drawn from normal distributions and incorporates constraints of cost payments for the decision maker not meeting the client demand. They plot the expected gross refining margin against value at risk; the authors emphasise the dramatic improvement of the optimisation when replacing variance. Three independent scenarios are generated and the optimisation summed over one-period at a time; where one period is separate to another, hence a static optimisation is represented; computational details of the mathematical program are not provided but no alternative to procedural programming is described.

3.1.4. Multi-period refinery planning

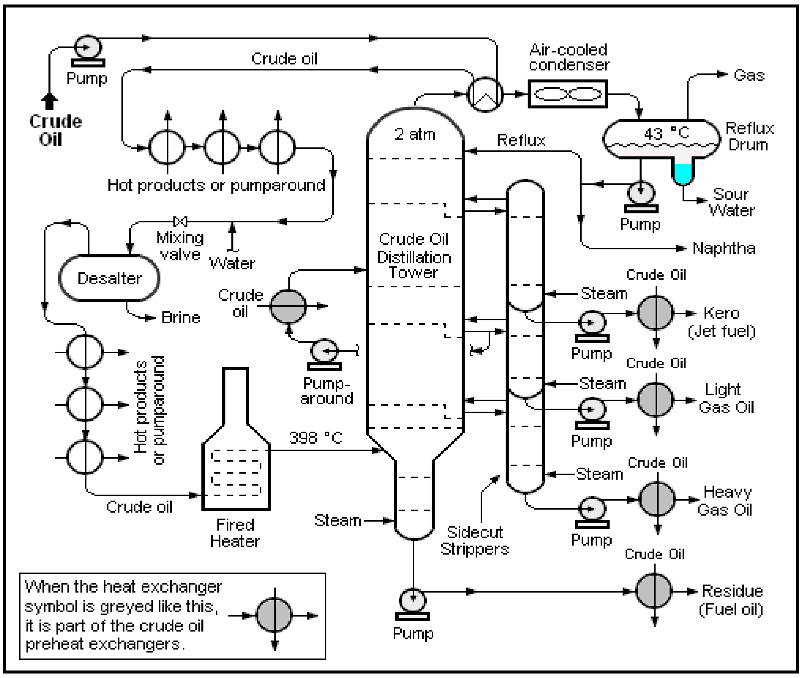
As mentioned in the previous section, Mathematical programming (MP) models applied to refineries are mostly deterministic; although, often the one-period approaches mentioned are used in practice by the oil majors for decision making. A more realistic model for valuation is a multi-period program. The multi-period approach is much more difficult due to the computational issues that arise. Authors consider the entire supply chain or a network of refinery capacities rather than approaching the difficulties with the state space or the stochastic nature of the commodity prices.

Benyoucef (2010) applies a multi-period stochastic program to refinery capacity management up until the year 2030. Random commodity price variables are not included, and the author optimises a network of refineries; algebraic modelling language (AML) details are again avoided. Marianthi et al. (1997) provide a multi period optimisation with uncertainty for short term scheduling, it is not at the granularity we require but is useful in the way the authors introduce risk to the refinery scheduling. Bernardo (1999) et al introduce a specialised cubature technique for solving high dimensional problems but it is only suitable for Gaussian distributions. The authors claim computational performance gains over efficient sampling techniques and is applied to engineering processes. This paper highlights even more the problems faced by researchers when there is a large state space but no alternative approach to optimising over such a state space using an AML is provided.

In the next section, we introduce the physical constraints present in this particular refinery model, along with the relevant data. In section three we discuss the multi-period construction. In section four, we introduce uncertainty via the random spot commodity prices and derive how they are related to the forward curve. In section five, we solve the optimisation refinery model using a unique AML representation. In section six, we discuss the results and their implications, and finally in chapter four we compare the previous refinery valuations to a strip of crack spread options.

3.2. STOCHASTIC PROGRAMMING FOR OIL REFINERIES

The construction in Ravi and Reddy (1998) is utilised to formulate the foundations of a linear deterministic program. As described in chapter one, the petroleum flows and streams in a refinery plant are complicated movements through each of the units present at a particular category of refinery. Figure 3.1 below depicts the raw crude being pumped into the distillation tower after being desalted, and on the right the exit, producing a typical set of marketable products.



(Figure 3.1: Crude oil distillation and fractionation)[41]

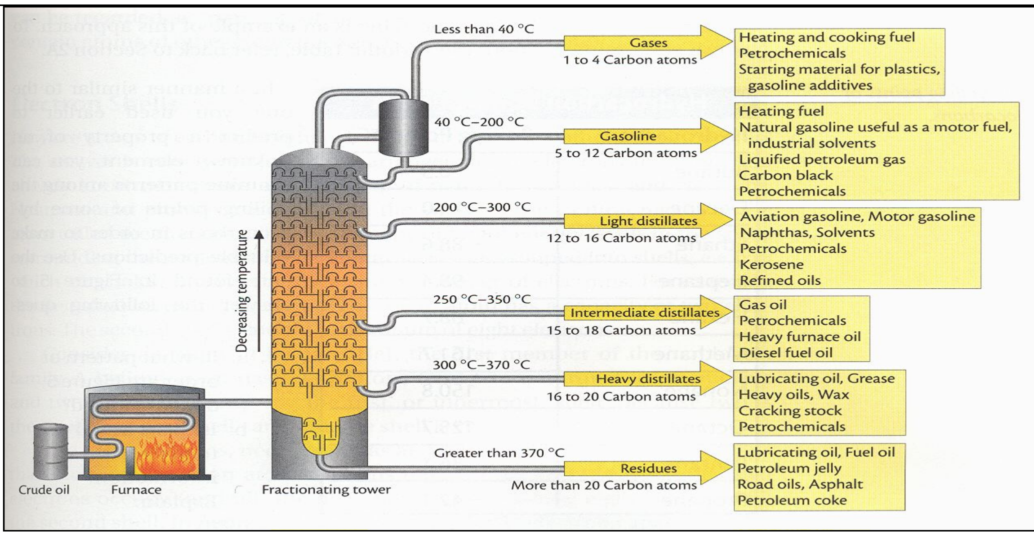
In any oil refinery there are a number of restrictions on the flow of crude and refined products that move through the crude oil distillation units (CDUs). Strict limits are present on how much hydrocarbon liquid can be held in storage within the refinery and the amounts that can be released onto the refined markets; see Moro and Pinto (2000) for a detailed description. The following section describes the structure of the linear program.

3.2.1. Refinery physical flow constraints

It is necessary to provide some background on the processes that occur when refining is carried out on crude oil to precursor the construction of the refinery model. This illustrates how in the real world the refiner's decisions come to affect the marketable products. Crude oil is refined into

[41] https://rbnenergy.com/complex-refining-101%E2%80%93distillation-no-test-and-no-math-guaranteed

useable products such as gasoline; hence gasoline has a higher cash value on the market. Extracting this monetary value is a complicated process that involves isolating mixtures according to their boiling point range. Gasoline boils from 40 to 200 degrees Celsius; figure 3.2 shows the boiling point range of the various fractions. Crude oil distillation separates the hydrocarbons into the fractions; the separation occurs in a large tower operated at atmospheric temperature. Light materials like naphtha are removed in the upper section; heavier materials such as residual fuel oil are withdrawn from the lower section. Residual fuel oil can be further separated into vacuum gas oil and vacuum residua - the vacuum gas oil goes to the catalytic cracking unit. In a modern refinery, the cracker carries out the most significant process; it cracks long chain hydrocarbons into shorter chain molecules. It enables the refiner to convert raw material into gasoline and distillate; if one compares the prices of raw fuel oils to those of gasoline, the advantage is conspicuous.



(Figure 3.2: Fractionation of crude oil)[42]

The refined products are obtained through blending of various fractions of crude within the refinery units. Petroleum products have many specifications; some linear programs are constructed to

[42] http://www.ril.com/downloads/pdf/about_jamnagar.pdf

optimise only the blending process in detail, rather than optimising the cash flows generated through the refining process. The linear program in this case would need mathematical terms for the properties of the hydrocarbon fluids. For instance, gasoline has the following specifications:

- Octane - a measure of harmful components, too high and it wastes power in a combustion engine.
- Viscosity - how easily the fuel flows.
- Sulphur - important for regulation reasons

In Ravi and Reddy (1996), the variation in the types of crude that can be purchased are not considered, one type of crude is assumed, this can be unrealistic for some refineries and the option for the decision maker to select from a number of crudes would not be a difficult extension. We adhere to the most suitable crude to elicit value from, in a model capturing cash flow value rather than yields from blending error.

The aim of the following section is to depict the dynamics of the commodity spot prices and forward curves to insert into the revenue refining model. It is standard practise to use current spot and forward prices to mark-to-market a book of physical and/or financial contracts; the forward curve providing information about the market perception of future spot prices. However, the prices observed today can be wildly inaccurate when used for decisions about future activity; a set of future prices are however required for valuation purposes.

The goal is not to provide a forecasting method but to apply a fundamental model of commodities to obtain information of possible future prices to be used within an optimisation model. In applying real market data to represent the spot price, S, of each refined product, our concern is threefold:

- To find the most appropriate mathematical structure for S, i.e. the type of process

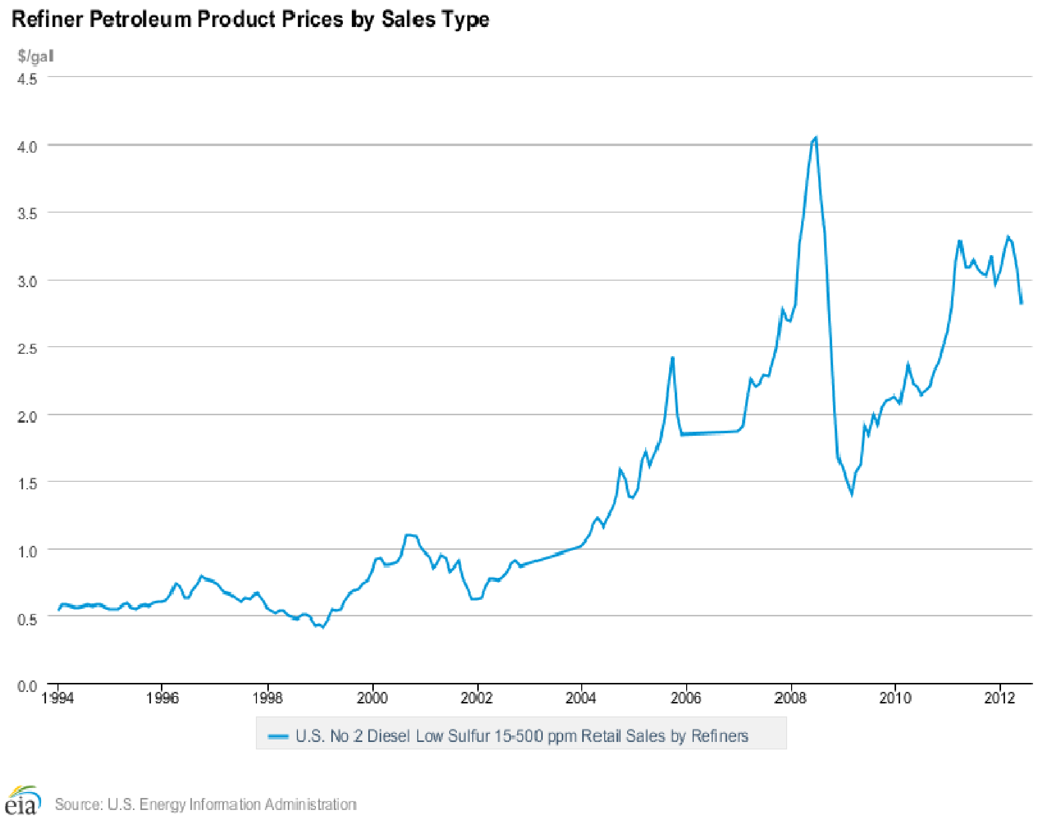
- Once we have chosen *S(t),* a stochastic process, there will be parameters that can be estimated from liquid markets and clean data
- From a *trajectories* standpoint, the Monte Carlo methods and/or trees, on average must generate data that looks like the observed ones; from a statistical standpoint - the moments of the distribution of *S(T)* (for T>t) must coincide with the empirical moments, at least for the first four

In comparison to stock prices, which grow on average, commodity prices do not exhibit trends over long periods. The evolution of the futures and spot prices in the next section may show sharp rises during short periods, over a long period they revert to "normal levels". This is a consequence of mean reversion with spikes caused by shocks in the supply/demand balance. In creating a realistic model of the stochastic process it is important to observe the time series over the required time period to analyse the specific features and details important to the particular commodity. The refined products future prices from NYMEX used in this chapter and the notation for formulating the mathematical model are presented in the following section.

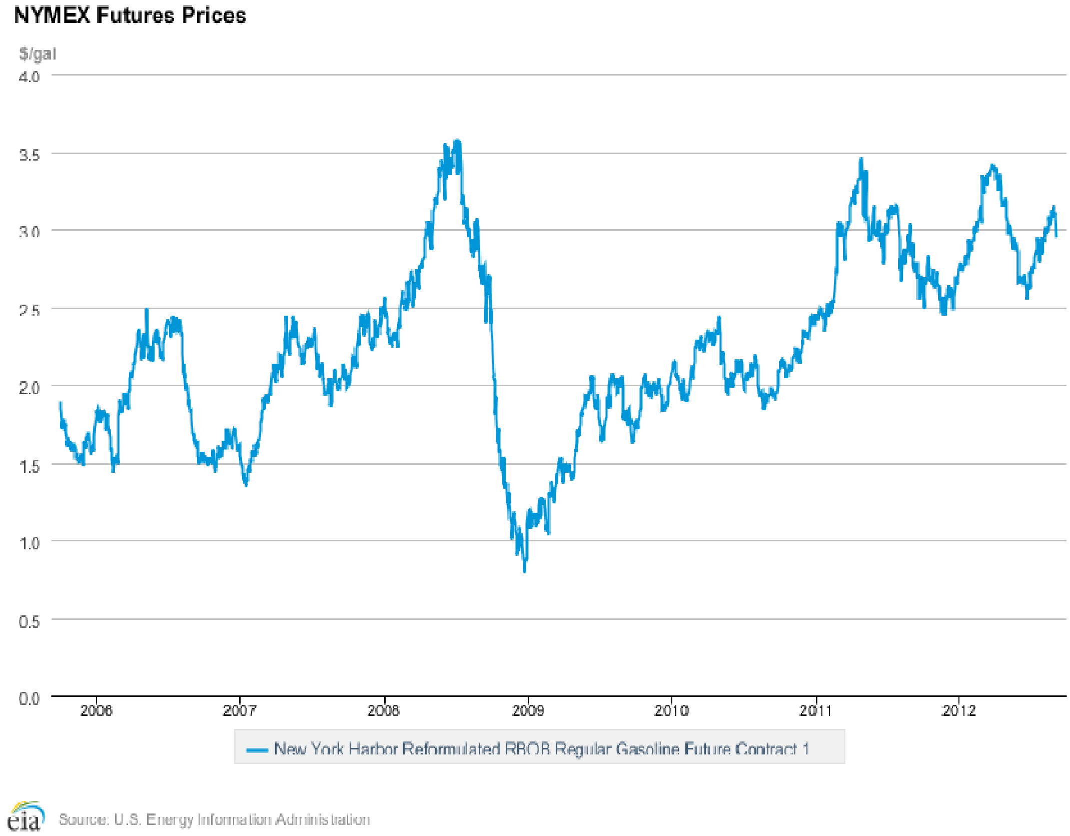
3.2.2. Refined oil products market data

The following data was obtained using the EIA data repository.

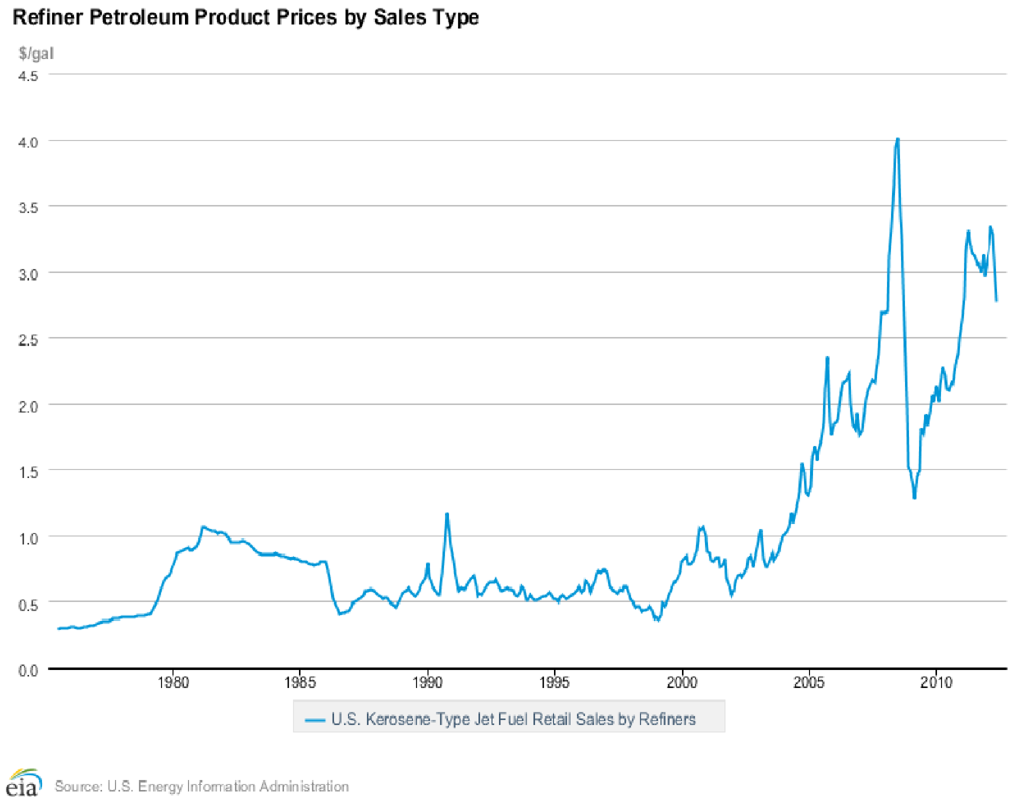
(<u>Figure 3.3:</u> WTI crude front month contract historical price series)



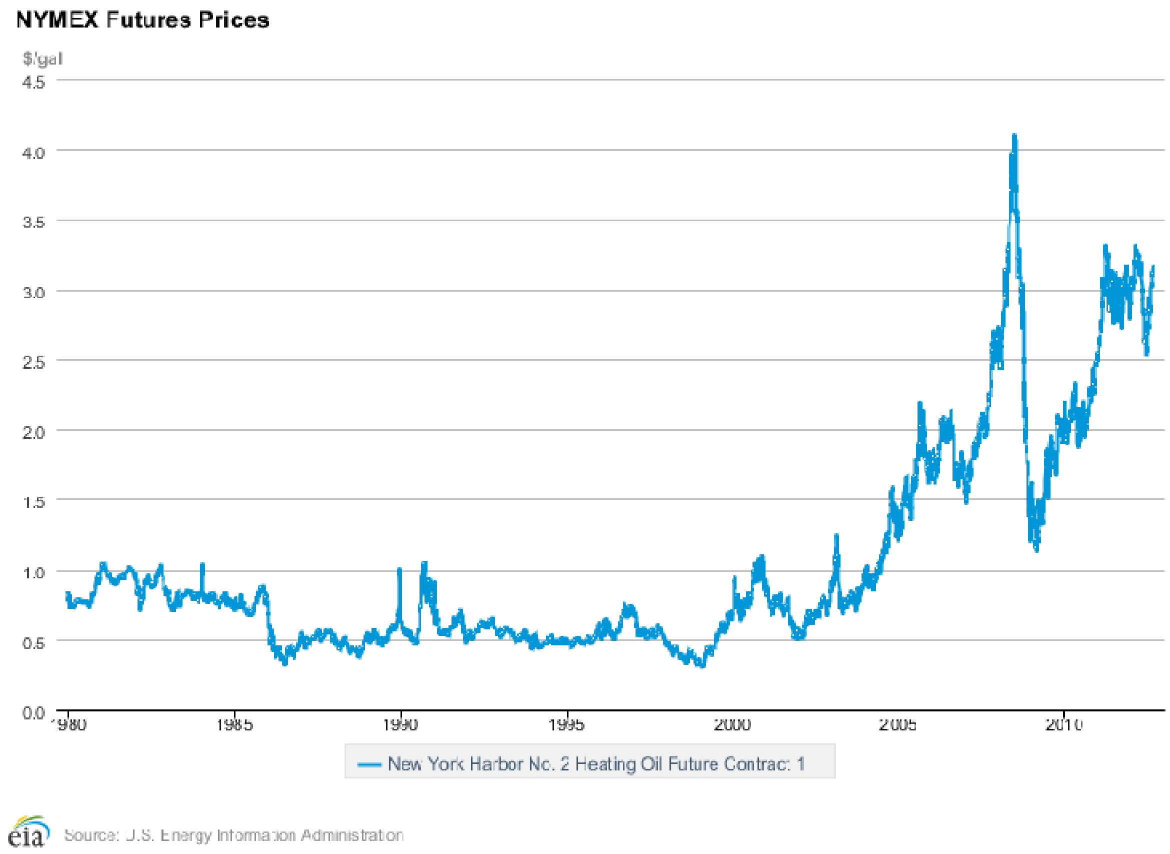
(Figure 3.4: Diesel fuel low sulphur historical price series)



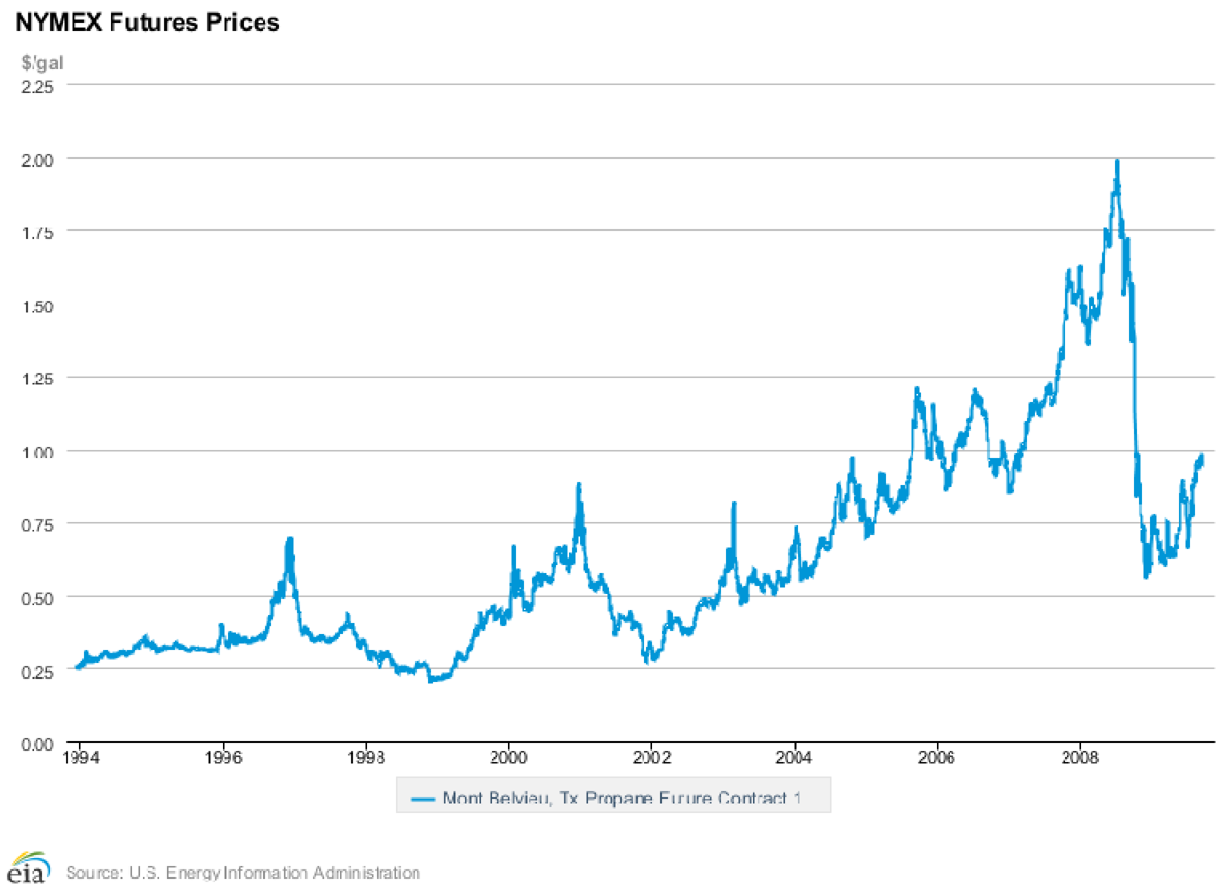
(Figure 3.5: NY RBOB gasoline front month contract historical price series)



(Figure 3.6: Daily jet fuel retail historical price series)



(Figure 3.7: New York Harbor No. 2 daily heating oil historical price future contract series)



(Figure 3.8: Propane front month future historical price series)

The above figures depict the mean reversion despite the noted increase in price series value as the commodities approach 2006; there is a increasing trend which then dips aggressively in 2008 due to the financial credit crunch. There is clear random variation and the correlation amongst the different petroleum products above is conspicuous. The figures for correlation and mean reversion are shown in the following sections. There is a clear difference in behaviour in comparison to stock price movements and it is this characteristic commodity behaviour that must be accurately captured in the model; see Schwartz (1997) for a thorough description.

In the next section, we describe the foundation of the refinery model, including the relevance of the optimisation time period.

3.2.3. The producer midterm planning framework

Refining is modelled at the level of detail common in tactical mid-term refinery planning, with a granularity of one month at its finest and the start-up and shut-down costs considered not significant. The model's decisions are considered in monthly periods consequently, a set of decisions remain constant for a month's refining – this is known as a refinery run in the market; each review period corresponds to a futures price maturity. The refinery consists of a number of units, the Crude Distillation Unit (CDU), the cracker and the Vacuum Distillation Unit (VDU). For example, Neiro and Pinto (2003) model the units of the refinery and their flows, using nodes and arcs, constructing a one-period model to capture the uncertainty. This requires that the uncertainty in the yields that are transferred from the CDU to the rest of the refinery are modelled using the data inside the CDU; there is an uncertain amount of error associated with the volumes chosen by the refinery owner.

In this work however, the uncertainty is attached to the prices of the outputs rather than the volumes from the refining process itself and the producer must schedule the production of each refined product. These decision variables of the refinery scheduling problem are stated below. Each of the variables represents the flow rate of a potential product within the refinery in units of tons/day over a refining run.

3.2.4. Decision Variables

$X_{i,t}$[43] [tonnes/day] : volume of product *i* bought or sold on this particular day

- $X_{1,t}$: Crude Oil

[43] In the multi-period setup these decision variables become per node and per time period: $X_1(n, t)$

- $X_{2,t}$: Gasoline
- $X_{3,t}$: Naphtha (after the splitter)
- $X_{4,t}$: Jet Fuel
- $X_{5,t}$: Heating Oil
- $X_{6,t}$: Fuel Oil
- $X_{7,t}$: Naphtha stream exiting the PDU
- $X_{8,t}$: Gas Oil
- $X_{9,t}$: Cracker Feed
- $X_{10,t}$: Residuum
- $X_{11,t}$: Gasoline after splitting of Naphtha exiting the PDU
- $X_{12,t}$: Gas Oil after the splitter
- $X_{13,t}$: Gas Oil stream entering the fuel oil blending facility
- $X_{14,t}$: Cracker Feed after the Splitter
- $X_{15,t}$: Cracker Feed stream entering the fuel oil blending facility
- $X_{16,t}$: Gasoline stream exiting the cracker unit
- $X_{17,t}$: Stream exiting the cracker unit into the splitter
- $X_{18,t}$: Heating oil stream after splitting of cracker output
- $X_{19,t}$: Cracker output stream

In the final model's objective function, only the first six and the 14th decision variables defined above are included - the rest are auxiliary variables required to represent the structure of a real refinery and present within the mass balance or constraint equations. As the variables in the objective function are the volumes that can be sold, the other variables aid in the representation of the flows within the refinery and contribute to the mass balance (weight into a refinery unit must equal the weight out), fixed blends (assuming fixed cuts from a barrel of crude) and unrestricted balance equations (without the auxiliary variables, an unrealistic set of volumes from the cracking process would be represented). The set of constraint equations that represent the flows shown in the next section, take on a variety of forms dependent on whether the structure is that of a topping, cracking, hydro skimming or coking refinery. The main reason why the mathematical framework applied here was selected is that there are a large number of studies on the topping refinery within the engineering literature that can be used for comparison purposes; additionally, it is trivial to alter these equations given a different type of refinery complex. The values assigned to the decision variables must satisfy the following constraints that describe the physical restrictions on this particular refinery:

- The capacity constraints
- Raw Material Availability

The mass balance constraints are required for the following units:

- Primary Unit
- Cracking unit

The increasing oil price has over the years encouraged refiners to search for solutions to extract more from the dregs of the barrel. Refiners are increasingly using solvent deasphalting and debottlenecking of existing vacuum and coking units. For example, the ROSE process, invented by KBR uses special internals and designs that permit extraction of maximum high-quality oils and fuels from vacuum residues and other heavy petroleum feedstock. Without discussing the internal mechanics, this would alter the yield specification on the model, hence impacting the optimal volumes of refined products; additional blending technology in the modelling process is usually implemented in a finer granularity refining model than required in midterm planning.

In this model, we represent the yields emanating from each separate units: the CDU, the VDU and the cracker, as fixed according to the equations below, (the below set-up is taken from Ravi and Reddy (1996) where the X variables are defined as above in section 3.2.4. as the volumes of the commodities, but here the objective function is converted to a multi-period problem where each node is, n, and each time period, t) - for another type of refinery, e.g. a Coking refinery, the following equations would be still be required but would consist of different coefficients.

The equations below essentially capture the primary distillation unit and a middle distillation cracker. The primary unit splits the crude oil into the following refined products: naphtha (13% yield), jet fuel (15%), gas oil (22%), cracker feed (20%) and residue (30%).

3.2.4.1. Fixed Yields

Primary unit:

$$-0.13\, X_1(n, t) + X_7(n, t) = 0 \qquad \forall n \in N \qquad (3.1)$$

$$-0.15\, X_1(n, t) + X_4(n, t) = 0 \qquad \forall n \in N \qquad (3.2)$$

$$-0.22 X_1(n, t) + X_8(n, t) = 0 \qquad \forall n \in N \qquad (3.3)$$

$$-0.20 X_1(n, t) + X_9(n, t) = 0 \qquad \forall n \in N \qquad (3.4)$$

$$-0.30\ X_1\ (n, t) + X_{10}\ (n, t) = 0 \qquad \forall n \in N \qquad (3.5)$$

Cracker:

$$-0.05X_{14}\ (n, t) + X_{20}\ (n, t) = 0 \qquad \forall n \in N \qquad (3.6)$$

$$-0.40X_{14}\ (n, t) + X_{16}\ (n, t) = 0 \qquad \forall n \in N \qquad (3.7)$$

$$-0.55X_{14}\ (n, t) + X_{17}\ (n, t) = 0 \qquad \forall n \in N \qquad (3.8)$$

With the higher utilisation of heavy crude oils, refiners often encounter higher residual loads with higher levels of contaminants, increased aromatics content, and more often than not, higher acids content in their feeds. On the other hand, the gas oil content of the new feeds will be lower, creating a potential loss of feed downstream such as in the FCC units. Due to the chemical composition of the petroleum crude, there is a limit on how much of each product can be blended within the complicated production of refinery outputs. We state below the equations that define these fixed blending compositions:

All the constraints below are in the form of equalities; there are three types of constraint: fixed yields, fixed blends and unrestricted balances. Unless there is a shutdown of the plant or adverse storage movements the right hand side of the balance constraint is always zero. Naphtha and jet fuel products are straight run, heating oil is a blend of 75% gas oil and 25% cracked oil. Fuel oil can be blended from the primary residue, cracked feed, gas oil and cracked oil in any proportions. Yields for the cracker are flared gas 5%, gasoline blend stock, 40%, and cracked oil 55%. This information along with the flow describes the physical system.

3.2.4.2. Fixed Blends

Gasoline blending:

$$0.5X_2(n,t) - X_{11}(n,t) = 0 \qquad \forall n \in N \qquad (3.9)$$

$$0.5X_2(n,t) - X_{16}(n,t) = 0 \qquad \forall n \in N \qquad (3.10)$$

Heating oil blending:

$$0.75X_5(n,t) - X_{12}(n,t) = 0 \qquad \forall n \in N \qquad (3.11)$$

$$0.25X_5(n,t) - X_{18}(n,t) = 0 \qquad \forall n \in N \qquad (3.12)$$

3.2.4.3. Unrestricted balances

Naphtha:

$$X_7(n,t) + X_3(n,t) + X_{11}(n,t) = 0 \qquad \forall n \in N \qquad (3.13)$$

Gas Oil:

$$-X_8(n,t) + X_{12}(n,t) + X_{13}(n,t) = 0 \qquad \forall n \in N \qquad (3.14)$$

Cracker feed:

$$-X_9(n,t) + X_{14}(n,t) + X_{15}(n,t) = 0 \qquad \forall n \in N \qquad (3.15)$$

Cracked oil:

$$-X_{17}(n, t) + X_{18}(n, t) + X_{19}(n, t) = 0 \qquad \forall n \in N \qquad (3.16)$$

Fuel oil:

$$-X_{10}(n, t) + X_{13}(n, t) + X_{15}(n, t) + X_{19}(n, t) + X_{6}(n, t) = 0 \qquad \forall n \in N \qquad (3.17)$$

Each refinery complex has a maximum amount of petroleum that it can physically store. These containers are at a size that is economically optimal for the region the refinery operates in. The below equations represent the limits on this refinery's production in tons per petroleum product per node and per month:

3.2.4.4. Raw material availability constraints

Gasoline:

$$X_{2}(n,t) \leq 2700 \text{ / month} \qquad \forall n \in N \qquad (3.18)$$

Naphtha:

$$X_{3}(n,t) \leq 1100 \text{ / month} \qquad \forall n \in N \qquad (3.19)$$

Jet fuel:

$$X_{4}(n,t) \leq 2300 \text{ / month} \qquad \forall n \in N \qquad (3.20)$$

Heating oil:

$$X_{5}(n,t) \leq 1700 \text{ / month} \qquad \forall n \in N \qquad (3.21)$$

Fuel oil:

X_6 (n,t) ≤ 9500 / month $\forall n \in N$ (3.22)

The refinery owner is interested in maximising the net profit, which is simply the costs subtracted from the revenues; this objective is a simple function over one period.

3.2.5. Deterministic objective function

The deterministic objective function for profit of *Ravi and Reddy (1996)* represents the profit in dollars over one time period, where the decision variables are as explained in section 2.4 and are in tonnes:

Maximise

$$\text{Profit} = -8.0\,X_1 + 18.5\,X_2 + 8.0\,X_3 + 12.5\,X_4 + 14.5\,X_5 + 6.0\,X_6 - 1.5\,X_{14} \quad (3.23)$$

One serious issue with this function in the chemical engineering literature, the random commodity prices are replaced with their mean prices using a set of normal distributions, when in fact it is known that commodity price series are not normally distributed, see Moro et al. (2007) for details. Further, the above model is deterministic and does not consider the risk or the uncertainty of the commodity prices in the next period when they are actually sold. The move to a multi-period model introduces the time series themselves and the analysis of the refined products statistically. We leave for others the examination of the cointegration via an Engle-Granger test; we aim to model the commodities from continuous stochastic equations rather than approaching them via econometrics using a Euler discretisation of integrating the separate commodities. Our approach is along the lines of Schwartz (1997) where the author discretises the continuous mean reverting equations and finds the parameters via the Kalman Filter.

In practice many refiners use Futures contracts or other derivatives to hedge their price risk. We use data for our oil refinery that does not use crack spread contracts to hedge, as is common in the US; in the region where this refinery is located there is an absence of liquid crack spreads. An oil Future contract is a standardised contract between two parties to buy or sell a specified type of oil of standardised quantity and quality for a price agreed today. In the case of crude oil, the main Futures exchanges are the New York Mercantile Exchange (NYMEX) and the Intercontinental Exchange (ICE) where the West Texas Intermediate (WTI) and North Sea Brent crude oil are traded.

The delivery periods in the model are fixed to each of the 12 calendar months $(M_1, M_2, \ldots, M_{12})$ for each year. We ignore differences in trading and delivery period in the model and assume each lot sold is for one month ahead using the financial market rather than the physical spot market. To capture the market risk we introduce a number of risk measures within our optimisation, with the goal of finding the most effective measure whilst retaining tractability – a valuation however does not require this addition. In terms of decision making over time this is a significant step for a number of reasons; one is that optimising with variance is computationally inefficient and is also not used in practice as it is now considered an unrealistic representation of financial risk. Secondly, this enables an investigation of the most realistic risk-reward model when we extend to a multi-period setting. The first step was to consider how to introduce the uncertainty in the price series into the static linear program.

3.2.6. Solving the deterministic objective function

Adopting the deterministic objective model in (3.23) using today's futures prices in tons, as shown in Ravi and Reddy, and not the returns of the price series of crude and the refined products gives:

Maximise

Profit $-385X_1 + 726X_2 + 385\,X_3 + 471X_4 + 520X_5 + 251X_6 - 72X_{14}$ (3.24)

The negative values are the purchasing and operational costs, the positive values are the saleable product prices. This includes the same constraints as model (3.23), shown above for one time period. This is a static linear optimisation problem, and so the CPLEX algorithm was used to solve it within the optimisation program, General Algebraic Modelling System (GAMS). The solution to equation (3.24), the deterministic objective function, under the constraints defined in section 3.2.5, is given below:

Model 1: Objective Value = $285,418 per day; optimal crude purchased: 1500 tons; time to solve: *negligible*

The solution is trivial computationally: there are 22 constraints, all equations including the objective are linear and there are a total of 18 decision variables. Now we extend the static optimisation incorporating the uncertainty in the commodity price series. This can be done in a number of ways; forecasting any random variable in practise is a vast area of research. In practice it is done by drawing a random number from a recommended probability distribution for the data generating process. When cash flow forecasting, the chemical engineering literature applies deterministic values, which misses the range, sensitivity and behaviour of the cash flows – we discuss this issue in detail in section four where we describe our choice of the set of stochastic processes. In the following section we detail how the one period model can be reformulated and consolidated with future periods to build the multi-period model including commodity price uncertainty.

3.3 MULTI-PERIOD OPTIMISATION

The legacy of Markowitz Portfolio Theory (MPT)[44] means that optimisation problems are often categorised by how the efficient frontier can be drawn; this was added to MPT in 1958 by James Tobin[45]. Each of the underlying assumptions has been challenged in various papers; delving into all of these is beyond our scope. The three main assumptions that are investigated or utilised in this refinery model are:

- Investors are price takers, and their actions do not influence external market prices
- The correlations between assets are always fixed and constant
- Returns on assets are normally distributed

There have been various extensions to deal with these assumptions; studies have, for example, managed to capture fat tails and asymmetry and included these within their optimisations to address assumption (c). In this model, ultimately we wish to maximise risk expected return for a given level of risk. The statistical concept of covariance captures this risk by implying that an investor should choose assets that do not crash together. In terms of the refinery this is difficult, due to all assets being heavily correlated, hence exposure to the crack spread[46] is the main risk to manage. In portfolio optimisation, we often consider *variances* and *covariance matrices* in our decisions, i.e. a multivariate joint probability density function (pdf) of returns is required; this proxy for risk is valid if asset returns are jointly normally distributed, however there are problems. If returns are not

[44] Markowitz, H.M. (March 1952). "Portfolio Selection". *The Journal of Finance* **7** (1): 77–91.

[45] Tobin, James (1958). "Liquidity preference as behaviour towards risk". *The Review of Economic Studies* **25** (2): 65–86.

[46] The crack-spread is a term used in the industry for the differential between the price of the refined products and the raw crude oil; indicating the profit margin that can be extracted from "cracking" crude.

normal then variance is no longer a valid risk measure. Despite this issue, a mean-variance model used to maximise returns and minimise risk, is a good place to begin. A modern portfolio optimisation approach is thus characterised by the following desirable features: firstly, it enables for realistic return distributions, secondly, it builds on a realistic risk measure, thirdly, it is computationally tractable, and finally, it enables for a parsimonious robust formulation. We contribute to the literature by proposing a refinery portfolio model that encompasses all of these features. Our approach is based on derived distributions from a discretised set of stochastic differential equations; it can rely upon a quantile risk measure, and leads to, after an approximation, a convex optimisation problem, enabling robustness checks that confirm its stability.

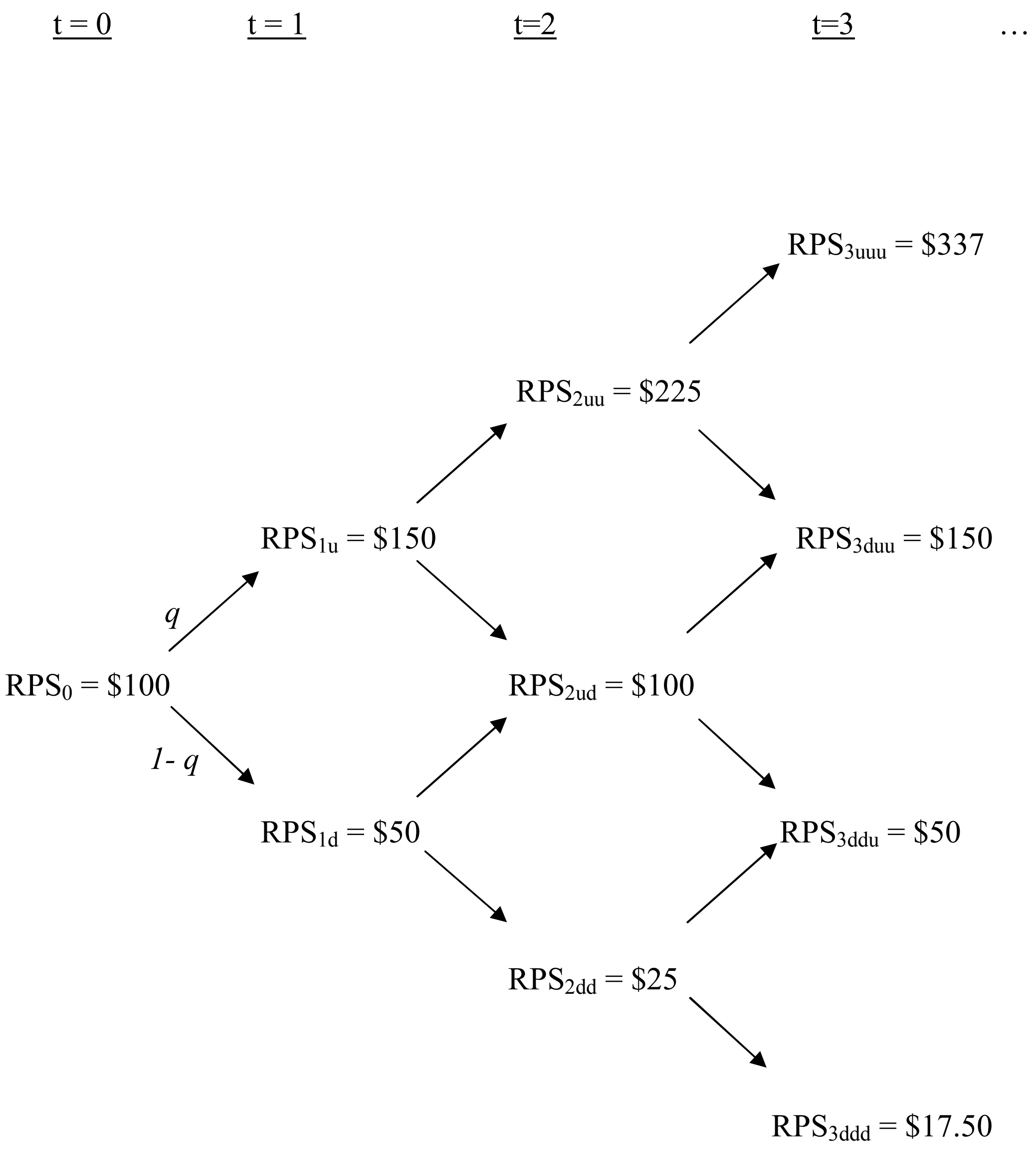
To build our argument, we start with a simple example using future cash flows to define optionality. The example requires that cash flows should be discounted using (a) a constant risk-free rate of interest of 5% (which would be unrealistic for a period of one month but is used to illustrate the value of optionality over a period of time) and assumes (b) a constant cost of crude oil, for example, crude is priced at $74 per barrel; these will both be made stochastic in the final model's implementation. Although in our final results we will use continuous compounding for the sake of simplicity we show here the calculations using simple compounding. In the next section, to extract optionality from this behaviour, we analyse how one period's refining decisions are related to the next.

3.3.0.3. N-period (N+1 dates) optimisation

By adding one more period of price uncertainty, our refining problem becomes more complicated. There are now five different investment strategies that can be implemented. Optimally, we could (a) refine now; (b) wait a month and refine if it has risen, but never refine if it has decreased; (c) wait a month and invest if the price has risen, but if it decreases wait another month and then refine if it has risen; (d) wait two months and only refine if it has risen twice in a row; or (e) never refine.

Which strategy to undertake depends on the initial cost of a barrel of crude and the initial set of refined products' prices, and the value of the refining option should be calculated for each path.

The other complicating factor here is that while we can still compute the value of the investment option by constructing a risk-free portfolio, the makeup of that portfolio will not be constant over the two months; we will have to alter the number of crack-spread contracts in the short position after the price of the refined petroleum products changes at t=1. Keeping the refining portfolio riskless by changing its composition is known as dynamic hedging. Without calculating the option value explicitly we state that the value of the option to invest as function of initial products prices has a piecewise-linear function. The, F_0, is always a convex function of the RPS_0, and is greater than or equal to the net payoff from exercising the option today, $RPS_0 - Costs_0$. This is a powerful result as we allow investment to fluctuate in continuous time and this means we require stochastic processes, see section 3.3.



(Figure 3.11: Three period optimisation)

Approaching this valuation using option pricing methods; we wish to calculate the option to refine at t=0, F_0, as a function of the initial RPS_0, as well as the optimal refining rule. To do this we can work backwards solving two separate investment problems looking forwards from t=1, first for RPS_{1u} and then for RPS_{1d}. In each case we determine F_1, the value of the option at t=1, by calculating a risk free portfolio and calculating its return. Given two possible values for F_1, we then move to t=0 and determine, F_0 by calculating a risk free portfolio. The descriptions above exemplify the fact that the refinery owner has optionality and that the value of this real option contains path dependence. The terminal value depends on the value of the underlying commodity prices, not only at this time, but also at prior points in time. Specifically, the refinery option's terminal value depends upon the "path" taken by all the underliers over the life of the refinery. This refinery has another condition, as the value of the option to refine today depends on the quantity of crude available, which itself depends on the last time the owner refined.

In the next section, we describe a deterministic trading strategy to define the intrinsic value of the oil refinery, and follow this by describing a stochastic dynamic program that captures the optionality.

3.3.2. Foundation of the oil refinery intrinsic value

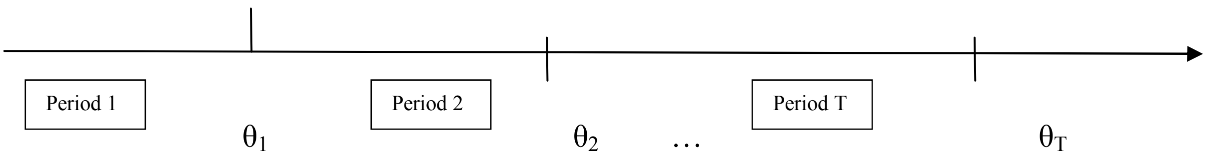
We now define the intrinsic value as the expected value of the refinery assuming risk-neutral dynamics, conditional on following the best initial deterministic set of decisions. We start out with the set of initial forward curves and find today's intrinsic plan. From this decision set we lock in the refining decision set today. If the decision is to purchase an amount of crude, the cost of purchasing follows from the current crude spot price and the volume of crude in its storage container. Additionally, if there are decisions to sell, the refining units are activated and the revenue flows from the current forward curves on the market for the petroleum products. If the decision is to not refine, there is no cash flow and the refinery is unaltered.

Next we simulate a realisation for tomorrow's forward curve conditional on today's forward curve and the given appropriate price dynamics. For this new forward curve, tomorrow's intrinsic plan is determined. This problem is updated for the refining decision set that is already locked in as well as the passage of time (one day closer to the terminal date). This plan is used to lock in the storage decision regarding tomorrow. We continue this procedure until we reach the terminal date of our refining problem. This procedure gives us one possible realisation of daily cash flows from refining over the relevant time horizon. We then repeat the procedure above to obtain the desired number of possible cash flow realisations. The value of the refinery is given by the average net present value of simulated daily cash flows. The problem with this intrinsic approach is that it misses the full flexibility available at each decision date being a deterministic approach; to capture this flexibility we need to also consider the fact that the financial markets provide us with alternative stochastic decisions at each point in time.

The stochastic dynamic programming approach is similar, but simulates possible spot price path realisations for each commodity. Each path is mapped into a discrete spot price state space. Based on the discrete spot price path realisations, we can solve the problem to optimality by stochastic dynamic programming. We consider the problem with a six-year horizon and our objective is to obtain the value of the cash flow from operating the refinery.

With the approach chosen, we now describe this method for *N* periods, including definitions for the refinery owner's specific choice variables and conditional random variables.

Adopting a discrete time setting, with a finite horizon, the decision periods are denoted θ_i, $i=1,\ldots\ldots, T$ (months):



(Figure 3.12: N period optimisation)

On each decision date, spaced apart by one month, a number of decisions are required: $q_1 - q_7$, these are the volumes in tons of the crude to buy, and refined products to sell both decided on the same date. We can represent the vector of *qs* as a control vector, U. To model the uncertainty over these future time periods a stochastic equation representative of the uncertain prices is selected. The refinery's current status is described by a state variable x, the current value, x_t, being known, but future values, x_{t+1}, x_{t+2}, random variables, are not. Assuming this process is Markov, all the information relevant to the determination of the probability distribution of future values are summarised in the current state, x_t.

At each period t, the volumes of commodities are available as choices to the refinery, and we represent them by a control vector *U*. The value, U_t, of the control at time, *t*, must be chosen only using the information that is available. The state and control at time, *t*, affect the refinery's profit flow, which we define as, $\pi_t (x_t, U_t)$. We let $\Phi(x_{t+1} \mid x_t, U_t)$ be the cumulative probability distribution function of the state next period, conditional upon current information (state and control variables).

The discount factor between any two periods is, $1/(1+\rho)$, where, ρ, is the discount rate. The ultimate goal is to choose the set of controls, $\{U_t\}$, over time that maximise the expected net present value of the payoffs; where the termination payoff function is defined by $\pi_T (x_T)$.

At date t, the refinery owner chooses control variables U_t, the immediate profit flow is, $\pi_t (x_t, U_t)$. In the following period, (t+1), the state will be x_{t+1}. Optimal decisions onwards will extract value, $V_{t+1}(x_{t+1})$. This is random as viewed from time t, so the expected value is required, $E_t[V_{t+1}(x_{t+1})]$. This is known as the continuation value, formally we have:

$$E_t[V_{t+1}(x_{t+1})] = \int V_{t+1}(x_{t+1}) d\Phi_t(x_{t+1} \mid x_t, U_t) \qquad (3.31)$$

Discounting this back to period t, the sum of the immediate profit from refining and the continuation value is:

$$V_t (x_t) = \pi_t (x_t, U_t) + 1/(1+\rho)\, E_t[V_{t+1}(x_{t+1})] \qquad (3.32)$$

The manager will choose the control vector, U_t, to maximise the sum of these two functions, and the result will be the value $V_t(x_t)$.

Hence, we write the value of the refinery at date t as:

$$V_t(x_t) = \max_{U_t}\{\pi_t(x_t, U_t) + \frac{1}{1+\rho} E_t[V_{t+1}(x_{t+1})]\} \tag{3.33}$$

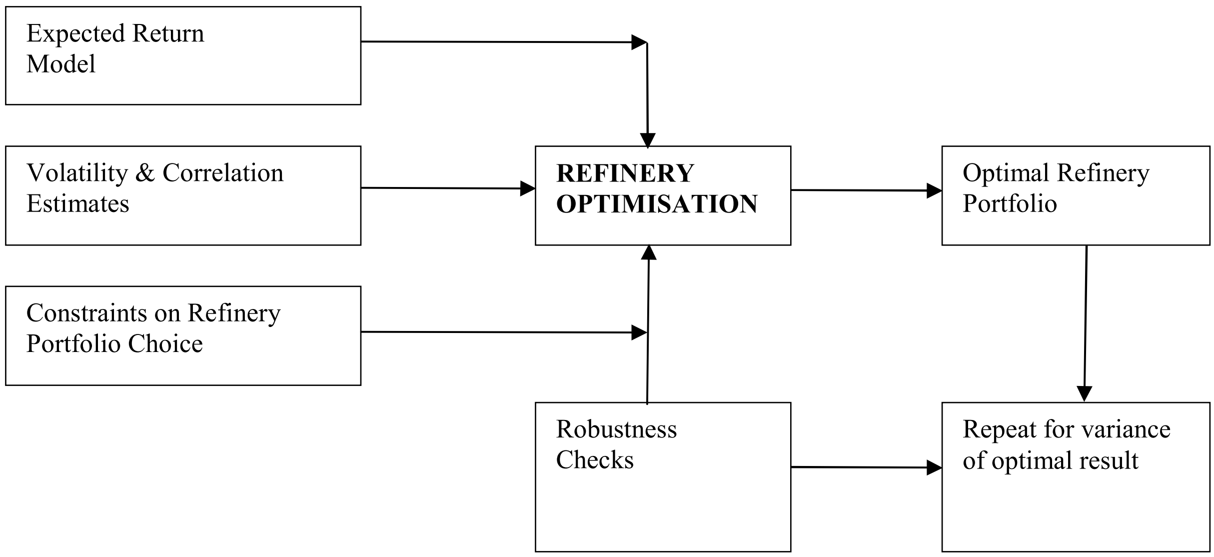
This is also known as the Bellman equation; see Bellman (1957) for a more rigorous derivation. Here we have a fixed horizon of six years, and knowing that we have an independent optimisation at the maturity, we start at the end of the horizon and optimise backwards. We assume that at the end of the horizon, the refinery gets a termination payoff, $\pi_T(x_T)$. Therefore, the value in the period previously is:

$$V_{T-1}(x_{T-1}) = \max_{U_{(T-1)}}\{\pi(x_{T-1}, U_{T-1}) + \frac{1}{1+\rho} E_{T-1}[V_T(x_T)]\} \tag{3.34}$$

Knowing the value function at, T-1, enables us to solve the maximisation problem for, U_{T-2}, leading to the value function, $V_{T-2}(x_{T-2})$, and we continue with this process until the present, t=0.

In equation (3.34) the value process does not consider the risk, in the published paper within the appendix we combine the mean-risk approach into the value process for decision making purposes. In this work we have left to others, the rigorous mathematical proofs of existence and uniqueness of solutions; we treat the limit to continuous time in a heuristic way; see Fleming and Rishel (1975) for a detailed background. In the following section we describe the data flow required to capture a financial valuation or create a decision process considering the risk.

3.3.2.1. Optimisation data flow chart



(Figure 3.13: Flow chart process behind the refinery optimisation)

As the figure above depicts, the ultimate goal is to have the optimum amounts of decision variables; in a multi-period setting, this is known as the policy function. Not only is variance a symmetric risk measure but it is also time inconsistent, we discuss this property in section 5.2., see Geman and Ohana (2008) for a proof. In referring to the literature, when including random product prices into an optimisation model, it is done in one of three ways:

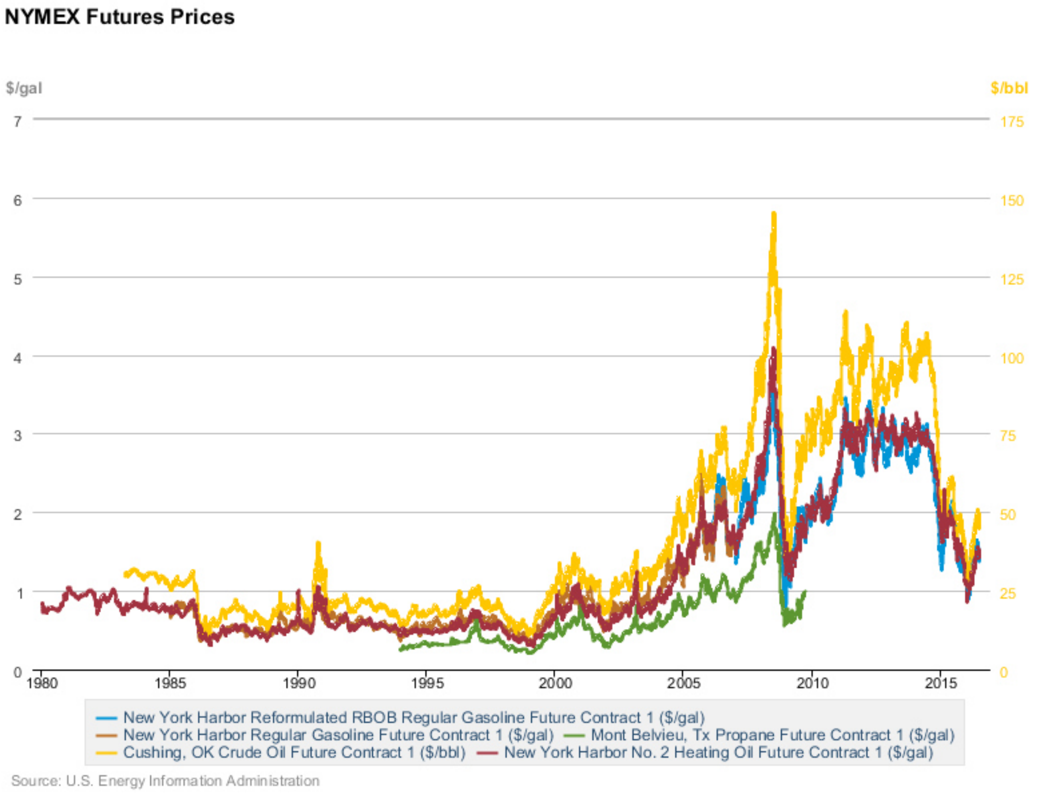
- Mean values replace the stochastic variables
- Continuous distributions are utilised
- Scenario generation (using discrete distributions)

Other moments are replaced using the historical data but most literature in this area assumes a normal distribution. The prices of refined products are made stochastic and the problem solved using scenario generation, type (3) above. However, the refinery model is required over *N* time periods, and therefore a stochastic differential equation (SDE) is more realistic, enabling a dynamic probability distribution to be modelled. In section four we discuss why we use SDEs, and describe the spot price process used to capture uncertainty in the commodity prices. In the following section we introduce the valuation methodology in a simplified setting.

3.3.3. Stochastic commodity behaviour

A model should be robust, in that the parameters should not move erratically from one day to the next, leading to a complete change in the refinery portfolio value. We make a distinction between: (a) the probability that the price of oil, *S(T),* in for example, July 2014 goes over \$120. This number represents the probability, p, of that event, ω, under the real measure denoted as, $\boldsymbol{P}$, and (b) the price market practitioners are willing to pay today to receive \$1 if the market price of oil at date T=1, July 2014, if the market price does go over \$120. We refer to this number as the probability of the event, ω, under the pricing measure denoted as, $\mathbb{Q}$. This is clear if one considers for example, a market player considering now whether to purchase for, July 2014 a virtual oil refinery, getting the refined products without owning or running the complex; the player is interested in the probability of the same event under the pricing measure, $\mathbb{Q}$; this number includes the risk aversion component as well as human judgement about how much the market is willing to pay to secure refined products for July 2014.

Crude and refined product prices are intertwined in a complex way, as seen in figure 3.14 below, which shows daily spot prices for crude and the refined product futures prices quoted on NYMEX for the last decade.



(Figure 3.14: NYMEX future price data on refinery commodities)

In this figure prices seem to revert towards a long term mean in the first half; after the financial recession in 2008, this mean seems to exist again but at a higher level. This mean reversion is often evidenced in the literature (see, for instance, Geman and Nguyen, (2003), for the case of agricultural commodities; and Pindyck (2000), for energy commodities). Specifically, these commodity prices fluctuate with a daily volatility of over 3.5% on average, which is high; yet, it is clear that they are not emanating a random walk as present in stock prices evolution. Furthermore, there is no clear seasonality here as shown in natural gas prices for example; evidence from simple tests shows that there are strong and weak months of the year for crude and gasoline for example, but high year-to-year volatility by month inhibits exploitation. Any acceptable commodity-linked valuation model must (a) account for mean reversion and (b) value the optionality given the volume and other refinery constraints. An additional feature that is often evidenced is the fact that the returns series exhibit tails that are fatter than those of a normal distribution; this easily seen by obtaining a QQ-plot of the level or log returns. If attempting to model the fat tails, one can consider jumps within the stochastic process or use GARCH modelling introduced by Bollerslev (1986). Our strategy is to develop a one factor tree model, since the one factor here has constant volatility it does not replicate some movements of the futures curve, this is not perfect, however since the state space is colossal; this one factor approach is more tractable. This will sidestep the dynamic volatility structure considerations; yet, one factor tree structures still capture optionality. Valuing contingent claims using multifactor models indicates that the more complexity the less likely an analytical solution exists. In the next section, we examine a continuous time price model of spot and forward prices. The chosen process is then extended into a discrete time model to be used in a pricing tree.

3.3.3.1. The Ornstein Uhlenbeck process

To model the commodity spot price we require a continuous time stochastic formula that exhibits mean reverting behaviour, for example:

$$S(t + dt) = S(t) + a(b - S(t))dt + \sigma dW_t \tag{3.3.1}$$

Which, at date t, is an affine function of dW_t; this is also known as the Ornstein Uhlenbeck process, where the: *a, b* and *σ* are positive constants, for a thorough explanation see Vasicek (1995).

Where:

b : the long term mean level – all future trajectories will evolve around a mean in the long run

a : the speed of mean reversion – it characterises the velocity at which such trajectories will regroup around b in time.

σ : instantaneous volatility – measuring instant by instant the amplitude of randomness entering the system.

The nice feature here is that the model can represent upward-sloping, downward-sloping or slightly humped shapes of the future price term structure. Hence, S(t+dt), is like dW_t, normally distributed, but it may take negative values, which is obviously an undesirable feature for commodity prices. The above is a one factor equilibrium model; a no-arbitrage model would fit the initial term structure. Gibson and Schwartz (1990) describe a two factor stochastic model where the spot price of the commodity and the instantaneous convenience yield are assumed to follow a joint stochastic process; with Brownian motions under the objective measure and correlated in time. This not only enables parallel but also twisting movements of the forward curve to be captured. This is more realistic, and as applied and discussed in the case of crude oil, gold and copper by Schwartz (1997), two factor models exhibit much difference in the long term over one factor models, when used for pricing contingent claims or the valuation of real assets; they are also more accurate in depicting the term structure of the future prices. A single factor model has very different implications about the volatility of future price returns as the maturity of the contracts increase, and is incapable of describing the volatility of the futures data. This has important knock-on effects for valuation when the models are applied for longer term real assets. The author's analysis includes stochastic interest rates on futures price data, calibrated with the Kalman filter; despite the improvements over a single factor model, for shorter maturity contracts the difference is minimal, considering we require a valuation, the extra complication associated with the additional stochastic convenience yield was considered adverse to our objective; a two factor model is introduced in chapter four as an enhancement to the valuation method used here. In the next section we discuss the relevant properties of the stochastic one factor model.

3.3.3.2. The Samuelson effect, mean reversion and positivity

Futures prices can be used as a proxy for spot prices, however, they can be unavailable for various reasons - please note where we construct a continuous future time series we apply the standard Panama approach to manage rolling effects[47]. The Schwartz (1997) model is a popular framework for energy commodity futures prices that resembles the geometric Brownian motion (GBM), whilst introducing mean-reversion to a long-term value *θ,* and ensuring positivity by using a logarithm function on the spot price. Due to commodity prices neither growing nor declining on average over time; they tend to mean-revert to a level which may be viewed as the marginal cost of production. Unfortunately, it is non-trivial to define the price of petroleum commodities exhibiting mean reversion or showing non-stationarity, see Geman (2000) for a discussion. Many tests exists, none of which dominate to conclude that mean reversion is representative. Newman and Yin (1996) improve upon stationarity tests supporting stationarity and auto regression using monthly data on oil. Based on these analyses we define the commodity process as:

$$d(InS_t) = a[\theta - lnS_t]dt + \sigma dW_t \qquad (3.3.2)$$

Where S_t is the asset price process, t is time, and W_t, a standard Brownian motion; *a*, is the force of mean reversion, θ the long-run equilibrium level - the above model prevents negative values of the commodity price. The above is a one factor stochastic differential equation: in terms of dynamics, the level moves are captured. A higher number factor model would allow the volatility or convenience yield to be captured and ensue in more accurate dynamics; however this would also increase the complexity of building a valuation from a discrete tree to represent the correlated commodity series. Changing variables we have:

$$Z_t = \text{In } S_t \qquad (3.3.3)$$

Applying Ito's lemma leads to:

---

[47] https://www.quantstart.com/articles/Continuous-Futures-Contracts-for-Backtesting-Purposes

$$dZ_t = a(\theta - \frac{\sigma^2}{2a} - Z_t)dt + \sigma dW_t \quad (3.3.4)$$

Now introduce the variable $X_t = e^{at} Z_t$ and use Ito's lemma again to obtain:

$$dX_t = e^{at} dZ_t + ae^{at} Z_t dt \quad (3.3.5)$$

$$dX_t = a(\theta - \frac{\sigma^2}{2a})e^{at} dt + \sigma e^{at} dW_t \quad (3.3.6)$$

Integrating this equation between dates t and T (where T > t) provides:

$$X(T) = X(t) + (\theta - \frac{\sigma^2}{2a})(e^{aT} - e^{at}) + \sigma\sqrt{\frac{e^{2aT} - e^{2at}}{2a}} W(T-t) \quad (3.3.7)$$

At date t, X(t) is observed and X(T) is an affine function of the normal variable W(T-t).

Hence:

$$L[X(T) \mid F_t] = \mathrm{N}(X(t) + (\theta - \frac{\sigma^2}{2a})(e^{aT} - e^{at}); \sigma\sqrt{\frac{e^{2aT} - e^{2at}}{2a}}) \quad (3.3.8)$$

And the law of $Z(T) = e^{-aT} X(T)$ is immediately derived as a normal variable with an adjustment of the mean and variance parameters of X (T):

$$L[Z(T) \mid F_t] = \mathrm{N}(e^{-a(T-t)} Z(t) + (\theta - \frac{\sigma^2}{2a})(1 - e^{-a(T-t)}); \sigma\sqrt{\frac{1 - e^{-2a(T-t)}}{2a}})$$

(3.3.9)

Forward prices, F(t,T) = E[S(T)|F_t], if we assume that the rational expectations hypothesis holds or that computations were, conducted under the pricing measure $\mathbb{Q}$:

$$F(t,T) = \exp\{e^{-a(T-t)} \text{In} S_t + (1-e^{-a(T-t)})(\theta - \sigma^2/2a) + \sigma^2/4a(1- e^{2a(T-t)})\} \quad (3.3.10)$$

The above equation is key in the building the forward curve in the numerical approach introduced later in the chapter. It is clear that as T goes to infinity only the second term in (3.3.10) remains:

$$F(t,+\infty) = \exp\left\{\theta - \frac{\sigma^2}{2a}\right\} \quad (3.3.11)$$

Differentiation provides the volatility of F(t,T):

$$\sigma_F(t,T) = \sigma e^{-a(T-t)} \quad (3.3.12)$$

Here we state the mean and variance of the above continuous process at date 0:

$$E_0[Z(T)] = e^{-aT} Z(0) + (1 - e^{-aT})(\theta - \frac{\sigma^2}{2a}) \quad (3.3.13)$$

$$Var_0[Z(T)] = \frac{\sigma^2}{2a}(1 - e^{-2aT}) \quad (3.3.14)$$

This is a single factor model, which can still introduce different volatilities as contracts mature at different dates. We use equations (3.3.13) and (3.3.14) extensively in our calibration procedure in section 4.5. Furthermore, equation (3.3.11) suggests that forward contract volatility decreases for longer maturity contracts, which is in agreement with the *Samuelson effect* - where he argues that futures should exhibit increased volatility as they approach maturity but at any point in time volatility should decline with time-to-maturity. An issue with this model is that the volatility goes to zero for long times to maturity, a property which is clearly a limitation, since it is not observed in practice. Introducing a second factor for stochastic volatility would better capture the properties of the forward curve; conversely, the state space would be increased by another dimension, enhancing the complexity of the optimisation problem, see Steiglitz (1998) for a discussion. Forward curves under the one factor model will exhibit level shifts, but cannot capture twists or inversions as commonly found by researchers using principal components analysis; despite this, correlation between forward curves can still be portrayed. A balance that was important to consider when implementing the model, was tractability. The tree, upon which the optimisation was applied, was calibrated on seven correlated forward curves. Using the equation above ensured all commodity prices were mean reverting, and log normal in distribution, whilst making calibration less strenuous; enabling the successful implementation of the optimisation calculation over many time periods. In the next section we discuss the calibration of the above SDE.

3.3.4. Two period - three dates, refinery optimisation

We have shown in the previous section how to simulate correlated continuous commodity prices, and we now discuss how we use these prices in our refinery model. If we use only the information available at date 0 to optimise the value over the entire lifetime of the refinery we have solved a static problem. We would choose the amount of crude to refine at each date by using today's scenario tree, an intrinsic value calculation; a dynamic optimisation requires that we rebuild our scenario set at each date until maturity, hence capturing extrinsic value. We begin with a description of the static problem and then describe the steps to ensure a dynamic set of decisions are made.

To simplify ideas, we concentrate on just two refined products, where at each date i, the spot price of a commodity is, $S_{j,i}$, for instance, naphtha - $S_{2,i}$, and kerosene – $S_{3,i}$, are refined from raw crude oil - $S_{1,i}$. We are ultimately interested in a valuation and hence interested in pricing the refinery. Therefore, we are in a risk neutral setting – at date 0 the crude price is known, whereas the refined product prices in one month's time and on into the future, are not. We can make this assumption despite using market data as we can assume that producers and consumers balance out the market premiums on either side of a transaction, see Parsons (2013) for details including a two-factor calibrated tree; in other words for the sake of model building the risk-neutral and objective measures are assumed equal. We firstly look at the decision process over two decision steps by constructing a simple recombining trinomial scenario tree - recombining is computationally simpler. This is not always used with trinomial trees but it greatly simplifies the complexity of the numerical scheme: it leads to a recombining tree, which has the number of nodes growing polynomially with the number of levels rather than exponentially. Boyle (1986) originally constructed the trinomial over the binomial to enhance the speed and accuracy for pricing purposes; as shown in the authors description the probabilities are specified so as to ensure that the price of underlying evolves as a martingale, while the moments - considering node spacing and probabilities, are matched to those of the distribution of the process. At date 0, we decide the optimal volumes of crude to refine, and products to sell forwards for date 1 – we use today's, *i*, price of crude and the prompt month prices to sell forwards, *i+1*, for the refined products. When we next move to date 1, *i+1*, the spot price of crude will again be known, but the product prices at date 2, *i+2*, will not, so we again sell forwards using the *i+2* prices. Upon reaching the end of the horizon, date 2, we no longer decide to refine, but sell all remaining products on the spot market – the prices will be known. In summary, our refining profit at each date where random prices are denoted by, $\tilde{S}_{j,i}$, known commodity prices as,

$S_{j,i}$, and decision variables by, $X_{j,i,z}$, - the volume of product, j, to buy or sell at date, i, decided on date, z, is:

$$\begin{aligned}
&Date\ 0:\ \tilde{S}_{3,1}X_{3,1,0} + \tilde{S}_{2,1}X_{2,1,0} - S_{1,0}X_{1,0,0} \\
&Date\ 1:\ \tilde{S}_{3,2}X_{3,2,1} + \tilde{S}_{2,2}X_{2,2,1} - S_{1,1}X_{1,1,1} \\
&Date\ 2:\ S_{3,2}X_{3,2,2} + S_{2,2}X_{2,2,2}
\end{aligned} \qquad (3.3.15)$$

Previously, we created a continuous-time model of spot and forward petroleum prices that exhibited realistic properties. This section describes the discrete-time version of that model for a simplified case of only three stochastic variables to be used in a pricing tree for valuing the refinery and the component optionality. The simplified discrete-time price model is as follows:

3.3.4.1. Discrete process for crude oil

$$In(S_{t+\Delta t}) = e^{-a\Delta t} In(S_t) + (1 - e^{-a\Delta t})(\theta - \frac{1}{2}\alpha^2_{t,\Delta t}) + \tilde{B}_{1,\Delta t} \qquad (3.3.16)$$

3.3.4.2. Discrete process for naphtha

$$In(S_{2,t+\Delta t}) = e^{-a_2\Delta t} In(S_{2,t}) + (1 - e^{-a_2\Delta t})(\theta_2 - \frac{1}{2}\beta^2_{t,\Delta t}) + \tilde{B}_{2,\Delta t} \qquad (3.3.17)$$

3.3.4.3. Discrete process for kerosene

$$In(S_{3,t+\Delta t}) = e^{-a_3\Delta t} In(S_{3,t}) + (1-e^{-a_3\Delta t})(\theta_3 - \frac{1}{2}\gamma_{t,\Delta t}^2) + \tilde{B}_{3,\Delta t} \quad (3.3.18)$$

Where:

Δt = the time step in years,

$$\alpha_{t,\Delta t}^2 = \sigma_{S,t}^2 \frac{(1-e^{-2a(\Delta t)})}{2a}, \quad (3.3.19)$$

$$\beta_{t,\Delta t}^2 = \sigma_{S_2,t}^2 \frac{(1-e^{-2a_2(\Delta t)})}{2a_2}, \quad (3.3.20)$$

$$\gamma_{t,\Delta t}^2 = \sigma_{S_3,t}^2 \frac{(1-e^{-2a_3(\Delta t)})}{2a_3}, \quad (3.3.21)$$

$$\tilde{B}_{1,\Delta t} \sim N(0, \alpha_{\Delta t}^2), \quad (3.3.22)$$

$$\tilde{B}_{2,\Delta t} \sim N(0, \beta_{\Delta t}^2), \quad (3.3.23)$$

$$\tilde{B}_{3,\Delta t} \sim N(0, \gamma_{\Delta t}^2), \quad (3.3.24)$$

$$\tilde{B}_{1,\Delta t} \perp \tilde{B}_{2,\Delta t} \perp \tilde{B}_{3,\Delta t} \ . \quad (3.3.25)$$

The above discretisation of the continuous process is achieved by ensuring that the correlation property is correctly represented. The independent Brownian motions are obtained by a Cholesky decomposition of the correlated Brownian motions; explained in section four, see Geman (2005) chapter 12 for a thorough description. In the next section we describe how the probabilities on the tree are obtained to construct a consistent pricing tree.

3.3.5. Transitional Probabilities

Formally we let:

(t, i, j, k) = a quadruple of indexes in the tree spanning time, spot price level of crude and spot price level of two refined products,

$S_{t,\,i}$ *= the crude spot price associated with being at spot price level index i for node (t, i, j, k),*

$S_{t,\,j}$ *= the naphtha spot price associated with being at spot price level index j for node (t, i, j, k),*

$S_{t,\,k}$ *= the kerosene spot price associated with being at spot price level index k for node (t, i, j, k),*

i'ε { the set of indexes for the three successor crude spot price levels emanating from node (t, i, j, k) },

j' ε { the set of indexes for the three successor naphtha spot price levels emanating from node (t, i, j, k) },

k'ε { the set of indexes for the three successor kerosene spot price levels emanating from node (t, i, j, k) },

$p(S_{t+\Delta t,i'}|\ t, i, j, k)$ *= the transition probability to crude spot price level* $S_{t+\Delta t,i'}$ *conditional on being at node (t, i, j, k)*

$p(S_{2,t+\Delta t,j'}|\ t, i, j, k)$ *= the transition probability to naphtha spot price level* $S_{2,t+\Delta t,j'}$ *conditional on being at node (t, i, j, k)*

$p(S_{3,t+\Delta t,k'}|\ t, i, j, k)$ *= the transition probability to kerosene spot price level* $S_{3,t+\Delta t,k'}$ *conditional on being at node (t, i, j, k)*

When transitioning to successor nodes whilst developing the tree, the transition probabilities are solved, and must satisfy the following for the crude and refined product prices, we therefore have three sets of conditions, (in this simplification we have four dimensions i, j, k and t, but in the full model we will need eight dimensions, one for time and the rest for the stochastic petroleum product prices):

3.3.5.1. Crude price probability conditions

$$E(In(S_{t+\Delta t})\,|\,t,i,j,k) = \sum_{\forall i'} p(S_{t+\Delta t,i'}\,|\,t,i,j,k)(In(S_{t+\Delta t,i'})) \quad ,$$

$$\alpha_{t,\Delta t}^2 = \sum_{\forall i'} p(S_{t+\Delta t,i'}\,|\,t,i,j,k)$$
$$\times\ (\mathrm{In}(\mathrm{S}_{t+\Delta t,i'}) - E(In(S_{t+\Delta t})\,|\,t,i,j,k))^2$$

$$1.0 = \sum_{\forall i'} p(S_{t+\Delta t,i'}\,|\,t,i,j,k), \qquad (3.3.26)$$

3.3.5.2. Naphtha price probability conditions

$$E(In(S_{2,t+\Delta t})\,|\,t,i,j,k) = \sum_{\forall j'} p(S_{2,t+\Delta t,j'}\,|\,t,i,j,k)(In(S_{2,t+\Delta t,j'}))$$

$$\beta_{t,\Delta t}^2 = \sum_{\forall j'} p(S_{2,t+\Delta t,j'}\,|\,t,i,j,k)$$
$$\times\ (\mathrm{In}(\mathrm{S}_{2,t+\Delta t,j'}) - E(In(S_{2,t+\Delta t})\,|\,t,i,j,k))^2$$

$$1.0 = \sum_{\forall j'} p(S_{2,t+\Delta t,j'}\,|\,t,i,j,k), \qquad (3.3.27)$$

3.3.5.3. Kerosene price probability conditions

$$E(In(S_{3,t+\Delta t}) \mid t,i,j,k) = \sum_{\forall k'} p(S_{3,t+\Delta t,k} \mid t,i,j,k)(In(S_{3,t+\Delta t,\, k'}))$$

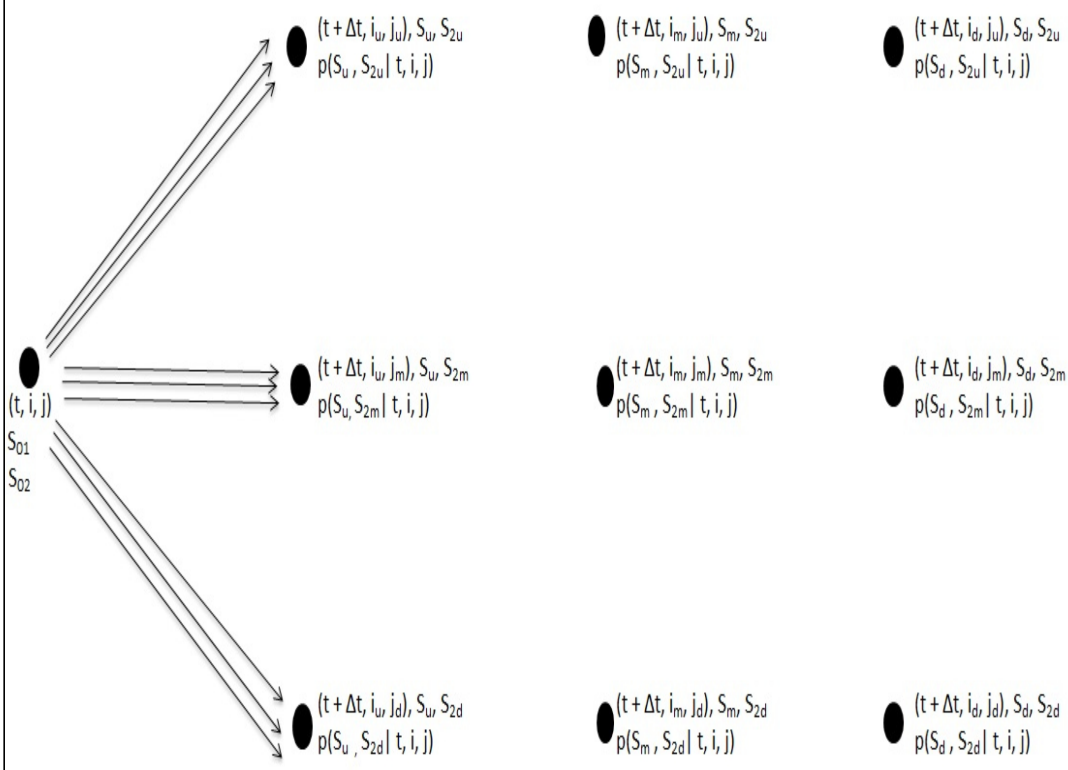
$$\gamma_{t,\Delta t}^{2} = \sum_{\forall k'} p(S_{3,t+\Delta t,\, k'} \mid t,i,j,k)$$
$$\times\ (\mathrm{In}(\mathrm{S}_{3,t+\Delta t,\, k'}) - E(In(S_{3,t+\Delta t}) \mid t,i,j,k))^{2}$$

$$1.0 = \sum_{\forall k'} p(S_{3,t+\Delta t,\, k'} \mid t,i,j,k), \qquad (3.3.28)$$

Following Hull & White (1997) and using the probability condition equations as shown above it is trivial to build a consistent scenario tree. The author's provide the probability equations for transition to an up, middle or down node in a risk neutral setting – in section four we build the tree consistent with all seven refined products.

The figure below depicts the simplified price processes for two stochastic variables moving to nine successor nodes from date 0; at node, (t, i, j), in this case three dimensions. If there are three stochastic processes as shown above then for node (t, i, j, k) there are, 27 successor nodes, therefore in the complete refinery model we will have 3^7 successor nodes, as there are seven refined products, this gives 2,187 successor nodes – rendering a valuation computationally intractable, based on standard solving times on trees as stated in Hull & White (1997); hence we introduce an appropriate approximation in section five.

FIGURE XII Node transition in the pricing tree



The figure shows transition from an initial node to its nine successors in the pricing tree. Both the intial crude price S_0, and the refined price, S_i are depicted as emanating from the same nodes; this is equivalent to there being two pricing trees. The value of these variables, along with their corresponding joint probabilities are shown. Both spot and forwards prices may transition to a high, medium or lower value state.

(Figure 3.15 Node transitioning on the trinomial tree)

Implementing a relevant approximation, with a constant interest rate, r, and, a change in the time increment, Δt equal to 1/12, as it takes one month to refine the crude, we can solve the above optimisation problem using standard non-linear program solvers. We now define the steps required for a dynamic solution to the above problem.

3.3.6. Steps for a two period dynamic optimisation

- At time 0, build an event tree with three decision steps and three sequential branchings representing the possible evolution of the forward curve at times 1 and 2. (See figure 3.15 above)

- Optimise the criterion (maximise the profit over the scenario tree) and implement the date 0 optimal decision. The profit has to be calculated over all possible paths through the tree from root to leaf.

- At time 1, build an event tree whose first node is information of date 1; this tree has two decision steps and represents the possible evolution of the forward curve over one time period to date 2.

- Optimise the criterion at date 1 and implement the date 1 optimal decision.

- Continue until date 2; here at maturity sell the remaining products on the spot market.

The above procedure is carried out 1,000 times to obtain robust optimal decision values, further, these steps can be applied to an infinite period problem.

In section five we will add three components to the above set up: (i) we will increase the number of periods from two to 72, (ii) we will increase from three state variables to seven, and finally (iii) we will apply an approximation to the model to ensure a valuation is obtainable on a standard quad core machine. In the following section we consider the optimisation problem along the nodes of the pricing tree.

3.3.7. Refinery valuation using the backward recursion methodology

In this section we discuss the backward recursion used to value the complicated embedded optionality in a refinery complex.

Considering optimal trading and price-taking actions at each node on the pricing tree, we begin at the terminal nodes. We proceed by initially calculating two values for each node whilst adhering to the specific refinery mass balance and capacity constraints already defined:

- The value of immediately trading a volume of crude oil at that node's spot price, ensuring that the volume choice is within the mass balance and capacity constraints present at that particular node
- The continuation value of holding the new crude oil inventory going into the following period

The optimal spot trade for each node is the one that maximises the sum of these two values. The preceding discussion implies that the optimal values at each node must be calculated for the seven sets of possible inventories held and spot volumes traded, which are continuous.

We now state this optimisation mathematically. Let each node in our one-factor tree be denoted by a octuplet of indexes spanning time, a spot price level of crude and a spot price level of all refined products as previously introduced. We use the following notation:

L = the set of hydrocarbon liquid inventory levels,

Q_l = the permissible set of trades at the given inventory level l,

q_j = a set of spot trade volumes that is positive for buys and negative for sales aggregated for all products,

r_t = the risk-free rate over time, (for simplicity assumed constant)

The optimal value at node *(t, i, j, k, l, m, n, o)* and inventory, l, is:

$$V_{(t,i,j,k,l,m,n,o)}(l) = \max_{q \in Q_l}[V^{spot}_{(t,i,j,k,l,m,n,o)}(q) + e^{-(r_{t+\Delta t}\Delta t)} E(V_{(t+\Delta t,i',j',k',l',m',n',o')}(l+q) \mid t,i,j,k,l,m,n,o)] \quad (3.3.29)$$

Where:

$V_{(t,i,j,k,l,m,n,o)}(l) =$ the optimal value at node *(t, i, j, k, l, m, n, o)* with current inventory l ϵ L,

$$V^{spot}_{(t,i,j,k,l,m,n,o)}(q) = \sum_{rp=2}^{7} q_{rp} F_{t,rp} - q_1 S_{t,1} \,, \qquad \text{ignoring transaction costs}$$

The dynamic program above ensures that the refinery owner weighs different actions against one another with the aim of maximising the direct payoff plus expected future payoffs. On average, an action with a high direct payoff (selling all the gasoline) has a lower expected future payoff (lower inventory of gasoline for the next period). The refinery owner needs to make a decision whilst considering the continuation values; see Boogert and De Jong (2006) for an example. In terms of solving, the recursion starts at the terminal period nodes, where all values in the expectation term above are zero; thus, the optimised value for each inventory that is possible at each of these final nodes is the value of selling as much of the refined products and left over crude oil inventory as is possible. To complete the refinery valuation, we have to calculate the above equation recursively, going backward through each node in the tree, until the initial period is reached. The value of our refinery is the value corresponding to the current inventory level in the initial period. This exact same recursion is described in the Clewlow-Strickland (1999) - a one-factor natural gas storage model, and is here present in the form of a model of seven oil products within a whole valuation model reliant upon a different set of restrictions - the authors also have a much smaller state space to manage and avoid the computational implementation discussion. Additionally, this approach is an extension of the recursion for a similar but less complex set of contracts, named swing options, used in Jaillet et al (2004) - again the authors do not discuss the difficult implementation when the state space is large. In the next section we investigate how the price series data is used to obtain a valuation.

3.4. A CONTINUOUS TIME MODEL OF COMMODITY SPOT PRICE PROCESS EVOLUTION

3.4.0.1. The calibration of the spot price process

Heretofore, we have not detailed the choice of the price model with respect to any measure context, objective or risk-adjusted, and the spot and forward prices must be related in this model for calibration purposes. In the risk-neutral world forward prices are expected spot prices. For crude oil and the refined products, storable commodities, in which the standard no-arbitrage arguments do not apply, we assume such behaviour exists in the real world as well. If there were a premium in forward prices over spot prices, it should be miniscule; since, consumers and producers are influenced by the contango or backwardation in the market. In backwardation, producers are selling refined products at high prices and consumers are willing to pay the premium due to the scarcity of the commodity. In contango, the discount benefits the consumer and the producer is willing to sell at the reduced price. Since both groups are well-dispersed and highly competitive there is no bargaining power over the other in the long run; under risk aversion and large refining frictions, if forward prices are martingales in the real world, then the real and risk-neutral world coincide. This makes real world historical data for parameter estimation extremely conducive to calibrating the model. In the next section we discuss how the correlation information was implemented on the scenario tree.

3.4.1. Generating correlated commodity forward price series

It is well known that the flexibility value of an option to exchange one asset for another is zero if the two asset prices are perfectly correlated, see Margrabe (1978). Despite the fact that the refined petroleum products are highly correlated it is not perfect. With the chosen continuous time price model of spot prices, we have a single factor mean reverting equation for each commodity:

Gasoline: $d(InS_{2t}) = a_2[\theta_2 - InS_{2t}]dt + \sigma_2 dW_{2t}$ (3.4.1)

Naphtha: $d(InS_{3t}) = a_3[\theta_3 - InS_{3t}]dt + \sigma_3 dW_{3t}$ (3.4.2)

Fuel oil: $d(InS_{4t}) = a_4[\theta_4 - InS_{4t}]dt + \sigma_4 dW_{4t}$ (3.4.3)

Heating Oil: $d(InS_{5t}) = a_5[\theta_5 - InS_{5t}]dt + \sigma_5 dW_{5t}$ (3.4.4)

Jet Fuel: $d(InS_{6t}) = a_6[\theta_6 - InS_{6t}]dt + \sigma_6 dW_{6t}$ (3.4.5)

Cracker Feed: $d(InS_{7t}) = a_7[\theta_7 - InS_{7t}]dt + \sigma_7 dW_{7t}$ (3.4.6)

Suppose as above, that a commodity, *i*, has a particular spot price, S_i, and spot dynamics described by the following:

$$d(InS_{it}) = a_i[\theta_i - InS_{it}]dt + \sigma_i dW_{it} \qquad (3.4.7)$$

Also, F_i, is the forward price of the ith asset, $i = 2 \ldots, 7$, and a_i and σ_i are the speed of mean reversion and volatility of that asset respectively and dW_i is the increment in Brownian motion. We can also think of dW_i as a random number drawn from a Normal distribution with mean zero and standard deviation $dt^{1/2}$ so that:

$$E(dW_i) = 0 \quad \text{and} \quad E(dW_i^2) = dt \tag{3.4.8}$$

And the random numbers dW_i and dW_j are correlated according to:

$$E[dW_i dW_j] = \rho_{ij} dt, \tag{3.4.9}$$

We use for example, $dW_1 dW_2 = \rho_{12}dt$, and $dW_2 dW_3 = \rho_{23}dt$, and so on until the seventh commodity. An interesting extension here would be to carry out the entire valuation with a comparison of alternative underlying stochastic processes. Here we have used a version of the one factor model of Schwartz (1997) to represent price uncertainty, but how would the valuation results differ in terms of robustness, if this process were to be changed to a GBM or a multivariate normal or even an affine stochastic set of equations?

Continuing, we state, ρ_{ij}, as the correlation coefficient between the ith and jth random walks. The symmetric matrix with ρ_{ij} as the entry in the ith row and jth column is named the correlation matrix. For example, here we have n = 7 and the correlation matrix is as follows:

$$D = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} & \rho_{15} & \rho_{16} & \rho_{17} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} & \rho_{25} & \rho_{26} & \rho_{27} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} & \rho_{53} & \rho_{36} & \rho_{37} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 & \rho_{45} & \rho_{46} & \rho_{47} \\ \rho_{51} & \rho_{52} & \rho_{53} & \rho_{54} & 1 & \rho_{56} & \rho_{57} \\ \rho_{61} & \rho_{62} & \rho_{63} & \rho_{64} & \rho_{65} & 1 & \rho_{67} \\ \rho_{71} & \rho_{72} & \rho_{73} & \rho_{74} & \rho_{75} & \rho_{76} & 1 \end{pmatrix} \tag{3.4.10}$$

Using NYMEX data the correlation matrix is calculated, but is not positive definite. The matrix is made positive symmetric definite using an approximation to ensure that, $y^T Dy \geq 0$. This is described in the following section.

3.4.1.1. Approximation for generating a symmetric positive definite matrix

We use the correlation structure inherent in the prices to construct the seven dimensional trinomial tree. However, the historical correlation matrix shown below was not in a form that could be used to build the scenario tree.

A covariance matrix in the following form was derived:

$$\Sigma = \begin{pmatrix} (\sigma_1^2) & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 & \rho_{14}\sigma_1\sigma_4 & \rho_{15}\sigma_1\sigma_5 & \rho_{16}\sigma_1\sigma_6 & \rho_{17}\sigma_1\sigma_7 \\ \rho_{21}\sigma_2\sigma_1 & (\sigma_2^2) & \rho_{23}\sigma_2\sigma_3 & \rho_{24}\sigma_2\sigma_4 & \rho_{25}\sigma_2\sigma 5 & \rho_{26}\sigma_2\sigma_6 & \rho_{27}\sigma_2\sigma_7 \\ \rho_{31}\sigma_3\sigma_1 & \rho_{32}\sigma_3\sigma_2 & (\sigma_3^2) & \rho_{34}\sigma_3\sigma_4 & \rho_{35}\sigma_3\sigma_5 & \rho_{36}\sigma_3\sigma_6 & \rho_{37}\sigma_3\sigma_7 \\ \rho_{41}\sigma_4\sigma_1 & \rho_{42}\sigma_4\sigma_2 & \rho_{43}\sigma_4\sigma_3 & (\sigma_4^2) & \rho_{45}\sigma_4\sigma_5 & \rho_{46}\sigma_4\sigma_6 & \rho_{47}\sigma_4\sigma_7 \\ \rho_{51}\sigma_5\sigma_1 & \rho_{52}\sigma_5\sigma_2 & \rho_{53}\sigma_5\sigma_3 & \rho_{54}\sigma_5\sigma_4 & (\sigma_5^2) & \rho_{56}\sigma_5\sigma_6 & \rho_{57}\sigma_5\sigma_7 \\ \rho_{61}\sigma_6\sigma_1 & \rho_{62}\sigma_6\sigma_2 & \rho_{63}\sigma_6\sigma_3 & \rho_{64}\sigma_6\sigma_4 & \rho_{65}\sigma_6\sigma_5 & (\sigma_6^2) & \rho_{67}\sigma_6\sigma_7 \\ \rho_{71}\sigma_7\sigma_1 & \rho_{72}\sigma_7\sigma_2 & \rho_{73}\sigma_7\sigma_3 & \rho_{74}\sigma_7\sigma_4 & \rho_{75}\sigma_7\sigma_5 & \rho_{76}\sigma_7\sigma_6 & (\sigma_7^2) \end{pmatrix} \qquad (3.4.11)$$

Table 3.1 below, shows the correlations of the futures front month prices using historical data from NYMEX; these are proxies for the spot price. The log returns of the prices are used between two consecutive days to calculate the correlation; these are found using the Pearson moment correlation coefficient. To generate the correlated forward prices tree, an approximation is applied to the correlations; the below matrix is not positive definite. The standardised t-statistic is: $t = (r - \rho)/(\sqrt{(1-r^2)/(n-2)})$, where r is the correlation coefficient, sample size n, and ρ is the true correlation coefficient, with a null hypothesis that this is zero. All t-stats are between 3 and 10 showing significance at the 5% level.

(Table 3.1: Correlations of commodity price return series – 10 years of futures prices NYMEX)

	WTI	Gasoline	Naphtha	Jet Fuel	Heating Oil	Fuel Oil	Cracker Feed
WTI	1						
Gasoline	0.6862	1					
Naphtha	1	0.7862	1				
Jet Fuel	0.767	0.885	0.767	1			
Heating Oil	0.8723	0.885	0.8723	0.843	1		
Fuel Oil	0.837	0.843	0.834	0.984	0.8901	1	
Cracker Feed	1	0.8862	1	0.827	0.8723	0.837	1

Generating the set of correlated trinomial trees is done by following Iman and Conover (1982). The authors describe a method for inducing a desired rank correlation matrix on a set of multivariate random variables for use in a simulation study such as the one present for the refinery. The method is simple to apply and is distribution free, and preserves the marginal distribution of the stochastic commodity variables. The authors apply their method to study the geologic disposal of radioactive waste. Without re-describing the author's construction the main idea is the following: suppose that a random row vector X has a correlation matrix I. The elements of X are uncorrelated. Let C be the desired correlation matrix of some transformation of X. Due to C being positive definite and symmetric, C may be written as C =PP' where P is a lower triangular matrix (Scheuer and Stoller, 1962). Then the transformed vector XP' has the desired correlation matrix C. This is the theoretical foundation for the trinomial tree simulation.

We apply a method to convert the above matrix into the following, which is positive symmetric definite; enabling the scenario tree to be constructed - including correlation information. This is done using an algorithm that minimises the root mean squared error between the current matrix and the nearest positive definite matrix obtained by increasing and decreasing the correlation values by a very small amount; a matrix is found that is the nearest positive definite one available. The matrix used in the valuation is shown below:

(Table 3.2: Correlations of commodity price return series – positive definite)

	WTI	Gasoline	Naphtha	Jet Fuel	Heating Oil	Fuel Oil	Cracker Feed
WTI	1.0000						
Gasoline	0.6932	1.0000					
Naphtha	1.0000	0.8932	1.0000				
Jet Fuel	0.7520	0.8593	0.7520	1.0000			
Heating Oil	0.8570	0.8587	0.8570	0.7989	1.0000		
Fuel Oil	0.8471	0.8604	0.8471	0.7671	0.8524	1.0000	
Cracker Feed	1.0000	0.8932	1.0000	0.8520	0.8570	0.8471	1

For an in depth description of this algorithm, see Higham (2002). We next describe the framework and motivations behind the process of discretising a particular stochastic model onto a trinomial tree.

We first diagonalise the above covariance matrix above and write it as:

$$\sum = \begin{pmatrix} \cos\theta & -\sin\theta... \\ \sin\theta & \cos\theta... \\ \vdots & \vdots \\ & \end{pmatrix} \begin{pmatrix} \lambda_1 & 0... & \\ 0 & \lambda_2... & \\ \vdots & & \\ & & \lambda_7 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta... \\ \sin\theta & \cos\theta... \\ \vdots & \vdots \\ & \end{pmatrix}^T \qquad (3.4.12)$$

Where:

$$\lambda_1 = \frac{1}{2}\left(\sigma_1^2 + \sigma_2^2 + \sqrt{(\sigma_1^2 + \sigma_2^2)^2 - 4(1-\rho^2)\sigma_1^2\sigma_2^2}\right)$$
$$\vdots \qquad (3.4.13)$$
$$\lambda_{n-1} = \frac{1}{2}\left(\sigma_{n-1}^2 + \sigma_n^2 + \sqrt{(\sigma_{n-1}^2 + \sigma_n^2)^2 - 4(1-\rho^2)\sigma_{n-1}^2\sigma_n^2}\right)$$

$$\theta = \arctan\left(\frac{\lambda_1 - \sigma_1^2}{\rho\sigma_1\sigma_2}\right) \qquad (3.4.14)$$

We can now write our vector of Brownian motions as:

$$\begin{pmatrix} \sigma_1 W_1(t) \\ \vdots \\ \sigma_n W_n(t) \end{pmatrix} = \begin{pmatrix} \cos\theta\sqrt{\lambda_1}B_1(t) - \sin\theta\sqrt{\lambda_2}B_2(t) \\ \vdots \\ \cos\theta\sqrt{\lambda_{n-1}}B_{n-1}(t) - \sin\theta\sqrt{\lambda_n}B_n(t) \end{pmatrix} \qquad (3.4.15)$$

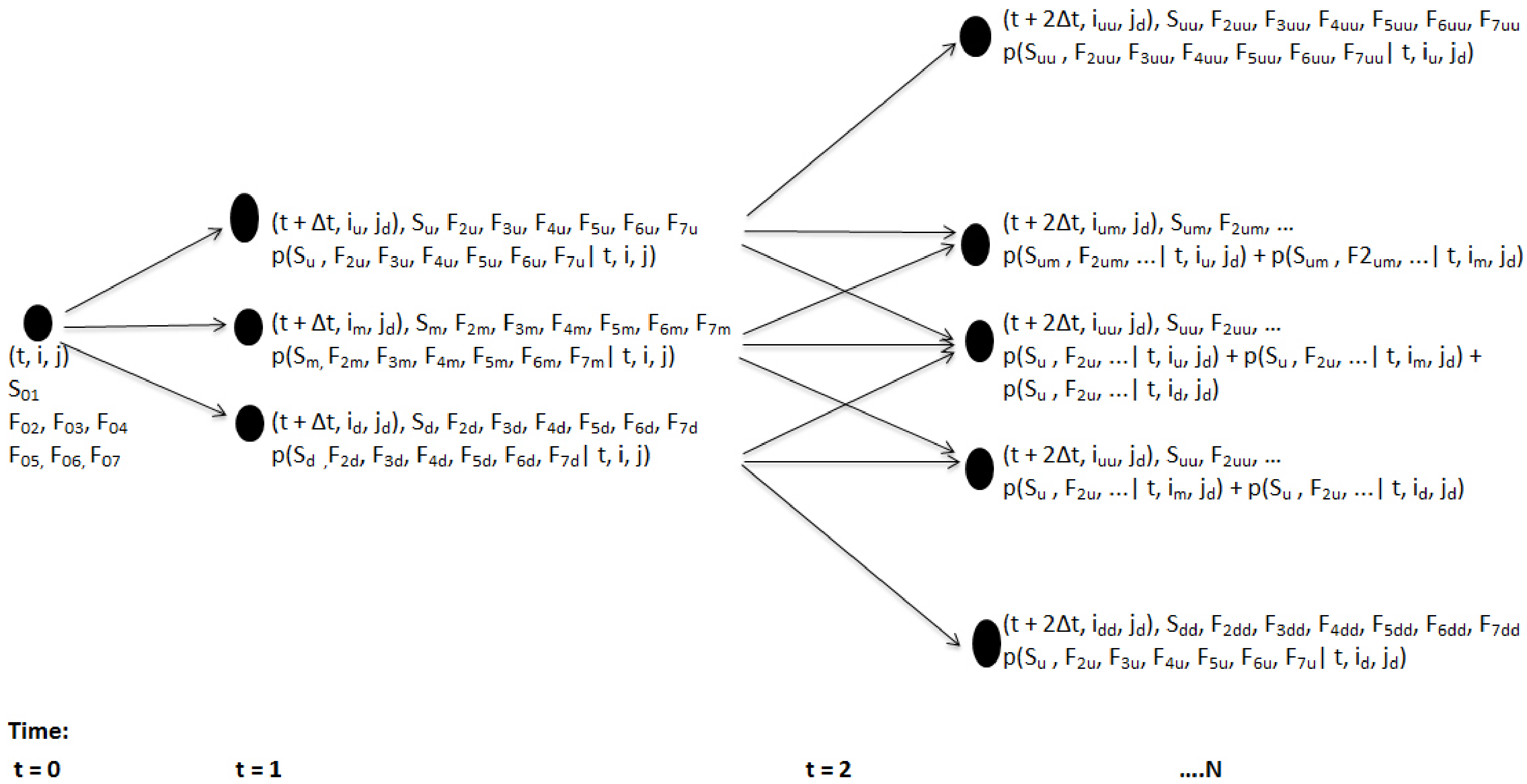
This is the basis for simulating the correlated commodity processes; where the $(B_1,\ldots,B_n)$ is a seven-dimensional standard Brownian motion. The discretisation of the diffusion is then equivalent to the discretisation of seven independent standard Brownian motions. Let $X_1,\ldots,X_7$ be seven independent and identically distributed random variables taking three values (-h, 0, h) with probabilities (p_u, p_m, p_d) respectively. The idea is to use $(X_1,\ldots,X_7)$ to approximate the increments $(B_1(t+dt) - B_1(t),\ldots, B_7(t+dt) - B_7(t))$. Due to the independence of $(X_1,\ldots,X_7)$ the probabilities of getting to each one of the 21 new nodes follow. From here, the discretisation of the price process is completed in the same manner as in a one dimensional tree. We simulate the diffusion up to the terminal date and we optimise for the refinery owner at this point. By backward induction we compute the price at the root of the tree and also optimise to find the value of the refinery and the optimal decision variables.

3.4.2. The discrete-time spot price process

Often closed-form mathematical solutions are unavailable when a contingent claim or valuation is subject to several sources of uncertainty – this is when discrete dynamic programming can provide the solution. Using stochastic variables that take on finitely many variables enables us to avoid measure theory. Lattices are too complex when dealing with multiple uncertainties, path-dependence or complex options, see Clewlow and Strickland (1998) for examples. These problems can be managed more effectively with scenario trees than lattices. Another example by authors pricing real options using trees, but with off the shelf software, is Brandao and Dyer (2005); the authors instead use a binomial tree with risk neutral probabilities to price a project using NPV and they apply a set of GBM processes – they leave for further research the solutions using trinomial trees and due to the third party software do not analyse the structure of the implementation program which certainly renders our current problem intractable. Tree data structures and industrial solvers liked IBM's CONOPT3 are not available in DPL, a specialist decision tree software package; restricting the range of problems that can be solved. Another example would be the pricing of compound options, which can be modelled by adding an additional decision node to a binomial tree. If the primary uncertainty associated with an asset is thought to be mean reverting, as in the case of petroleum related product prices, then Hahn and Dyer (2008) show how a binomial tree may be used to approximate such mean-reverting models as the one factor Ornstein-Uhlenbeck process or the two-factor Schwartz and Smith (2000) process; see Schwartz and Smith (2000) for a detailed explanation of the calibration. However, the extra degree of freedom provided by a trinomial tree is more realistic for mean reversion. We build the tree as seven-dimensional with the dimensions spanning the spot and forward prices, and the time to maturity, with the trinomial branches for each dimension. Thus, each node leads directly into 21 successor nodes. The discretisation of the continuous process is described formally by use of the Arrow-Debreu prices (see Hull and White 1996), where we assume that as $\Delta t \rightarrow 0$, we obtain δt.

Figure 3.16 below, illustrates the above concepts utilised to construct the seven correlated prices onto the trinomial tree; this time increment had to be at least one month over six years, otherwise valuation would be impossible.

FIGURE XII Node transition in the pricing tree

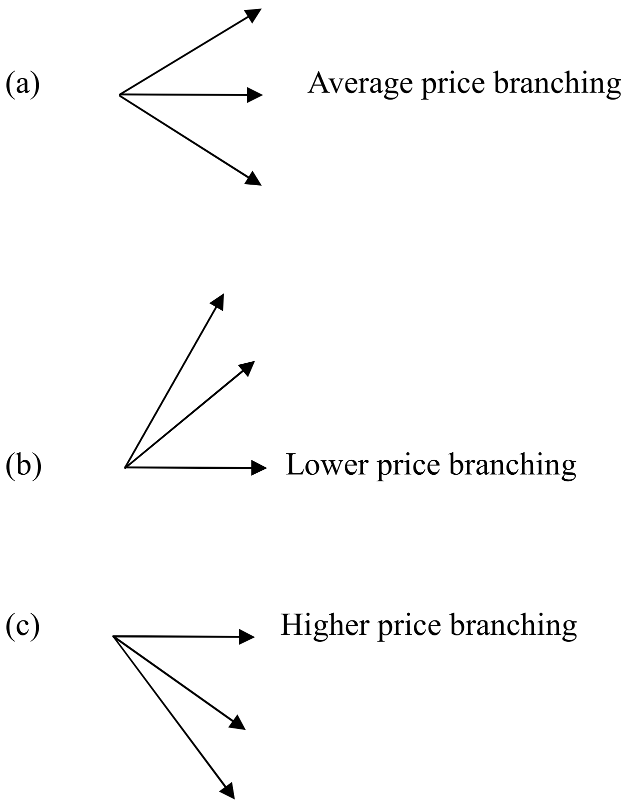


The figure shows transition from an initial node to its successors in the pricing tree. Both the intial spot price S_0, and the forward prices, F_i are depicted as emanating from the same nodes; this is equivalent to there being seven pricing trees. The value of these variables, along with their corresponding joint probabilities are shown. Both spot and forwards prices may transition to a high, medium of lower value state.

(Figure 3.16 Node transitioning on the trinomial tree with probability inputs)

3.4.2.1. Trinomial tree building procedure: Stage one

This procedure exactly follows Hull and White (1996), the number of branches at each period is variable dependent upon the number of time slots but is a function of the commodity price; it is important to note that we alter the branching whenever the probabilities would otherwise be negative, this is implemented within the code. If the simulated process becomes negative the branching is altered this is also true if the price becomes too large. The stochastic process in equation (4.3) is *discretised* onto a trinomial tree formally shown in *section 5.3.2.*; we now discuss the general procedure. The choice of the InS_t as the underlying is more computationally efficient than using just S_t; it is also convenient to use a trinomial instead of a binomial tree as it provides an extra degree of freedom, incorporating mean reversion as a feature on the tree. This framework is equivalent to using the explicit finite difference method, for a thorough description, see Munk (2011). We use alternative branching based on the whether the forward price is high, average or low. Alternative branching therefore also proves practical for incorporating mean reversion:



(Figure 3.17: Branching for the stochastic process along the trinomial tree)

We assume that, on the discretised tree, a function of S_t, $f(S_t)$, can be represented by the stochastic process shown below:

$$df(S_t) = [\theta - a\, f(S_t)]dt + \sigma dW_t \qquad (3.4.16)$$

This originates from the form of the spot price process chosen for the commodities; where we have chosen the function ***f***, to be a logarithm function. Hence, we start by setting $Z_t = \text{In}(S_t)$, (see Ross (1995) for an alternative to the logarithm function), so that we now have:

$$dZ_t = [\theta - aZ_t]dt + \sigma dW_t \qquad (3.4.17)$$

The tree building process starts by constructing a tree for Z* where the process is symmetrical about

$Z^*=0$. We define:

$$dZ_t = [\theta - aZ_t]dt + \sigma dW_t \qquad \text{and} \qquad dZ_t^* = - aZt^*dt + \sigma dW_t \qquad (3.4.32)$$

The first stage is to build a tree for Z_t^*, that follows the same process as, Z_t, except that $\theta_t=0$ and the initial value is zero.

For convenience in stability and convergence, the spacing between spot prices on the tree, ΔS, is set as:

$$\Delta S = \sigma\sqrt{3\Delta t} \qquad (3.4.18)$$

See Hull & White (1996) for proof of uniqueness and convergence. This choice for spacing also turns out to be good for error minimisation. The branching method at each node is then applied. After the geometry is constructed the transition probabilities are calculated.

Define (i,j) as the node where t=i Δt and S*=j ΔS. (The *i* variable is a positive integer representing a time increment and the *j* is a positive or negative integer representing a space increment). The branching method used at each node must lead to positive probabilities. We define j_{max} as the value of j where we switch from branching figure (a) to branching figure (c); and j_{min} as the value of j where we switch from branching figure (a) to branching figure (b). Hull and White (1994) show that the probabilities are always positive if j_{max} is set equal to the smallest integer greater than 0.184/(a Δt) and j_{min} is set equal to - j_{max}. We define p_u, p_m, and p_d as the risk-neutral probabilities of the highest, middle and lowest branches emanating from the node. The risk-neutral probabilities are chosen to match the expected mean and variance of the change in S* over the next time interval Δt.

The mean change in S^* in time Δt is: $-aS^*\Delta t$, and the variance change is, $\sigma^2\Delta t$. At node (i,j), $S^* = j\Delta S$. If the branching has the form in branching figure (a), the p_u, p_m, and p_d must satisfy the following three equations to match the mean, the standard deviation and sum to one:

$$p_u \Delta S - p_d \Delta S = -a\, j\, \Delta S\, \Delta t \quad (3.4.19)$$

$$p_u \Delta S^2 + p_d \Delta S^2 = \sigma^2 \Delta t + a^2 j^2 \Delta S^2 \Delta t^2 \quad (3.4.20)$$

$$p_u + p_m + p_d = 1 \quad (3.4.21)$$

Applying, $\Delta S = \sigma\sqrt{3\Delta t}$ the solution to these is:

$$p_u = \frac{1}{6} + \frac{1}{2}(a^2 j^2 \Delta t^2 - aj\Delta t) \quad (3.4.22)$$

$$p_m = \frac{2}{3} - a^2 j^2 \Delta t^2 \quad (3.4.23)$$

$$p_d = \frac{1}{6} + \frac{1}{2}(a^2 j^2 \Delta t^2 + aj\Delta t) \quad (3.4.24)$$

If the branching is as in figure (b), the solution becomes:

$$p_u = \frac{1}{6} + \frac{1}{2}(a^2 j^2 \Delta t^2 + aj\Delta t) \quad (3.4.25)$$

$$p_m = -\frac{1}{3} - a^2 j^2 \Delta t^2 - 2aj\Delta t \quad (3.4.26)$$

$$p_d = \frac{7}{6} + \frac{1}{2}(a^2 j^2 \Delta t^2 + 3aj\Delta t) \quad (3.4.27)$$

If the branching is as in figure (c), the solution becomes:

$$p_u = \frac{7}{6} + \frac{1}{2}(a^2 j^2 \Delta t^2 - 3aj\Delta t) \tag{3.4.28}$$

$$p_m = -\frac{1}{3} - a^2 j^2 \Delta t^2 + 2aj\Delta t \tag{3.4.29}$$

$$p_d = \frac{1}{6} + \frac{1}{2}(a^2 j^2 \Delta t^2 - aj\Delta t) \tag{3.4.30}$$

We now describe the next stage of the discretisation.

3.4.2.2. Trinomial tree building procedure: Stage two

The second stage is to consists of displacing the nodes in the simplified tree to add the proper drift and to be consistent with the observed forward prices.

As shown in equation 3.3.10. the forward prices are described by the following equation:

$$F(t,T) = \exp\{e^{-a(T-t)} \text{In} S_t + (1-e^{-a(T-t)})(\theta - \sigma^2/2a) + \sigma^2/4a(1- e^{2a(T-t)})\} \tag{3.3.10}$$

With the calibrated parameters and the spot price at a point in time it is trivial to build the forward curve - see figure A.3 for a comparison of calibration results of the forward curve on WTI oil in 2014 using one and two factor models.

We can introduce the correct time varying drift by displacing the nodes at time $i\Delta t$ by an amount a. The a's are selected to ensure that the tree correctly returns the observed forward price curve. The value of Z at node (i,j) in the new tree equals the value of Z* at the corresponding node in the original tree plus a_i; the probabilities remain unchanged. The vital part here is to use forward induction and the state prices to ensure that the tree returns the current prices.

This is all accomplished by displacing the nodes on the S*-tree so we define:

$$\gamma(t) = S(t) - S^*(t) \quad (3.4.31)$$

It follows that:

$$d\gamma_t = [\theta - a\gamma_t]dt \quad (3.4.32)$$

The solution to this is:

$$\gamma(t) = F(0,t) + \sigma^2/2a^2(1-e^{-at})^2 \quad (3.4.33)$$

The F(0,t) is the initial forward market curve; that can easily be fit with the Hull & White tree. The $\gamma(t)$'s are calculated recursively and we define $\gamma_i(t)$ as $\gamma(i\Delta t)$, the value of S at time $i\,\Delta t$ on the S-tree minus the corresponding value of S* at time $i\,\Delta t$ on the S*-tree. We define $S_{i,j}$ as the present value of a commodity if node (i,j) is reached and zero otherwise. If the underlying stochastic process were to include two factors or include a time varying mean or volatility, this tree building procedure can be altered to accommodate these features - in chapter four we do just this. This does however, introduce non-stationarity into the stochastic model. We now discuss the procedure for calibrating the model to market price data.

3.4.3. Forward curves calibration procedure using Levenberg-Marquardt (1944)

Often when calibrating parameters academics are using a likelihood function or the Kalman filter; here we introduce a more efficient calibration method designed during the second world war

(1944). A convenient tool suited to dealing with markets where the state variables are unobservable, but are generated in a Markov manner, is known as the state space form. Reformulating the model into this form means the Kalman filter can be used to find the parameters that best estimate the model, hence the time series of the unobservable state variables. Frequently, the spot price of a commodity is so volatile that a near futures contract is used as a proxy instead, see Harvey (1989) for a survey of state space form algorithms. The problem with using the Kalman filter is that it assumes a linear dynamical system, and that measurements have a multivariate Gaussian distribution. Although, there are applicable extensions that could be applied to this refinery model set-up; we chose an algorithm that is simplistic, static and required less computational effort.

For a description of an application in simulation using trees see Hoyland and Wallace, (2001). The authors construct a three period tree using volatility clumping and mean reversion; however, the tree quickly becomes “over specified” when the number of periods is increased over three. The calibration of the parameters goes hand in hand with the generation of the scenarios; hence, the Hull and White trinomial tree (1994) is utilised. To calibrate the initial stochastic forward curve model to the tree, and therefore to the real world, we find the parameters a, θ and σ, using the *Levenberg-Marquardt* (LM) method fitting the first two moments of the Forward curve model to the NYMEX future prices data. This algorithm optimises a least squares error and is effective even if an initial set of starting parameters are omitted. It does however, only find a local and not a global minimum; see Levenberg (1944) for the foundations. The mean and variance were equally weighted, and the parameters chosen to enable the model to fit the data with minimum error. In general, LM fitting significantly outperforms gradient descent and conjugate gradient methods when used on non linear least squares applications. The formula below minimises a measure of distance between the mean and variance of the constructed model, defined in equations, (3.3.13) and (3.3.14), and the specified values from the historical future price data:

$$\min_{\sigma,\theta,a} \sum_{T} \sum_{i \in S} w_i (\hat{f_i}(\sigma,a,\theta,T) - S_{VALi})^2 \qquad (3.4.34)$$

Where $f_i(\sigma,a,\theta)$ is the mathematical formula for the statistical property being minimised, here i=1 for the mean, or i=2 for variance, T are the maturity of the future price contracts, σ is the volatility in the model, w_i is the weight for statistical moment, S_{VALi} is the specified value of the statistical property i of S, and a is the mean reversion speed in the forward curve model. Once the "*a*", "*θ*" and the "*σ*" are obtained from the stochastic process calibration, all the equations needed to create the trinomial tree can be calculated, equations (3.4.25 – 3.4.34). In reality on NYMEX only the first five contracts are very liquid and so we use these within the least squares above. The parameters are those that generate the lowest squared sum difference. GAMS is used for parameter calibration and after 14 iterations it takes five minutes to converge, here the tolerance was $10e^{-4}$, at which point the results are obtained. There is no guarantee here that the parameters are a unique and global solution; therefore we alter upper and lower bounds for the initial values and use different sets, the values are insensitive to these tests and so we conclude that the parameters are locally optimal. The results of the above calibration for crude are shown in the following section and appendix A.4.; standard errors are in brackets:

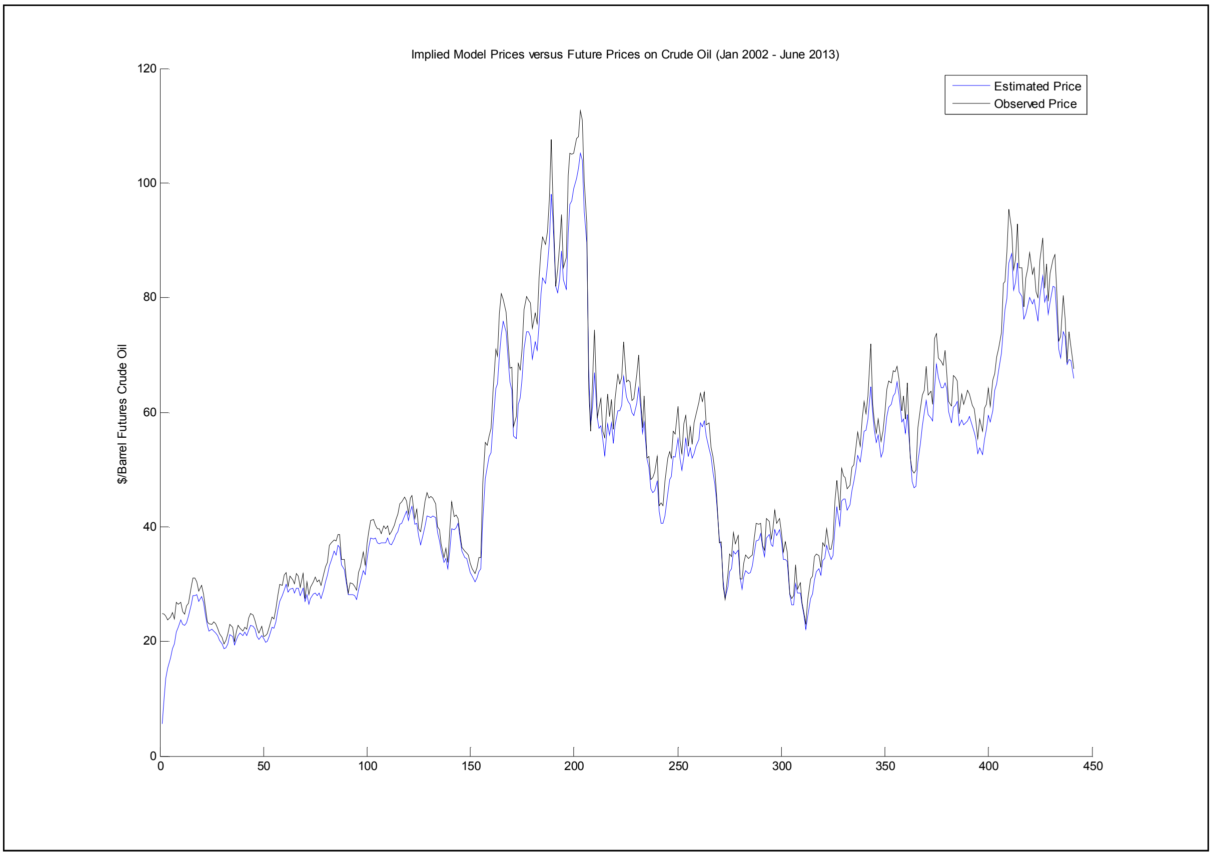
3.4.3.1. Results of forward curves calibration

One Factor Stochastic Model: Crude Oil
Period : 1/2/2002 to 01/06/2013
Contracts : F1, F3, F5, F7, F9
NOBS : 120
a_1 : 0.561 (0.08)
θ_1 : 3.343 (0.346)
σ_1 : 0.2955 (0.03)
$RMSE_1$: 0.9993

(*Table 3.3:* The calibration results of the stochastic one factor model over ten years using crude future contract prices from NYMEX.*)*

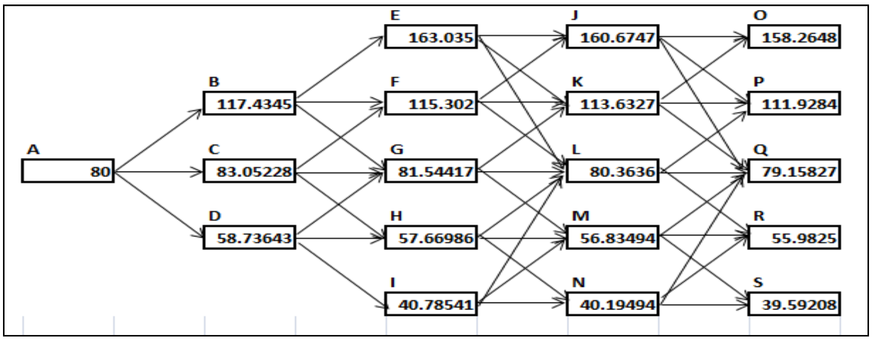
The forward contracts in the above table from months one to nine are utilised again in the next chapter for calibrating a two factor model. We could have chosen a daily granularity as the data was available but one should model at the same detail as price estimations are required - due to the state space being ten years and decisions being made monthly we choose monthly contracts. It is useful to note the term structure of the forward curve under the different models which is shown in appendix A 3.0.

An example fitted forward curve for the stochastic crude process is shown below:



(Figure 3.18: Crude oil model prices versus futures market prices, NYMEX)

Crude oil is calibrated using historical data, and the correlated prices are generated for all the refined product prices; see the appendix for the results. By taking into account the historical correlations, a multidimensional trinomial tree is constructed. Firstly, all the SDEs are calibrated and discretised onto the scenario tree, next the decision set for commodity volumes is optimised over the entire horizon in one sequence; further, the entire process is repeated 1,000 times and finally, averages of the optimal values recorded. For example, we show prices for four periods; spot and futures prices for petroleum products are given below for each node on the tree:



(Figure 3.19: Four period (months) trinomial tree of spot crude oil market prices.) *72 periods were calculated in practice.

The values for four months of all oil related products are shown in table 3.4:

(Table 3.4: The prices of crude and refined products at each node in $ per barrel)

NODE	F_1	F_2	F_3	F_4	F_5	F_6	F_7
A	80	185	80	125	145	60	15
B	117	296	117	200	232	96	24
C	83	236	83	155	185	76	19
D	58	176	58	114	138	57	14
E	163	386	163	260	301	128	35
F	115	322	115	222	253	106	28
G	82	264	82	179	207	86	25
H	57	201	57	138	161	68	23
I	41	145	41	96	114	48	17
J	160	411	160	278	324	135	13
K	113	350	113	238	277	116	35
L	80	291	80	197	230	95	30
M	56	231	56	157	183	78	24
N	40	170	40	116	136	55	19
O	158	475	158	319	269	152	15
P	112	395	112	265	304	127	40
Q	79	330	79	223	259	107	33
R	56	269	56	183	211	88	26
S	40	211	40	142	168	69	22

Table 3.5 below is constructed using the equations (3.4.25) to (3.4.34) to show the calculated

probabilities of the forward curve dynamics over two years; in the final results up to at least six years are required. It is important to note that in the C program used to carry out the real option valuation, all $/barrel values were converted to $/ton as the rest of the program is defined in these units, and as carried out by refinery practitioners.

In practise we simulate the spot and forward prices up to 72 months forwards. We calculate optimal decisions at each node, along each path by using specialist software package, GAMS.

3.4.3.2. Delta of the oil refinery

One factor stochastic models are applied on trading desks, giving realistic pricing features; however, when considering the hedging of an option or a real asset there are serious issues. In practice, pricing can be carried out with a one factor model but hedging requires multifactor models – this is also known as outside model hedging. For example, calculating how sensitive the value of the refinery is to a movement in the set of prices, a type of multidimensional delta, should allow for many movements in the term structure. This is however outside the scope of our work and we leave it for further research - a Monte Carlo simulation can be investigated for this purpose. In the next section we discuss how to solve the Bellman equation on the discretised trinomial tree along with a relevant approximation.

(Table 3.5: *Transition trinomial probabilities for 2 years is shown below; going from a starting node to the next node, with probability up, pu, probability middle pm, and probability down pd.)*

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6.06104E-15
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.07693E-15	-9.22067E-14
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.08469E-15	-5.6892E-14	6.83084E-13
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.64402E-15	-3.88671E-14	3.90075E-13	-3.26119E-12
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.58795E-15	-2.94164E-14	2.44663E-13	-1.72881E-12	1.12376E-11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.91874E-15	-2.45855E-14	1.68843E-13	-9.9881E-13	5.54195E-12	-2.96884E-11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.83204E-15	-2.24175E-14	1.28686E-13	-6.28063E-13	2.96305E-12	-1.36228E-11	6.2423E-11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.92584E-15	-2.1491E-14	1.09798E-13	-4.26648E-13	1.72118E-12	-6.72917E-12	2.66684E-11	-1.07022E-10
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.094E-14	-1.89627E-14	1.10168E-13	-3.01995E-13	1.10144E-12	-3.55257E-12	1.22469E-11	-4.24518E-11	1.52317E-10
0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.45E-14	-3.67388E-15	1.49621E-13	-1.7749E-13	8.43142E-13	-1.90318E-12	6.16817E-12	-1.77819E-11	5.63881E-11	-1.81584E-10
0	0	0	0	0	0	0	0	0	0	0	0	0	6.772E-14	7.74E-14	3.38803E-13	1.51343E-13	1.07464E-12	-5.42334E-13	4.03118E-12	-7.05413E-12	2.3187E-11	-6.07815E-11	1.85415E-10
0	0	0	0	0	0	0	0	0	0	0	0	2.362E-13	5.118E-13	1.3E-12	1.76106E-12	3.32517E-12	3.08772E-12	6.69486E-12	2.04299E-12	1.59689E-11	-1.47076E-11	6.65767E-11	-1.49307E-10
0	0	0	0	0	0	0	0	0	0	0	1.066E-12	3.133E-12	7.145E-12	1.19E-11	1.86048E-11	2.43956E-11	3.29188E-11	3.66385E-11	4.89491E-11	4.22404E-11	7.52195E-11	1.4335E-11	1.81068E-10
0	0	0	0	0	0	0	0	0	0	6.394E-12	2.227E-11	5.098E-11	9.013E-11	1.39E-10	1.91317E-10	2.47382E-10	2.99604E-10	3.53418E-10	3.94793E-10	4.46578E-10	4.62418E-10	5.39277E-10	4.61467E-10
0	0	0	0	0	0	0	0	0	5.256E-11	2.006E-10	4.684E-10	8.465E-10	1.31E-09	1.82E-09	2.35533E-09	2.87524E-09	3.36804E-09	3.81514E-09	4.22039E-09	4.56534E-09	4.8822E-09	5.11463E-09	5.3913E-09
0	0	0	0	0	0	0	0	6.06423E-10	2.401E-09	5.651E-09	1.024E-08	1.583E-08	2.199E-08	2.83E-08	3.44958E-08	4.02871E-08	4.55579E-08	5.0249E-08	5.43423E-08	5.78718E-08	6.08645E-08	6.34067E-08	6.54951E-08
0	0	0	0	0	0	0	9.83389E-09	3.868E-08	8.972E-08	1.603E-07	2.446E-07	3.358E-07	4.281E-07	5.17E-07	5.99364E-07	6.7372E-07	7.39347E-07	7.9629E-07	8.45041E-07	8.8633E-07	9.2101E-07	9.49927E-07	9.73929E-07
0	0	0	0	0	0	2.10726E-07	7.91628E-07	1.76469E-06	3.05E-06	4.526E-06	6.075E-06	7.601E-06	9.039E-06	1.03E-05	1.15124E-05	1.25272E-05	1.33988E-05	1.41388E-05	1.47612E-05	1.5281E-05	1.57125E-05	1.6069E-05	1.63625E-05
0	0	0	0	0	5.057E-06	1.74481E-05	3.63464E-05	5.95052E-05	8.455E-05	0.0001096	0.0001332	0.0001548	0.000174	0.000191	0.000204965	0.000217055	0.0002272	0.000235651	0.000242654	0.000248429	0.000253177	0.000257068	0.00026025
0	0	0	0	0.0001084	0.0003287	0.000620145	0.000939619	0.001255965	0.0015506	0.0018143	0.0020442	0.002241	0.0024072	0.002546	0.002661617	0.002756983	0.002835435	0.002899766	0.002952387	0.002995345	0.003030361	0.003058868	0.003082053
0	0	0	0.0017574	0.0044404	0.007311	0.009988561	0.01232762	0.014302345	0.0159386	0.0172803	0.018374	0.0192625	0.019983	0.020567	0.021039057	0.02142149	0.021731013	0.021981529	0.022184299	0.022348435	0.022481309	0.022588884	0.022675984
0	0	0.0202778	0.0395417	0.0542541	0.0649118	0.072583288	0.078158317	0.08226856	0.0853441	0.0876766	0.0894664	0.0908533	0.0919368	0.092789	0.093462832	0.093998092	0.094424813	0.094766014	0.095039493	0.095259121	0.095435785	0.095578073	0.095692795
0	0.1666667	0.2205556	0.2370333	0.2408764	0.2405019	0.238851455	0.236965962	0.235203477	0.2336701	0.2323783	0.2313078	0.2304285	0.22971	0.229125	0.228648702	0.228262219	0.227948615	0.227694282	0.227488092	0.227320972	0.227185542	0.227075807	0.226986898
1	0.6666667	0.5183333	0.4433352	0.4006416	0.3738829	0.355877782	0.343142667	0.333816688	0.326818	0.3214724	0.317336	0.3141038	0.3115591	0.309544	0.30794136	0.306661831	0.305637471	0.304815534	0.304154821	0.303622935	0.303194254	0.302848426	0.302569225
0	0.1666667	0.2205556	0.2370333	0.2408764	0.2405019	0.238851455	0.236965962	0.235203477	0.2336701	0.2323783	0.2313078	0.2304285	0.22971	0.229125	0.228648702	0.228262219	0.227948615	0.227694282	0.227488092	0.227320972	0.227185542	0.227075807	0.226986898
0	0	0.0202778	0.0395417	0.0542541	0.0649118	0.072583288	0.078158317	0.08226856	0.0853441	0.0876766	0.0894664	0.0908533	0.0919368	0.092789	0.093462832	0.093998092	0.094424813	0.094766014	0.095039493	0.095259121	0.095435785	0.095578073	0.095692795
0	0	0	0.0017574	0.0044404	0.007311	0.009988561	0.01232762	0.014302345	0.0159386	0.0172803	0.018374	0.0192625	0.019983	0.020567	0.021039057	0.02142149	0.021731013	0.021981529	0.022184299	0.022348435	0.022481309	0.022588884	0.022675984
0	0	0	0	0.0001084	0.0003287	0.000620145	0.000939619	0.001255965	0.0015506	0.0018143	0.0020442	0.002241	0.0024072	0.002546	0.002661617	0.002756983	0.002835435	0.002899766	0.002952387	0.002995345	0.003030361	0.003058868	0.003082053
0	0	0	0	0	5.057E-06	1.74481E-05	3.63464E-05	5.95052E-05	8.455E-05	0.0001096	0.0001332	0.0001548	0.000174	0.000191	0.000204965	0.000217055	0.0002272	0.000235651	0.000242654	0.000248429	0.000253177	0.000257068	0.00026025
0	0	0	0	0	0	2.10726E-07	7.91628E-07	1.76469E-06	3.05E-06	4.526E-06	6.075E-06	7.601E-06	9.039E-06	1.03E-05	1.15124E-05	1.25272E-05	1.33988E-05	1.41388E-05	1.47612E-05	1.5281E-05	1.57125E-05	1.6069E-05	1.63625E-05
0	0	0	0	0	0	0	9.83389E-09	3.868E-08	8.972E-08	1.603E-07	2.446E-07	3.358E-07	4.281E-07	5.17E-07	5.99364E-07	6.7372E-07	7.39347E-07	7.9629E-07	8.45041E-07	8.8633E-07	9.2101E-07	9.49927E-07	9.73929E-07
0	0	0	0	0	0	0	0	6.06423E-10	2.401E-09	5.651E-09	1.024E-08	1.583E-08	2.199E-08	2.83E-08	3.44958E-08	4.02871E-08	4.55579E-08	5.0249E-08	5.43423E-08	5.78718E-08	6.08645E-08	6.34067E-08	6.54951E-08
0	0	0	0	0	0	0	0	0	5.256E-11	2.006E-10	4.684E-10	8.465E-10	1.31E-09	1.82E-09	2.35533E-09	2.87524E-09	3.36804E-09	3.81514E-09	4.22039E-09	4.56534E-09	4.8822E-09	5.11463E-09	5.3913E-09
0	0	0	0	0	0	0	0	0	0	6.394E-12	2.227E-11	5.098E-11	9.013E-11	1.39E-10	1.91317E-10	2.47382E-10	2.99604E-10	3.53418E-10	3.94793E-10	4.46578E-10	4.62418E-10	5.39277E-10	4.61467E-10
0	0	0	0	0	0	0	0	0	0	0	1.066E-12	3.133E-12	7.145E-12	1.19E-11	1.86048E-11	2.43956E-11	3.29188E-11	3.66385E-11	4.89491E-11	4.22404E-11	7.52195E-11	1.4335E-11	1.81068E-10
0	0	0	0	0	0	0	0	0	0	0	0	2.362E-13	5.118E-13	1.3E-12	1.76106E-12	3.32517E-12	3.08772E-12	6.69486E-12	2.04299E-12	1.59689E-11	-1.47076E-11	6.65767E-11	-1.49307E-10
0	0	0	0	0	0	0	0	0	0	0	0	0	6.772E-14	7.74E-14	3.38803E-13	1.51343E-13	1.07464E-12	-5.42334E-13	4.03118E-12	-7.05413E-12	2.3187E-11	-6.07815E-11	1.85415E-10
0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.45E-14	-3.67388E-15	1.49621E-13	-1.7749E-13	8.43142E-13	-1.90318E-12	6.16817E-12	-1.77819E-11	5.63881E-11	-1.81584E-10
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.094E-14	-1.89627E-14	1.10168E-13	-3.01995E-13	1.10144E-12	-3.55257E-12	1.22469E-11	-4.24518E-11	1.52317E-10
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.92584E-15	-2.1491E-14	1.09798E-13	-4.26648E-13	1.72118E-12	-6.72917E-12	2.66684E-11	-1.07022E-10
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.83204E-15	-2.24175E-14	1.28686E-13	-6.28063E-13	2.96305E-12	-1.36228E-11	6.2423E-11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.91874E-15	-2.45855E-14	1.68843E-13	-9.9881E-13	5.54195E-12	-2.96884E-11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.58795E-15	-2.94164E-14	2.44663E-13	-1.72881E-12	1.12376E-11
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.64402E-15	-3.88671E-14	3.90075E-13	-3.26119E-12
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3.08469E-15	-5.6892E-14	6.83084E-13
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.07693E-15	-9.22067E-14
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6.06104E-15

3.5. NUMERICAL PROCEDURE FOR THE VALUATION

Our model is related to the literature on the valuation of commodity and energy real options, for instance, see a natural gas storage valuation in Thompson et al (2002). Often the valuations are based on low-dimensional representations of the evolution of the spot price. In contrast, we create a discrete-time stochastic dynamic programming model, using a high-dimensional representation of the evolution of the underling spot prices. Jaillet et al. value natural gas storage as an extended swing option. The author's method implements a low-dimensional model of the gas spot price evolution, and restricts the policy space. Our model uses a high-dimensional approach of the spot prices processes, and does not apply constraints on trading decisions. Haugh and Kogan generalise the work of pricing American options. The refinery is much more difficult than this valuation because it features an inventory of hydrocarbon liquid that is absent in American option valuation. Secomandi et al. (2008) compare optimal and heuristic approaches for natural gas storage valuation. The authors apply a re-optimised deterministic model, restricting the spot price evolution to a one-factor mean-reverting spot price model, which we have also applied but on seven state variables rather than one.

The quantity of interest in this model is the value of a given oil refinery portfolio at the time of inception. This value depends on how the petroleum product prices change over time as a refinery owner acts dependent on these changes as follows: buying crude and refining it at a given point in time, storing and doing nothing if prices are unfavourable, selling refined products at a later point in time. Such a portfolio can be valued as the discounted risk-neutral expected value of the cash flows from optimally operating the refinery during its tenure, whilst respecting its operational constraints.

The optimisation and valuation take place under the volumetric risk: the volumes of crude and refined products; these are related to the exogenous non-traded variables such as weather and price risk. Decision theory has paid attention to the problem of dynamic choice under uncertainty. The notation applied to the refinery optimisation problem is given below:

i – time period

q_{ij} – volume of oil product, j, purchased or sold (decision variables) at period i

ξ_i - a vector of spot and forward prices of the petroleum products

$f(q_{ij}, \xi_i)$ – a function relating the dependent variables to the cash flow

$G = (G_i)_{i=1,\ldots,T}$ – sequence of the discrete time cash flows,

$(\theta_i)_{i=1,\ldots,T}$ – decisions are occurring at these times

$\omega \in \Omega$ – states of nature

$V_i(G, \omega)$ – the value measure (represents the value of the sequence of cash flows)

NB: Cumulative cash flows, as considered in the coherent risk measure literature, can be viewed as value processes.

In this framework, the date θ_i value is assessed recursively by aggregation of the current cash flow, G_i, and the discounted value measure, V_{i+1}, seen from date θ_i. It is important to state that here, V_i, is a $\Im_{\theta_i}$ adapted process.

Where $q_{1,...,T}$ are the decision variables, ξ_i the vector of random prices, $f(q_{ij,}\ \xi_i)$ is a function relating the dependent variables to the cash flows, i is the time period, Q is a set of feasible volumetric decisions ensuring non-negativity and capacity constraints, it also captures the state across periods.

A set of production decisions at each date θ_i will result from the optimisation of the chosen value measure. This optimisation not only yields the first set of decisions, $q_{10,}....q_{70}$ at date θ_0, but a whole set of decision planning for *all* subsequent dates up until date θ_T.

Consider the cash flow sequence on these dates depending on decisions and a multi-dimensional process: $(\xi_{ji})_{1\le i\le T} : G_i := f(q_{ji}, \xi_{ij})$, here ξ_{ij} is a vector of spot prices for the oil products, q_{ij}, is the set of all production variables, and f is a reasonably behaved function relating these two variables to the cash flow at each date i.

The state is x_i, on which decisions depend at time θ_i. It is supposed that, after a set of decisions, q_{ij}, is made at time θ_i, the state x_i, leads to another state: $x_{i+1} = h(x_i, q_{ij}, \in_{i+1})$, where h is a deterministic function, and ε is a normally distributed random variable. We define the optimisation problem related to a dynamic value measure, V_i, at each time period i as:

$$J_i(x_i) = \underset{(q_k)_{k\ge t} \in A(x_i)}{Max} V_i(G_i) \qquad (3.5.1)$$

Here, J_i, is the optimal value of maximising the value function, V_i, dependent on cash flows G_i, the, $A(x_i)$ are the set of admissible strategies; the set of qs are calculated to maximise the, $J_i(x_i)$ over the finite time horizon. The discrete time model has a finite time horizon, $i=1,\ldots\ldots,T$ (months) and a set of decision dates $(\theta_i)_{i=1,\ldots\ldots,T}$. In addition the commodity is supposed to be traded, stored and consumed in the same location. Refining product j, decision variables q_{ij}, corresponding to period i, are subject to the standard physical refinery constraints defined in section 3.2.5, in algebraic modelling language (AML) described as:

$$0 \leq q_{i1}^{buy} \leq Q_{i1}^{buy}, \quad 0 \leq q_{ij}^{sell} \leq Q_{ij}^{sell}, \qquad i \geq 1, \qquad 2 \leq j \leq 7 \qquad (3.5.2)$$

$$L_0 = L_{init}, \quad L_i = L_{i-1} + q_{i1}^{buy} - \sum_{j=2}^{7} q_i^{j,sell}, \quad 1 \leq i \leq T, \qquad 2 \leq j \leq 7 \qquad (3.5.3)$$

$$L_{min} \leq L_{ij} \leq L_{max} \qquad \forall i = 1,...,T, \quad L_T \geq L_{end}, \qquad 1 \leq j \leq 7 \qquad (3.5.4)$$

Where L_{ij} is the liquid volume of hydrocarbon for product *j* at time period *i*, other notation utilised is defined as follows:

L_0 – Overall refinery complex liquid volume at date 0

L_{init} – initial capacity level defined

L_{min} – minimum refinery complex capacity

L_{max} – maximum refinery complex capacity

L_{end} – boundary requirement for optimisation of the refinery

Q^{buy} – The maximum volume of crude that can be bought in any period

Q^{sell} – The maximum volume of particular refined oil that can be produced in any period

Each $\omega \in \Omega$ represents a realisation of the process $\xi_i = (F_{i,i,1}, F_{i,k,j,\ k>i})$, *i=1,….,T* – here the index *k* must be greater than *i* as products are sold forwards, and it takes a minimum of one month to refine and deliver to a client. Only forward contracts are considered where we denote by $F_{i,k}$[48] the forward price of the commodity end product quoted during period *i* for delivery in period k ($k \geq i$) and S_i the post price of a commodity, where $S_i = F_{i,i}$. All processes will be a discrete time-finite number of states of the world. The assumption that the state space is discrete ensures the convexity of any optimal result. The total cash flow during period, *i,* is denoted as G_i:

[48] Here, $F_{i,j}$ can be considered as the average price over all the quotation dates in the period i on forward contracts for delivery in period k

$$G_i = \sum_{j=2}^{7} e^{-r(\theta_k - \theta_i)} q_i^j F_{i,k}^j - q_i^1 S_{1i} \ , \tag{3.5.5}$$

Where *j,* is refined product index, *1* is the crude oil index and other symbols are as defined previously. Note that it was assumed that there is an absence of credit risk, hence equality between forward and futures prices.

We denote, $A(x_i)$ as the set of (F_k) measurable admissible strategies $(q_k)_{k>i} = (q_{ik}^j, q_i^1)$ from state: $x_i = g(L_{i-1}, \xi_i, \varepsilon_i)$.

The fundamental Bellman equation is here:

$$V_i(q_{i-1}, \xi_{[1,i]} \mid \zeta_0) = \max_{q_i \in A_i(q_{i-1}, \xi_i)} [f_i(q_i, \xi_i) + V_{i+1}(q_i, \xi_{[1,i+1]}) \mid \zeta_0] \tag{3.5.6}$$

Assuming between period independence holds for the stochastic processes that generate the cash flow, G_i, and at every period, i, we maximise, G_i, then at the last period above, the value function, $V_T(G_{T-1})$, is the optimal value of the above problem. The optimal value calculated recursively until the present day is *myopic*.

If both the decision set and the objective function are convex, and also closed, then any local maximum is a global maximum.

The optimal value $J_i(x_i)$ of the portfolio satisfies the dynamic programming equation:

$$J_i(x_i) = \underset{(q_k)_{k \geq t} \in A(x_i)}{Max} \{ G_i(\underline{\underline{\boldsymbol{q_i}}}) + \beta_i (E_i (J_{i+1}(x_{i+1}))) \} \tag{3.5.7}$$

The state x_{i+1}, is given by the state transition equation:

$$x_{i+1} = (L_{i-1} - \sum_{j=2}^{7} q_i^j + q_i^1, g(\xi_i, \epsilon_{i+1})) \tag{3.5.8}$$

Trinomial trees are often applied in real option valuation, they are simple for management to understand and robust, see Hahn (2008). A tree like this determines the refinery prices and associated probabilities over each period we select until the lifetime of the refinery. American options can be priced in this way once the state space is judicious. To understand the valuation of the refinery we emphasise the primary differences. Like a standard approach our tree is recombining; all transitional probabilities are positive and add to one; our choice of steps is tractable; the mean of the trinomial distribution is equal to the mean of the stochastic process we have chosen; the variance also and finally, it spans enough time periods to value the asset. American options are often priced using lattice methods, but the number of periods is usually much less than six years and there is only one underlying, additionally there are very different constraints in place for the refinery.

In the next section we discuss the properties that will ensure the optimisation is realistic, and we will describe how the above formulation is solved computationally; introducing our approximation and reason for faster than standard solving time.

3.5.1. Numerical Approximation using stages to obtain a valuation

We can scale down our problem using a parameter, k, within the GAMS code. The number of periods is too high for computation; hence we start with k=1 and increase until our solution takes over two hours, which is at k = 11. This period of time was chosen as the maximum that a user would be willing to wait for valuation purposes. This staging shortcut was only possible due to the mapping shortcut described in section 3.6.5.

3.5.2. Methods available to solve the dynamic program

To computationally solve the problem, we use a slightly alternative stochastic programming technique. The set of realisations of commodity spot prices and the forward curve is represented on an event tree with nodes n ε N, the decisions, $q(t, \omega)$, are indexed on the nodes of the tree, and the terminal-date objective maximised numerically with respect to all decisions $(q_n)_{n \varepsilon N}$ using a large scale non-linear solver. It is important to describe the problem in AML as many linear programs cannot in fact be solved or tested due to specific properties of the problem in question.

The payoff to the refinery owner is defined as the value of refining today, plus the discounted expectation of refining over the life of the oil refinery, except at maturity. The option to choose the optimal refining times have no analytical solution in this setting as there are too many state variables. Having discretised the stochastic differential equations onto a trinomial tree, we have the equivalent of an *explicit finite difference grid***.**

One way to think of this problem is to assume that we are again in continuous time, as if we could refine continuously, and to approximate all the derivatives by finite differences. The space domain can be partitioned using a mesh of $x_0,\ldots,x_j$ and the time domain using a mesh $t_0,\ldots,t_N$. Assuming a uniform partition in both space and time, the difference between two consecutive space points is h and between two consecutive time points is k. The points become:

$V(x_j, t_n) = v_j^n$, this represents the numerical approximation of $V_{j,n} = f(x_j, t_n)$.

Maximum points are: T = N Δt, and X = M Δx

The value of the refinery written in continuous time is the following:

$$V_t(x_t) = \underset{(q_k)_{k \geq t} \in A(x_t)}{Max} \{ \int_t^{t+1} G_t(\underline{\underline{q_t}}, x)dt + V_{t+1}(x_{t+1}, q_{[t,t+1]}) \} \qquad (3.5.9)$$

After some manipulation we have:

$$V_t(x_t) = \inf_{(q_k)_{k \geq t} \in A(x_i)} H(-\tfrac{dV_t}{dx}, x_t, q_{[t]}) \tag{3.5.10}$$

Following arguments from (Dixit and Pindyck and Willmott) we can form the dynamic programming equation as a HJB equation (3.5.10) with the one factor SDE for each commodity as a vector. To solve the set-up that would follow, one would have to construct a finite difference grid with appropriate boundary conditions; we leave this approach for others.

Due to the non-analytical solution of the above HJB equation, we develop a model of dynamic programming in a setting where uncertainty is modelled using discrete-time Markov processes. Time plays a very important role for investment decisions. Dynamic programming is a tool that is particularly applicable when considering uncertainty. It fragments a sequence of decisions into two parts: the initial decision, and a value function that encapsulates the consequences of all preceding decisions, starting with the position that results from the current decision. In a finite horizon problem as we have here, the last decision at its end has nothing afterwards, and can be found using a deterministic calculation. Obtaining this solution then provides the valuation function appropriate to the penultimate decision. This enables a decision two periods from the end, and this continues until today's decisions can be found.

The key problem with the oil refinery and the reason why this optimisation has been avoided up until now is the huge dimensional space that is present; solving over this number of periods for this size state space is rare in engineering applications and even more so in the finance literature. The mathematical complications connected to the real option theory originates from the idea that the general issue cannot be solved without a probabilistic solution to a firm's optimal investment decision policy, not only at date 0 but for all periods in time until the end of the real asset's horizon. For example, Brandao (2002) provides an example of the real option valuation application using a binomial tree; hence one less degree of freedom than ours, valuing a highway project in Brazil that includes 20 time periods. The author's decision tree has 5.2e5 scenarios, in this refinery valuation there are 72 time periods; meaning we have in total 7.5e33 scenarios. Solving times and computational analysis are not described by the authors. As described in Hull (2003) in detail, solving a real option problem on a binomial tree with periods n = 30, means computation takes approximately 1000 minutes to solve; we find a stochastic value utilising a trinomial tree with 72 time periods in just over 140 minutes.

3.5.3. Mathematical issues with the valuation

Many authors have successfully described the convergence properties of discrete tree applications of their stochastic processes. In Heston and Zhou (2000) table 3.1 provides the order of the local error and the rate of convergence of multinomial models when the payoff functions are smooth enough. The step size is Δ, which is proportional to 1/n in an n-period multinomial period. The refinery portfolio problem is 72-periods; a step size is approximately 1/72; the trinomial tree, assuming payoff functions are smooth enough, converges in at least $O((1/72)^2)$. However, the authors Albanese and Mijatovic (2006) go further showing the rate of convergence if the PDF of a continuous-time Markov chain to a Brownian motion with drift, is of the same order $O(\Delta^2)$. Convergence is assured here as shown by Hull & White (1997) that when the step size is as shown in section 3.4, convergence is assured and is optimal. In the next section we provide the pseudo code for the solution of the valuation.

3.5.4. Pseudo Code for the static refinery valuation

A given refinery simulation consists of the following steps:

- Set up the number of stages; ***k = 1 to 72*** (*The maximum possible input is the lifetime of* *the refinery*)

- Repeat the below steps 1000 times; ***for j = 1 to 1000*** (*1000 iterations is enough for convergence for valuation of a real option, but usually not for a financial option*)[49]

- At date 0, build an event tree incorporating each commodity price in time, with 72 decision steps and 21 sequential branches at each node; representing the possible evolutions of the forward curve; (The model can be rerun each day)

  call buildTrinomialTree(j)

- Starting at Maturity, using the terminal boundary condition, optimise the criterion - by deciding on optimal production volumes; (At the terminal date the optimisation is static)

  for t=T to t=0 step -1,

  call maxObjective(j, t, k);

  (*The above function considers all possible paths at each node*)

- Move backwards recursively, deciding the previous iteration's decision variables, until ***date 0*** is reached (*The problem is path dependent*)

- Record all optimal decision variables for the optimal path, and the valuation on each iteration of j; (*Optimal values of:* $\boldsymbol{q_{ji}}$, ***V(i),*** will be obtained)

  Next j

- Plot numerical results;

  end;

  call plotGraphs();

[49] See Albanese (2006) for convergence of a random walk with drift to its continuous-time counterpart

The issue with the above steps is that the valuation is carried out on today's information only and is therefore a deterministic valuation. In the following section we discuss how to convert this to a dynamic valuation.

3.5.5. Pseudo Code for the dynamic refinery valuation

Following Geman and Ohana (2008), we construct the following method to value a dynamic paper refinery: the realised one factor spot process of the refined products differs from the paths described in the event tree as we have to update the optimal strategy at each time step, i.e. we have to re-optimise the decision variables taking into account the new state of the portfolio and information that has become available. To analyse the robustness of our model with respect to this new information we create a simulation which dynamically reproduces strategies under price scenarios which are independent of the ones used for the optimisation. Extending the steps explained in section 3.3.4., under a dynamic setting, the model is re-optimised at each new date *k*, based on new information, ζ_k , equation (3.5.9) becomes:

$$V_i(q_{i-1}, \xi_{[1,i]} \mid \zeta_k) = \max_{q_i \in A_i(q_{i-1}, \xi_i)} [f_i(q_i, \xi_i) + V_{i+1}(q_i, \xi_{[1,i+1]}) \mid \zeta_k] \qquad (3.5.11)$$

A given dynamic optimisation has the following stages:

- At time 1, build an event tree with 72 decision steps and 21 sequential branching at each node representing the possible evolution of the one factor spot price process.
- Optimise the criterion from the terminal date and implement the date 1 optimal decision.

- Simulate the new spot price process at date 2 from equations (ones that have S(t-1)) using the Gaussian vector with zero mean and covariance matrix Σ.
- At time 2, build an event tree whose first node is the information of date 2; this tree has 71 decision steps.
- Optimise the criterion on the remaining horizon and implement the time 2 optimal decision.
- And so on until date 72.

The above procedure is performed 1000 times, leading to 1000 simulations of a six-year refinery management optimisation.

In the table below the data shows the leaves at the end of the trinomial tree and their corresponding wealth and probability; there is uncertainty associated however with these cash flows hence we next show how the risk is considered in this calculation.

The previous studies that consider valuing a refinery consider at most two refined products due to the colossal state space and the issues with optimising over this number of variables. The embedded optionality means the owner can choose an alternative product if another is not performing as expected within the physical constraints. This increases the value of the refinery as losses can be minimised and profits maximised more efficiently. The problem has been avoided up until now due to the curse of dimensionality present; we alleviate this issue by using a solver in GAMS that can optimise a dual or quad core computer whilst leveraging off IBM's CONOPT algorithm. This state space means closed form solutions are rarely available and efficient numerical methods must be applied. The choice of a trinomial tree enables dynamic programming to be implemented ensuring that the owner's alternative decision choices are represented accurately and realistically. A limitation here is that the correlation is not investigated in a stochastic or more complex manner and we leave as an extension the effects on the refinery valuation due to a change in the correlated commodities to others.

*(**Table 3.5.1:** Refinery valuation across a two year period on the trinomial tree final leaves; along with the associated probabilities.*

Leaf Number	Refinery Value ($ millions)	Probability
1	772.6571757	6.06104E-15
2	761.2796757	-9.22067E-14
3	749.9021757	6.83084E-13
4	738.5246757	-3.26119E-12
5	727.1471757	1.12376E-11
6	715.7696757	-2.96884E-11
7	704.3921757	6.2423E-11
8	693.0146757	-1.07022E-10
9	681.6371757	1.52317E-10
10	670.2596757	-1.81584E-10
11	658.8821757	1.85415E-10
12	647.5046757	-1.49307E-10
13	636.1271757	1.81068E-10
14	624.7496757	4.61467E-10
15	613.3721757	5.3913E-09
16	601.9946757	6.54951E-08
17	590.6171757	9.73929E-07
18	579.2396757	1.63625E-05
19	567.8621757	0.00026025
20	556.4846757	0.003082053
21	545.1071757	0.022675984
22	533.7296757	0.095692795
23	522.3521757	0.226986898
24	510.9746757	0.302569225
25	499.5971757	0.226986898
26	488.2196757	0.095692795
27	476.8421757	0.022675984
28	465.4646757	0.003082053
29	454.0871757	0.00026025
30	442.7096757	1.63625E-05
31	431.3321757	9.73929E-07
32	419.9546757	6.54951E-08
33	408.5771757	5.3913E-09
34	397.1996757	4.61467E-10
35	385.8221757	1.81068E-10
36	374.4446757	-1.49307E-10
37	363.0671757	1.85415E-10
38	351.6896757	-1.81584E-10
39	340.3121757	1.52317E-10
40	328.9346757	-1.07022E-10
41	317.5571757	6.2423E-11
42	306.1796757	-2.96884E-11
43	294.8021757	1.12376E-11
44	283.4246757	-3.26119E-12
45	272.0471757	6.83084E-13
46	260.6696757	-9.22067E-14
47	249.2921757	6.06104E-15

3.5.1. Risk

There are many issues with value at risk but our aim here is to capture some level of risk for investing in the alternative refined products in the volumes claimed maximising by the optimisation.

In table 3.6 where we examine the valuation and computation time for a different number of paths for the refinery simulation we also include a calculation of the value at risk to the owner of the refinery. This is how much value we expect to lose in a worst case scenario over one day with 90% confidence - this is calculated using a Monte Carlo simulation and finding the 90th percentile of final wealth of the refinery. Since we are in an incomplete market and portfolio replication is impossible our valuation is simply a range of values over the leaves of our tree; implying a risk-neutral assumption on the stochastic equations that are representing the returns series of the commodities. Since we are not attempting to predict values this is considered a reasonable approach; dynamic programming along a trinomial tree is considered a realistic approach to obtain a valuation for an illiquid real asset. This is because of the state space that can be captured with seven correlated trees and the embedded optionality calculated using dynamic programming. There are many examples of real option valuations in an incomplete market evident within the literature but not for a single oil refinery; see Rollins and Insely (2003) for an optimal harvest comparison. Here the authors construct a Markov decision process over a two factor simulation. They claim ignoring the optionality present to management will result in incorrect forest valuation as we find with the refinery. The option to stop, change or increase production is missed by static valuation methods. The refinery is a similar problem in that the exercise price is the operational and management costs of the refinery; embedded in the refinery is the opportunity to refine at the optimum time based on refined product volumes and prices (despite the returns being modelled), as well as the option to abandon production or delay if the refined product prices are too low. Our problem is comparable in that it is essentially an explicit finite difference approach.

Despite the limitation of using a constant correlation matrix to generate the trinomial trees the fact that correlation would fluctuate giving higher and lower valuations over time; the valuation is capturing the alternative choices available to owner realistically and enabling the numerical method to be solved in a logical manner. If the correlation of the refined products were to increase the valuation would decrease and vice versa but we leave as an extension a correlation study.

3.6. DYNAMMIC PROGRAMMING USING NODES

In terms of valuing an asset, a multi-period solution is required. Multi-period valuation for an oil refinery has not been accomplished within the literature. For the optimisation to be classified as dynamic, the decisions at each point in time must have an effect on the next time period's available decision set, and a solution must contain a set for the entire planning time horizon. This is captured by the equation (3.5.4) in nodal form:

$$L_0 = L_{init}, \quad L_n = L_{parent(n)} + q_{n1}^{buy} - \sum_{j=2}^{7} q_n^{j,sell} \quad \forall n \in N, \quad 2 \le j \le 7 \qquad (3.5.4)$$

This equation ensures that the volume of liquid hydrocarbons held at the refinery complex across time is managed within physical limits. This can be regarded as the state equation for the refinery.

The refiner is assumed to be a price taker, i.e., not able to influence the crude or refined products market price, which is exogenous to the model. We do not consider hedging the producer's portfolio by utilising Futures contracts. We now define the nodal formulation, essential to solving the problem with a relevant non linear solver.

3.6.1.0. Set notation for optimisation

3.6.1.1. Decision Variables

- q_{n1} : The amount of crude purchased at decision node n
- q_{nj} : The amount of refined product j sold at decision node n

- CW_n : The cumulative wealth at node n
- $E[\pi]$: Expected profit for the entire planning period
- π_n : profit for a particular node

3.6.1.2. Sets and Indices

- N: Set containing indexed nodes *n* of the event tree
- Parent(n): Index of predecessor node to n
- t(n) : Index of time period that corresponds to node n
- N_{end} : Set containing all nodes that correspond to the last period
- k(t, n): Set containing all the stages corresponding to a node and a time period

3.6.1.3. Parameters

- p_n : probability of being at node n
- L_{max} : Maximum refinery hydrocarbon capacity in tons
- L_{min} : Minimum refinery hydrocarbon capacity in tons
- L_{end} : hydrocarbon capacity level required at the end horizon
- L_{start} : hydrocarbon capacity level at the beginning of the time horizon
- T_t : Years from now until period t
- F_{nj} : The forward price of the refined product j sold at decision node n
- S_{n1} : The price of crude purchased at decision node n
- pr_n : The production rate per month in days (refinery run rate)

3.6.2. Expected profit maximisation

The nodal mathematical formulation to maximise expected profit is given by:

$$\underset{q}{Maximise}$$

$$E[\pi] = \sum_{n \in N} p_n \left[(1+r)^{-T_{t(n)}} [\sum_{j>1}^{7} F_{j,n} q_{j,n}] - S_{1,n} q_{1,n} \right], \tag{3.6.1}$$

3.6.3. Maximisation of dynamic objective function

$$\underset{q}{Maximise}$$

$$\sum_{n \in N_{END}} p_n \left[(1+r)^{-T_{t(n)}} [\sum_{j>1}^{7} F_{j,n} q_{j,n}] - S_{1,n} q_{1,n} \right] \tag{3.6.2}$$

Subject to

$L_n = L_{parent(n)} + q_{n1}^{buy} - \sum q_{nj}^{sell}$ $\quad \forall n \in N \quad 2 \leq j \leq 7 \quad n \neq n_1$ (3.6.3)

$L_n = L_{end,\, n}$ $\quad \forall n \in N_{end}$ (3.6.4)

$L_{n1} = L_{start}$ $\quad \forall n \in N$ (3.6.5)

$L_{min} \leq L_n \leq L_{max}$ $\quad \forall n \in N$ (3.6.6)

$0 \leq S_n$ $\quad \forall n \in N$ (3.6.7)

$0 \leq F_{nj}$ $\quad \forall n \in N \quad 2 \leq j \leq 7$ (3.6.8)

(All other required constraints for each product are given in sections 3.2.4.1 - 3.2.4.4)

3.6.4. Solving the multi-period refinery portfolio model

We adopt a discrete time setting, with a finite horizon. Decision variables corresponding to period, θ_i, are subject to constraints based on mass balances and capacity constraints, see section 3.2.4. for details. We consider on the tree, spot and future prices of crude to buy and refined products to sell, where $S_t = F(t,t)$. The crude price is known today and the saleable product prices are not – hence obtained from the simulated forward curve.

Remark: Cash-flows due to forward trading are in this paper registered at transaction date and discounted from delivery date at the risk free rate *r*. We adopt this idea because we want cash-flows at dates θ_i, to only depend on date θ_i decisions and not on previous ones, as would be the case if cash-flows from forward transactions were registered at delivery date. This only has an effect on the final wealth situation. We also assume that we have liquid spot and one month ahead forward markets; prices then remain the same for refining over the month in question.

The decision maker's optimisation problem as given in equation (3.5.9) was:

$$J_i(x_i) = \underset{(q_k)_{k \geq t} \in A(x_i)}{Max} \{\rho_{i+1|\xi[1,i]}[G_i(\underline{\underline{q_i}}) + \beta_i(E_i(J_{i+1}(x_{i+1})))]\} \tag{3.5.9}$$

To solve the problem above, the literature commonly uses a procedural function based Bellman approach - even if recursion is applied, procedural optimisation is much slower than the computation applied here. We use a node based approximation along with a Bellman recursive formulation starting at the end of the scenario tree. The set of prices on the scenario tree represent the information on the monthly commodity spot and forward prices. For multi-period programming this is the gruelling part of the calculation. Obtaining a realistic *discretised* process on a tree is a wily task. Each path from the root to a leaf of the tree corresponds to one scenario, over which the multi-period optimisation can be calculated. The stochastic model is constructed in terms of the nodes {1,..., n, …., N} on the scenario tree and the tree structure is described by giving each node the probability p_n, $1 \leq n \leq N$. The planning horizon is divided into ***k*** stages, to relax the computational onus (a number of periods can be aggregated into a stage), where each stage is k, $1 \leq k \leq K$, and is associated with the time horizon of the real asset, T_k and to the associated set of nodes N_k. Note that the model can be extended to varying time lengths and is optimised over nodes per period. It is assumed to begin with, that we have a few stages and at maximum, that there are as many stages as there are months, as it takes approximately one month for a barrel of crude to arrive, be refined and ready to ship to the client. We assume that the initial storage of hydrocarbon liquid is the same at the end of the horizon.

The decision variables $q_{i,n}$ denote the net positions. The spot price of crude is denoted by $S_{1,n}$, and the forward prices of the refined products are $F_{i,n}$, these are the stochastic parameters within the formulation. The objective function representing the producer's profit and financial risk consisted of product sales minus crude oil costs and a term representing a composite risk measure on profits. Assuming that the working and operational costs of the refinery were not significantly impacted by the stochastic parameters.

The nodal model was constructed with r, the risk-free interest rate, *CW(n),* the cumulative wealth at node n, and we also run the model with, *U(CW)* an increasing concave utility function of wealth describing the producer's risk aversion. Due to the tree having a trinomial structure, it can be imagined that there are three realisations or states of the economy at each node. The model constructed finds values of the decision variables $q_{i,n}$, using: the stages set - $k \varepsilon K$, the products set - $j \epsilon J$, the time period set - $\theta_i \epsilon T_k$, the node set - $n \epsilon N_k$, and finally, p_n ε $[0,1]$ – the probability set.

3.6.5. Maximising of final wealth

One can also maximise the utility of final wealth, instead of expected profit:

$$Max \sum_{n \in N_{END}} p_n CW_n \qquad (3.6.10)$$

Subject to:

$$CW_n = CW_{Parent(n)} + (1+r)^{-T_{t(n)}} [\sum_{i>1} F_{i,\,n} q_{i,n}] - S_{1,n} q_{1,n} \qquad (3.6.11)$$

(And all constraints as those described in equations 3.6.3 – 3.6.8, and sections 3.2.4.1 - 3.2.4.4)

CW_n represents the discounted present value (PV) of cumulative wealth at each node n, given in the set of end nodes. The cumulative wealth is used at the leaf nodes; hence a simple sum is required over the nodes that belong to the last stage ($n \varepsilon N_{end}$). To some this will seem unorthodox but it is equivalent to the calculation being carried out along all scenarios from root to leaf.

The final wealth of the refinery CW_T, represents the expected profit by summing up the nodal profits, weighted by their nodal probabilities p_n. Each cash flow is discounted by the interest rate r, obtained from data on the current Indian yield curve[50]; we choose an end of horizon of six years for comparison reasons but also value up to ten years to find the computational time at this horizon - the term structure enables us to discount the relevant cash flows with a realistic discount factor. Additionally, each node has a particular conditional probability, an assigned crude price and an assigned set of forward product prices for its particular period $\theta_i(n)$.

The initial wealth at the start of the tree is zero:

$$CW_{0,\,0} = 0; \qquad (3.6.12)$$

---

50 Bloomberg: India Govt Bond Generic Bid Yield 10 Year, GIND10YR:IND, 7.25 0.06 0.82%

And non-negativity constraints are present:

$$q_{i,n} \geq 0, \qquad S_{i,n} \geq 0, \qquad t \geq 0, \qquad F_{i,n} \geq 0, \tag{3.6.13}$$

Apart from the first node, each node has a parent function pointing to its predecessor node. It is assumed that the operator can sell as much as produced of each refined product onto the oil market at each time period or each node as can physically be stored. Moreover, the other physical constraints of the refinery system are respected at each node (please see section 3.2.4. for details).

The objective function in code can be set to represent the expected utility of the final wealth or just the expected profit with composite risk as shown in equation (3.5.9) over the planning horizon.

To solve the original problem given in equation (3.5.9) we apply backward induction. Analytical backwards induction is applicable to a wide variety of situations, but can be infeasible if there are too many state variables or too many time periods. Therefore, we solve by a recursive calculation along the set of nodes, which are *mapped* to time periods starting at the end horizon; this is visually evident from the trinomial tree. There is also a stage mapping; these mapping constructs are the main contribution of the numerical approach designed for the refinery. From Moro & Pinto (2000) to Grossman (1998) to many other refinery optimisations; most computational or numerical procedures either do not state how they solve their objectives in enough computational detail within the state space or simply make approximations and solve using procedural programs; this is very inefficient. The approach utilised in this chapter is named dynamic set assignment, enabling the problem to be solved, when mapping with the stages in O(n) time as opposed to O(2^n), and it will replace many procedural based optimisers in the future; enabling larger and more complex state space optimisations to be solved.

3.7. Numerical Results

In this section we discuss the numerical results obtained by solving the multi-period stochastic model. Technically, the optimisation of this decision process contingent on the future events is more parsimonious when "reasoning under uncertainty" can be decoupled from the optimisation process itself. This occurs when the probability distributions describing future events are not influenced by the decisions selected by the agent; uncertainty is exogenous to the decision process. The idea is to exploit the finite scenario tree approximations in order to extract a good decision policy for the correlated continuous commodity distributions.

The simulation framework is based on GAMS release 23.9.1 x86, for modelling and solving the optimisation problem by the non-linear optimisation package, CONOPT3. This optimisation problem is solved using mapping objects or in dynamic set assignment. Suppose for example, that there is a simple map object, m, which maps each value of a function that has already been calculated to its own result, and then the function is modified to use it and correspondingly updated. The resulting function requires only O(n) time instead of exponential time – *in recursive operations without this map, exponential time is standard*. This technique of storing values already found is called *memoization* (1968). The bottom-up approach was applied where smaller values of the problem were found first and the larger values built from them. It takes constant O(1) space, in contrast to a Bellman equation solved from the top-down approach which requires O(n) space to store the map. We have used this methodology to obtain solutions to the numerical problem presented – this enabled a valuation in reasonable time.

In the optimisation procedure the tree of prices must firstly be constructed; generating the nodes and their values takes the longest fraction of solution time. C++ carries out all set assignments: time set, realisations set and probabilities; generating the files containing a seven-dimensional correlated scenario tree of discretised prices for input into GAMS. All trees also utilise a common computational data structure named simply as a *tree* structure. This enables the computation to gain more time as memory is utilised more efficiently. Due to the dynamic set assignment the problem becomes very similar to matrix multiplication in terms calculating numbers per map - essentially the algorithm is closer to O(log(n)) when summing the values along the tree.

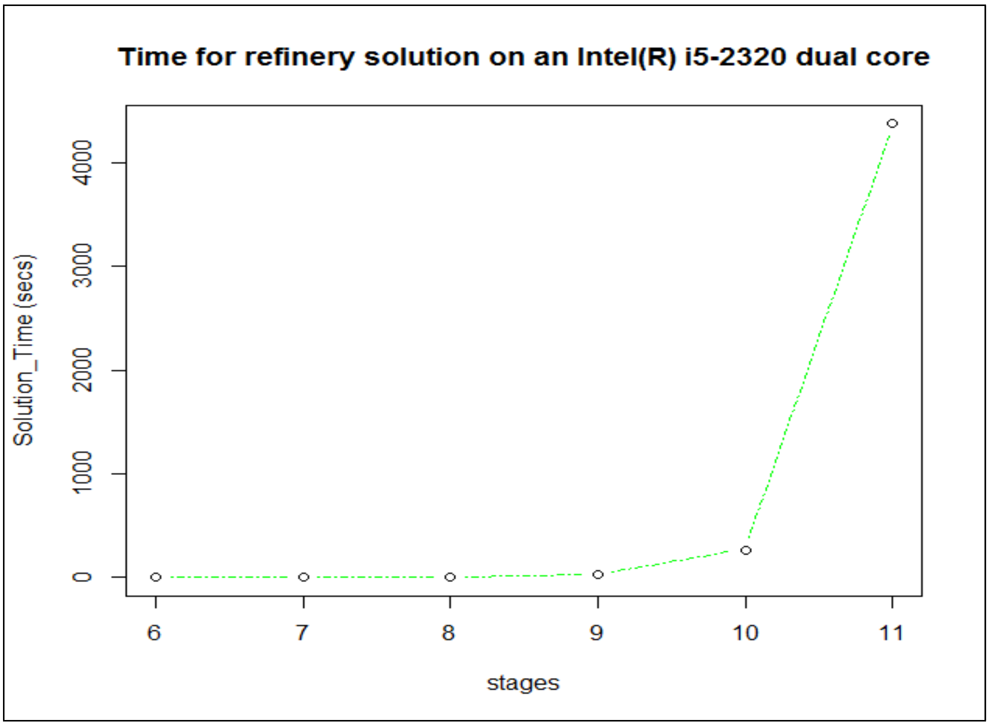
For example, due to the way the problem is constructed in GAMS, one year can be initially represented by four stages, which is equivalent to 360 days. This ability to defragment the problem into stages is another uniqueness to the numerical procedure; due to the mapping the program knows which stage belongs to which time, product price and all other set variables at any point in time. For instance, with four stages the time sets entered into the code can be represented by: $T_1 = \{t: 1 \leq t \leq 90\}$, $T_2 = \{t: 91 \leq t \leq 180\}$, $T_3 = \{t: 181 \leq t \leq 270\}$, and $T_4 = \{t: 271 \leq t \leq 360\}$. For four stages the nodes entered into code are the following sets: $N_1 = 1$, $N_2 = \{2,..., 4\}$, $N_3 = \{5 ,..., 9\}$, $N_4 = \{15,..., 19\}$. In the four stage one year tree, there are 19 nodes and 27 scenarios. In the final model implemented to value the refinery effectively, an 11 stage tree was constructed and calculated; in the library of solutions the maximum k used was 72 - taking days, an extremely long time to compute. With a reasonable maximum of k =11, it is possible to have a comparison to the static version. For 11 stages there were 59,049 scenarios, taking four hours and 22 minutes to solve on an Intel(R) Pentium M 2.00 *GHz* processor. It took approximately 73 minutes to solve the same problem on an *Intel(R) Core(TM) i5-2320 CPU@3.00 GHz dual core*. The structure of the tree reduces the states space and ensures the solution is tractable in both cases. This is evident from the number of scenarios and time taken to solve the model. All models contain prices that were simulated 1000 times to obtain an average representation of the uncertain commodity price set.

The model was solved by considering the sources of variability of the spot and refined product prices. The price dynamics were captured using the mean-reverting spot price processes. In terms of value over and above the intrinsic value we notice an increase in the total expected value of the refinery at time horizon T. Overall, the increase in the value of the objective function went from, \$560 million to, \$1131 million with optionality, over a six year period. Results are shown in Table 3.6 below:

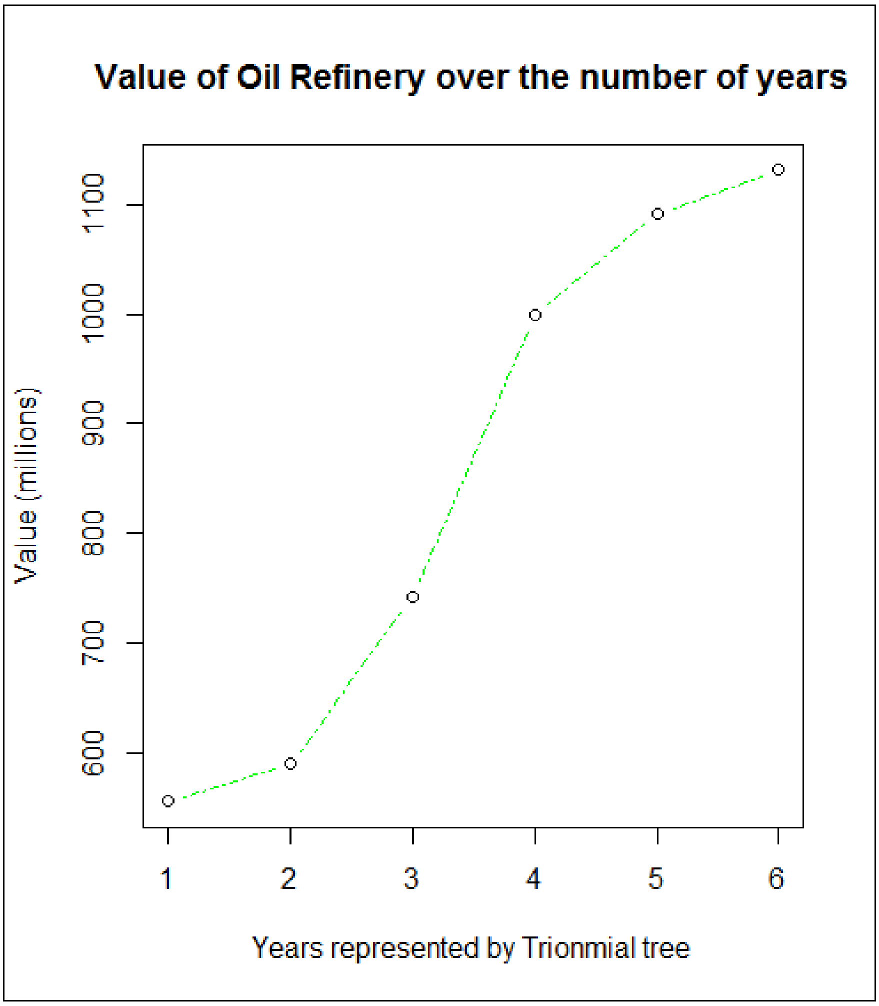
Number of Stages (k)	**Number of Scenarios** (root to leaf)	**Refinery Value** $ millions	**Time to Solve on an Intel(R) i5-2320 dual core Processor** (*seconds*)	**Risk** (VaR for whole period)	**Approx. Operational Years Represented by tree (T)**
6	243	556	0.616	78.45	1
7	729	590	1.257	83.20	2
8	2,187	743	4.119	99.08	3
9	6,561	999	34.761	109.12	4
10	19,683	1,091	267.793	124.66	5
11	59,049	1,131	4,375.453	136.80	6

*(**Table 3.6:** Summary of stochastic trinomial tree optimisation of a Topping refinery with risk values, including the number of scenarios, and the CPU solution times) *All values were calculated 1000 times and then averaged.*

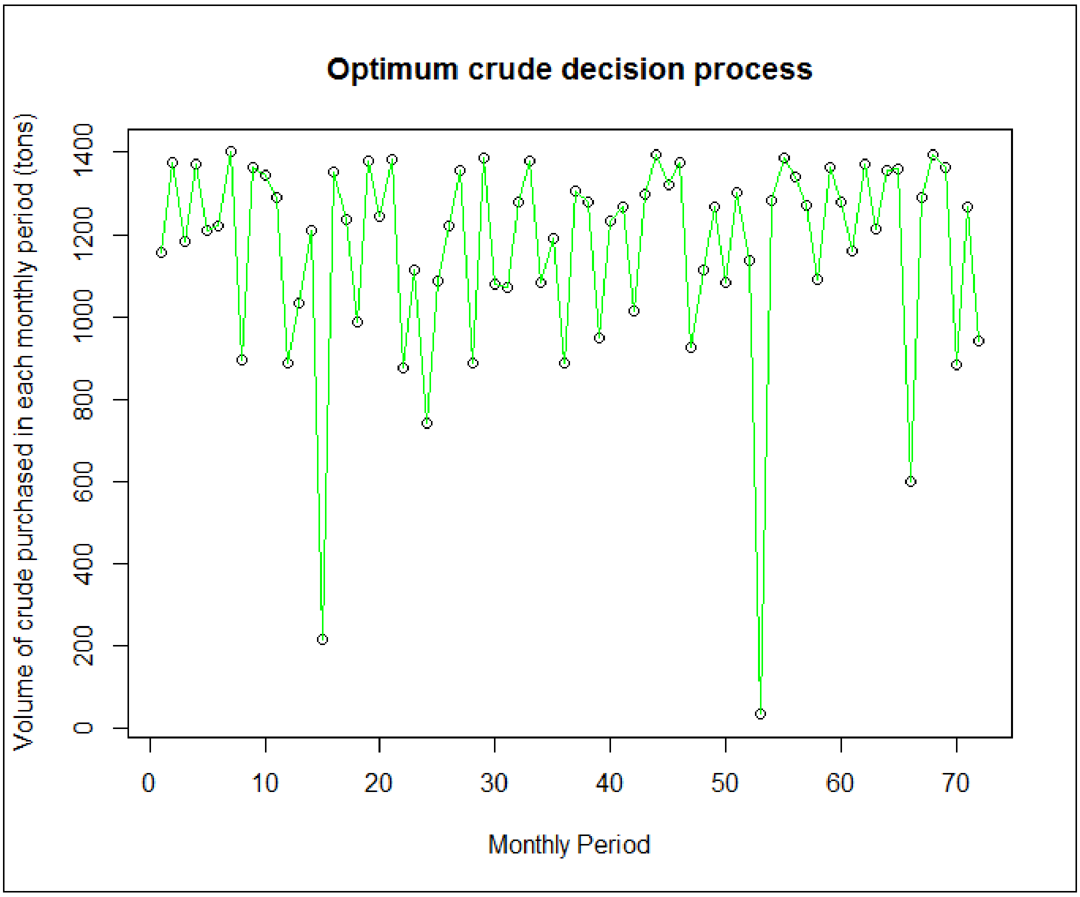
3.8. Graphical results



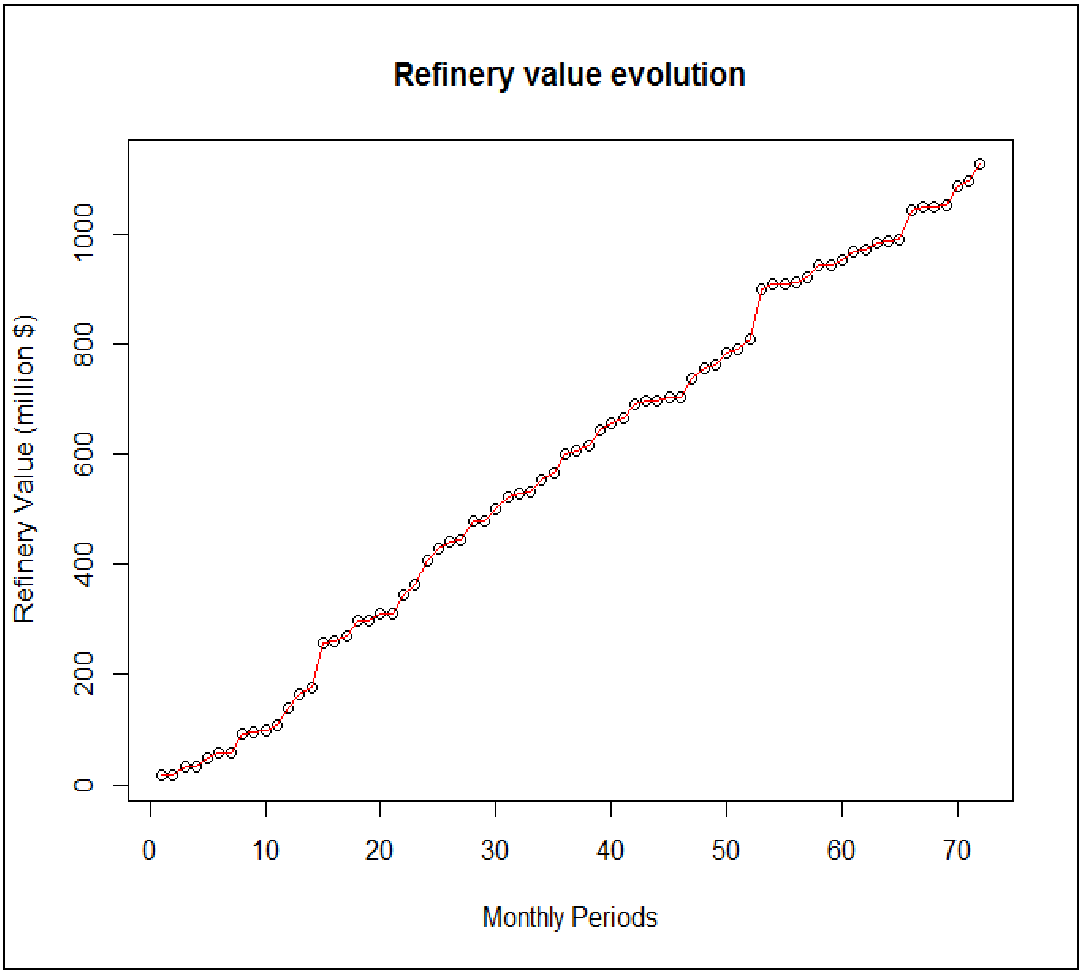
*(**Figure 3.20:** As the number of stages input into the program is increased the solution time increases in a non linear way)*



***(Figure 3.21:** As the number of years increases towards six the value begins to stabilise)*



*(**Figure 3.22:** The refinery owner chooses an amount of crude to refine at each monthly period over the lifetime of the refinery)*



*(**Figure 3.23:** At different periods the valuation increases more than others as the refinery manager optimises the decisions along the trinomial tree)*

The applicability of a portfolio optimisation approach to a real world problem does not only rely upon numerical tractability, but crucially also on the robustness of the model. Deviations in input data should not have major consequences on optimisation results; see Zhu and Fukushima (2006) for an analysis. We ran the optimisation at 100 to 100,000 times; the averages and standard deviations of the decision variable results of 1,000 simulations are shown in table 3.7 below:

3.9. Out of sample stability analysis

	Crude	Gasoline	Naphtha	Fuel Oil	Jet Fuel	Heating Oil	Feed
Average optimal portfolio	1350	2700	900	4210	2300	1650	135
Standard deviation in optimal allocation	76	161	88	191	206	113	9

***(Table 3.7:** Stability associated with scenario generation procedure decision variables. *Trees were generated and optimised 1000 times)*

In general, finite difference schemes offer more flexibility than say binomial trees when handling complicated discretisation problems. We use a trinomial tree, however, as with other explicit methods the approach suffers from stability constraints that restrict the time step size used in the numerical solution, and can in general extend the solution time. We have, however, shown that our method solves in a reasonable amount of time despite its explicit grid nature, this is due to the solving technique of *memoization* and the other computational constructs described in the previous section.

3.10 Analysis of the Valuation Model

The main problem with the set up is the fact that we are using a one factor model – this is not representative of the behaviour of refined product series. For example Koekebakker and Ollmar (2005) use ten factors to capture Electricity prices on Nordpool. However, the stability and speed with which the results are obtained supports the choice of a one factor model. In practice if the 11 stage optimisation was considered as too laborious the user can choose to drop down to a suitable time – they would miss some intrinsic value but there may be valid reasons for this shortcut. An interesting extension would be to use different stochastic processes and investigate how this would affect the valuation of the refinery. A two factor model would capture more realistic behaviour in the price series but strain the computational side of the solution. The separation of the simulated paths and the dynamic recursion is the advantage of this approach; hence mean reversion is captured and value extracted from the spot price processes.

Another piece that has not been addressed is the hedging calculation; this would entail creating a representative portfolio of the refinery form market products that are indicative of its value. This is difficult in our case as there is no accessible liquid market for financial contracts within India that correlates with our simulated prices. An approximation could be to use those contracts on CBOT that are representative of the required product – we discuss this and the other extensions in Chapter four further.

3.11. CONCLUSION

In this chapter we have introduced a one factor stochastic model for monthly midterm production planning of an oil refinery – enabling a valuation of the real asset to be obtained. We have shown a unique approach for solving a real-option valuation for an oil refinery that is consistent with the assumptions of the risk-neutral pricing of options. The approach is a straightforward and flexible model that can be implemented for other real-option valuations where correlated stochastic variables provide the uncertainty for the real asset. The embedded optionality represents a much higher price than DCF reveals which is inline with many other real option valuations within the literature. A limitation of the approach is that we are in an incomplete markets setting hence a single no arbitrage price is unavailable due to replicable contracts being unavailable. The embedded optionality is however captured by the correlated trinomial trees built so that dynamic programming can be applied. This usually means there is no closed form setting and the numerical approach considered provides a range of values with their associated probabilities. The model is a consistent multi-period optimisation, using dynamic programming to solve from maturity backwards until today. The profit comes from the direct production of the refined products being sold instantly onto the spot market, where the risk is the volatility of the commodity market prices but can also be set in code to be a composite risk function. We considered as the stochastic variables in the model, the spot prices of crude oil and the forward refined products prices. The method used to generate a finite horizon multi-period value for the refinery is unique; stochastic prices are modelled using a trinomial tree, which captures the idiosyncratic features of petroleum related products. The contribution in this chapter has been to obtain a real option valuation of a topping oil refinery; enabling a comparison to a static value to be made. To our knowledge there has yet to have been a dynamic programming valuation on an oil refinery in the literature. Another limitation is however that a constant correlation matrix is utilised to obtain the correlated tree of prices - this unrealistic but we leave for others the extension of investigating a stochastic correlation on the refinery valuation. After the scenarios are generated on a trinomial tree, the decision problem is solved using non linear optimisation within GAMS to obtain a finite horizon value of \$1,131 million. The model considered 72 monthly periods, whilst adhering to financial and physical constraints at each point in time, combined with a simple intuitive decision rule by repeatedly maximising the intrinsic value of the refinery. Our model captures more of the true management flexibility than DCF analysis, which gives a value of \$556 million, and to our knowledge is the only dynamic real-option valuation of an oil refinery within the literature using computational advantages. The contribution is three-fold; firstly, we have used a nodal representation of the tree structure within code – meaning that the data structures enable memorization, and a very quick solution time for this state space. Secondly, we have used a one factor SDE from Schwartz (1997) that is extremely tractable and can manage shifts of the forward curve. Thirdly, to our knowledge this is the only asset valuation approach for the pricing of an entrepreneurial refinery using a Bellman equation, with all seven refined product stochastic processes being represented along the nodes of the tree, that applies dynamic set assignment ensuring that the mapping of the stages is completed in O(n) time and not exponential time.

An interesting extension would be to run the optimisation with a different set of generation processes for the stochastic variables along the tree; an alternative formulation for uncertainty, and analysing how this impacted the valuation.

Chapter 4

A CRACK SPREAD OPTION REFINERY VALUATION

"*We are all in the gutter, but some of us are looking at the stars.*"

Oscar Wilde

In this chapter we develop a valuation methodology for an oil refinery that is tractable, solves in reasonable time and captures intrinsic and extrinsic value. We make standard no arbitrage assumptions, and use European call options to suggest that a refinery asset is equivalent to a strip of daily options over its time horizon subject to physical and flow constraints. Using linear program techniques and GAMS we incorporate stochastic equations on the main commodity prices and apply dynamic programming to obtain a value for today that is unique locally and globally. We produce a static and dynamic optimisation that can be altered to incorporate additions to the stochastic model or applied to other real assets that are reliant upon a spread based valuation. There are two major benefits of our calculations: one, using option valuation techniques for the commodity price we have provided a long term valuation that solves in seconds and is realistic in comparison to actual option prices, secondly we have developed an optimisation decision volume set, that is more beneficial than a 'go with the flow' refinery decision set resulting in economically improved decisions for a refinery owner.

INTRODUCTION

Auditors, Investment bankers, and entrepreneurial investors all need to value oil refineries and we have created an extensive numerical approach in the previous chapter. This calculation is however expensive in terms of time, CPU and memory usage. The motivation for this chapter is that a closed form solution for pricing a real asset where optionality exists, with the assumption that the underlying prices follow correlated geometric Brownian motions, does not exist, see Eydeland and Geman (1998) for a thorough description – additionally, the valuation is required in a timely manner. Many practitioners use closed form approximations as they are fast and accurate enough for estimating values for spreads whether within trading institutions or hedging at a refinery itself, but in terms of a valuation, we require an accurate calculation that captures as much of the crack spread value as possible. Thus we want to produce a numerical approach that will provide a financial valuation of the refinery complex quickly enough for the various relevant professionals. We repeat here for clarity that an average refinery owner will have experience of the oil market in his or her's region and usually make decisions based on their experience in addition to information generated from a computer model. Many of these simulations produced by professional refinery software companies such as ASPEN TECH or Honeywell, take many hours to solve and do not consider the full dynamics of the underlyings over the complete time horizon due to the computational and dimensional hurdles. Instead the LP program managers are concentrating on optimising the mass balance flows or a different problem altogether, for example, supply chain optimisation. If the oil market is in a steady and low volatility state, a refiner owner can "go with the flow" and use momentum indications of which product to sell more of – this does not consider the knock on effects to the other refined products and a full linear program solution allows a more statistically and consistently correct decision set to be made. Over time, a computerised decision set will outperform the owners and one that solves in a timely manner is more likely to be applied in practise; see chapter three for a detailed analysis.

In this chapter we propose a tractable method based on the crack spread and compare it to existing spread valuation methods. We make standard no arbitrage assumptions, and use European call options to suggest that a refinery asset is a strip of daily options over its time horizon subject to the same physical and flow constraints described in the previous chapter. We are assuming here that owning the refinery complex is equivalent to the value that can be extracted from the volumetric comparison in crack spread option contracts. We repeat here for clarity that the owner has a decision set to execute on each day whether to refine or not and how much of each refined product to produce at the end of an assay run. This chapter attempts to achieve what chapter three does in a faster time by not building a simulation with trinomial trees and by not using recursion to capture the optionality. Instead we replace this huge state space with option contracts - a simplified approach that will reduce the realism, but also simplify the computation. The choices of which methods to use to price spread option contracts follows logically from the literature where the following methods are either practically used more frequently or are regarded in the literature as the go to method for spread option pricing. Bachelier, Kirk and Alexander and Venkatramanan are described extensively in most spread option analyses included in the excellent and complete literature review of spread option pricing in Carmona and Durrelman. The method used in this chapter is trivial but very effective; the refinery optimisation framework from chapter three remains but the simulation and trees are removed and replaced with option contracts.

We first use Bachelier's method assuming an arithmetic Brownian Motion as it allows a closed form formulae, on which we can base our numerical method. It does not include mean reversion and has other practical issues, but is still used by traders to this day on the commodity trading floor, next to Kirk's as estimators of options on commodity spreads. Hence, we include and compare it to Kirk's method, the very efficient Alexander and Vankatramanan approximation procedure and the famous two factor Schwartz and Smith (2000) commodity based model used to calculate option prices. We compare our method to these four methods, in addition to actual option prices within a linear program to determine the effect on the refinery valuation. We compare our procedure, which we call the *CSORV* or crack spread oil refinery valuation method, to the four other valuations by solving the existing Linear Program and replacing the crack spread option contract in the objective function with the relevant methods in the literature, including: Bachelier (1900), BA, Kirk (1995), KI, Alexander and Venkatramanan (2011), AV, and Schwartz and Smith (2000), SS. The first three methods: BA, KI and AV are considered closed form results, SS and CSORV are numerical models. Kirk's approximation is ubiquitous on trading floors throughout the commodity sector, but it is restricted to low strikes. Bachalier's, BA, is a useful approximation for many spread options being trivial to calculate and assuming a log normal distribution for the underlying spread – this is also a disadvantage. Both BA and KI suffer from assuming that the underlying prices that make up the spread are bivariate lognormal distributions, which is quite unrealistic for most commodity price series. AV's is essentially closed form after the optimisation and calibration of the strike convention, whereas SS is used frequently on commodities but requires calibration and then simulation. The CSORV is numerical but still solves in rapid time and is more accurate than the above four approaches – this is due to the stochastic equation capturing more of the mean reversion over the time period, and using the same strike convention as in AV's paper, enhancing its accuracy. A drawback is that stochastic volatility and jumps are not considered. Correlation frowns rather than smiles are a constant feature of the crack spread option market; our method's aim is to value real assets, and our focus is not to capture this feature as AV11's and SS's both do.

For our purposes of valuation constant correlation is a necessary simplification due to the computation and enables a comparison to be made to closed form and numerical spread methods. To our knowledge the method in this chapter is the first attempt at valuing an oil refinery using an optimisation augmented with a discretised and simplified stochastic price set. In this chapter we have a valuation methodology for an oil refinery that is tractable, solves in reasonable time and captures intrinsic and extrinsic value. Using linear program techniques and GAMS, we incorporate stochastic equations for the main commodity prices and apply dynamic programming to obtain a value for today that is unique locally and globally and solves in a timely manner.

The formulation and modelling of real options that accurately incorporate managerial flexibility are unlikely to result in closed form solutions; hence the vast number of complex numerical algorithms available in the refinery literature, see Elkamel et al (2011) for an in-depth summary. In the previous chapter we have constructed a numerical method that approximates the cash flows generated by the refinery asset under relevant constraints. The initial choice of selecting all seven dimensions to represent the commodity prices captured all state variable value; we now choose to reduce this to three state variables. The most significant reasons for investigating more efficient solutions are that practitioners are unlikely to apply a method that is not nearly instantaneous – especially on the commodity trade floor. This would be relevant if for example a commodity desk wished to hedge or speculate on oil refinery asset outcomes. In terms of financial valuation; a timely manner is not as high a priority. However, obtaining a fair valuation is fraught with issues; the three main difficulties are:

1 - Which stochastic equations truly represent the commodity price series realistically?

2 - How to reduce the dimensionality of the problem to make the calculation solvable?

3 - What is the most efficient way to depict the refinery asset problem within a mathematical construct?

The method constructed in this chapter is reducing the number of dimensions, this tremendously speeds up the valuation and enables other avenues of interest to be pursued more easily, for instance, the associated risk measures can be investigated, see appendix for a publication including a risk optimised objective function. Opportunities to the owner of the refinery can be seen as embedded optionality, but capturing and simplifying these decisions enables the valuation to be tractable and realistic. The core underlying value of the refinery is the crack spread. The crack spread as defined in the previous chapter is the primary difference between the purchase and sale of refinery products; the main cost being crude oil and the two most significant selling products are the heating oil and the gasoline. The crack spread is the primary risk for the refiner, and at the time of deciding which products to refine, the prices of the outputs are unknown. The production is not taking place instantaneously; the time at which purchasing and selling decisions are made cannot be successfully made without being computer aided due to the sheer number of factors involved.

Once committed to refining the recently purchased crude, the owner is locked in to producing a set volume of hydrocarbon outputs one month from the cracking initiation. The next day represents a new decision timeline, regardless of the fact that the process began for a batch the day before. In trading petroleum related contracts listed on an exchange that manage the differential risk between crude and its products, the refiner can lock in the margin to ensure that either long term contractual obligations are met, or that the crack risk is hedged effectively along the horizon. Decisions made in a timely manner make or break the refinery asset, if a hedge is not pursued before cracking and the spread heads southwards – profits for that assay run will be a disaster. The New York Mercantile Exchange, NYMEX, offer a number of contracts on refinery spreads that are purchased frequently by refineries throughout the US; we use these despite this being less realistic for a refinery based in India. In making these complex nested decisions the owner must be well versed in the fundamental and technical effects that impact the values of the various commodity prices. Table 4.1 below, describes some of the factors that practitioners consider when deciding how to manage a hedge or a speculative trade on the crack spread's value along their decision timeline. We organise the chapter as follows: firstly we introduce crack spread options and discuss their nuances. In section 2.1 we introduce the various methods implemented to value the 3-2-1 gulf coast contract. In section 2.2 we discuss the data, and in section 2.3 explain the static valuation. In 2.4 we present the various methods chosen to price a crack spread option; we discuss the stochastic processes choices in section 2.5, whilst in section 3 we analyse the results. In section 4 we compare all four methods to ours, and in section 5 we conclude.

4.1.1 Crack spread effects

There are many different contracts that refinery owners can utilise for hedging and speculation purposes – capturing the spread's value. This value is often termed the "paper refinery" by petroleum specialists: it is of interest when marking to market or when implementing investment decisions.

*(**Table 4.1:** Factors affecting the crack spread value, assuming the other refined products remain the same value)*

Concern	Refinery Effects	Crack Spread Reaction
1. External geopolitical issues — politics, geography, demography, economics and foreign policy	Crude oil supply, *Crack Spread decreases initially* — higher crude oil prices relative to refined products	*Crack increases later,* as refineries respond to tighter crude oil supply and reduce product outputs
2. Slower economic growth GDP measures etc.	Decline in refined products demand	Crack weakness (value decreases)
3. Strong sustained product demand	High refinery utilisation	Crack strength (value increases)
4. Environmental regulation on tighter product specifications	Tightening of product supply	Crack strength (value increases)
5. Expiration of trading month	Cash market realities — long or short products	Cracks values can vary due to closing of positions
6. Tax increases after certain date	Increased sales in front of tax deadlines	Crack weakens (decreases) in front of tax deadline and strengthens post deadline
7. Summer seasonality	Increase in gasoline demand	Crack strength (value increases)
8. Winter seasonality	Increase in distillate demand	Crack strength (value increases)
9. Refinery maintenance	Decline in product production	Crack strength (value increases)
10. Currency weakness (In $)	Crude oil price strength (price increases)	Crack weakness (value decreases)

The next stage in assessing the financial worth of the oil refinery was to develop a more tractable real option valuation; chapter three's algorithm cut previous calculation time down to under two hours. In some professional circumstances this would be impractical; interested parties would wish to calculate hedging formulas and the valuation itself, in minutes rather than hours. Valuing assets that rely upon commodities is an actively researched and practicable arena for many reasons, for example, investment, divestment, hedging and speculation considerations. An oil refinery is similar to a chemical plant - a producer with exposure to a differential spread, subject to a set of physical and flow constraints. The real complexity in an option approach is capturing the value the refinery owner has embedded within its complex decision process. The choice value associated with real assets fall into one of the following categories: the option to expand, the option to abandon and the option to defer. The oil refinery owner has the choice to vary *intertemporal* purchasing and production decisions of the crude and refined products, i.e. the option to defer production. The risk-neutral valuation approach, ubiquitous within the financial mathematics literature, can be extended to elicit a real option valuation of land, plant, equipment and assets. This approach does not require risk adjusted discount rates; yet a set of market price parameters must be calibrated. Many investments are underpinned by the uncertainty connected to future prices. Throughout the programming routines developed in this thesis, the commodity state variables or inputs, were defined as sets, see the appendix for an example - these state prices can be extracted explicitly to estimate the risk-neutral stochastic processes involved in the refinery calculation, and avoid the requirement to obtain a market price of risk parameter for the random variables exemplar set. These sets are the standard data structure for computer programs named algebraic modelling systems (e.g. GAMS). The advantages in using GAMS and constructing the problem in this manner is that computational data structures do not require development as GAMS has the required constructs already available in its library - dependent on the problem be solved. We design a program in C++ to generate the prices and probabilities on the a trinomial tree for all methods requiring discretisation, which is then injected into GAMS as a table data structure (essentially a computational trinomial tree structure). GAMS can utilise this as input in set notation. These sets enable group/set mathematics to solve calculations instead of procedural calculations - a much faster and efficient method than common in the literature; it is standard to do this in a procedural way without data structures or applying *memoization*.

After the refinery manager's initial purchase decisions, the knock on effects alter the remaining decision set due to the limitations on the particular crude assay being refined; this is indicative of the path dependency present in the calculation. For example, producing more gasoline means less heating oil, and yet the CDU must maintain the mass balances at all times, therefore certain decisions are not physically possible, e.g. producing all kerosene or no naphtha at all. If for instance, the refiner knew that residual fuel oil would be the highest valued product in one months' time, then as much of this fuel oil as could be physically stored would be processed – this decision can be implemented, but only as far as the capacity and flow constraints allow; a difficult number to generate without inputting this into the LP. The refiner has flexibility available - an embedded option, which is strictly constrained due to the chemical processes that are viable under the complex path dependency present in the refinery. Capturing this mesh of decisions is the aim of any model solution and a numerical approximation should be tractable and remain financially accurate.

4.1.2. Calculating the crack spread value with traded contracts

The dependency of a set of decisions on the next time slot is investigated in this chapter, where we use a real data set of commodity future prices and correlations with a simplification where we reduce the number of free parameters. In other words, the 3-2-1 Gulf Coast crack spread contract is represented on three trinomial trees consolidated into one; instead of requiring seven. This is the amount of value in a barrel of WTI crude, refined into its two most valuable products: gasoline and heating oil. We firstly calculate a very basic option value, attempting to value the refinery with data from NYMEX[51] – the refinery is considered as a strip of 3-2-1 crack spread options over a finite horizon; Geman and Eydeland (1999) apply this concept of a daily strip of options on the spread, but apply it to a virtual power plant, where the output is electricity and the fuel is natural gas. At first the authors apply a very different production approach in a two parameter family to price power options using a power stack function. Our approach is closer to the daily option spread valuation approach used later in their article, where they select a stochastic jump diffusion equation to represent electricity by leveraging Merton's 1976 formula – the authors present the framework, but omit the calculation, which is fraught with practical difficulties for a refinery complex over this length of duration. On each day of activity, if the refining cash margin is positive, then we should refine, else, we should switch the refinery off. We will discuss this method's limitations in the following analysis.

The calculation that underpins the optimisation described later in the chapter is to take the value on an option on gasoline, and the value of an option on heating oil, consolidate them in the correct multiples and thus obtain the value of the crack spread option.

Fully integrated oil majors like Exxon Mobil and Chevron have a natural hedge against the refining margins as they own commodity focussed assets and financial contracts in large capacities, but independents like Valero and Tesoro do not. The corresponding trading of financial contracts at these firms follow alternate strategies. The payoff profile of a refinery *spread* call contract on a crude oil future, (F_{crude}) and *one* refined product, (F_{prod}), with exercise price, K, is given by:

$$\max\ \{(0.42.F_{prod} - F_{crude}) - K, 0\} \qquad (1)$$

(The 0.42 is due to there being 42 gallons in one barrel of crude)

$$\max\{(\text{crack-spread value - refining costs}), 0\} \qquad (2)$$

[51] NYMEX purchased call contracts over ten years from 2000 Dec to Dec 2010

This type of contract is often utilised by the vertically integrated oil majors to speculate, whereas for the independents, it is a hedging instrument. Applying basic microeconomic arguments, the crack spread that is underlying the above contract should be a mean reverting process, despite the numerous studies arguing the opposite, in recent months it is clear that crude has again reverted to an average or lower. The spread itself is usually captured using an American style options contract when listed on a financial exchange, see Carmona and Durrelman (2003) for an extended discussion and summary of spread options related literature. In the next section we obtain descriptive statistics for the underlying value to support the choice of a mean reverting process, and describe the characteristics of the crack spread data to build the argument for simplifying the discretised stochastic process in the way chosen in this chapter.

4.1.3. Descriptive statistics for commodity future price returns

Statistics for daily crude oil, heating oil, and gasoline futures return data are shown below, where we label CR as crude oil, HO as heating oil, and GA as gasoline; each number in the brackets is the maturity month of the contract. There are a total of 3,914 observations from January 2000 to December 2012, and the data is collected whenever a price is available (refineries continue cracking during weekends).

***(Table 4.2:* Descriptive statistics for future price returns)**

*The Jarque-Bera test of normality. * indicates significance at the 1% level.*

Contract	Drift (μ)	Volatility (σ)	σ (annual)	Skewness	Kurtosis	Jarque-Bera
CR(1)	0.0015	0.0240	0.2387	-0.16	5.15	782.09*
CR(2)	0.0015	0.0259	0.2679	-0.27	5.81	1337.08*
CR(4)	0.0014	0.0276	0.2855	-0.29	6.48	2114.90*
HO(1)	0.0015	0.0245	0.2466	-0.02	4.97	652.33*
HO(2)	0.0014	0.0247	0.2804	-0.04	4.71	485.89*
HO(4)	0.0013	0.0281	0.3037	-0.02	4.51	383.76*
GA(1)	0.0016	0.0254	0.2599	0.29	8.51	4872.93*
GA(2)	0.0016	0.0270	0.2860	0.15	7.08	2830.33*
GA(4)	0.0016	0.0287	0.3028	0.08	6.00	1369.97*

We generate a crack-spread price for the ten-year period with the commodity future price data for the prompt month contract collected above.

The most significant and tricky feature to capture in this modelling process is the correlation of the saleable hydrocarbon prices as this tremendously affects the profit differential; missing from the above table. GARCH can represent the volatility clustering present in financial returns, but analytical results in continuous time for spread valuation capturing this behaviour, without large approximations, remain elusive, see Borovkova (2007). Another reason for the issues in calculating spread option values is that we have a weighted sum of variables. In Borovkova (2007) the state variables are assumed log normal, but the sum of log normal variables is not log normal. It is well known that spread option prices calculated via the *Bachelier method*[52] applied by Shimko (1994) for instance are also inaccurate compared to those from a Monte Carlo simulation, which can capture different stochastic underlying processes. To alleviate this issue, Borovkova et al (2007) apply many log-normal distributions for the underlying spread value. This is beneficial as it ensures that the state prices in the calculation do not become negative yet they exhibit negative skewness. In reality modelling the spread or the difference between the two commodities with a single Brownian motion misses a huge chunk of structure that can be represented with alternative multi-factor

[52] The distribution of the spread is assumed normal but gives vastly different numbers for real options

stochastic equations; we pursue an approach to represent the structure of the underlying prices in our method.

4.1.4. Assumptions underlying the crack spread valuation

At any period in time, without purchasing additional financial contracts, the refiner is short crude oil, as it needs buying, and is long the refined products, as the owner has the ability to produce saleable hydrocarbons. Therefore, the refiner is long the crack spread; in practice, traders associated with the refinery will construct trades on the constituents of the crack spread based upon their estimation of the spread being under or overvalued. For example, if a dispersion trader feels the crack spread is overvalued, then it will be sold short and vice versa. The hedging replication argument enables the refinery to be valued as a strip of crack spread options; a comparison can then be made using various existing option calculations.

We create a model in which the set of underlying commodity price dynamics are realistic; the stochastic processes should be mean reverting, and for an in depth modelling approach to be valid, the following important conditions should additionally be met, see Geman (2005):

1. The crack spread option must have a starting date and a maturity T.

2. The underlying prices must be clearly identified: S_1, S_2, S_3 or F_1, F_2, F_3, if using futures.

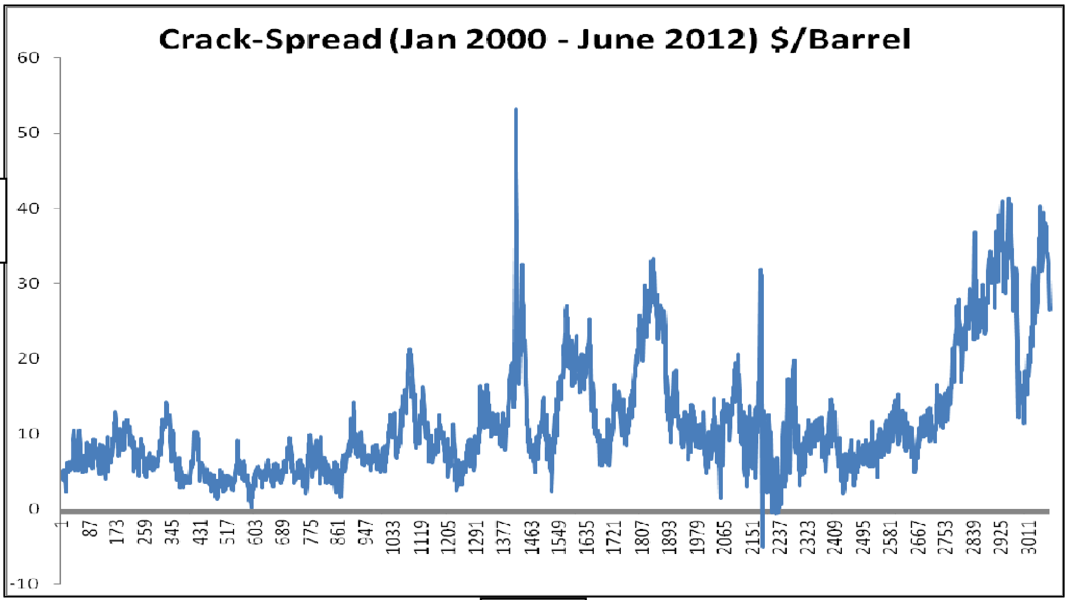
3. The crude, S_1, and refined products S_2 to S_3, must be traded in continuous time in liquid markets, to allow for dynamic hedging, the cornerstone of valuation by arbitrage. (This is not present in our case of the Vadinar refinery, Gujurat India, but we will assume for the modelling case here that it is)

4. We should be able to exhibit appropriate stochastic processes for the evolution of $S_1 - S_3$, in particular, because data series of past values are available. (This includes in the case of the refinery the correlation between the saleable product prices)

5. The type of option must be recognised: European versus American versus compound; since it will obviously impact the price obtained for the physical asset. In practice, this choice is rarely unique, and one can make a compromise between the tractability and accuracy of the representation.

6. The stochastic processes for S_1-S_3 should lead to market completeness; otherwise there will not be uniqueness of the valuation, which is a problem if one is paying hundreds of millions of dollars to acquire a physical asset, or marking to market a set of complex derivatives.

In figures 2 and 3 the crack spread calculation over ten years is depicted; mean reverting behaviour is conspicuous, as are the dips below the $0 profit level, occurring three times within this period; despite becoming less so in the latter parts of the graph; on the right hand side it can be seen that the price deviates greatly from its mean; this trend has recently reversed and again the case for reversion has strengthened.

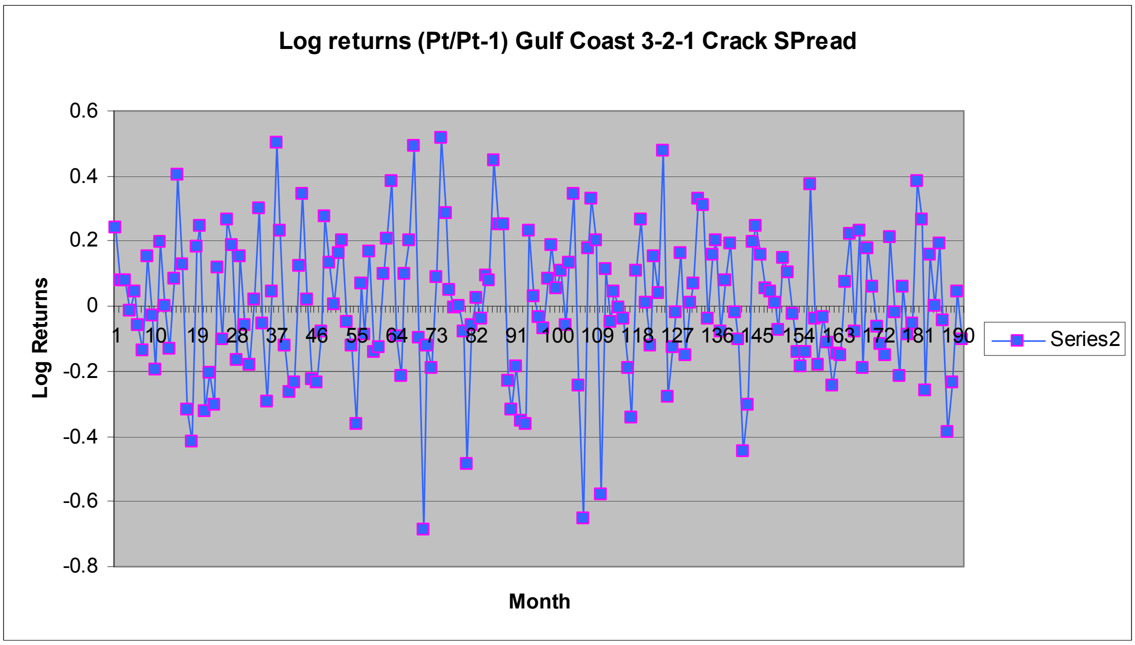


$ per barrel

Days

(Figure 4.2:** Daily calculated crack spread prices, Gulf Coast 3-2-1 underlying in $/barrel)* *Prices for the crack spread are calculated as standard with the 3-2-1 contract; 1 heating oil barrel spot price + 2 gasoline barrels - 3 crude barrels, all divided by 3.(remembering to multiply gallon prices by 42).*

The calculation for the above figure is standard for the Gulf Coast contract as done in the literature; negative values for the crack spread are possible - reasons for this are major source of research with some academics claiming the spot price of crude has lag effects and when the value for the refined products is low this can cause a negative crack spread.



***(Figure 4.3:** Daily log return crack spread prices, Gulf Coast 3-2-1 underlying)*

We calculate the log returns on the crack spread price tomorrow divided by today's. The scale of spikes below varies between twice the original price and the negative of this value, but on average it is moving by approximately 30%.

4.1.5. Crack spread modelling

Some authors suggest modelling the two separate prices (bivariate models) that constitute the spread, others consider the difference, hence a univariate model. Dempster, Medova and Tang (2008) suggest that the one factor Schwartz is unlikely to capture the mean reversion on multi asset models as well as single asset modelling of the spread. They provide Engle-Granger and ADF tests to support their claims; despite this, the structure of the commodity series will be less accurate than modelling each commodity with a univariate model. The authors provide an in-depth study of long term crack spreads and the possible valuation models. Arithmetic and geometric Brownian motions are popular for these purposes; in terms of testing the spread's behaviour a Jarque-Bera test quickly reveals that the returns of gasoline, heating oil and crude oil are not normally distributed; there is also weak skewness to the right, and there is excess kurtosis: more evidence for non normality. A Dickey Fuller test on each of these returns shows a unit root, but after first differences all become stationary series. Duan and Pliska (2004) find a negative vega (where the value of the option decreases with an increase in volatility) evident in crack spread options, a peculiarity underdeveloped at this time, and this cannot be retrieved by univariate modelling – the authors use a Taylor expansion, and determine the vegas and deltas, by considering two commodity prices and their corresponding standard deviations. By considering the co-integration of the commodity prices they find a negative vega, where the partial derivatives of the spread option value to each one of these four variables is calculated. In practise this is a serious drawback for univariate methods as many hedging and speculation strategies use negative vega as a signal for executing a trade; as mentioned the vega is the sensitivity of the option price with respect to a change in the underlying volatility. Another strong incentive for models with the information from all contributing commodities, a multivariate model, is that without them, the Greeks are calculated with a huge loss in accuracy; see Carmona and Durrleman (2003) for an example. Additionally, if volatility is to be implemented within the stochastic equation the numerical procedure becomes much less tractable - in terms of pricing and hedging implications, the accuracy is however a priority. Authors often capture the volatility structure using a GARCH feature if constant volatility is considered too unrealistic. Mahringer and Prokopczuk (2010) describe the issues with the crack spread options being of American type and they capture this behaviour by generating a simulation feeding a calculation of the optimal stopping time. The authors simulate 50,000 paths to calculate crack spread option prices in a multivariate setting, and it takes in total 2-3 hours for the calculation to complete in the bivariate case. The pricing model has a particular GARCH structure on the volatility, and all option prices are estimated using Longstaff and Schwartz (2001). These studies support the case for a multivariate, numerical and tractable model that not only calculates crack spread contract values, but also outputs the Greeks in reasonable time - we leave the Greeks calculation for future research. Dempster and Hong (2000) developed a Fast Fourier transform approach that captures stochastic volatility and the jumps for spread options; a very computationally intensive calculation. More recent studies have captured the time varying volatility and co-integration behaviours of the spread; see Duan and Pliska (2004). Maturity effects are emphasised in Theriault (2007); the Samuelson effect and the volatility effects are important features that any model of crack spreads should exhibit. If we are to include these effects the simulation of futures prices for crude oil, heating oil and gasoline should be considered only, as the other prices will introduce intractability. The pricing framework that we are in is one of multiple state factors, definite mean reversion, path dependency, and American option features. Simulation using a multivariate model is the natural method to pursue. We start by comparing market option prices to Kirk's approximation as it is the most widely applied analytical solution for spread contracts in the energy markets, see C. F. Lo (2013); the authors provide a concise derivation of the Kirk formula. Li et al. (2008) provide a number of approximations for the case of multiple prices where underlying commodities are distributed jointly normal. They provide analytical formulas for the case of a Geometric Brownian Motion and a log Ornstein-Uhlenbeck process – these formulas are fast to calculate and consequently enable the authors to produce the Greeks. Lower and upper bounds are shown in Carmona and Durrleman (2005); there is an assumption of Geometric Brownian Motion, but the lower and upper bounds calculated are useful, and as described by the authors, studies including baskets of underlyings are underdeveloped generally. Trading a spread option is said to be equivalent to trading the correlation between the two prices. Kirk (1995), Mbanefo (1997) and Alexander and Scourse (2004), all describe the correlation of the prices in a spread option calculation as being highly volatile, hence a constant correlation would seem inappropriate; we use a constant correlation despite the issues as the options calculated are for no longer than monthly periods. In the next section we introduce the available contracts for futures trading on the crack spread.

4.1.6. Types of crack spread options listed on the NYMEX exchange

In table 4.3 below, common crack spread related contracts on the NYMEX exchange are detailed; contracts for spreads on location, calendar and production are common. There is a tremendous selection of products on NYMEX; most are heavily traded with volumes in the 100s and 1000s.

(Table 4.3: Crack spread options listed on the NYMEX exchange*There are days when open interest and volumes on all of these products is extremely high)(Data accessed on 05/03/2013)

Clearing	Globex	Floor	ClearPort	Product Name	Product Group	Subgroup	Category	Sub-Category	Cleared As	Exchange	Volume	Open Interest
RM	ARE	RM	RM	RBOB Gasoline Crack Spread Futures	Energy	Refined Products	North American	Crack Spreads	Futures	NYMEX	0	2,259
HK	AHL	HK	HK	Heating Oil Crack Spread Futures	Energy	Refined Products	North American	Crack Spreads	Futures	NYMEX	210	10,226
CHY	CHY	CH	CH	Heating Oil Crack Spread Options	Energy	Refined Products	North American	Crack Spreads	Options	NYMEX	0	0
3U	A3U	3U	3U	European Gasoil Brent Crack Spread Average Price Option	Energy	Refined Products	European	Crack Spreads	Options	NYMEX	0	350
1ND	1NA	1ND	1ND	Singapore Mogas 92 Unleaded (Platts) Dubai (Platts) Crack Spread Futures	Energy	Refined Products	Asian	Crack Spreads	Futures	NYMEX	0	0
SFC	SFC	SFC	SFC	Singapore Fuel Oil 180	Energy	Refined Products	Asian	Crack Spreads	Futures	NYMEX	0	591
FO	FO	FO	FO	3.5% Fuel Oil Barges FOB Rdam (Platts) Crack Spread Futures	Energy	Refined Products	European	Crack Spreads	Futures	NYMEX	608	41,960
3Y	A3Y	3Y	3Y	RBOB Gasoline Crack Spread Average Price Options	Energy	Refined Products	North American	Spreads	Options	NYMEX	0	0
3W	A3W	3W	3W	Heating Oil Crack Spread Average Price Options	Energy	Refined Products	North American	Spreads	Options	NYMEX	0	216

The contract choice above will be significant in representing the valuation as a set of underlying values, but even more vital to the consolidated refinery value process is the correlation. Ideally, due to the commodity prices affecting the refinery's profits in a dynamic and complex way, the correlation structure should be integrated into the model; we omit a time varying correlation component for the sake of computational ease. The option to adjust production over very short time periods is captured by the real option approach, and the grid of the time horizon should fully reflect the corresponding coarseness – here we assume decision variables cover monthly periods, but the optionality available to the refinery owner is of daily granularity. In practise when long term contracts have to be met the refiner does not always have this level of flexibility as a certain level of production will have already been committed. In that case, the long term contract can be thought of as a hedging instrument, we leave this further analyses for others.

Valuing the refinery

4.2.1. Data

The NYMEX heating oil and gasoline crack spread options are the most liquid exchange traded crack spread options on the market. We collect end of day prices for gasoline and heating oil crack spread contracts directly from NYMEX; prices for these options are recorded at the end of the open-outcry session of each trading day. Whereas, trading of the underlying futures contracts is generally on an electronic platform. The sample of gasoline and heating oil crack spread call options used in our valuation comprises price observations covering the period from December 2000 to December 2012.

We apply one main exclusion criteria to the data where we considered option contracts which were exactly one month in maturity length. We also excluded all options with prices below \$0.375 to reduce the price discreteness related biases. In total we ended up with 10,500 gasoline crack spread options and 11,600 heating oil crack spread option prices. The average price of all gasoline crack spread options was \$2.37 and \$2.41 for heating oil crack spread options.

***(Table 4.4:** Sample crack spread option prices)*

This table provides the summary statistics of NYMEX crack spread option prices over the period Dec 2000 to Dec 2012.

	No. of OTM price observations	Average price of end of day OTM option prices	No. of ATM price observations	Average price of end of day ATM option prices	No. of ITM price observations	Average price of end of day ITM option prices
HOO(1)	2285	2.16	660	2.39	1460	2.87
GAO(1)	1975	2.03	672	2.21	1290	2.38

To obtain the correct option price representing the refiner's optionality we use a calculation to replicate the 3-2-1 contract by always aggregating two gasoline crack spread options and one heating oil crack spread option; this addition represents an approximation of the total crack spread option value for the refinery. In table 4.5 below are the results of the pricing of an RBOB gasoline option from NYMEX across all methods:

(Table 4.5: Gasoline Crack Spread Option prices)

Underlying: Gasoline RBOB Crack Spread Futures, Date: 29/05/2014, NYMEX quoted price = $3.946

Strikes K	**NYMEX**	**KI**	**BA**	**AV**	**HW (N=40)** ***S.E. 0.04***	**SS (N=40)** ***S.E. 0.03***
15.50	3.94	7.89	7.06	5.18	5.12	5.16
15.75	3.69	7.67	7.99	3.93	3.87	3.81
16.00	3.44	7.56	7.76	3.79	3.66	3.62
16.25	3.19	7.43	7.57	3.49	3.53	3.59
16.50	2.69	6.94	6.33	2.92	2.82	2.87
16.75	2.44	6.42	6.97	2.79	2.66	2.77
17.00	2.19	6.19	6.64	2.48	2.36	2.43
17.25	1.94	5.35	5.87	2.29	2.09	2.14
17.50	1.69	5.06	5.46	1.96	1.88	1.97

We now describe the model to value the refinery with crack spread option price data.

4.2.2. A strip of European crack spread options

Shown in the below formula is the value of the refinery at time 0: the discounted sum of the crack spread options over the lifetime of the oil refinery complex (the discount factors are contained within the value C of the option itself). The Linear Program (LP) valuation of the refinery complex at period zero is given by, V_0^{LP}, based on the information we have at date zero (ζ_0). It is equal to the maximum profit, revenue minus costs, or saleable products, heating oil volume (q_2) and gasoline volume (q_3) minus crude volume (q_1) – all in units of tons, over the time horizon (T) in years. The crack spread option $C_{i,j}$ represents this profit, where the, i, is the start date of the option and the, j, is the maturity of the option. Each option is assumed European and has a maturity of one month; we collect heating oil and gasoline real option prices from NYMEX to represent this value. The total hydrocarbon liquid within the refinery complex at any time is L, and this acts as the state variable in an equivalent final Bellman equation. As explained in detail in chapter three, we represent the refinery as a strip of daily crack spread options, where the decision variables are the inputs and outputs of the refinery: q_1, q_2, q_3, and L, and the operational costs are assumed constant as found in chapter two at \$2.29 per barrel. The valuation formula is:

$$V(0) = \sum_{i=1}^{N} \sum_{j>i}^{N} C_i^j(0) q_{i,j,1} \qquad (3)$$

The *V(0)* is the value of the refinery asset at date zero, where, the sum over i=1 to N periods, where the $q_{i,j}$ is the notional amount of crude refined per period in tons, *i* is the starting period, *j* is the sale date of the refined products (assuming instant sales after being refined), and *C(0)* is the option value at the date of the valuation today. We include the above equation (3) in the Linear Program to value the refinery by replacing the option price, C(0), with different approximations where the goal is to compare existing analytical and numerical approaches to our numerical simulation. If the variable costs of refining were zero, and the refined product set considered as one price, we would have an option to exchange one commodity for another – for which there is a closed form formula derived by Margrabe (1978), unfortunately with a spread option and a non-zero strike or variable costs, no such analytical shortcut exists. The oil refinery *i-j* crack spread, crude purchased on date *i* and refined outputs sold on date *j*, has the following payoff for injecting one unit of crude oil at time T_i, using the interest rate curve of the local region for δ, extracting refined products at time T_j, and the ζ is the information available at date 0:

$$C_i^j(0) = \delta^i E[\max(0, \delta^{j-i} \sum_{z=2}^{3} F_{z,i,j} - F_{1,i,i} - O \& M_i) | \zeta_0] \qquad (4)$$

The time, T_0, value of such an option is the difference between the sum of refined product futures prices, $F_{i,j}$ and the crude spot price, $S_i = F_{i,i}$, adjusted for the costs of refining associated with one barrel of crude, operational and management costs, O & M, – in this implementation all units are converted to tons, with information at date zero, ζ_0.

The model we introduce in this chapter, works with a set of crack spread options, the model assumes the production choice volumes from a ton of crude as defined in section 2.4 – more simply, we no longer have seven volume parameters. Purchasing of crude and sales of refining products are associated with maturities *0, 1,…., N-2,* and *1,2,…., N-1*, respectively. The aim in this model is to, at time t = 0, build a set of refinery spread options, and calculate the decision variables within each period; giving an optimal solution to a linear program, and consequently the output value. The decision variables in this linear program are the notional amounts of crude, and the inventory level associated with refining over a given period; which indirectly describes the volumes of gasoline and heating oil. The linear program we investigate is subject to the same physical and flow constraints defined in the previous chapter; the value at date 0 of the set of crack spread options is: $V_0^{LP}(\zeta_0)$, where as in the previous chapter, *L* is the total hydrocarbon liquid held at the refinery – the inventory level, *q* is the volume of the product in question here for products 1-3 and, *C* is the option price on each day. We now define the mathematical requirements.

4.2.3. The Linear program valuation construction

In equation (5) the Linear program (*LP*) is defined at date *0* with information, *ζ*, where the refinery owner wants to maximise over the decision variable, *q,* volume of crude oil and the total petroleum liquid *L*, until maturity of the refinery *T* - as already explained, throughout this chapter we replace the option price C in equation (5) below with a value from one of the five methods: BA, KI, AV, SS, CSORV or the NYMEX option real prices:

$$V_0^{LP}(\zeta_0) := \max_{q,L} \sum_{i \in [0,\dots,T]} \sum_{j \in [1,\dots,T], i<j} C_0^{i,j}(\zeta_0) q_{i,j,1} \quad (5)$$

This objective function is subject to the same physical and mass balance constraints defined in the previous chapter. These are not repeated as they are not the important focus of this chapter. In equation (6) below, we define the non-negativity constraints for crude $q_{i,j,1}$ and heating oil $q_{i,j,2}$ and gasoline $q_{i,j,3}$, including the upper volumetric limitation at the refinery, defined as Q, where the amount of liquid bought Q_{buy} or sold Q_{sell}, has an upper limit at the refinery complex:

$$\text{s.t.} \quad 0 \le q_{i,j,1} \le Q_i^{buy}, \qquad 0 \le q_{i,j,z} \le Q_{ij}^{sell}, \qquad i \ge 1, \qquad i < j, \quad 2 \le z \le 3 \qquad (6)$$

In chapter three seven variables represented the saleable products, here we use three, and they all aggregate at any one time to equal the total hydrocarbon liquid present at the refinery as L_i, where L_0 is on date zero and L_{init} is a parameter entered into our program to allow a minimum level of hydrocarbon liquid on the start date:

$$L_0 = L_{init}, \qquad L_i = L_{i-1} + q_{i,j,1} - \sum_{z=2}^{3} q_{i,j,z}, \qquad 1 \le i \le T, \qquad 2 \le z \le 3 \qquad (7)$$

In equation (8) below we state the lower and upper bounds on the hydrocarbon liquid at any time in the refinery, this is input into the LP, it is equal to the current total liquid present plus the crude bought minus the gasoline and heating oil sold on each day, i, and the end of horizon liquid can also be set in the program – analyses of changing this parameter is not central to our problem:

$$L_{min} \le L_i \le L_{max} \qquad \forall i = 1,...,T, \qquad L_T \ge L_{end}, \qquad (8)$$

[A*ll refinery plant constraints defined as in chapter 3 section 3.2.4.*] (9)

The objective function in (5) is the value of the portfolio of spread options over the refinery's lifetime, in contrast to the previous chapter the objective function is altered, although the remaining linear program remains unedited. The profit is not explicitly defined in the objective function as seven outputs minus crude as the input – value is instead emulated within the option contract approximation; by crude oil, heating oil and RBOB Gasoline, hence the inventory balance has only three liquid commodities instead of seven - representative of the Gulf Coast contract. The constraints sets (7) and (8) express inventory balance and boundary conditions, respectively. Constraint sets in (6) and (9) enforce capacity and mass balance constraints respectively. Constraint (8) also poses non-negativity conditions on the decision variables, and the end of horizon inventory level. There are no closed form formulae for the spread option values within the objective function shown in (5), they can however be computed numerically, or they can be calculated using approximate closed form formulae. We compare five different methods to obtain the solution to (5)-(9) above - all approaches approximate the crack spread option values before optimising the LP. We assume that these are European call options on the historical crack spread value minus a strike for refining costs – the five methods considered are: (a) Bachelier's method (BA) (b) Kirk's approximation (1995) (KI) (c) Alexander and Vankatramanan's (2011) approach (AV), (d) Schwartz and Smith (2000) two factor model on the commodity price simulation, and finally (e) our reduced trinomial tree method for just three commodity products: WTI crude, RBOB gasoline and heating oil, required in formula (1) – we call this the *CSORV* method. This tree valuation converges to the true solution as is shown in Hull (2003) for option valuation under assumptions of risk neutrality and mean reversion; acceptable levels of convergence require at least 100 simulations – we apply 10,000. Real options should be approximated using European options as we can only exercise the physical option when we physically hold the underlying commodities[53] - despite this, if there were American option features present, trinomial trees would provide a framework in which to capture them.

When modelling spreads, the majority of the literature regards the underlying prices as log-normal in distribution; this is a result of most studies being applied to equity prices. The positivity restriction on equity indices does not apply to the commodity spreads themselves, and as one can see in figure 2 of the crack spread calculation, both positive and negative values can and do occur. A number of articles propose arithmetic Brownian motions for the dynamics, this permits closed form formulae for options on the spread option contract, see Shimko (1991) for the details. We begin with the classical setting - there is a riskless bank account with constant interest rate r, the arbitrage-free model contains two commodity prices at time t denoted by, $S_1(t)$ and $S_2(t)$. We assume that under the risk neutral measure the joint dynamics represent any stochastic differential equations (SDE) that are Gaussian in distribution, where their SDEs are driven by volatilities σ_1 and σ_2 respectively, with W_1 and W_2 as two Brownian motions and constant correlation ρ. In the case of a spread option on two commodity spot prices, the instantaneous convenience yields do not appear in the equations. To reiterate, we know that a European spread option value on two underlying prices, with spot prices S_1 and S_2 and strike price K, has a payoff at maturity of:

[53] Raymond Cheng and Walt Tyrrell, "Using options theory for commodity spreads", EnergyRisk, October, 2006.

$$C_T = \max[S_{2T} - S_{1T} - K, 0] \quad (10)$$

We use the following notation for the formulae required to value crack spreads, where the parameters in the model: *α, β, γ, δ and κ* are defined as follows:

$$\alpha = S_2(0)e^{-rT}, \quad \beta = \sigma_2\sqrt{T}, \quad \gamma = S_1(0)e^{-rT}, \quad \delta = \sigma_1\sqrt{T}$$
$$\text{and} \quad \kappa = Ke^{-rT} \quad (11)$$

Where, r is a constant interest rate, T is the time to maturity, $S_1(0)$ and $S_2(0)$ the spot prices of two commodities at date 0, σ_1 and σ_2 their volatilities and, K the strike price, here representing the costs of refining a barrel of crude. We next introduce the basic methodology created by Bachelier in his PhD thesis from 1900 and apply it to obtain crack spread option prices; see Wilcox (1990) for a thorough description.

4.2.4. Crack spread option pricing methods

Poitras (1998) discusses arithmetic Brownian motions when pricing spread options - simplifying to this level enables analytical pricing but numerical approaches are found to be more accurate. An arithmetic Brownian motion approach follows.

4.2.4.1 The Bachelier method applied to the crack spread option

According to Bachelier we can approximate the spread, S, by the following, where the other parameters are as above:

$$S = \alpha e^{\beta S_1 - \beta^2/2} - \gamma e^{\delta S_2 - \delta^2/2} \quad (12)$$

For the underlying to match the first two moments, the mean and variance, we need:

$$S \sim N(E\{S\}, \mathrm{var}\{S\}) \tag{13}$$

Standard computations give these moments as:

$$E\{S\} = \alpha - \gamma \tag{14}$$

$$\mathrm{var}\{S\} = \alpha^2(e^{\beta^2} - 1) - 2\alpha\gamma(e^{\rho\beta\delta} - 1) + \gamma^2(e^{\delta^2} - 1) \tag{15}$$

If the value of the crack spread at the terminal date is assumed to have a Gaussian distribution, then Bachelier's European call option price approximation is the following, assuming continuous trading, perfect markets, and absence of arbitrage we have, where all symbols are defined above and the cumulative and density functions of a normal distribution are utilised:

$$C^B = (\alpha - \gamma - \kappa)\Phi(\frac{\alpha - \gamma - \kappa}{\sigma^B}) + \sigma^B \varphi(\frac{\alpha - \gamma - \kappa}{\sigma^B}) \tag{16}$$

Where the Bachelier volatility is:

$$\sigma^B = \sqrt{\alpha^2(e^{\beta^2} - 1) - 2\alpha\gamma(e^{\rho\beta\delta} - 1) + \gamma^2(e^{\delta^2} - 1)} \tag{17}$$

It is well known that the Bachelier method gives poor results for spread options when either correlation is close to 1 or -1, or the strike is near 0.

The correlation for most of the refined products is very high as already discussed, to improve upon this, we also carry out the optimisation using the Kirk approximation of the spread, see Lo (2013) for an analysis of equation (17) shown below, and an analysis of why it is still very relevant within energy trading today.

4.2.4.2. Kirk's Approximation applied to the crack spread option

The Kirk method is the market standard for commodity and energy related spread option contracts; under particular circumstances it is accurate and remains efficient; for a European call option price we have the following:

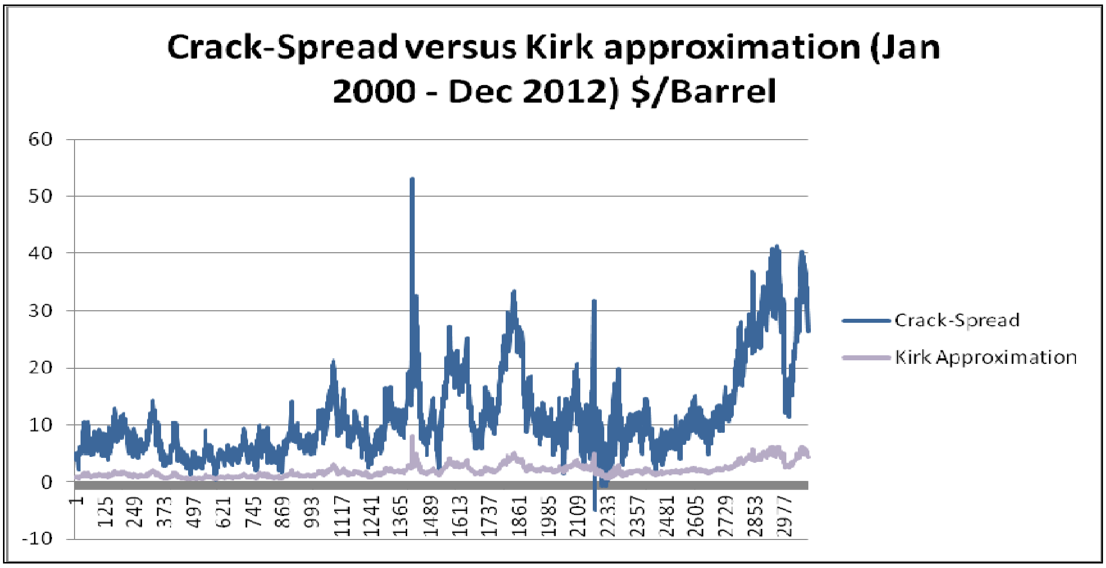
$$C^K = \alpha\Phi\left(\frac{In(\frac{\alpha}{\gamma+\kappa})}{\sigma^K} + \frac{\sigma^K}{2}\right) - (\gamma+\kappa)\Phi\left(\frac{In(\frac{\alpha}{\gamma+\kappa})}{\sigma^K} - \frac{\sigma^K}{2}\right) \tag{18}$$

Where the Kirk volatility is:

$$\sigma^K = \sqrt{\beta^2 - 2\rho\beta\delta\frac{\gamma}{\gamma+\kappa} + \delta^2\left(\frac{\gamma}{\gamma+\kappa}\right)^2} \tag{19}$$

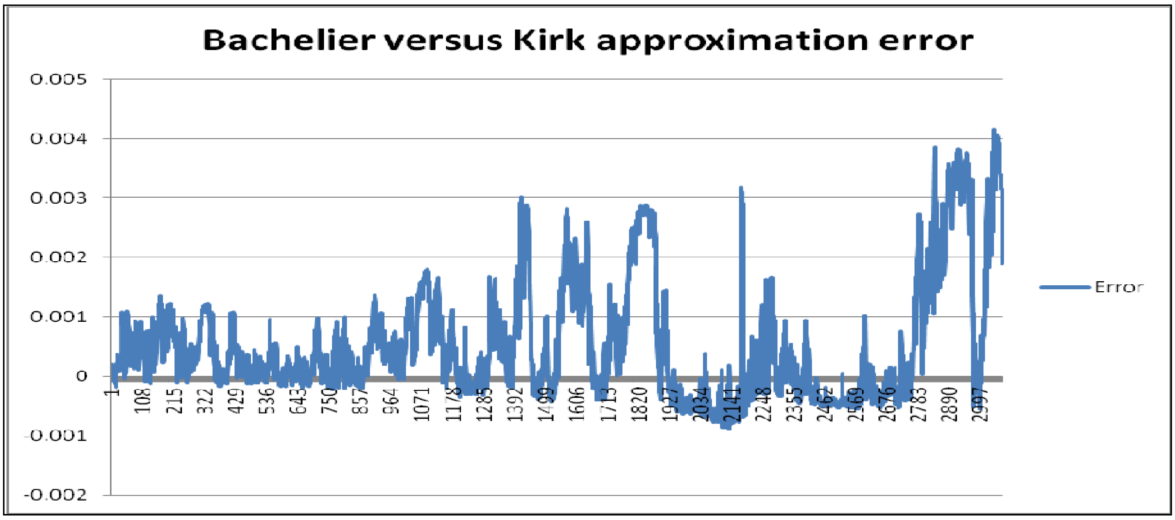
See Kirk (1995) for the derivation - it has been found in numerous studies that the Kirk call option approximation price on a spread is very accurate in respect to pricing, but it suffers when calculating the *Greeks* required for hedging purposes or when the volatility is high enough – in this valuation model we are more interested in the pricing capabilities, see Carmona and Durrelman (2003) for an alternative method when hedging is required. It also suffers when the strike is high, even when the option is at the money (option underlying price equals the strike price); a much more consistent price under these circumstances is given by AV, which works well across all strikes and is shows more stability even under various maturity contracts. The authors sum the prices of two compound exchange options, and then apply Margrabe (1978) producing an analytical solution to the exchange of two products. In terms of considering an accurate representation of the spread over many significant considerations, the model of Alexander and Scourse (2004) captures observed implied volatility skews, market correlation frowns and is consistent and accurate across all strikes; however AV11's model is more tractable to implement.

In the figure 4.4 below we compare the calculated crack-spread over the relevant period using equation (1), and the option price using the Kirk approximation in equation (17), to show the option price relationship with its underlying - the crack spread is the underlying of this option price, but this shows the leverage available to a refinery owner who wishes to manage the crack spread risk:



(Figure 4.4: Calculated crack spread prices, versus Kirk approx. of the option value $/barrel, strike at $2.44 per barrel)

Although Kirk's model is applied on most commodity trade floors, in figure 4.4 above it emphasises the information contained in the option price - it is highly correlated to the underlying. In figure 4.5 the difference between the BA and KI methods in calculating the value of the crack spread option is portrayed. This difference would be more pronounced if say, the maturity of the option were over a longer period of time; we are however, always managing a daily option with one month to maturity. This is an absolute error; it is significant when trading in huge volumes as is practised on commodity floors – for a single refinery valuation this error is less evident.



(Figure 4.5: Calculated option price difference between Bachelier and Kirk $/barrel)

4.2.4.3. Alexander and Venkatramanan's approximation

The authors discuss the pricing and hedging of European spread options on the spread of two underlying log normal prices using a new analytical approximation. They present the prices and the behaviour of the Greeks of the options with formulae that utilise compounded exchange options; a unique approach. They show that market implied volatility frowns are captured more accurately than Kirk's, and option values are consistent across all strikes. According to the authors, the risk neutral price of a European spread option may be expressed as the sum of risk neutral prices of two compounded exchange options:

$$C^t = e^{-r(T-t)}(E_{\mathbb{Q}}\{[w[U_{1T} - U_{2T}]]^+\} + E_{\mathbb{Q}}\{[w[V_{1T} - V_{2T}]]^+\} \quad (20)$$

Where U_{1T} and V_{1T} are pay-offs to European call and put options on commodity price number 1; and U_{2T} and V_{2T} are pay-offs to European call and put options on commodity price number 2, respectively. The general framework, where the above call options follow Black Scholes is applied, see the author's derivation on page 9. We apply these formulas to compare our numerical method; we leave the Greek risk numbers analysis for further work.

We compare the five approaches chosen by altering the objective function within the LP described in (5) – (9); only the value of the crack spread option price is altered within each simulation, C(0) - the rest of the LP is identical. For the CSORV approach, the linear program developed above, is solved by representing the option values on a discretised trinomial tree, but with a reduced dimensional set. Instead of calculating each commodity price, we only simulate those that are required to calculate the 3-2-1 Gulf Coast crack spread, i.e. gasoline, heating oil and crude oil – whilst considering the correlation structure and all remaining constraints. At each node on the tree, there are three random prices; the gasoline, heating oil and crude spot prices at each time period *i*. We obtain the forward prices ahead, by simply using the corresponding period's vector of prices over the corresponding probabilities. The values, of the crack spread options in the above objective function, are elicited through evaluating equation (1) at each time period on the tree, and the intrinsic value of the option on the same node using, the difference in the maximum of the spread and the discounted value of the next periods potential spread. We describe this calculation in more detail in the next section.

Both the SS and CSORV calculate daily crack spread option prices via trinomial trees; the underlying future prices are first calibrated and then consolidated to obtain a final daily crack spread daily price. Finally, for each day a forty step trinomial tree is simulated forwards and then dynamic programming is applied to find the crack spread option price, a minimum of 40 steps is required to converge on actual NYMEX prices. We name these daily forty step trees, sub-trees and we describe these calibrations in the next section and in the appendix.

4.2.4.4. Schwartz and Smith(2000) estimation applied to the crack spread option

We build and calibrate the Schwartz and Smith two factor model on the crack spread by calibrating it to the commodities that constitute the 3-2-1 Gulf Coast contract. This is calculated using EIA future price data over the ten years studied in this thesis; this is similar to the calibration we carry out in the next section with the Hull & White two factor model. We also use the crack spread option data to imply the volatility over the time horizon; the results of this are displayed in the appendix.

Allow S_t to be the commodity price at time t:

$$In(S_t) = \xi_t + \chi_t \qquad (21)$$

Where ξ_t is the long term equilibrium level and χ_t is the short term movement from the equilibrium price. The long term variable is modelled as a Brownian motion with a drift of μ_ξ and volatility σ_ξ :

$$d\xi_t = \mu_\xi \, dt + \sigma_\xi \, dz_\xi \qquad (22)$$

In the short term the model applies a mean reversion variable, where the mean-reversion coefficient is κ and the volatility is σ_χ :

$$d\chi_t = -\kappa\, \chi_t\, dt + \sigma_\chi\, dz_\chi \tag{23}$$

where the dz_χ and the dz_ξ are correlated standard Brownian motion movements; where $dz_\xi\, dz_\chi = \rho_{\xi\chi}dt$. Due to the valuation of the refinery being a risk-neutral one the dynamics of the commodities must also be. Hence, the equations for the two factor price processes become:

$$d\chi_t = (-\kappa\, \chi_t - \lambda_\chi)dt + \sigma_\chi\, dz^*_\chi \tag{24}$$

$$d\xi_t = (\mu_\xi - \lambda_\xi)dt + \sigma_\xi\, dz^*_\xi \tag{25}$$

Where here the λ_ξ and λ_χ are risk premiums; here the short term process reverts to $-\lambda_\chi/\kappa$ instead of zero; it now represents the real commodity process.

We discretise the stochastic processes onto a trinomial tree and calculate the crack spread at each node by applying a forty step simulation. Having the final consolidated crack spread on one tree means we can calculate the option price using dynamic programming. This method is mirrored in the next section with the Hull & White two factor model once the prices for the crack spread are calculated.

4.2.5. Hull and White Two factor model

Extending our one factor model we describe the behaviour of the three underlying spot price processes by a new two factor model – with the goal of discretising them onto trinomial trees. The derivation and calibration can be found in the appendix; the two-factor Hull & White model is:

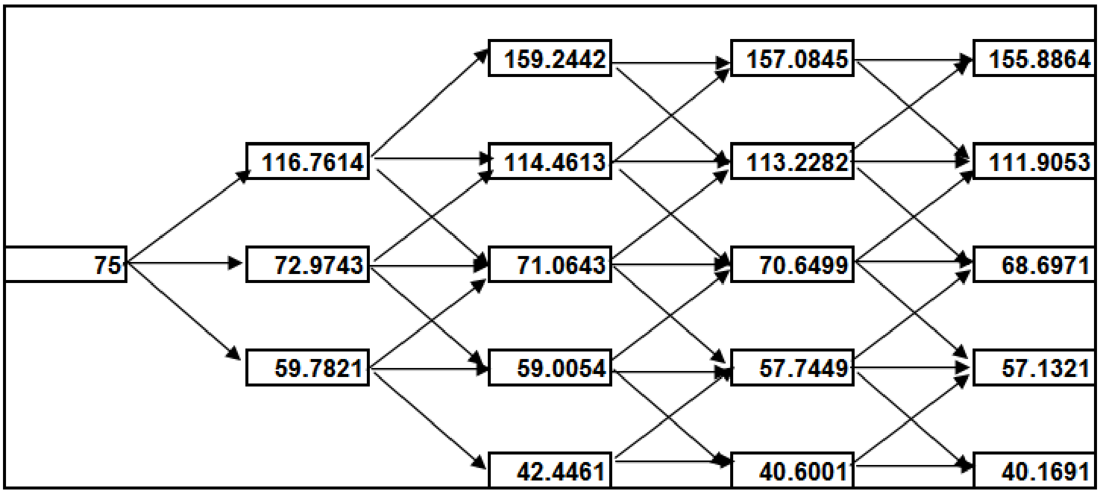
$$dS_t = (\theta_t + U_t - aS_t)dt + \sigma_1 dW_1 \tag{26}$$

$$dU_t = -bU_t dt + \sigma_2 dW_2 \tag{27}$$

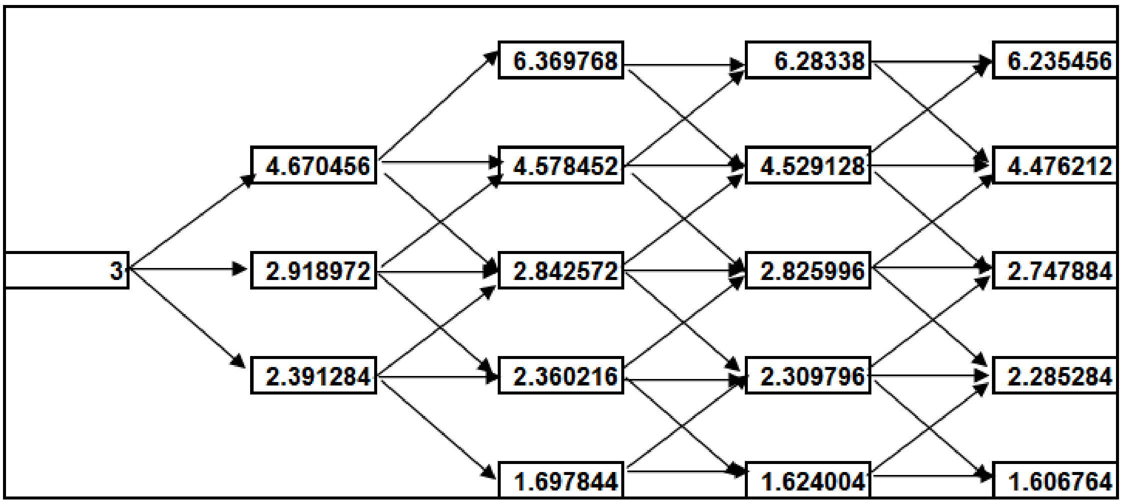
$$\rho dt = dW_1 dW_2 \tag{28}$$

Despite the individual processes being able to go negative in this model, we regard this as irrelevant in our calculation as we wish to value the spread which can become negative. The spot price process of the commodity reverts to a long-term rate that includes the stochastic process, U_t. The volatility of the spot price process and the mean reversion component, U_t, are σ_1 and σ_2, respectively. There are two mean reversion coefficients: a is the short-term mean reversion speed and b is the long-term mean reversion speed. When a is high and b is low, this model produces a forward curve where short term prices are more volatile than long-term prices, and vice versa. Despite the additional computational complexity, the empirical evidence does support the need for additional dimensions of the forward curve shifts represented by two factor models. Principal components analysis shows that 90% of the price changes can be explained by parallel shifts, yet changes in the slope of forward curve are also significant. Chantziara et al. (2008) describe the difficulties of forecasting in the oil market applying a PCA analysis to crude oil, gasoline and heating oil futures. Their first three component factors capture 93% of the crude oil structure supporting the need for not only parallel but slope shifts. Below we show a sample of prices for each commodity with the two factor model. The final step to obtain the option prices is to consolidate the commodities onto one final trinomial tree by calculating the crack spread and dynamic programming is then applied to obtain the option prices.

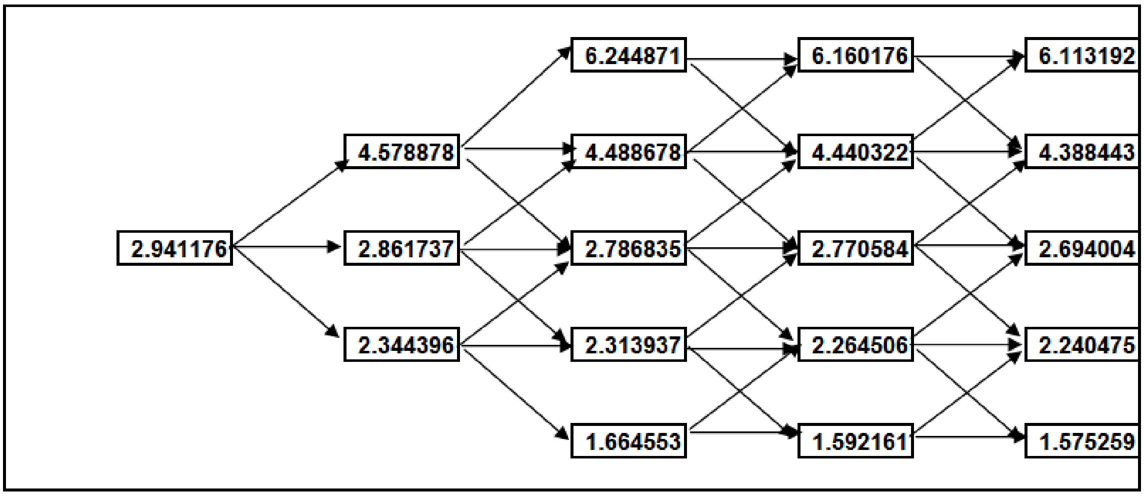
__(Figure 4.6:__ Crude oil trinomial price series $/barrel)



__(Figure 4.7:__ Gasoline trinomial price series $/Gallon)



*(**Figure 4.8:** Heating oil trinomial price series $/Gallon)*



In appendix 3.0 we depict the different forward curves in time showing the one factor model from chapter three, and the both of the two factor models calibrated here: SS and H&W. It is clear that the one factor model cannot capture the term structure as accurately as the two factor models can, which are both very close to each other. Practitioners of pricing are unlikely to use a model that does not fit the initial forward curve, yet for valuation purposes this is less significant.

Table 4.6 below depicts all option model price errors against the actual NYMEX option prices. Our CSORV model does very well in competing with the more established spread formulae and is very effective across all error measures.

*(**Table 4.6:** European call crack spread option price errors)*

Model	**MPE**	**MAPE**	**RRMSE**
Heating Oil crack spread options			
Kirk	-25.75%	59.64%	51.13%
Bachelier	-25.46%	58.06%	50.45%
Alexander	-4.32%	18.84%	19.91%
Schwartz	-4.66%	19.12%	21.34%
CSORV	-4.26%	18.63%	19.91%
Gasoline Oil crack spread options			
Kirk	-26.74%	52.89%	52.27%
Bachelier	-26.06%	52.05%	52.43%
Alexander	-4.48%	19.30%	21.10%
Schwartz	-5.20%	19.94%	21.38%
CSORV	-4.41%	19.09%	21.08%

In section 3., we state the solving time for each methodology and the final static refinery valuation at date 0.

4.3. The refinery valuation results[54]

Static linear program valuation approach	Time to solve on a dual core i5 processor (seconds)	Oil Refinery Valuation (million $)
Kirk's Approx. (KI)	0.606	634
Bachelier's method (BA)	0.651	761
Alexander and Venkatramanan (AV)	0.764	753
Schwartz & Smith (SS)	3.39	758 (2.96)
CSORV. (100,000 simulations)	3.29	729 (1.53)
Actual NYMEX Call option Spread Values*	0.469	739

***(Table 4.7:** Oil refinery valuation methods with 100,000 simulations, values in million $. Standard error in valuation in brackets.* NYMEX call crack spread option prices over ten years)*

After collecting over ten years worth of data, we assume that the 3-2-1 crack spread option price is represented by equation (4); equivalent to approximating the refinery option spread value as: $C_{i,j}$. Solving the above LP using GAMS, choosing *ILOG CPLEX optimiser,* we obtain a value of $729 million for the refinery. It is slower than the analytical methods, but closer to the real option price data, yet it remains fast in comparison to a complete numerical solution as described in chapter three. Shown in table 4.7 is a comparison of the various valuations – Bachelier's method is higher

[54] The details of the GAMS Linear programming including analysis are given in chapter three

than Kirk's as it is essentially a one factor model trying to directly estimate the difference, $F_2(T) - F_1(T)$, as such it cannot capture the structure of the actual spread's movement completely. Alexander's is higher than Kirk's and this is due to the volatility structure being more accurately depicted. Our numerical approximation is higher than Kirk's, but lower than AV's, SS's and BA's as it captures more mean reversion than the other approaches possibly can with the assumptions of the underlying's behaviour – assuming normality or another process that is without mean reversion will have drawbacks on accuracy.

The above numerical approach, is however static: we utilise the information available today, ζ_0, to construct a solution for the entire lifetime of the oil refinery. Consequently, leveraging the approach described in chapter three, section 5.5., we next re-optimise the LP model, and solve daily, with the new information that becomes available – until maturity.

4.3.1. The Updated linear program valuation results applied to the refinery[55]

As in the updated optimisation procedure described in chapter three, section 5.5., a linear program with new information, ζ_k, becoming available at each new date: *k*, equations (5)-(9) now becomes:

$$V_k^{LP}(\zeta_k) := \max_{q,L} \sum_{i\in[k,...,T]} \sum_{j\in[k,...,T],i<j} C_k^{i,j}(\zeta_k) q_{i,j,1} \tag{29}$$

s.t. $0 \le q_{i,j,1} \le Q_i^{buy}$, $0 \le q_{i,j}^{z} \le Q_{ij}^{sell}$, $i \ge 1$, $i < j$, $2 \le z \le 3$ (30)

$L_0 = L_{init}$, $L_i = L_{i-1} + q_{i,j,1} - \sum_{z=2}^{7} q_i^{j,z}$, $k \le i \le T$, $2 \le z \le 3$ (31)

$L_{min} \le L_i \le L_{max}$ $\forall i = k,...,T,$ $L_T \ge L_{end}$, (32)

[55] The details of the GAMS Linear programming, including analysis, are given in chapter three

[A*ll refinery plant constraints defined in chapter 3 section 3.2.4.*] (33)

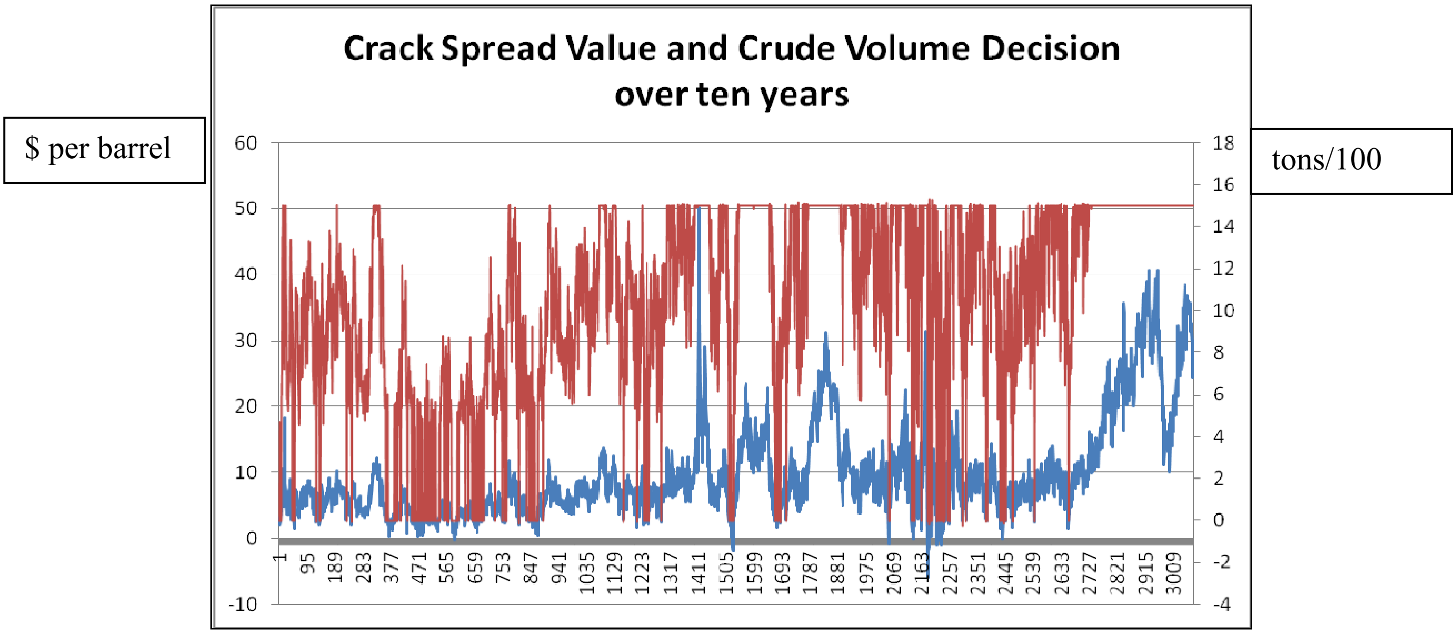
The LP approach in this section can only be valid if the refinery is in a market of liquid crack spread contracts, which is certainly not the case for the Vadinar refinery in Gujarat.

Updated linear program valuation approach	**Time to solve on a dual core i5 processor (seconds)**	**Oil Refinery Valuation (million $)**
Kirk's Approx. (KI)	256	764
Bachelier's method (BA)	302	863
Alexander and Venkatramanan (AV)	461	853
Schwartz & Smith (SS) (100,000 simulations)	1,397	859 (7.94)
CSORV (100,000 simulations)	1,432	787 (4.67)
Actual NYMEX Call option spread values*	270	817

***(Table 4.8:** All methods with an updated oil refinery valuation, values in million $ Standard errors of MC in brackets. *NYMEX call option crack spread prices)*

Below in figure 4.9 are the resultant crude volumes in tons over the lifetime of the refinery that should be selected to maximise the profit.

(Figure 4.9: Daily crack spread value ($ per barrel) versus the crude oil purchase decisions (tons); (the left y axis is $ per barrel, the right y axis is tons/100).



The blue line is the left axis, with the crack spread valued in $ per barrel, and the red line is the crude volume in tons/100 – the right axis. The decision process is highly volatile and detailed; despite this we can see that when the spread value is high we have a maximum volume decision choice, and vice versa, when the spread value is low the volume choice set is low. It is averaging approximately 1000 tons of crude per day, the owner would not in practice follow the optimisation program exactly, but would let it inform the general decision process.

4.3.2. Analysis

In practice, the valuation of the refinery will involve a mixture of positions in each saleable commodity using forward/futures contracts to secure future revenues, while allowing for maximisation in the next period, and always taking into account the physical and flow constraints.

The above formulation in (29)-(33) does not specifically take into account the alternative commodity production set prices the refinery owner is exposed to per period, but assumes that the simplified Gulf Coast crack spread calculation represents all of the optionality. The value of trading the mean reversion of each commodity price adds to the extrinsic component of the valuation – with a more simplified price representation, there is less extrinsic trading value opportunities. It is however very close to the linear program with actual option prices and AV11's analytical approximation. This approach in the full state space defined in chapter three is defined as:

$$V(0) = e^{-rt} \max_{Q} E[\sum_{j=1}^{7} \sum_{i=1}^{T} (0, Q_j F_j(i) - Q_1 S_1(i) - O\,\&\,M_i)] \qquad (34)$$

To solve the incipient optimisation problem, several methods were available:

1. The most elementary consists of using the current forward curve to find the optimal portfolio of long and short forward contracts attached to the refining period of one month for up to 10 years in a legitimate financial valuation. The corresponding values represent the intrinsic value of the refinery facility – this is similar to our static approach.
2. More complex, would be to consider a stochastic optimisation on a tree or through Monte Carlo simulations: the quantity of cumulative production denoted by Q_t, working backwards in time to solve the optimisation problem – enabling the extrinsic value to be captured accurately - this is in fact our dynamic approach.

The latter is precisely the method that was developed in chapter 3 to calculate the value of the refinery complex; this is an accurate representation of intrinsic and extrinsic value, and the uncertainty in the stochastic prices facing the refinery owner. In table 4.9 we compare all methodologies applied to the refinery in this thesis.

4.3.3. Comparison of all oil refinery valuation methods

Valuation approach	Refinery Value ($ millions)
DCF Analysis (Static, Chp. II)	570
Stochastic Dynamic Program (Chp. III)	1001
Strip of Crack Spread Options (Static, Chp. IV) (CSORV)	729
(Chp. IV) Kirk (1995)	764
(Chp. IV) , Wilcox et al (1990) Bachelier (1900)	863
(Chp. IV) Alexander and Venkatramanan (2011)	853
(Chp. III) *Out of Sample Re-optimised Stochastic Program (Chp. III)*	*1,206*
Re-optimised strip of crack spread options (CSORV) (Chp. IV)	790
Re-optimised SDP (Chp. III)	1,131

*(**Table 4.9:** Alternative valuation approaches to the oil refinery: five methods from chapter 4, three from chapter 3, one from chapter I and one from chapter II)*

4.3.4. Refinery valuation research extensions

The issues from this chapter's valuation arise due to the underlying processes that we have selected; in effect to simplify the main difficulty in valuing the refinery, which is the spread option contract value itself. The three commodity prices that are simulated are done so with a two-factor factor model; this could be enhanced by choosing a more complex continuous stochastic equation to capture the characteristics that the two-factor model cannot. For example, it is well known that stochastic volatility enables a more accurate price process as does a stochastic interest rate; this is due to smiles and skews being more accurately depicted. Dempster and Hong (2000) use a Fast Fourier Transform method to provide very accurate spread option prices, values and computational times are given and compared to a Monte Carlo simulation; yet the authors assume that prices are GBM, and it is computationally intensive. The authors also assume that the characteristic function of the multiple assets is known, which is feasible in a setting of normal distributions; varied or realistic distributions cannot be assessed in this way. However, a study into the joint characteristic function of the three refinery assets in this chapter could be carried out with the goal of finding an analytical approximation that is realistic, or consequently where a Fast Fourier Transform or similar numerical algorithm could be applied. Including an extension to capture accurately transaction costs would also improve the model for practitioners - where the correlation could be modelled so could the transaction costs behaviour whilst the refinery is in operation. The approximations for the crack spread options that represent the refinery optionality are accurate even for a simplified valuation as shown by the resulting error calculations versus NYMEX contracts. We have shown that a numerical approximation using a trinomial tree, CSORV, converges, incorporates mean reversion, and captures more optionality than Kirk's and Bachelier's whilst being within 'shooting distance' of an explicit method like Alexander's model, and close to another two factor model Schwartz and Smith. Finally, due to it being closer to actual NYMEX option values it is comparable to the more complex model created in chapter three providing a more accurate value for the refinery asset than current available approaches. However, there are a number of avenues that have either been left for further research or not considered in this chapter:

1 - Neither financial risk, for instance a constraint on variance, nor robustness measures, for example conditional value at risk (CVAR) were included, as the aim was to obtain a tractable and

efficient value - including the refinery owner's decision set at each period in time. This extension would not be an extensive undertaking as only the objective function would need to be reconstructed with an additional term inserted for risk. After re-optimising a valuation would be viable and include a consideration for a level of risk aversion in the strip of decisions made. The overall analysis shows that this simplification, despite taking longer than the other approaches, is more realistic and indicates that there are further directions that can be investigated in terms of spread options.

2 - Another extension to this chapter could include a calculation for the Greeks and other risk sensitivities associated with the oil refinery; either inserted into the objective function or as a set of new constraints - this would interest practitioners as the refining climate has morphed dramatically in recent years.

3 - Another angle that could be investigated would be alternative operational research techniques to either enhance the computation or enable more dimensions for prices to be captured whilst remaining practically viable; interrogating the duals for these optimisations would be an interesting avenue, as they provide information about shadow costs and can further the economic understanding of the resultant refinery decision set.

4 - Further research may consider the modelling of the dynamics of the entire forward curve for valuation purposes, further increasing the state space means this is a tremendously difficult optimisation. An alternative would be to include stochastic volatility into the numerical approach to capture the smiles and skews apparent in the market structure. Other methods like stochastic optimal control could be implemented with a similar appreciation for the complexity of decisions that need to be executed at each refining period.

5 - An approximation with stochastic optimal control, where the problem is perceived from the Hamilton Jacobi Bellman point of view would be interesting. Developing the complexity of the model would enhance the accuracy, but would contribute to the computational difficulty; this trade-off is of the utmost importance when concerning practitioners.

6 - Gibson and Schwartz's two factor model would be of interest in this framework as would other stochastic volatility models. Computational shortcuts would also provide enhancements, where the trinomial tree could be reduced or the recursive function made more efficient. If a further general enhancement were to include regulatory considerations; these could be inserted into the linear program constraint equations. As more regulatory actions are imposed on refiners, the developed world refineries will need to simulate the possibilities in methods to enhance profits whilst remaining within the environmental constraints – the strategies to maximise profit may well behave in a more complex way that only a computer program can replicate.

7 - Another extension could apply to the pricing algorithm of the related 1:1:1 soybean crush spread which is one part soybean meal minus one part soybeans or the 2:1:1 cattle crush spread which is one part feeder cattle plus one part corn minus two parts live cattle – obtaining data for the options would be half the battle; the price series continuous stochastic equations would have to be selected and calibrated accordingly.

In chapter three we built the entire state space, and found a numerical structure upon which to make the solution tractable, whereas, chapter four has simplified that structure, and focussed in on the daily spread value by estimating the option contract's value within the rigidities already defined. This allows a comparison to existing methods for valuing spread options, and gives a resultant value that is more accurate than standard spread contract prices within the linear program developed.

4.4. Conclusion

Valuation of oil refineries has been given extra impetus by the recent glut of crude oil not seen in 30 years. There has been a glut since OPEC began an oil price war with the US shale industry in 2014. It is in both sides benefit to keep oil pumping despite the political conflicts. Estimates are common in the market of oil reaching prices as low as $30 per barrel in January 2016 – with Iran supporting production at this price. US shale production owners have reduced output to cope with falling profits and the IMF has halved 2016 estimates of Saudi Arabian GDP. Measures will be taken by some to stem the fall, but a set price is difficult to forecast. Global demand from the biggest consumers: US, China and Japan has stagnated; prices could be skewed by longer term investment funds in commodities – yet we will likely see the federal reserve increasing interest rates in the US, the Bank of England with an smaller interest rate increase or the European Central Bank introducing a quantitative easing program. Prices for WTI crude oil have moved dramatically in recent years, the closing low of $33.87/bbl on December 18th 2015, and the intraday extreme for 2008 of $32.40/bbl before the price tripled into 2011 with a high of $114.83/bbl. Refineries are the tool by which these countries control the price of the refined products on the markets - understanding their value will enable a clearer picture of how to untangle and decipher the stranglehold countries and in turn companies have over crude oil. Obtaining fair value of these vital real assets is a hugely under researched area requiring defragmenting and analyses.

In this chapter we have constructed a much simplified linear program which replaces chapter three's extensive simulation using trinomial trees with option contract prices. The choices of which methods to use for the spread option prices was evident from the literature as they are the de facto methods used in practice and within academia. The trivial optimisation solves in under five minutes and represents a realistic indication of the refinery's financial worth. The limitations of the approach are that a huge amount of the state space has been removed and we are again as in chapter three working within incomplete markets with no single no arbitrage price available. Yet we produce a range of values utilising actual market data and leveraging off current spread option pricing methods. The Linear Program is solved by replacing the crack spread value with a European Call Option - as we view the refinery as a strip of options. We next replaced this option value with different values from five industry standard calculations to compare the results. We find that BA, KI and AV are all extremely quick and give solid results but that when compared to actual NYMEX option values our CSORV, which is very similar to the SS model, is more accurate - despite being much slower.

In comparing the oil refinery valuations a number of interesting avenues to examine are revealed. Chapter three provides the most accurate value but takes the longest; hence chapter four's CSORV being faster but less accurate. The stochastic dynamic program in chapter three is computationally intensive, but allows an accurate representation of the profit available by altering the production set at each period in time. For example, a financial dispute on fair value could be investigated using this method – with accurate data the algorithm can be altered to suit the auditor's needs - alternative refinery types could be examined with our method, risks can be easily incorporated and different time horizons considered. In contrast, the static strip of crack spread options valuation in chapter four is computed very quickly, and provides some of the intrinsic value, enabling it to be more suitable for a financial trade on the crack spread.

Thesis Conclusion

This thesis is an investigation into the financial value of an illiquid real asset - a topping oil refinery. In the first chapter we provide a background and history of the oil industry; including the refineries distinct characteristics and their marketable products. A statistical analysis, including the first four moments and a time series analysis of all seven refinery based petroleum products: crude oil, naphtha, gasoline, heating oil, kerosene, cracker feed and fuel oil, is constructed. The technical and fundamental drivers of the oil refinery downstream market are described and the purpose of the research justified - a better statistical knowledge of these refined products enables an owner to make decisions that are more accurate and have computer aided support. Although we collect data for the Indian Gujurat refinery, we also examine refineries like those at Fawley owned by ExxonMobil - the complexity of the refinery has an impact on the products it can supply and hence how far afield it can sell its outputs. The history we describe implies that not only is the oil industry a majorly significant economic creator, but it will remain so for the foreseeable future.

In the second chapter, we create a discounted cash flow valuation for a typical topping oil refinery. Growth rates in the calculation are assumed constant, as are interest rates. Sales and costs project forwards at a rate researched by oil refinery consultants KBH, at this time and accountancy calculations for revenue, and EBITDA is verified by an ACA qualified accountant. By analysing this standard and trivial calculation we find that the outputs under-value the refinery massively by ignoring the *'optionality'* the owner has - for example the refinery can be switched off if required, expanded or even production deferred until a later date. The time value of money means that all future revenues and costs can be discounted back to today - capturing the behaviour of the petroleum products is vitally important in financially valuing an asset that is underpinned by their performance. To ignore the dynamics or complexities of the time series misses a huge chunk of information that a refinery owner of experience will price into his decision making from the very beginning of the refining assay. The value obtained in chapter two assuming very standard taxation and accountancy rules is £570 million. We investigate refineries being bought and sold from the financial press to find if there is a relevant comparable number. Due to so few refineries being bought and sold, and the time between when there were transactions being so many years; there was no value that could be used as an indication, hence the programs developed in later chapters of the thesis.

In chapter three, a dynamic programming approach to valuation is sought; where the optimisation is carried out on a no arbitrage trinomial tree, that encapsulates the crack spread option value over a period of ten years. This is done by choosing a mean reverting, forward curve capturing, and Markov stochastic process for each commodity relevant to the refinery. We choose to represent the seven commodities with Hull & White (1990) single factor processes, this includes the very high correlation between each time series, hence are able to be discretised onto a tree that matches the means, variances and the constant volatility. The optimisation along the tree, finds at each node the optimal volume of product to buy or sell along with the relevant pricing on that node; the valuation is then calculated from the option value at the terminal node, discounted and compared to the intrinsic value, recursively backward to the current time period. The speed of calculation is extremely fast due to the program GAMS already having trinomial tree structures built into its library; a C++ program can then be used to input all the tree pricing data into the linear program required for valuation - the nodal formulation applied in GAMS is unique and enables memoization to be utilised in a unique way. In this LP, we define all relevant capacity and flow constraints, and all positivity and hydrocarbon liquid state equations; boundary conditions on the liquid flow are necessary to obtain a convergence. Convergence is achieved in approximately two hours, despite calculating over the huge dimensional state space. An average valuation of roughly £1131 million is found. Because the constraints are all being linear, and the objective function is piecewise and convex, GAMS can utilises the CONOPT3 IBM solver to obtain a feasible and unique solution.

Both static and dynamic valuations are provided in chapter three. These rely upon: a risk free curve for discounting, the NYMEX data for future commodity prices, and the topping oil refinery linear program construction introduced at the beginning of the chapter. The reasons for choosing this construction are described along with the benefits of such a simple linear program that can be manipulated in code to support our choice of stochastic equations. The valuation is carried out under the risk neutral measure, is optimal and unique locally and globally, is carried out under all relevant physical and flow constraints, and finally, converges in under two hours for a huge state space due to the computational shortcuts and the choices made to construct the model using a nodal formulation.

In chapter four, a simplified version of chapter three's LP is created that enables the optionality to be captured, not by the simulation of calibrated stochastic equations, but instead relevant crack spread option contract prices are injected into the objective function and the speed of valuation is cut down to five minutes – enabling a tractable and still realistic calculation to be obtained. Three spread option pricing methods common in practise are utilised within the LP to compare the methods to our choice of using the Hull & White two factor model. There is also the advantage that there is a large amount of code available for Hull & White methods. We justify the use of the two factor model as having a representation of the forward curve, enabling a mean reversion to be captured and a more complicated volatility term structure than possible with a one factor model. One can also prevent negative values of the commodity prices to occur individually by changing the branch structure each time the probability of a negative price may occur. We find that our method, the CSORV method, is extremely accurate and very fast, providing a valuation that is very close to actual spread options data from NYMEX – giving a final dynamic valuation of £787 million, whereas the actual options within the LP give £817 million.

The difference in the valuation of the refinery compared to the DCF value seem large but this is also found by Brandao (2002) and Copeland (2001/2004), where real option approaches dramatically increase the potential value of an asset in comparison to static calculations.

The overall analysis and calculations provide an estimation of a refinery's value based on the probable prices of its constituent petroleum products over time. A real option valuation is presented in both chapters three and four, which could be useful to investors, entrepreneurs, speculators and refinery hedge traders. It is in fact very difficult to find a methodology that enables the huge state space to be managed - there are 10 years of monthly periods, along with seven refined products, which can move in three different directions. A stochastic equation that fits appropriately into a Hamilton Jacobi Bellman equation and solves analytically could have been selected, yet unfortunately there are no relevant equations to capture the complexities required for all seven refined products in such a formula. There are many numerical approaches that can be considered on which to build the valuation; from Gaussian quadrature, to binomial trees for the price series. A central issue in the thesis is the balance between accuracy and tractability. For instance, increasing the number of stochastic factors in the mean reverting equation in the chapter three program, where all seven prices were simulated, would massively increase the computing time. Hence a one factor was chosen in this chapter. Whereas, in chapter four, using just the crack spread, Gulf Coast Contract commodity prices two factor stochastic processes can be included in the valuation, while obtaining a value very quickly. A risk neutral, optimal value is found that can be extended in a number of directions in an easy way due to the LP's construction.

The analysis in this thesis opens various interesting research question that may be worth investigating.

In this thesis the stochastic processes for prices are given by Hull & White. These have the advantage that they are Markov and fit the forward curve. These have the disadvantage that prices can go negative, though the code prevents that. There many other possible types of stochastic equations for the commodity price series – such as those suggested by Gibson and Schwartz (1990). These equations would need to allow for stochastic volatility or included jumps, as the forward curve must be included and maintain the Markov property.

A different numerical construction could be used to simulate the processes; a high dimensional tree, or another method for dealing with the high-dimensional property of the problem. For instance, a manipulation of quadrature that could incorporate complex commodity price dynamics or a multi-dimensional integration incorporating the relevant dynamics. There are a variety of additional methods that can be used if one assumes the process is Gaussian.

There may be alternative optimisation procedures than those available in GAMS that are more accurate and realistic. They would still need to implement the relevant parts of the solution; the commodity prices and the physical/mass balance constraints of the refinery.

The valuation procedures developed in chapters 3 and 4 do not generate the Greeks that would be used by the refinery owner who wishes to hedge his/her position. It may be possible using Monte Carlo methods to extend the procedure to generate the Greeks.

Other possible extensions could include, transactions costs, taxation and regulatory restrictions (e.g. limitations on sulphur emissions) within the LP formulation. The analysis in the thesis has assumed risk neutrality in the calculation of the objective function - this could be relaxed. The thesis has considered the case of a topping refinery, where the LPs were available, but it could be extended to a cracking refinery, were the LPs to be available.

APPENDIX

A.1.1. Calibration of gasoline

One Factor Stochastic Model: Gasoline
Period : 1/2/2002 to 01/06/2013
Contracts : F1, F3, F5, F7, F9
NOBS : 136
a_2 : 0.451 (0.095)
θ_2 : 1.003 (0.046)
σ_2 : 0.4048 (0.105)
RMSE : 1.13

(***<u>Table A1:</u>*** The calibration results of the stochastic gasoline one factor model over ten years using market prices from NYMEX.

A.1.2. Calibration of naphtha

One Factor Stochastic Model: Naphtha
Period : 1/2/2002 to 01/06/2013
Contracts : F1, F3, F5, F7, F9
NOBS : 136
a_3 : 0.0722 (0.055)
θ_3 : 0.8617 (0.046)
σ_3 : 0.4069 (0.025)
RMSE : 2.06

(***<u>Table A2:</u>*** The calibration results of the stochastic naphtha one factor model over ten years using market prices from NYMEX.

A.1.3. Calibration of heating oil

One Factor Stochastic Model: Heating Oil
Period : 1/2/2002 to 01/06/2013
Contracts : F1, F3, F5, F7, F9
NOBS : 136
a_4 : 0.5200 (0.049)
θ_4 : 0.8417 (0.062)
σ_4 : 0.3775 (0.078)
RMSE : 1.87

(***Table A3:*** The calibration results of the stochastic heating oil one factor model over ten years using market prices from NYMEX.

A.1.4. Calibration of kerosene

One Factor Stochastic Model: Jet Fuel
Period : 1/2/2002 to 01/06/2013
Contracts : F1, F3, F5, F7, F9
NOBS : 136
a_5 : 0.5968 (0.054)
θ_5 : 0.8183 (0.323)
σ_5 : 0.4408 (0.184)
RMSE : 1.64

(***Table A4:*** The calibration results of the stochastic kerosene one factor model over ten years using market prices from NYMEX.

A.1.5. Calibration of cracker feed

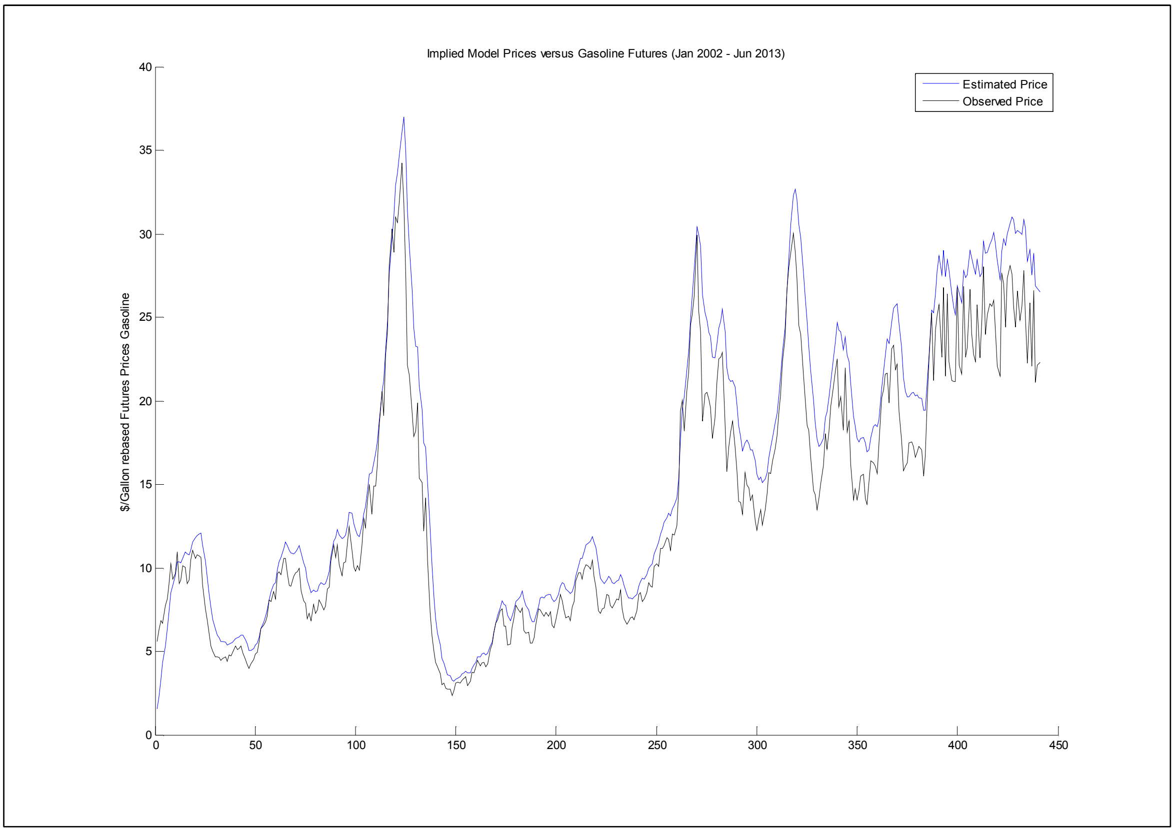
One Factor Stochastic Model: Cracker Feed
Period : 1/2/2002 to 01/06/2013
Contracts : F1, F3, F5, F7, F9
NOBS : 136
a_6 : 0.301 (0.135)
θ_6 : 3.093 (0.306)
σ_6 : 0.334 (0.095)
RMSE : 2.54

(***Table A5:*** The calibration results of the stochastic cracker feed one factor model over ten years using market prices from NYMEX.

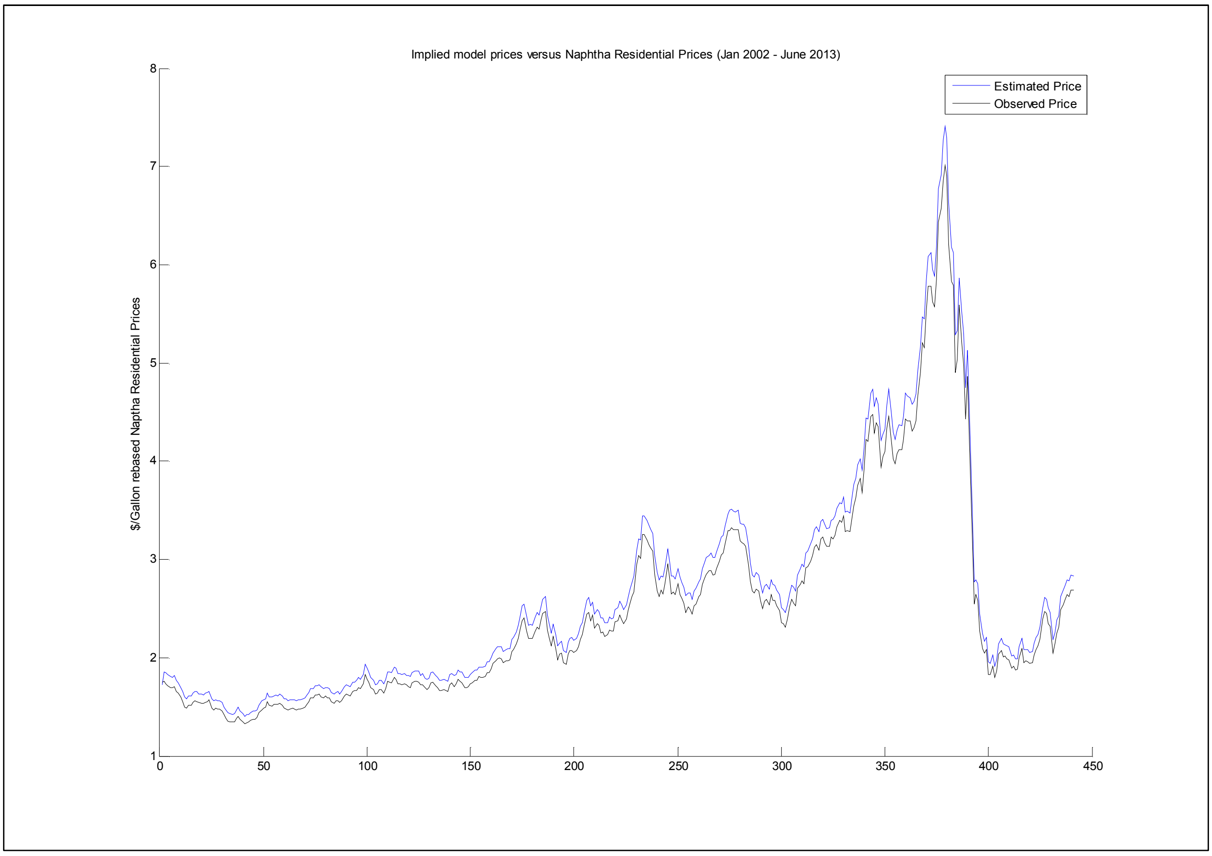
A.1.6. Calibration of diesel

One Factor Stochastic Model: Diesel
Period : 1/2/2002 to 01/06/2013
Contracts : F1, F3, F5, F7, F9
NOBS : 136
a_7 : 0.2095 (0.042)
θ_7 : 0.3270 (0.071)
σ_7 : 0.4069 (0.016)
RMSE : 3.64

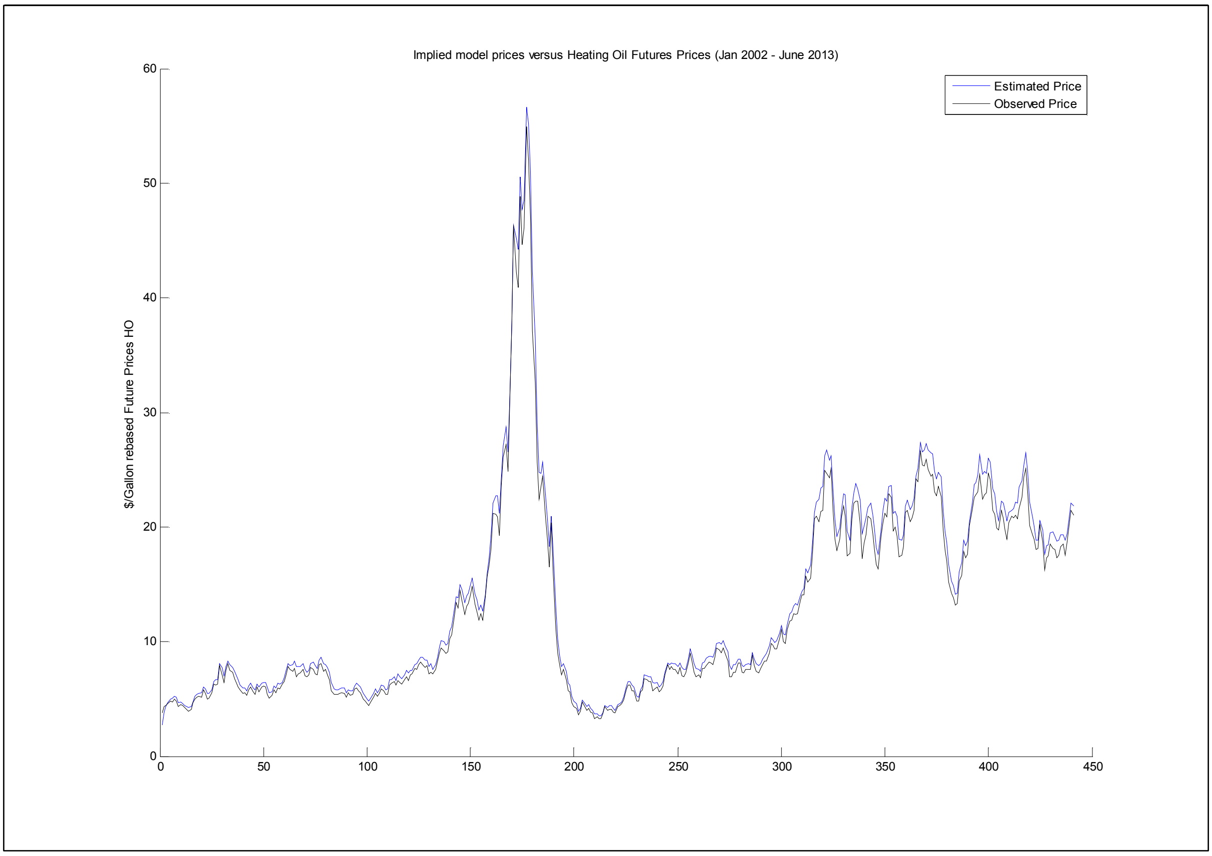
(Table A6: The calibration results of the Diesel stochastic one factor model over ten years using market prices from NYMEX.



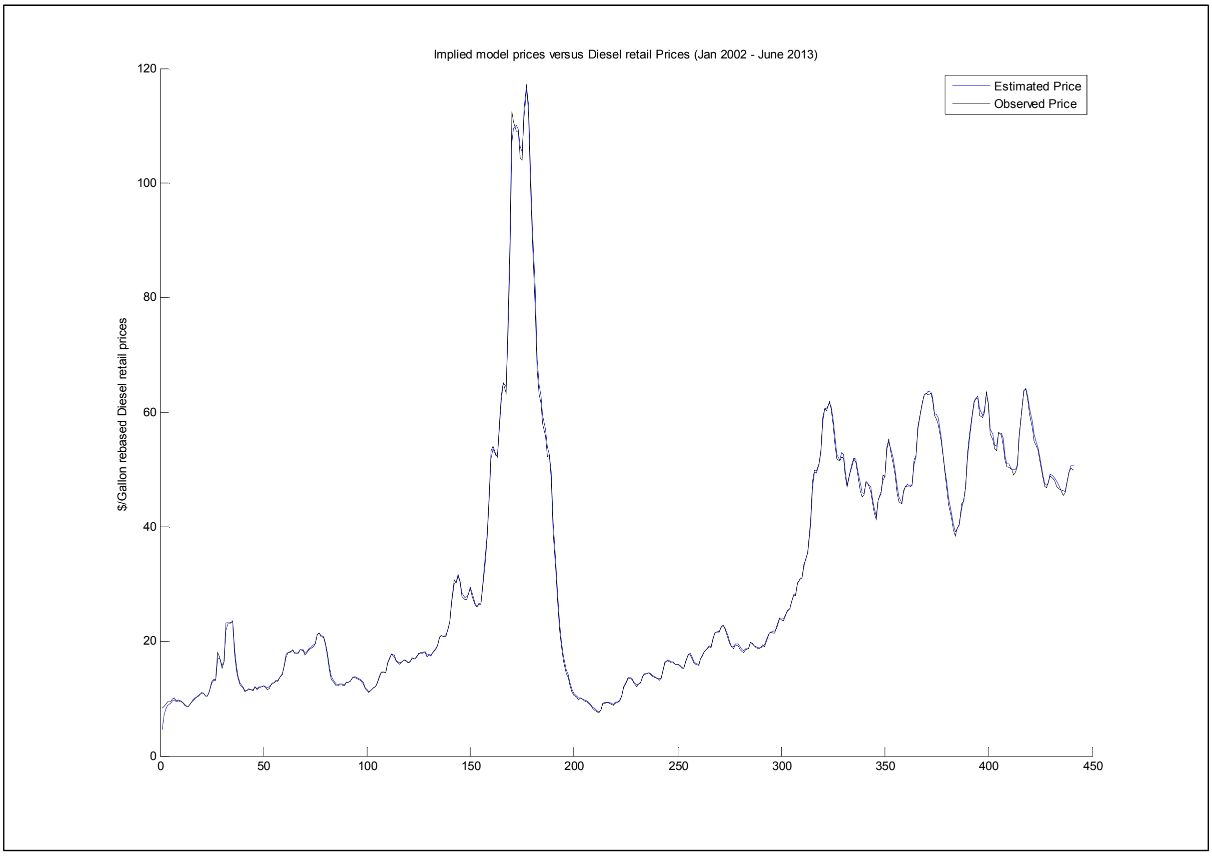
(Figure A.1. Gasoline Model prices versus Market Futures prices, NYMEX)



(Figure A.2 Naphtha model prices versus market residential prices, EIA)



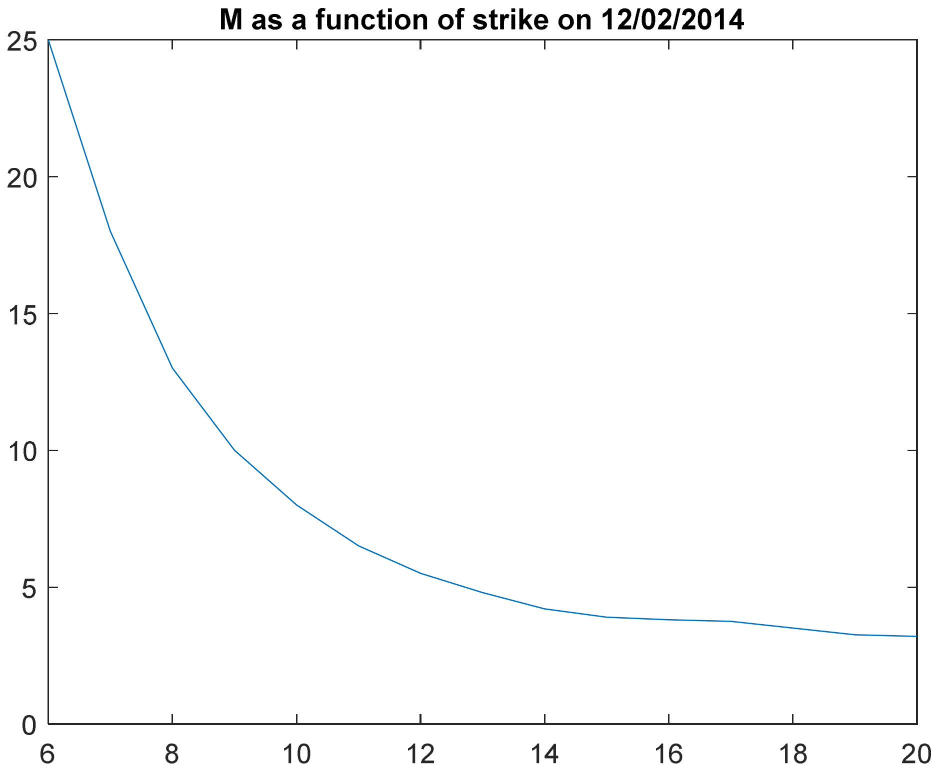
<u>(Figure A.3 Heating oil model prices versus market futures prices, NYMEX)</u>



(Figure A.4. Diesel model prices versus futures market prices, NYMEX)

A.2.0. Calibration of Alexander and Venkatramanan (2011)

We construct and calibrate this model exactly as described by the authors but with different dates on their data set. In the below figure we show the resultant m factor used in the model as a function of strike of the crack spread options:



(Figure: A.5. m as a function of strike K on the 12/02/2014, for option data from NYMEX)

A.2.1. Calibration of Hull and White two factor stochastic process

After the tree has been constructed for the process x, the spot process S can be defined at S(t) = x(t) + α(t), where α is a deterministic function. It is calculated by applying the Arrow-Debreu node prices and the market price of future prices.

We denote by $Q_{i+1, j, k}$ the present value of a commodity that pays 1 if the node (i+1, j , k) is reached and zero otherwise. These are all calculated recursively, knowing α_i and $Q_{i, h, l}$ for all (h, l), by:

$$Q_{i+1, j, k} = \Sigma_{h,l} \; Q_{i, h, l} \, q_i(h, l, j, k) \exp\{-(\alpha_i + x_{i, h, l}) \, \Delta t_i\} \quad (1)$$

Where $q_i(h, l, j, k)$ is the probability of moving from (i, h, l) to (i +1, j, k).

The α_{i+1} is found by solving:

$$S_M(0, t_{i+2}) = \Sigma_{i,j} Q_{i+1, j,k} \exp\{-(\alpha_{i+1} + x_{i+1, j, k}) \Delta t_{i+1}\} \quad (2)$$

i.e.

$$\alpha_{i+1} = (1/ \Delta t_{i+1})\ln[(\Sigma_{i,j} Q_{i+1, j,k} \exp\{ x_{i+1, j, k} \Delta t_{i+1}\})/ S_M(0, t_{i+2})]$$

The initial values for α and Q are: $Q_{0,0,0} = 1$ and $\alpha_0 = -\ln(S_M(0, t_1))/t_1$.

A.2.2. Trinomial tree construction for Hull and White two factor process

We have already seen how to construct a trinomial tree for a single factor process of the type $dx(t) = -a\, x(t)dt + \sigma(t)dW(t)$, we now use this technique for the two-factor process that represent the spread.

Firstly, we consider the process x verifying the same equation as S, with θ = 0:

$$dx(t) = [U(t) – a\, x(t)]dt + \sigma_1 dW_1(t), \quad x(0) = 0 \quad (3)$$

$$dU(t) = -bU(t)dt + \sigma_2 dW_2(t), \quad U(0) = 0 \quad (4)$$

If we assume that a ≠ b, then the dependence of x on U can be removed by:

$$Y = x + (U / b – a) \quad (5)$$

Hence,

$$dY(t) = -aY(t)dt + \sigma_3 dW_3(t), \quad Y(0) = 0 \quad (6)$$

$$dU(t) = -bU(t)dt + \sigma_2 W_2(t), \qquad U(0) = 0 \tag{7}$$

Where

$$\sigma_3^2 = \sigma_1^2 + \sigma_2^2/(b - a)^2 + 2\rho\sigma_1\sigma_2/(b - a) \tag{8}$$

And W_3 is a Brownian motion. The correlation between W_2 and W_3 is:

$$\rho_{uy} = [\rho\sigma_1 + \sigma_2/(b - a)]/\ \sigma_3 \tag{9}$$

The first step is to construct a tree for x with two trinomial trees, for both processes U and Y as above. Next we use the formula:

$$x = Y - U\ /\ (b - a) \tag{10}$$

The tree obtained for x will be a two-dimensional trinomial tree, where every node will have nine branches, a combination of branches U and Y.

At time t_i, we have nodes Y(i, h) and U(i,1), hence the node for x is x(i,h,1). We define j the index of the middle branch in the tree for Y, emanating from y(i,h) with corresponding probabilities pu, pm, pd, and define k the index of the middle branch in the tree of U, emanating from U(i, 1), with probabilities qu, qm, qd. Next starting from x(i, h,1), the process moves to nine branches x(i, j + ϵ_1, k + ϵ_2), where ϵ_1 and ϵ_2 take values 0, 1 or -1. Finally, the probabilities associated with each node of the nine branches is required. When there is zero correlation between U and Y (i.e. ρ_{uy} = 0), the matrix of probabilities for the nine branches is simply:

	U-Move			
Y-Move		Down	Middle	Up
	Down	$p_d q_d$	$p_d q_m$	$p_d q_u$
	Middle	$p_m q_d$	$p_m q_m$	$p_m q_u$
	Up	$p_u q_d$	$p_u q_m$	$p_u q_u$

In correlated processes such as ours, the elements above are shifted so that the sum of the shifts in each row and column is zero.

A.2.3. Calibration results for Hull and White two factor process

We estimate the five parameters of the model ($a,\sigma_1,b,\sigma_2,\rho$) fitting given observed market data (NYMEX crack-spread futures volatility and the European Call option implied volatility surface on the crack-spread). Both sets of values are found by aggregating existing contracts; for the spread itself we use three futures, and for the option we combine two option contracts on RBOB and Heating oil.

In this example we use the NYMEX crack-spread futures (calculated using RBOB, Heating oil and WTI futures contracts) volatility. The calibration is performed by minimizing the sum of the squares of the percentage differences between model and market future prices. For this purpose, we used an optimization algorithm, within GAMS that combines interior point methods and quasi-Newton techniques. We give an example of our results in the below table:

Table A.2.3.1: *The results of the calibration to crack-spread futures volatility on 12/02/2014*

Maturity (months)	Implied Volatility	Our Volatility
1	0.1637	0.1629
2	0.1614	0.1618
3	0.1594	0.1613
4	0.1530	0.1546
5	0.1465	0.1477
7	0.1373	0.1354
10	0.1475	0.1436
15	0.1528	0.1539
24	0.1594	0.1582

These results are obtained with the following Hull/White parameters:

$a = 0.54937$

$\sigma_1 = 0.004453$

$b = 0.073944$

$\sigma_2 = 0.00499$

$\rho = -0.97311$

Table A.2.3.2: *At-the-money European crack spread options-volatility quotes on 12/02/2014*

	1M	2M	3M	4M	5M	6M	7M	8M	9M	10M
1M	0.1834	0.1645	0.1630	0.1510	0.1540	0.1590	0.1499	0.1424	0.1373	0.1470
2M	0.1813	0.1604	0.1690	0.1594	0.1520	0.1592	0.1465	0.1430	0.1300	0.1485
3M	0.1664	0.1550	0.1640	0.1540	0.1590	0.1454	0.1433	0.1400	0.1288	0.1360
4M	0.1637	0.1560	0.1560	0.1494	0.1540	0.1420	0.1495	0.1370	0.1250	0.1335
5M	0.1485	0.1482	0.1460	0.1440	0.1403	0.1478	0.1455	0.1330	0.1122	0.1201
7M	0.1373	0.1490	0.1435	0.1450	0.1412	0.1390	0.1378	0.1360	0.1150	0.1230
10M	0.1475	0.1409	0.1460	0.1430	0.1403	0.1360	0.1361	0.1360	0.1150	0.1241

We report the calibration results in the table below that depicts the fitted crack spread option volatilities as implied by Hull/White two factor model crack spread prices backed out of the AV crack spread option formula.

Table A.2.3.3: *At-the-money Hull & White two factor implied volatilities.*

	1M	2M	3M	4M	5M	6M	7M	8M	9M	10M
1M	0.1937	0.1647	0.1632	0.1514	0.1555	0.1594	0.1509	0.1504	0.1383	0.1491
2M	0.1914	0.1604	0.1690	0.1594	0.1520	0.1592	0.1465	0.1438	0.1310	0.1487
3M	0.1794	0.1569	0.1640	0.1540	0.1590	0.1454	0.1433	0.1423	0.1242	0.1312
4M	0.1730	0.1549	0.1562	0.1494	0.1540	0.1420	0.1495	0.1386	0.1262	0.1394
5M	0.1565	0.1494	0.1468	0.1440	0.1403	0.1478	0.1455	0.1391	0.1131	0.1234
7M	0.1373	0.1483	0.1435	0.1459	0.1412	0.1390	0.1378	0.1366	0.1163	0.1234
10M	0.1475	0.1449	0.1460	0.1433	0.1426	0.1373	0.1359	0.1342	0.1201	0.1202

$a = 0.83372, \quad \sigma_1 = 0.017353, \; b = 0.090124, \; \sigma_2 = 0.00845, \rho = -0.39876$

A.2.4. Sub-tree calibration for Schwartz and Smith (2000) two-factor process

The model has seven parameters (κ, σ_χ, μ_ξ, σ_ξ, $\rho_{\xi\chi}$, λ_ξ , λ_χ) that need to be estimated for crude oil, heating oil and gasoline futures listed on NYMEX; we follow the approach in Jafarizadeh and Bratvold (2012), enabling the program to avoid implementing the Kalman filter. We can define the hidden factors in terms of the model parameters whilst fitting the futures curve and implied volatility curves on futures along the required tenors. This will allow NYMEX market data and enable a risk neutral approach. As described in Schwartz and Smith (2000) if $\phi = In(F_{T,t})$ and the volatility of $In(F_{T,t})$ is $\sigma_\phi(t, T)$, the value of a European call option on a futures contract maturing at time T with exercise price K, and time t until the option expires is:

$$c = e^{-rt} \{ F_{T,t}N(d) - KN[d - \sigma_\phi(t, T)]\} \quad (1)$$

Where

$$d = \{In(F/K)/ \sigma_\phi(t, T)\}\ 1/2\ \sigma_\phi(t, T) \quad (2)$$

Here the N(d) is a cumulative probability distribution.

We have collected data for options on crude, heating oil and gasoline futures, C_t, for the underlying futures price $F_{T,0}$, time to maturity T, strike price K; the options were reported on NYMEX on 12/02/2014. The options contracts approximately matched the maturity of the futures contracts (t = T); it is then easy to use the inverse problem to find the volatility $\sigma_\phi(t, T)$ associated with each option. The $\sigma_\phi(t, T)$ is described in terms of the parameters of the Schwartz and Smith model:

$$\sigma_\phi(t, T) = e^{-2\kappa(T-t)} (1 - e^{-2\kappa(t)})(\sigma^2_\chi / 2\kappa) + \sigma^2_\xi + 2\, e^{-2\kappa(T-t)}(1 - e^{-\kappa t})(\rho\xi\chi^\sigma\chi^\sigma\xi/ \kappa) \quad (3)$$

Assuming that the options expire at the same time as the futures then $e^{-2\kappa(T-t)} = e^{-\kappa(T-t)} = 1$. As T approaches infinity, the implied annualised volatility of the futures contracts will approximately be equal to the long term factor volatility:

$$\sigma_{\phi}(T, T)/\sqrt{(T)} \approx \sqrt{(e^{-2\kappa(t)}\sigma^2_{\chi} + \sigma^2_{\xi}T + 2e^{-\kappa T}\rho_{\xi\chi}\sigma_{\chi}\sigma_{\xi})} \quad (4)$$

However, for near maturity contracts it can be shown that the volatility becomes:

$$\sigma_{\phi}(T, T)/\sqrt{(T)} \approx \sqrt{(\sigma^2_{\xi})} \quad (5)$$

The deviation from the equilibrium at time 0 is χ_0 and the risk premium for the short term λ_{χ}. The log of the current spot calculation of the commodity price is the sum of ξ_0 and χ_0; hence we can write:

$$\chi_0 = \text{In}(S_0) - \xi_0 \quad (6)$$

Note that the risk premiums for the short and long terms cannot be estimated as they are not observed. In this risk neutral approach only the drift factor for the long term need be estimated eliminating the need for the market price of risk in this case. We calibrate all three underlying commodities separately and consolidate them to build the final trinomial tree for the crack-spread.

***(<u>Table A.2.4.1</u>:** Resultant Calibration results for Schwartz and Smith (2000) on WTI crude oil for Dec 2000 - Dec 2010)*

Parameter	Description	Value
χ_0	Short term deviation of the log spot price	0.22
ξ_0	Long term deviation of the log spot price	4.87
σ_{ξ}	Volatility of the long term	14%
σ_{χ}	Volatility of the short term	29%
μ_{ξ}	Risk Neutral drift rate for the long term	-7.27%
λ_{χ}	Risk premium for the short term	0
κ	Mean reversion coefficient	0.87
$\rho_{\xi\chi}$	Correlation coefficient	0.276

*(**Table A.2.4.2:** Resultant Calibration results for Schwartz and Smith (2000) on No.2 heating oil for Dec 2000 - Dec 2010)*

Parameter	Description	Value
χ_0	Short term deviation of the log spot price	0.6462
ξ_0	Long term deviation of the log spot price	3.476
σ_ξ	Volatility of the long term	30.75%
σ_χ	Volatility of the short term	84%
μ_ξ	Risk Neutral drift rate for the long term	-11.21%
λ_χ	Risk premium for the short term	0
κ	Mean reversion coefficient	3.334
$\rho_{\xi\chi}$	Correlation coefficient	-0.5966

*(**Table A.2.4.3**: Resultant Calibration results for Schwartz and Smith (2000) on RBOB NYH Gasoline for Dec 2000 - Dec 2010)*

Parameter	Description	Value
χ_0	Short term deviation of the log spot price	0.16
ξ_0	Long term deviation of the log spot price	3.84
σ_ξ	Volatility of the long term	30.74%
σ_χ	Volatility of the short term	30.61%
μ_ξ	Risk Neutral drift rate for the long term	-15.5%
λ_χ	Risk premium for the short term	0
κ	Mean reversion coefficient	0.21
$\rho_{\xi\chi}$	Correlation coefficient	-0.84

There are four steps for calculating the parameters in the above process:

Step 1 - Estimating σ_ξ

Calculate the implied volatilities for a long maturity futures contract on the commodity using equation (1); then use equation (3) to calculate σ_ξ.

Step 2 - Estimating μ_ξ and κ

Construct the log of the futures curve from the observed futures price contracts; next we calculate the gradient of the log of the futures curve, and finally estimate the μ_ξ by subtracting (1/2 σ^2_ξ) from the gradient.

By calculating the half life from the futures curve = In(2) / κ; we can use it estimate κ.

Step 3 - Estimating σ_χ and $\rho_{\xi\chi}$

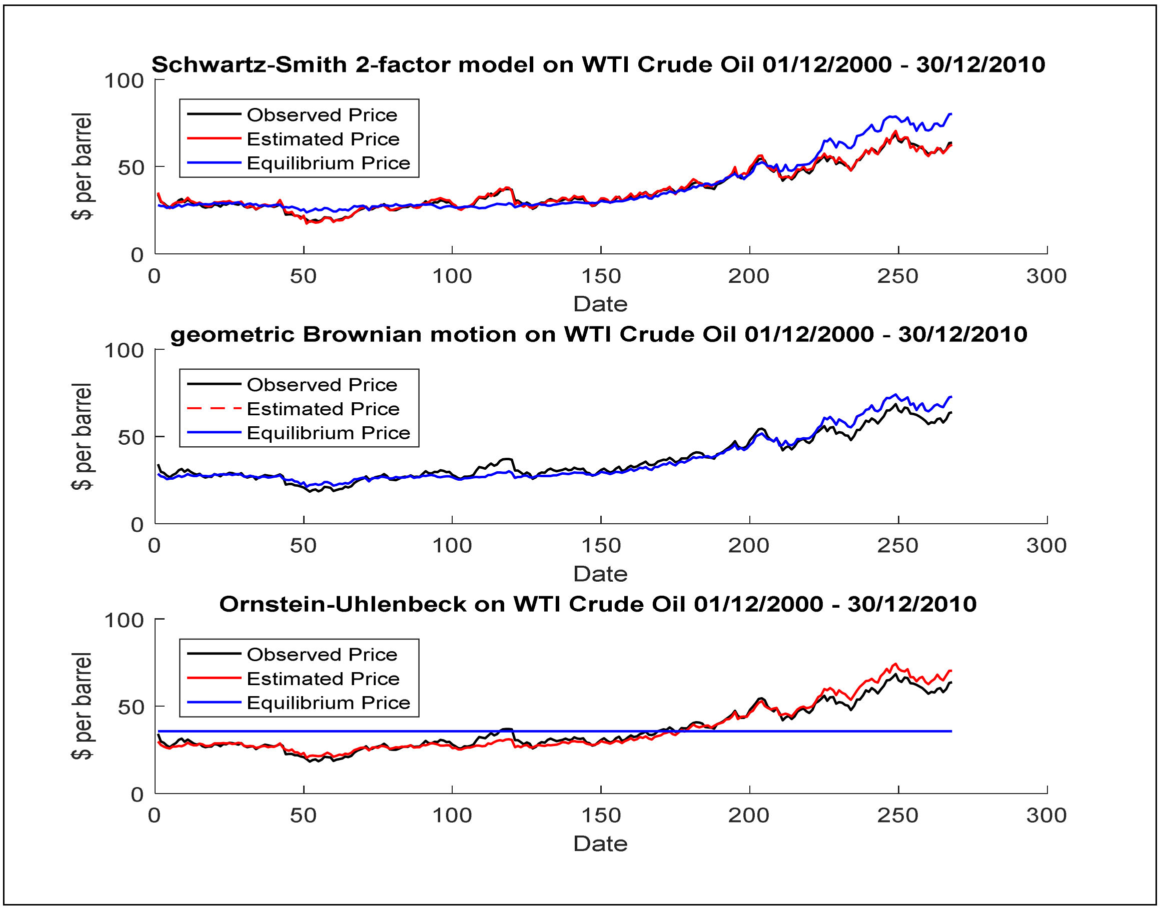
Using equation (1) we calculate the implied volatilities of two near maturity futures contracts on the spread and build a set of equations by inserting different implied volatilities into equation (4). Next we insert the estimated parameters from steps 1 and 2 into the set of equations and solve for σ_χ and $\rho_{\xi\chi}$.

Step 4 - Estimating ξ_0, χ_0, and λ_χ

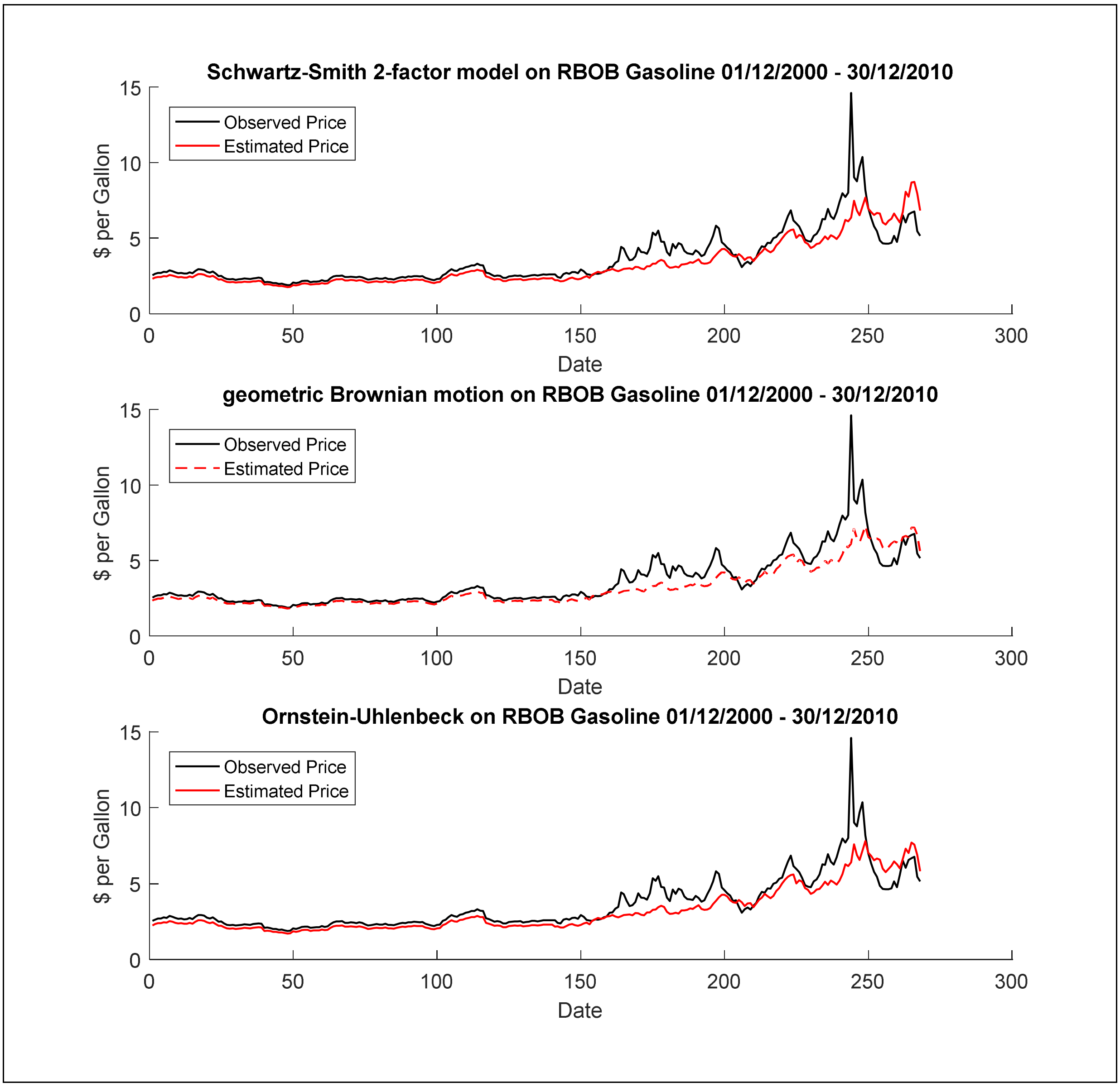
We set λ_χ= 0; using equation (5) and (6) we build a set of equations and solve them to find ξ_0 and χ_0.

Consolidating the results for all three commodities and applying the above processes in equations (2) and (3), we next construct a trinomial tree of the crack spread stochastic price process. We do this by following Hull & White's trinomial tree building process as already described in appendix 6.3; we do this for the long and short term processes and sum the results; we now have a crack spread price over 72 monthly periods; simulating 40 steps out from each node then enables dynamic programming to be applied resulting in the crack spread option prices. We name this approach a sub-tree simulation. The results of the individual commodity calibrations are shown below; alongside a GBM and Ornstein-Uhlenbeck process to emphasise the accuracy of the model.

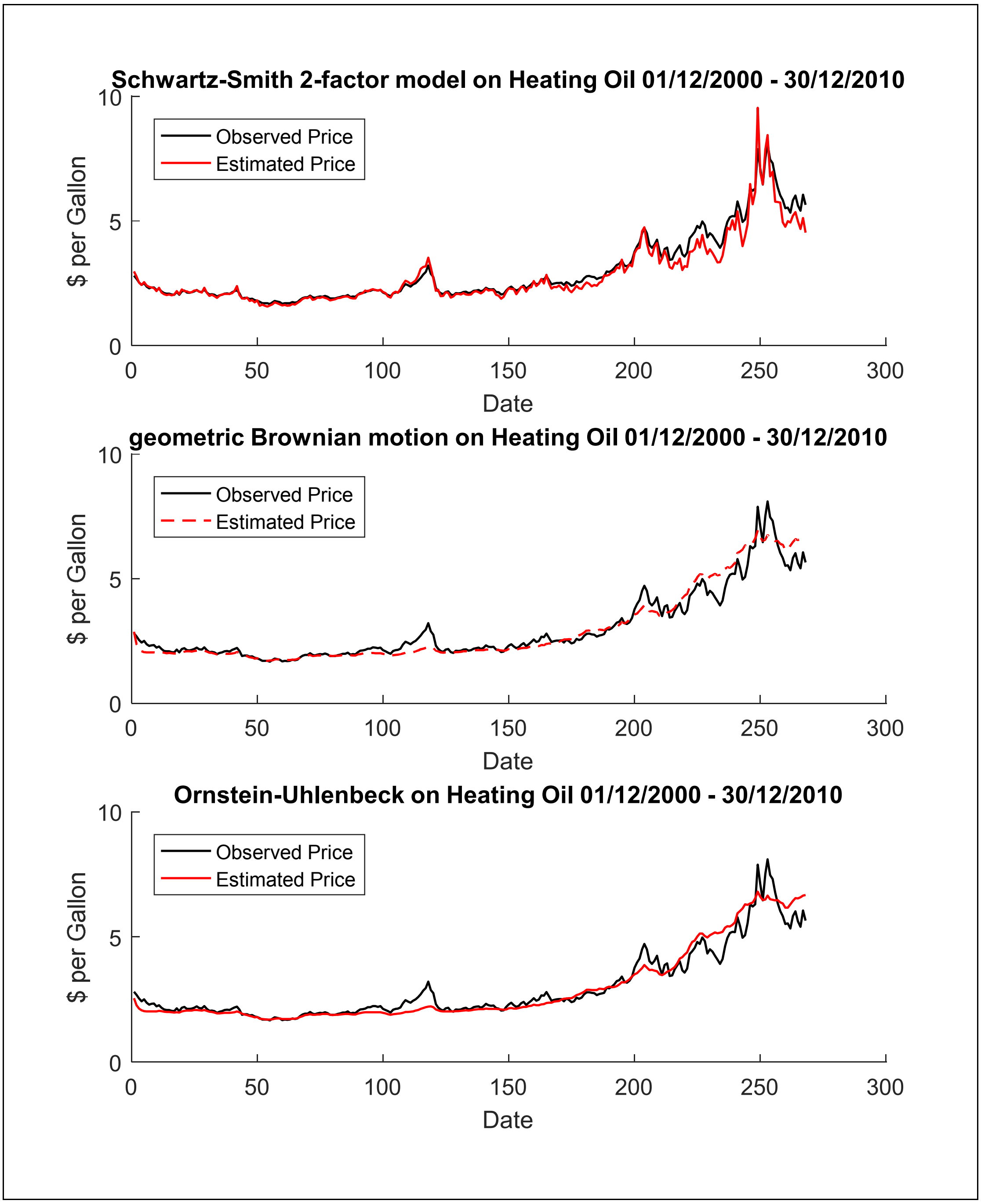
(Figure A.2.4.1 *WTI Crude Oil calibration results using Schwartz and Smith (2000)):*



(Figure A.2.4.2 *RBOB Gasoline calibration results using Schwartz and Smith (2000)):*



__(Figure A.2.4.3__ No.2 Heating Oil NYMEX calibration results using Schwartz and Smith (2000)):):



A 3.0 Comparison of forward curves for the one factor and two factor models

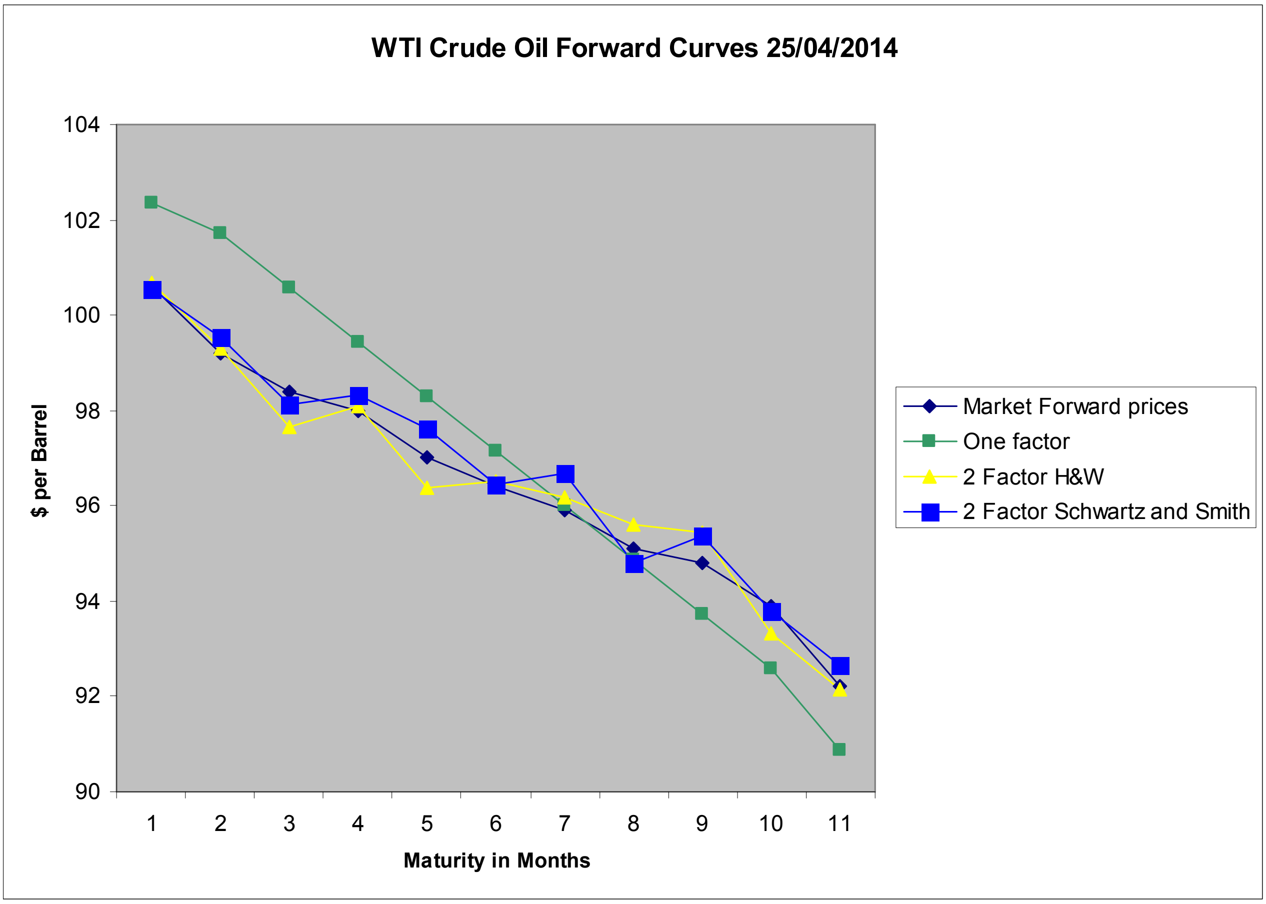


Figure A.3: Comparison of calibration results of the forward curve on WTI oil using one and two factor models.

Published Paper

A modern two-stage stochastic programming portfolio model for an oil refinery with financial risk management

Patrick Johnson O'Driscoll
Department of Economics, Mathematics and Statistics, Birkbeck,
University of London, Malet Street, London, UK, WC1E 7HX
podris01@mail.bbk.ac.uk

Abstract

The proposal which we wish to make is a two-stage stochastic programming model for a competitive oil refinery with stochastic crude and fuel prices. Most models for refineries are deterministic, and those considering the stochastic problem do so by utilising a Gaussian assumption on profits - implementing variance as the risk measure. Our model falls into the category of optimisation with coherent risk measures where robustness, rather than ambiguity, is the focus. The objective is to maximise the refiner's profit under raw material, product inventory constraints and a financial risk constraint. The two-stage model leverages off a unique discrete scenario generation technique alongside an admissible and computational tractable drawdown risk measure. The expected value of perfect information calculation of each model gives a value for the additional benefit, which the decision maker receives in considering the uncertainty inherent in the problem.

Keywords: Stochastic programming; Refinery planning; Optimization under uncertainty; Probability fitting; Mean-variance; Conditional Drawdown-At-Risk (CDaR)

1. Introduction

Refiners are exposed to high uncertainties due to the nature of the oil markets. Recently, they have had to contend with much lower refining margins. This is due to the crude oil market's lack of spare capacity at the refinery level; hence risk management then becomes a more pressing issue for refiner consumers and producers - knowledge of the product prices can help to reduce this risk. The refiner's portfolio includes his own production and a set of sales and costs over one period of time. In this paper we develop a stochastic portfolio model with the aim of maximising the refiner's gross refining margin (GRM).

The goal of such a model is to reduce the economic risk; this is connected to the fact that the oil spot price may be highly volatile due to various unpredictable causes. The basis risk factors include the wholesale spot crude and refined products prices. Optimisation models generally assume that market prices are unaffected by the decision of the utility manager, and by the uncertain yields within the units of the refinery. The model we propose differs from those typically discussed in the literature, since we want to concentrate on the financial risks; an example is shown in Xidonas, P. et al (2010). We ignore the daily detailed operational refinery flows as analysed for instance by Ribas et al (2012).

In the two-stage stochastic programming approach the first stage production variables are selected; being unrelated and independent to the uncertain events. The second term manages the expectation of the uncertain events. The stochastic literature on this subject differs in the way the expectation and risk are considered. In practice, the two stages are usually considered to be over one month, with the model being rerun after the uncertain events materialise and the second stage recourse decisions have been made. The objective function can minimise costs or maximise profits, we optimise profit.

In section 1, we introduce the current refinery planning and financial risk literature. In section 2 we formulate our two stage stochastic program, creating the foundation for the rest of the paper. In section 3 we implement the measures for financial risk management of the refinery. In section 4 we state and analyse the results for each model implemented. We conclude in section 5 and propose extensions.

1.1. *Refinery Planning*

Benyoucef (2010) presents a non-linear stochastic program representing a network of refineries in Algeria. Scenario generation is utilised to depict the demands of the refined products and uncertainties are represented by normal distributions. To ensure feasibility and maintain the model's realism, a penalty term is included in the objective function, where the aim is to minimise costs in the presence of necessary constraints. We focus on a similar objective as the author, but for a single entrepreneurial refinery where we explicitly capture the stochasticity of the oil prices. The author does however introduce a scenario tree attempting to solve the problem over a number of time periods, yet suffers the negative implications of using standard deviation as the risk measure. We model the prices rather than the demands, and fit distributions to capture more accurately the behavioural aspects of the time series' characteristics. Uncertainty is usually represented by the variation in the prices inherent in the problem. Leiras, A. et al. (2011) summarise the numerous categories of refinery planning tackled in the literature. They show that fewer authors formulate a model with stochastic variables to capture uncertainty.

Most models of petroleum refineries are linear and avoid the corporate planning requirement. Neiro and Pinto (2003) present the problem of managing multiple refinery operations using a mixed integer non linear program (MINLP). Their objective function maximises the net present value under raw material and product inventory constraints. These include mass balance and operating constraints for each refinery in their network. Instead of using the market prices for assets, demand scenarios are generated. The authors consider the refinery supply chain in their optimisation, however, financial risk measures are ignored, and scenarios given are drawn from normal distributions. The disaggregation techniques used here are useful; enabling the large scale MINLP to be solved. Overall there are 383 variables and 349 equations with the algorithm DICOPT used to solve the planning problem. Another example of applying a MINLP is constructed in Aldaihani, and Al-Deehani (2010); the efficiency of the Kuwait Stock exchange is examined in an optimisation run, where the scenarios of an equally weighted basket of stocks are simulated. The authors coerce mathematical programming with financial strategy, aiming to lower the risks for investors whilst reaching a particular level of financial return.

Ribas et al (2012) present a risk averse and risk neutral approach to maximising the profit of a refinery; uncertainty is introduced in the product prices, the oil supply and the capacity process. The authors test their model using data from the Brazilian refinery system, and measure robustness as the objective of the model rather than the profit. The scenarios are generated by requesting an expert's opinion on oil price uncertainty inputs, which gives an expected value of perfect information (EVPI) of 2.54%. An example of scenario generation using sample average approximation (SAA) is shown in Elkamel, A. and Al-Qahtani, K. (2011), the model emphasises the added value of a petrochemical network over the deterministic approaches previously defined by the authors – the model is tested on an industrial case of multiple refineries and a PVC complex. Tahar, R.B.M. and Abduljabbar, W.K. (2010) apply a new genetic based scenario simulation algorithm to short term scheduling - maximising the profit whilst considering the transportation and operational costs of the oil refinery which we omit.

1.2. *Financial Risk*

If we wish to measure the risks that the refiner faces, which measure function should we consider? In practice models that do not consider risk will not be used; Uryasev and Rockafellar (2002) present a number of optimising problems in linear and non linear settings considering different risk measures, which still remain relevant today. They show that by using Conditional Value at Risk, (CVaR), global optimums can be found, and tractability will be hugely improved over typical measures. Therefore, CVaR is noted as more conservative, yet more realistic. There are many different methodologies for modelling uncertainty within optimisation, Bertsimas et al (2010) show that if the underlyings have a known probability distribution, the ideas of ambiguity and robustness can be combined to provide a favourable optimisation. Our problem is similar in that, the data is assumed to have a probability distribution; the difference is that the authors approach the problem with a soft robust optimisation – whereas we model with the closely related convex set of risk measures. We then fit the data and generate scenarios using a unique approach focussing on uncertainty and not on ambiguity.

Our optimisation model includes, in the final case, the drawdown function as specified in Cheklov et al (2003), where distributions can be considered for the underlyings. The model is multidimensional, and the optimisation is carried out with discrete scenario values – which are conducive to our scenario generation. The drawdown function utilised satisfies the axioms of a deviation measure: non-negativity, insensitivity to constant shift, positive homogeneity and convexity.

2. The Two-Stage Stochastic Program

In the current competitive economic climate, global businesses are constantly confronted with making decisions in an uncertain environment. One motivation for refinery stochastic programs is to be able to quantify such decisions in terms of profit. For example, Mulvey's Towers Perrin-Tillinghast asset liability model (ALM) saved US West $450 million in opportunity costs in its pension plan. The two-stage stochastic program discussed thus far, aims to aid refineries in this process of dealing with uncertainty in a similar way. During each time period the decision maker in the refinery needs to decide how much crude to purchase, and how much of each refined product to release onto the market. The prices and decisions made today are assumed known, the "here and now" problem. In the next time period the prices will change, and second stage production variables will have already been made. The "wait and see" decisions are made after knowledge of the second stage prices has become known. The "stochasticity" appears in this second term in the form of uncertain exogenous variables. The aim in midterm planning is, to choose the first stage decisions in such a way that the second stage variables can be added to the first to give a higher expected profit. Gupta and Maranas (2000) were the first to use the scenario analysis approach; it has become a significant method in practice. For an example applied to real electricity options with stochastic simulation, see Culik, M. (2010). The author considers four different options available to the project manager, which are priced into the contracts and ultimately the net present value obtained for the project. One contribution of our work is to provide a unique way in which to generate the prices using probability distributions. Our key assumption here is that the price series can indeed be represented by a set of probability distributions. The representative scenarios are required to capture the uncertainty in the random price variables within the stochastic programming framework. These two-stage models are implemented at refineries as they give a detailed view of decisions in the short/mid-term.

2.1. The producer midterm planning model

The tactical midterm refining planning problem is very complex; hence the decision based systems used at refineries to aid owners, e.g. PIMS (AspenTech), which utilises mathematical programming, is present at most refineries in the world, see (Mann 2003). In this paper refining is modelled at the level of detail common in the literature - with a granularity of one day at its finest, with start-up and shut-down costs considered insignificant, whilst Wang (2013) manages this issue by utilising a finite planning horizon model with periodic preventative maintenance, and random failure. We start by introducing the model of the refinery with daily periods, although decisions are considered in monthly stages. The refinery consists of a number of units, the crude distillation unit (CDU), the cracker and the vacuum distillation unit (VDU). We model the refinery complex by considering the volumes bought and sold to maximise the decision maker's profit within each period. Few decide to approach this problem stochastically due to its complexity, and even fewer integrate the financial planning problem – operational planning is much more common.

The refinery producer must schedule the production of each refined product. This is expressed as the price of the particular product multiplied by the volume. As practised in the literature there is no restriction on the amount of crude oil that can be purchased except the capacity constraint, and no minimum amount that needs to be produced of any refined product. The decision variables of the refinery scheduling problem are stated below:

$X_{i,t}$ [tonnes/day] : volume of product *i* bought or sold on this particular day

7. $X_{1,t}$: Crude Oil
8. $X_{2,t}$: Gasoline
9. $X_{3,t}$: Naphtha (after the splitter)
10. $X_{4,t}$: Jet Fuel
11. $X_{5,t}$: Heating Oil
12. $X_{6,t}$: Fuel Oil
13. $X_{7,t}$: Naphtha stream exiting the PDU
14. $X_{8,t}$: Gas Oil
15. $X_{9,t}$: Cracker Feed
16. $X_{10,t}$: Residuum
17. $X_{11,t}$: Gasoline after splitting of Naphtha exiting the PDU
18. $X_{12,t}$: Gas Oil after the splitter
19. $X_{13,t}$: Gas Oil stream entering the fuel oil blending facility
20. $X_{14,t}$: Cracker Feed after the Splitter
21. $X_{15,t}$: Cracker Feed stream entering the fuel oil blending facility
22. $X_{16,t}$: Gasoline stream exiting the cracker unit
23. $X_{17,t}$: Stream exiting the cracker unit into the splitter
24. $X_{18,t}$: Heating oil stream after splitting of cracker output
25. $X_{19,t}$: Cracker output stream

Only the decision variables where *i=1 to 7*, and *i=14,* are required within the objective function. They represent the crude being bought and refined products being sold; the other variables are present in the constraints. The values assigned to the decision variables must satisfy the following constraints, which represent the physical restrictions on this particular refinery, (see Appendix A for details):

The capacity constraints

The mass balance constraints for the following units:

1. Primary Unit
2. Cracker
3. Fixed Blends
4. Unrestricted Balances
5. Raw Material Availability

The following general model is used for the purpose of monthly scheduling; it is the deterministic objective function of *Ravi and Reddy (1996))*:

Maximise

$$\text{Profit} = -\,8.0\,X_1 + 18.5\,X_2 + 8.0\,X_3 + 12.5\,X_4 + 14.5\,X_5 + 6.0\,X_6 - 1.5\,X_{14} \quad (1)$$

In equation (1) the authors replace the stochastic commodity prices with their mean prices from a set of normal distributions - in fact, it is well known that commodity price series are very far from Gaussian.

2.2. From deterministic to stochastic

Updating the above deterministic objective model to today's mean prices of crude and refined products gives ***Model 1***:

Maximise

$$\text{Profit} = -385X_1 + 726X_2 + 385\,X_3 + 471X_4 + 520X_5 + 251X_6 - 72X_{14} \quad (2)$$

The negative terms are the purchasing and operational costs. The positive values are the saleable product prices. This is a linear optimisation problem; hence the CPLEX algorithm is applied within the optimisation program, General Algebraic Modelling System (GAMS). The solution to equation (2), the deterministic objective function, is given below:

Objective Value = $285, 418 per day

This approach does not take into account the stochasticity of the prices; therefore it ignores the decision maker's choice to switch to selling alternative amounts of each product.

When including random prices in an optimisation model, one of the following three methods is implemented:

1. Mean value prices replace the stochastic variables
2. Continuous distributions are manipulated
3. Scenario generation (using discrete distributions)

Next, the prices of refined products and the crude oil are made stochastic and the problem solved using scenario generation (type 3 above). The expectation of the product prices in the next stage is the uncertain component of the profit. Evaluating uncertain functions and their expectations is the core involvement within stochastic programming. Clay et al (1998) used the certainty equivalent transformation, and Ierapetritou et al (2007) used quadrature in order to solve the expectation function. The expectation function can be made more complicated if the distribution of the profit is considered continuous. In which case, a more complex method is required to deal with the multi-dimensional integral. Our aim is to "discretise" the objective function as follows:

$$Z = \sum_{S=1}^{NS} p_s g(X_s) \qquad (3)$$

Where X_s denotes the random parameter vector under scenario S, $g(X_s)$ denotes the objective function, and $p_{s,}$ denotes the probability of scenario S. In a Monte Carlo approach, Z is estimated by random sampling. For our stochastic program (SP) the distribution of the profits are not considered directly, but instead historical data is analysed to generate possible scenarios for the uncertain parameters in the future. Prices are examined to determine a distribution of closest fit; rather than assuming normality. There is however no perfect fit, yet to model the expectation, a distribution is required to generate the scenarios. The "bracket-median" approximation was implemented to obtain scenario based values, and their corresponding probabilities.

2.3. Multivariate Fitting and Scenario Generation

We model the key uncertainties by cumulative distribution functions (CDFs) defined over uncertain quantities - these are the prices of the refined products and crude oil. In practise the CDF is obtained in one of two ways: judgemental assessments from an "expert", or from historical data.

This results in several points from each of the CDFs being calculated. Wettergren, T.A. and Baylog, J.G. (2014) confront a discrete search planning problem where the object that is being searched for determines the cumulative distribution functions. They find an optimal solution over a finite horizon using a greedy algorithm, with the uncertainty captured by perturbations on the known prior probabilities. In risk analysis, Monte Carlo simulations or triangular function fitting for three points are frequently applied.[56] In decision analysis for uncertain outcomes, the discrete distribution serves as a substitute for the entire continuous distribution. Without this simplification the computational difficulties can increase dramatically and solutions can be impossible to calculate.

Scenario generation is accounted for in SPs by using a scenario tree: internal, stratified or even random sampling can be used to improve computation. Scenario generation, by the very fact it has far fewer outcomes, means a computational advantage. The Basel Committee, for example, proposes using scenario based risk management for financial companies like Investment Banks. The four leading approaches within the literature for scenario generation are:

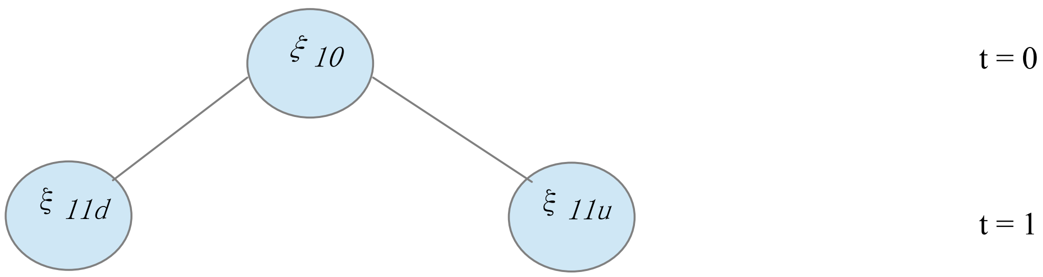
- Sampling
- Statistical Approaches
- Simulation
- Hybrids

Decision models like SPs are useful because of their ex ante and ex post properties. With sampling approaches, we are extracting numbers from a pdf, so that for a given X value we have the associated probability that this is the scenario value and its corresponding branch probability. However, as noted in Birge and Louveaux (1997), the scenarios are particularly useful when the optimal solution to a stochastic program varies considerably with changes in the value of the stochastic variables. The scenarios must comply with the *non-"anticipativity"* concept: we do not know the future and so decisions for the future can only be based on the past behaviour of the asset prices. These concepts are implemented in our models.

A scenario tree is illustrated in the figure below. At a point in time the nodes represent the state in the world - choices or decisions will be made at each node. The tree has branches for each value of

[56] Three-Point Approximations for Continuous Random Variables
Author(s): Donald L. Keefer and Samuel E. Bodily, 1983

the random vector representing refinery product prices, $\xi_t = \{\xi_{1t}, \xi_{2t}, \ldots \xi_{nt}\}$, in each time stage $t = 1,\ldots\ldots,T$.



(***Figure 1:*** A Scenario Tree example for product one)

It is known within the oil industry that the market structure has shifted during the last ten years. Questions about the properties of the price series being mean-reverting are constantly reoccurring see H. Geman (2005). In the past, fundamental supply and demand factors affected the crude oil price predictably. In the last decade any shock to the refining system has had a resonating impact on the price of crude than observed historically – this is due to refinery spare capacity reduction. In contrast to the previous decades and particularly in the US, many refineries have closed and there are now many financial markets trading paper or derivative contracts with crude oil as the underlying. Sources state that only 20% of the value in crude is through trading the raw physical product itself. Due to these intense refining markets and financial influences there is a strong motivation for refinery models with the necessary components. To capture the relevant crude dynamics that are now present; ten years of historical price data is chosen as input for the scenario generation.

The scenario's prices and probabilities are generated, whilst taking into account historical correlations, using the method described by Iman and Conover (1982), where the target correlation matrix is shown below in Table 2. We consider a monthly time period, from $t=0$ to $t=1$; the CDFs are fitted using historical data. Next, discrete probabilities are derived to calculate the expectation term. Often practitioners use a standard non-parametric bootstrap; however the problem of how to generate a correlation structure in the child nodes is not alleviated. We define five states of the world, where state one is the set of prices regarded as in a low economic state, and state five as the state where the set of prices are regarded as in a high economic state.

The selected CDF distributions are significant when fitted to the series data, and the correlations shown below are calculated from the same sample data. Copulas are used extensively within Finance to generate multivariate random variable scenarios. One of the problems with copulas is that they still require a rank correlation definition. We choose the Spearman's rank correlation method using a ranking procedure to generate a set of correlated commodity price series.

	WTI	**Gasoline**	**Naphtha**	**Jet Fuel**	**Heating Oil**	**Fuel Oil**	**Cracker Feed**
WTI	1						
Gasoline	0.9862	1					
Naphtha	1	0.9862	1				
Jet Fuel	0.967	0.985	0.967	1			
Heating Oil	0.9723	0.985	0.9723	0.743	1		
Fuel Oil	0.937	0.943	0.937	0.984	0.9901	1	
Cracker Feed	1	0.9862	1	0.967	0.9723	0.937	1

(***Table 2:*** Historical Correlations)

Asset	Expected Value	Standard Deviation	Skewness	Kurtosis	Resultant Fitted Distributions
WTI	249.15	242.68	1.76	6.89	Exponential (249.15)***
Gasoline	471.46	199.32	1.12	4.49	Gamma (2.6302, 123.27) ***
Naphtha	520.97	483.91	1.689	6.70	Exponential (513.06) ***
Jet Fuel	366.25	324.19	3.13	15.95	Pearson 5 (2.84, 677.86) ***
Heating Oil	251.31	270.20	3.74	23.644	Pearson 5 (3.1008, 677.85) ***
Fuel Oil	726.41	198.22	0.622	3.36	Gamma (14.112, 52.301) ***

Cracker Feed	72.09	34.17	1.25	4.57	Gamma (2.4438, 20.947) ***

(***Table 3:*** Statistical Properties derived from the fitted CDF distributions in units of tons per day.
****Significance at the 99% level.*

2.4. Objective function for the refinery including stochastic prices

The profit function can be split into each state, with its corresponding probability, and the expectation term can be made discrete by expanding the expectation as follows:

$$E[profit] = \sum_{products} \sum_{scenarios} p_s (Sales - Costs) X_i =$$

$$0.2\,(-271X_1 + 616X_2 + 271X_3 + 309X_4 + 407X_5 + 182X_6 - 51X_{14})$$

$$+\,0.2\,(-327X_1 + 652X_2 + 327X_3 + 387X_4 + 504X_5 + 195X_6 - 61X_{14})$$

$$+\,0.2\,(-385X_1 + 726X_2 + 385X_3 + 471X_4 + 520X_5 + 251X_6 - 72X_{14})$$

$$+\,0.2\,(-446X_1 + 866X_2 + 446X_3 + 573X_4 + 5279X_5 + 310X_6 - 84X_{14})$$

$$+\,0.2\,(-500X_1 + 1036X_2 + 500X_3 + 939X_4 + 586X_5 + 328X_6 - 94X_{14}) \qquad (5)$$

Where, X_i is the decision variable representing the volume of products to be bought and sold.

i = 1, 2, 3, 4, 5, 6, 14 ϵ (Products Index)

s = 1, 2, 3, 4, 5 ϵ (Scenarios or States of the world)

In the following approaches, different risk measures are considered to find the most effective construction. The literature varies on the choice of risk measure, and how the decision maker's preferences are taken into account.

The prices in each scenario are independent by definition, but within each scenario the product prices are conspicuously correlated with each other, and not independent as stated for example in

Bernardo et al (1999). Therefore, the prices are considered constant, and the amounts that go into each product as the random variables. The constraints for the models are defined in the Appendix, which describes the mass balances for the refinery units and physical limitations held onsite at the refinery complex. By changing the risk measure considered, *RM*, and the aversion to risk, α, and maximising the profit, Π, a Mean-Risk model is constructed as shown in equations 9-11 below.
In terms of mean variance, we now write the stochastic model as:

$$\textit{Maximise} \quad \pi_1 = E[\pi_0] - \alpha\sqrt{Var(\pi_0)} \tag{9}$$

Where constraints of the refinery are shown in the Appendix

Due to rules of *Programming* this can also be reformulated to the following:

$$\textit{Maximise} \quad \pi_1 = -\alpha\sqrt{Var(\pi_0)} \tag{10}$$

Such that $E[\pi_0] \geq$ A Target Objective function value

Where constraints of the refinery are shown in the Appendix

In this paper the construction modelled enabled us to define a target profit threshold, and solve for the decision variables:

$$\textit{Maximise} \quad \pi_1 = -\alpha\sqrt{RM(\pi_0)} \tag{11}$$

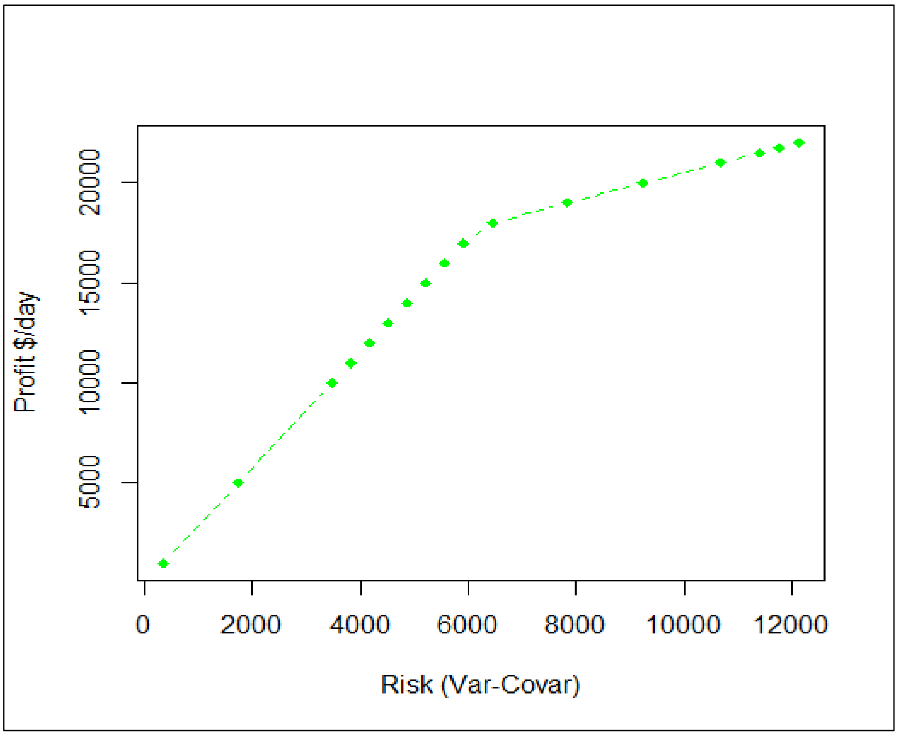
Such that $E[\pi_0] \geq$ A Target Objective function value

Where Constraints of the refinery are shown in the Appendix

The result of the construction is shown in five models detailed in section 4.

Equation (11) is clearly a non-linear and piecewise construction. In fact, if it is non-convex, a convex approximation is required, whereas the convex formulation is trivial to optimise.

Gupta and Maranas (2003) explain why any risk measure implemented should not be symmetric. If the investor or decision maker is considering minimising risk, it should be the downside only. Therefore, various risk measures are investigated. The following sections detail the alternative risk measure refinery models, which are generated on a sample of 1,000 draws from the fitted distributions. Shown in the figure below is the result of Model 2.



(***Figure 2:*** *The efficient frontier for the profit function in $ per day for the GRM (Model 2)*)

A maximum of $20,000 with a risk of $12,000 per day is achieved. By applying a utility parameter the profit and risk can be adjusted higher, however their underlying relationship does not change.

3. Risk Management

Risk management is key for oil refineries – market conditions are tougher than ever and managing the volatility of commodity related prices is an intricate job.

3.1. Value-At-Risk (VaR)

"A risk-taking institution that does not compute *VaR* might escape disaster, but an institution that cannot compute VaR will not." [57]

The value at risk measure was introduced by JP Morgan (RiskMetricsTM 1995) to measure market risk, although it can be used for credit and/or operational risk for example. It was invented as a predictive (ex ante) tool to prevent fund managers from exceeding specified portfolio policies. It relies upon three parameters: the time horizon; the confidence interval; and the highest amount of value that can be lost in the given time horizon, under normal market conditions. For the two-stage

[57] Aaron Brown (June/July 2008). " Private Profits and Socialized Risk ". GARP Risk Review.

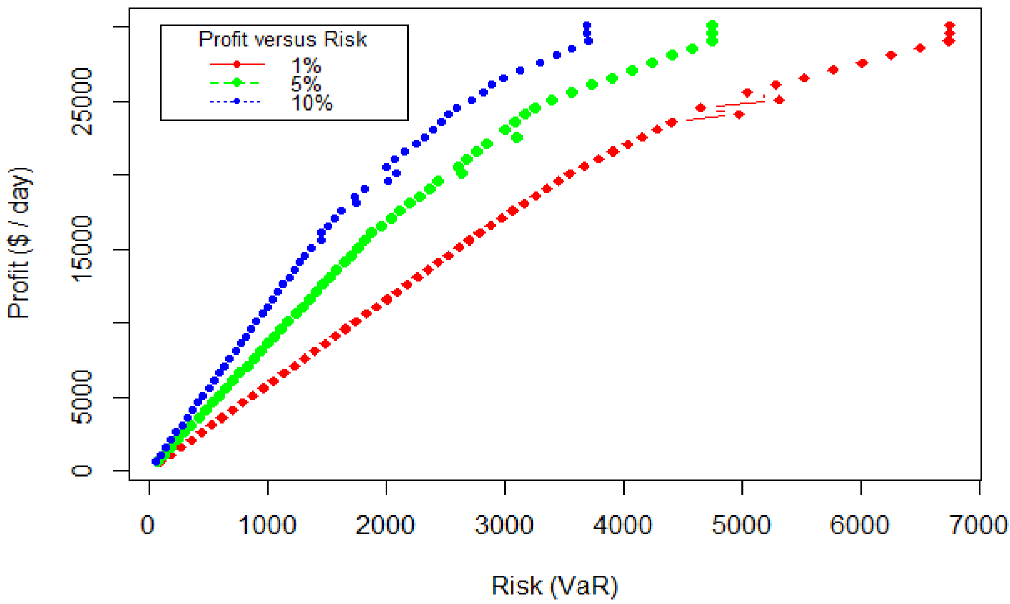
stochastic problem with a finite number of scenarios as in this paper, VaR is calculated by sorting the scenarios in ascending profit order, and then taking the profit value of the scenario for which the cumulative probability equals the specified confidence level. The definition follows:

Definition 1 (Value at Risk (VaR)).

For any confidence level α ϵ (0,1], the value at risk, denoted by VaR(α), for a random variable with payoff h, is defined as the 'critical value' at which the probability incurring a loss of no less than VaR(α) is at least α. That is,

$$VaR(\alpha) = \sup\{u : \Pr\{h \le h_0 - u\} \ge \alpha\}, \forall \alpha \in (0,1] \tag{12}$$

Where h_0 is the initial payoff amount. The different levels of α can be considered as differing levels of investor utility:



***(Figure 3:** The efficient frontier for the profit function, versus the value at risk (Model 3))*

As can be seen from the results, optimising with value at risk produces a non-convex shape, which exhibits local minima and is of combinatorial character. Computationally it is considered an unreliable risk measure and thus difficult to optimise. Basak and Shapiro (1999) show theoretically, that optimal decisions based on VaR, result in higher risk exposure than when decisions are based on expected losses.

3.2. Conditional Value-At-Risk (CVaR)

If for example, a portfolio is in a state of negative P&L; the magnitude of the losses is not captured accurately by *VaR*. Additionally, asset prices are rarely distributed normally; therefore, measures other than VaR should be applied. Common alternatives to VaR as a risk metric are expected shortfall, variance, and mean absolute deviation (MAD). Variance for example, penalises up movements of variance as well as downward movements, MAD does not penalise outliers, and VaR ignores large outliers in the distribution. For discrete distributions, *CVaR* is defined as the weighted average of those losses exceeding *VaR*. However, within the last decade, as stated earlier, the crude oil price series is no longer regarded as a mean reverting asset, and therefore, CVaR is the immediate progression over the models considering VaR within an optimisation. Furthermore, CVaR within the optimisation literature has not been applied to a single entrepreneurial oil refinery.

If returns are discrete and/or non-normal then sub-additivity is not necessarily present, and diversification of a portfolio may increase VaR. CVaR as a risk measure is coherent, as defined below, "measuring risk without sub-additivity is like measuring the distance between two points using a rubber band instead of a ruler "[58]:

Definition 2 (Coherence). *A Functional p(x)*, is a coherent risk measure if it has the following properties:

- Sub-additivity:

$$p(x+y) \leq p(x) + p(y), \forall x, y \in R \tag{13}$$

- Positive Homogeneity:

$$p(\lambda x) = \lambda p(x), \forall x, \lambda \in R \tag{14}$$

- Monotonicity:

$$if\ x \leq y \ldots then\ p(x) \leq p(y), \forall x, y \in R \tag{15}$$

- Translational Invariance:

$$p(x + \alpha r) = \alpha p(x), \forall x, r, \alpha \in R \tag{16}$$

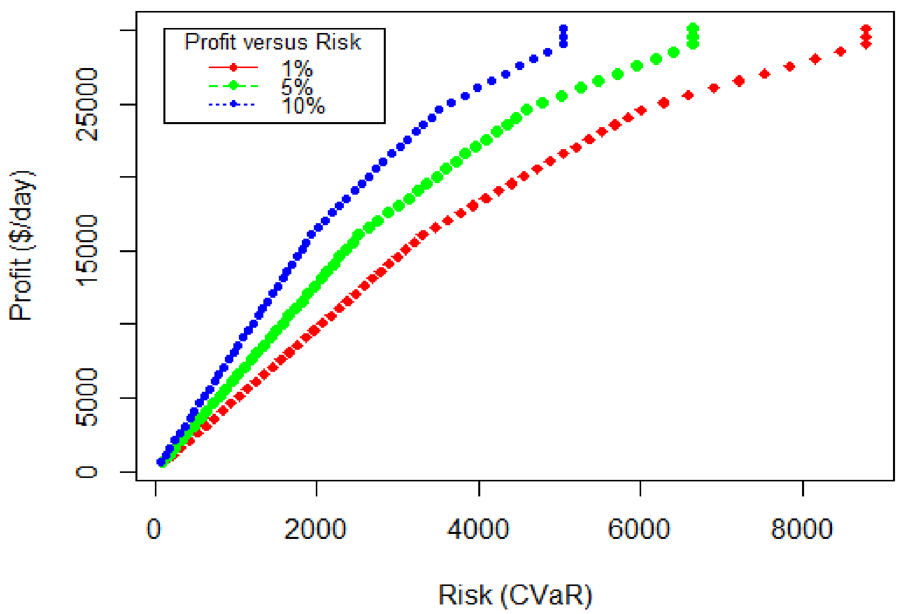
[58] Szego 2002

Definition 3 (Conditional Value-At-Risk) CVaR is the average of the losses greater than the VaR, stated with a particular confidence level, here α.

$$CVaR(\alpha) = E[h_0 - h \mid h_0 - h \geq VaR(\alpha)], \forall \alpha \in (0,1] \tag{17}$$

Where h is the random variable, which can for example represent the payoff of a portfolio. The α, is the confidence level, and h_0 the initial value for h.

Rockafellar and Uryasev (2002) showed that, CVaR is superior to VaR in all optimisation applications. Computationally, within an optimisation procedure due to the theory of linear programming, CVaR can enter into the objective function or into the constraints of the problem, producing an equivalent solution. We implement CVaR within the objective function as shown for example in (11). Furthermore, after implementing CVaR as the risk measure, the increased conservativeness emerges:



(***Figure 4:*** *The efficient frontier for the profit function versus the conditional value at risk of the refinery's GRM (Model 4)*)

3.3. Conditional Drawdown-At-Risk (CDaR)

CDaR, Uryasev (2010), in an aggregated format, is the number and magnitude of the portfolio drawdowns over a period of time. A drawdown is the drop in portfolio value, compared to the maximum achieved in the past. For a particular threshold α, the *α – CDaR*, is defined as the mean of the worst (1- α)*100% draw-downs experienced over a period of time. The definition of CDaR follows.

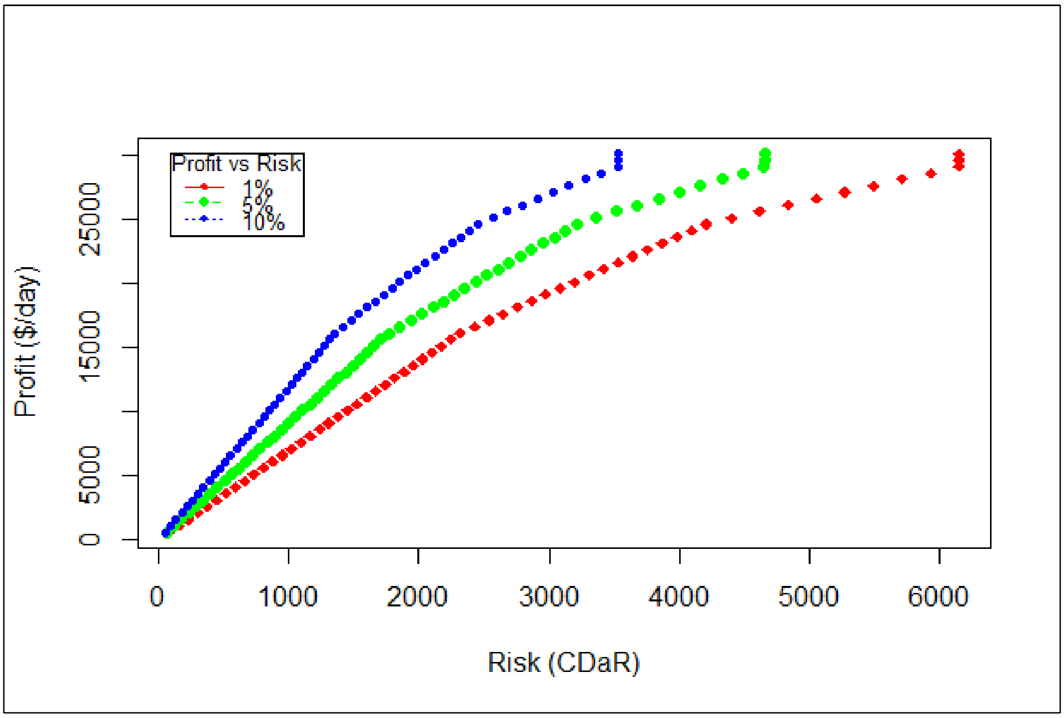
Definition 4 (CDaR) *Conditional DrawDown at Risk is a measure of risk at a specific level of confidence* α, *it is the average of a set of worst drawdowns at a specific threshold.*

$$\Delta_\alpha(\vec{x}) = \frac{1}{(1-\alpha)T}\int_\Omega f(\vec{x},t)dt \tag{18}$$

Where,

$$\Omega = \{t \in [0,T] : f(x,t) \geq \alpha(x,t)\} \tag{19}$$

Where Δ_α is the CDaR, α is the confidence level, *f(x, t)* is the drawdown function, which is the difference between the maximum of the profit function over history preceding the point *t*, and the value of this function at time *t*. The CDaR - optimisation problem is non-linear, convex and piecewise in structure, hence it can be reduced to a linear programming problem by using auxiliary variables.



*(**Figure 5:** The efficient frontier for the profit function versus the conditional drawdown at risk (Model 5))*[59]

Figures two-five, show that the CDaR model was the best performer from a risk/reward perspective. We have chosen α as 90%, 95% and 99% to compare the results across the models. Another important indicator for these types of models is defined below in section.

3.4. Expected Value of Perfect Information (EVPI)

The EVPI is an important measure for stochastic programming, Raiffa and Schlaifer (1961). Given an uncertain situation that is "on the cards"; EVPI provides an example of what value is possible to extract practically.

To calculate this figure we need the ***"here and now"*** value of the objective function, J_{HN}. Whereas the ***"wait and see"*** value is J_{WS}, is an expectation based on the outcome of the uncertain random variables.

Definition 5 (Expected Value of perfect information)

For a particular realization $\xi = \xi(w)$, $w \in I$, we consider the objective functional:

$$J(x, \xi) := C^T x + \max\{q^T y \mid Wy = h - Tx, y \geq 0\}, \quad (20)$$

The associated maximisation problem is:

$$\max_x E_\xi J(x, \xi) \quad (21)$$

The optimal solution of the above is sometimes referred to as the ***here-and-now*** solution.

We can denote the optimal value of our recourse problem by:

$$RP := \max_x J(x, \xi) \quad (22)$$

[59]Drawdowns for a set of chosen weights, are assumed static over the historical period in question

Another related maximisation problem is to find the optimal solution for all possible scenarios and to consider the expected value of the associated optimal value:

$$WS := E_{\xi} \max_x J(x, \xi) \quad (23)$$

This is called the ***wait-and-see*** solution. The difference between the optimal values of the here-and-now solution and the wait-and-see solution is the ***expected value of perfect information***:

$$EVPI := RP - WS \quad (24)$$

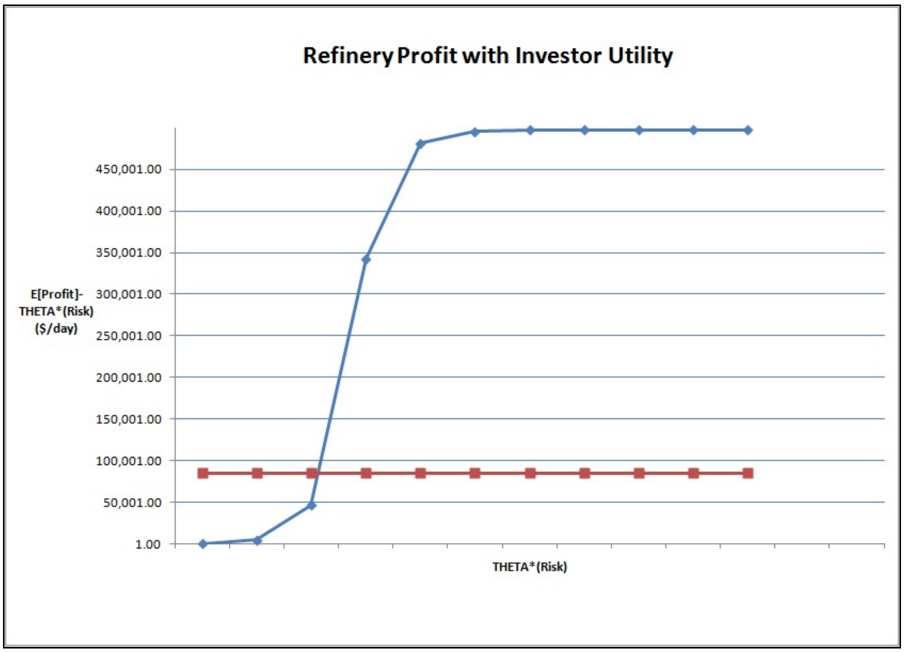
4. Results

The constrained optimisation results are shown in Table 4, see *Appendix A* for model constraints.

Model Name	**GAMS File Name**	**Profit** ($/day)	**Crude Oil Purchased**	**Risk** ($/day)	**Stochastic Robustness** (EVPI/WS)
1	DETERM_Today_LP	235,418	5,795	-	-
2	STOCH_VAR/COVAR_LP	329,000	5,781	81,195	15.64%
3	STOCH_COV_A_VaR_NLP	125,000	2,196	7,187	13.86%
4	STOCH_COV_A_CVaR_NLP	125,000	2,196	8,197	4.27%
5	STOCH_COV_A_CDaR_NLP	125,000	2,196	5,532	5.89%

(***Table 4:*** *Summary of all five Models using two-stage stochastic programming.*
NB: *All models are optimised on prices from the scenario generation averaged from 1,000 samples.)*

The results using the different risk measures show that there are large decision discrepancies for the refiner based upon which measure is selected. Models three to five reach a maximum of £125,000 per day and are solved instantly. Dependent upon the α, in front of the risk term, the model can also be considered at different utility levels – a risk averse investor would select the 10% significance model. These results show a huge gain in using alternative risk measures over variance. The lower the robustness measure, the more reliable the model; it is clear that the most appropriate model is either models four or five as one and two contain an unacceptable level of risk. If risk is the priority, five is the model to implement at the oil refinery, albeit three and four are not impractical. These models have a particular structure due to the constraints and optimisation utilised to construct them. Shown in Figure 6 below is a plot of Model 2, where a utility parameter has been used within the risk measure; all stochastic models follow this same behaviour. This illustrates the investor's level of risk aversion, hence the model structure.



(***Figure 6:*** *Comparing the efficient frontier for Model 2 and the deterministic profit function using a utility parameter from θ=1 to θ=0.00000001)*

The y-axis represents the objective function of profit with a level aversion to risk; whereas the x-axis is the profit risk for the refiner. We can optimise for risk from risk seeking to complete aversion. The red line above represents the deterministic result, and the blue line is the stochastic model. The diagram illustrates that at a particular level of risk there is no more profit to be attained, yet with too little risk seeking, a minimal profit level is reachable. This suggests that there is an optimum level of risk to be sought.

Analysing the model's robustness we calculate the EVPI as described in section three. For model five, we find the following:

WS = \$381,972 per day

RP = \$367,114 per day

Therefore, the average EVPI has an additional benefit of \$14, 858 per day to the owner of the oil refinery – this result shows that models considering uncertainty have additional value to the refinery owner.

5. CONCLUSION

In this paper we have introduced a two stage stochastic model with various risk measures for monthly midterm production planning of a typical Topping oil refinery. The output product prices are seen as exogenous to the refiner; meaning that there are enough refineries producing the same product that the failure of one has minimal impact on the market, i.e. the refiner is a price taker. The crude oil and refined product prices are considered stochastic by using a discrete scenario generation for five possible economic states – the last being the lowest set of commodity prices and the first the highest. To generate these five states a probability distribution is "discretised" for each series and the corresponding probability obtained. The risk measures are then assessed, and optimisation with a variance-covariance matrix is used. CVaR and CDaR are then introduced into the formulation and provide better computational properties than standard attempts using variance or VaR. We found that the quantile functions: CVaR and CDaR, were very efficient and simple to implement using GAMS, and the scenario generation method was efficacious. The expected value of perfect information indicates that there is an additional benefit of \$14,858 per day to the refiner, if the uncertain events are included in the calculation using CDaR. It is clear how the uncertainty can be managed by the refiner, and the tractability of these models would be appealing in practice. This work could be extended by considering the operational flow planning problem at a more granular level, or integrating the other issues at a refinery complex, e.g. the transportation costs or short term scheduling. Integrated together with the above stochastic formulation, the refinery unit yield uncertainties would also provide deeper insights. Another extension could be achieved by considering, the multi-stage version of this problem, which is yet to be formulated successfully. Multiple stages introduces computational difficulties, for example the "curse of dimensionality" - if successfully formulated, it would possibly leading to a financial value for the refinery. There is space for investigating optimisations with quantile functions; currently only the financial examples are persistent.

ACKNOWLEDGEMENTS

The author would like to thank Prof. Ron Smith and Dr. Alexander Karalis Isaac for helpful comments, and Ying Wu for her help and support.

REFERENCES

Aldaihani, and Al-Deehani (2010). Mathematical models and a tabu search for the portfolio management problem in the Kuwait stock exchange. *Int. J. Operational Research* **7**(4): 445 – 462.

ALEXEI CHEKHLOV, STANISLAV URYASEV, and MICHAEL ZABARANKIN (2005). DRAWDOWN MEASURE IN PORTFOLIO OPTIMIZATION. *Int. J. Theor. Appl. Finan.* **08**(13).

Basak S and Shapiro A (2001). Value-at-Risk Based Risk Management: Optimal Policies and Asset Prices. *The Review of Financial Studies* **14**(2): 371 – 405.

Benyoucef A (2010). Oil refining planning under price and demand uncertainties: case of Algeria. 29th USAEE/IAEE North America conference.

Ben-Tal, Bertsimas and Brown (2010). A Soft Robust Model for Optimization Under Ambiguity. *OPERATIONS RESEARCH* **58**(4): 1220 – 1234.

Bernardo FP, Pistikopoulos EN and Saraiva PM (1999). Integration and Computational Issues in Stochastic Design and Planning Optimisation Problems. *Industrial & Engineering Chemistry Research* **38**: 3056 - 3068.

Clay RL and Grossmann IE (1998). A disaggregation algorithm for the optimization of stochastic planning models. *Computers & Chemical Engineering* **21**(7): 7511 – 774.

Culik, M (2010). Flexibility and project value: interactions and multiple real options. *Int. J. of Operational Research.* **11**(4): 443 - 462.

Dempster MAH, Pedrón NH, Medova EA, Scott JE and Sembos A (2000). Planning Logistics Operations in the Oil Industry. *The Journal of the Operational Research Society* **51**(11): 1271 - 1288.

Elkamel, A. and Al-Qahtani, K. (2011). Integration and coordination of multisite refinery and petrochemical networks under uncertainty. *Int. J. Process Systems Engineering* **1**(¾): 237-265.

Eppen GD, Martin RK and Schrage L (1989). A Scenario Approach to Capacity Planning. *Journal of the Operations Research Society of America* **37**(4): 517 – 527.

Geman, H (2005). Energy Commodity Prices: Is Mean-Reversion Dead? *The Journal of Alternative Investments* **8**(2): 31 – 45.

Gupta, A and Maranas, CD (2004). Real-Options-Based Planning Strategies under Uncertainty. *Industrial & Engineering Chemistry Research* **43**: 3870-3878.

Ierapetritou MG and Jia Z (2007). Short-term scheduling of chemical process including uncertainty. *Control Engineering Practice* **15**(10): 1207–1221.

Kelly, JD and Mann, JL (2003a). Crude-oil blend scheduling optimization: An application with multi-million dollar benefits: Part I. Hydrocarbon Processing 47-53. http://www.hydrocarbonprocessing.com/Article/2599682/Crude-oil-blend-scheduling-optimization-an-application-with-multimillion-dollar-benefitsPart-2.html, accessed summer 2012.

Leiras, A., Ribas, G., Hamacher, S. and Elkamel, A. (2011). Literature review of oil refineries planning under uncertainty. *Int. J. Oil, Gas and Coal Technology* **4**(2): 156-173.

Markowitz H (1952). Portfolio Selection. *The Journal of Finance* **7**(1): 77-91.

Moro LFL and Pinto JM (2000). Planning and Scheduling for petroleum refineries using mathematical programming. *Computers & Chemical Engineering* **24**(9): 2259 – 2276.

Neiro SMS and Pinto JM (2005). Supply Chain Optimization of Petroleum Refinery Complexes. *Proceedings of the fourth international conference.*

Pongsadki A, Rangsunvigit P, Siemanond K and Bagajewicz MJ (2005). Financial risk management in the planning of refinery operations. *International Journal of Production Economics* **103**(1): 64 – 86.

Ravi V and Reddy PJ (1998). Fuzzy linear fractional goal programming applied to refinery operations planning. *Fuzzy Sets and Systems* **96**(2): 173 – 182.

Ribas et al (2012). Operational planning of oil refineries under uncertainty Special issue: Applied Stochastic Optimization. *IMA Journal of Management Mathematics.*

Rockafellar RT and Uryasev S (2000). Conditional value-at-risk: optimization algorithms and applications. *Computational Intelligence for Financial Engineering*, **14**: 49-57.

Tahar, R.B.M. and Abduljabbar, W.K. (2010). A simulation and optimisation model for the transport of refinery products. *Int. J. Oil, Gas and Coal Technology*, **3**(2): 160 – 169.

Wang, S (2013). Integrated model of production planning and imperfect preventive maintenance policy for single machine system. *Int. J. of Operational Research* 2013, **18(**2): 140 - 156.

Wettergren, T.A. and Baylog, J.G. (2014). Discrete search allocation with object uncertainty. *Int. J. Operational Research*, **20**(1): 1 – 20.

Xidonas, P., Mavrotas, G. and Psarras, J. (2010). Portfolio management within the frame of multiobjective mathematical programming: a categorised bibliographic study. *Int. J. Operational Research*, **8**(1): 21 – 41.

Zhang, Xu and Chen (2009). Technical Note — A Risk-Averse Newsvendor Model Under the CVaR Criterion. *OPERATIONS RESEARCH* **57**(4): 1040-1044.

6. APPENDIX

A.1 Mass Balance Constraints

[A Fixed Yields]

Primary Unit:

$$-0.13\ X_1 + X_7 = 0$$

$$-0.15\ X_1 + X_4 = 0$$

$$-0.22X_1 + X_8 = 0$$

$$-0.20X_1 + X_9 = 0$$

$$-0.30\ X_1 + X_{10} = 0$$

Cracker:

$$-0.05X_{14} + X_{20} = 0$$

$$-0.40X_{14} + X_{16} = 0$$

$$-0.55X_{14} + X_{17} = 0$$

[Fixed Blends]

Gasoline blending:

$$0.5X_2 - X_{11} = 0$$

$$0.5X_2 - X_{16} = 0$$

Heating Oil Blending:

$$0.75X_5 - X_{12} = 0$$

$$0.25X_5 - X_{18} = 0$$

[Unrestricted Balances]

Naphtha: $-X_7 + X_3 + X_{11} = 0$

Gas Oil: $-X_8 + X_{12} + X_{13} = 0$

Cracker Feed: $-X_9 + X_{14} + X_{15} = 0$

Cracked Oil: $-X_{17} + X_{18} + X_{19} = 0$

Fuel Oil: $-X_{10} + X_{13} + X_{15} + X_{19} + X_6 = 0$

A.2 Raw material availability constraints

Crude Oil: $X_1 \leq 15000$

Gasoline: $X_2 \leq 2700$

Naphtha: $X_3 \leq 1100$

Jet Fuel: $X_4 \leq 2300$

Heating Oil: $X_5 \leq 1700$

Fuel Oil: $X_6 \leq 9500$

REFERENCES

AANAND VENKATRAMANNAN & CAROL ALEXANDER, Closed Form Approximations for Spread Options (2011), Applied Mathematical Finance, Vol. 1, pp.1–26

Alexander, Carol and Scourse, Andrew , ‘Bivariate Normal Mixture Spread Option Valuation,’ Quantitative Finance, Vol. 4, No. 6 (2004), pp. 637–648.

Bachelier. L. , Th´eorie de l a sp´eculation. Annales Scientifiques de l’Ecole Nor-´ male Sup´erieure, III-17:21–86, (1900). Thesis for the Doctorate in Mathematical Sciences (defended March 29, 1900). Reprinted by Editions Jacques Gabay,´ Paris, 1995. English translation in The random character of stock market prices, Ed. P. Cootner, pp. 17-78, Cambridge, MIT Press, 1964.

Bellman, R.E. (1957). "Dynamic Programming." Princeton University Press, Princeton, NJ. Republished (2003): Dover, ISBN 0-486-42809-5

Benyoucef, A., “Oil refining planning under price and demand uncertainties: case of Algeria” (2010), 29th USAEE/IAEE North America conference

Bernardo, Fernando P. , Efstratios N. Pistikopoulos, Pedro M. Saraiva, “Integration and Computational Issues in Stochastic Design and Planning Optimisation Problems”, *Industrial & Engineering Chemistry Research*, Vol. 38, (1999), *pp. 3056-3068*

Black, F., M. Scholes. (1973). The pricing of options and corporate liabilities. J. Political Econom. 81 637–659.

Borovkova, Svetlana, Ferry J. Permana and Hans v.d. Weide, "A closed form approach to the valuation and hedging of basket and spread options." *The Journal of Derivatives*, (Summer 2007), Vol. 14, No. (4), *pp. 8-24*

Boyle. Phelim P. , " A Lattice Framework for Option Pricing with Two State Variables", Journal of Financial and Quantitative Analysis, Volume 23, Issue 01, (March 1986), pp 1-12.

Brandão, L. (2002). An application of real-option theory in discrete time to the valuation of a highway concession project in Brazil. Unpublished doctoral dissertation, Pontificia Universidade Catolica do Rio de Janeiro, Rio de Janeiro, Brazil.

Brandão, L., J. Dyer. (2005). Decision analysis and real options: A discrete time approach to real option valuation. Ann. Oper. Res. 135(1) 21–39.

Cao et al, March (2009),, Chance constrained programming models for refinery short-term crude oil scheduling problem, Volume 33, Issue 3, pp. 1696–1707.

Carmona, R. and Durrleman, V., "Pricing and hedging spread options in a log-normal model." *Working Paper, Department of Operations Research and Financial Engineering, Princeton University,* (March 16th 2003).

Carmona, R. and Ludkovski, M, “Valuation of Energy Storage: an Optimal Switching Approach. Quantitative Finance,10 (4), (2010) 359–374.

Chantziara, Thalia and Skiadopoulos, George. "*Can the dynamics of the term structure of petroleum futures be forecasted? Evidence from major markets"*. Energy Economics, Vol.30 (No.3). (2008) pp. 962-985. ISSN 0140-9883

Clewlow, L., C. Strickland. (1999). "Implementing Derivative Models". John Wiley & Sons, Chichester, UK.

Copeland, T., V. Antikarov. (2001). "Real Options". Texere LLC, New York.

Copeland, T., P. Tufano. (2004). "A real-world way to manage real options". Harvard Bus. Rev. 82(3) pp. 90–99.

Cox, J., S. Ross, M. Rubinstein. (1979). Option pricing: A simplified approach. J. Financial Econom. 7, 229–263.

Dempster and Hong, July (2000), Spread Option Valuation and the Fast Fourier Transform, RESEARCH PAPERS IN MANAGEMENT STUDIES.

Dempster , M.A.H. , Medova, Elena, Ke Tang, “Long term spread option valuation and hedging. Journal of Banking & Finance 32 (2008) 2530–2540.

Dixit, A., R. Pindyck. (1994). Investment Under Uncertainty. Princeton University Press, Princeton, NJ.

Duan and Pliska (2004), An Empirical Model Comparison for Valuing Crack Spread Options, ICMA Centre Discussion Papers in Finance DP2010-01

Elkamel et, Al-Quataini, Literature review of oil refineries planning under Uncertainty, Int. J. Oil, Gas and Coal Technology, Vol. 4, No. 2, (2011).

Eydeland, A., and Geman, H., "Pricing Power Derivatives", *RISK Books*, (September 1998)

Eydeland, A., and Geman, H., "Fundamentals of Electricity Derivatives." *Chapter of the book: Energy Modelling and the Management of Uncertainty*, RISK Books, (June 1999)

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

Geman, H., "Energy Commodity Prices: Is Mean-Reversion Dead?", *Journal of Alternative Investment,* (2005).

Geman, H., and Ohana, S., "Time-consistency in managing a commodity portfolio: A dynamic risk measure approach", *Journal of BANKING & FINANCE, (32),* (Jan. 2008), *pp.1991 - 2005*

Geman, H. and V.N. Nguyen (2003), " Soybean Inventory and Forward Curve Dynamics", MANAGEMENT SCIENCE Vol. 51, No. 7, July 2005, pp. 1076–1091

Graham, J. and H. Campbell, "Theory and Practice of Corporate Finance: Evidence from the Field," *Journal of Financial Economics*, 60 (2001): 187-243.

Guoming, L., Margot, F., Secomandi, N. “An approximate dynamic programming approach to benchmark practice-based heuristics for natural gas storage valuation”, (Sep 2008), *Tepper working paper*.

Hahn, W., J. Dyer. "A discrete-time approach for valuing real options with underlying mean-reverting stochastic processes." European Journal of Operational Research. Volume 184, Issue 2, 16 January 2008, pp. 534–548.

Harrison, J. Michael and Kreps, David M., "Martingales and Arbitrage in Multiperiod Securities Markets", *Journal of Economic Theory,* 20, (1979), *pp. 381 – 408*

Higham, Nick (2002) Computing the nearest correlation matrix - a problem from finance; *IMA Journal of Numerical Analysis* **22**, 329–343.

Hull, J. 2003. Options, Futures and Other Derivatives. Prentice Hall, Upper Saddle River, NJ.

Hull, J. A.White, Numerical procedures for implementing term structure models I, *Journal of Derivatives, Fall (1994).*

Hull, J., A.White, Numerical procedures for implementing term structure models II, *Journal of Derivatives, Winter (1994).*

Iman Ronald L. & Conover W. J., "A distribution-free approach to inducing rank correlation among input variables", Journal Communications in Statistics - Simulation and Computation Volume 11, (1982) - Issue 3.

Jafarizadeh, Babak; Bratvold, Reidar B, "Two-Factor Oil-Price Model and Real Option Valuation: An Example of Oilfield Abandonment,", (2012), SPE Economics & Management, 4(3): 158–170

Jensen, Michael C, "Agency costs of free cash flow, corporate finance and takeovers". American Economic Review (1986). **76** (2): 323–329).

Kelly, J. D. and Mann, J. L, "Crude-oil blend scheduling optimization: An application with multi-million dollar benefits: Part I." Hydrocarbon Processing (2003a) June, *pp 47-53*.

Kirk, E. "Correlation in the Energy Markets, in Managing Energy Price Risk.", *London: Risk Publications and Enron.* (1995)

Lin et al, An interval-fuzzy two-stage stochastic optimization model for regional energy systems planning under uncertainty, Volume 37, Issue 3, March (2009), pp. 868–878.

Lo, C.F. "A simple derivation of Kirk's approximation for spread options", *Applied Mathematics Letters*, (April 2013), Vol. 26, *pp. 904-907*

Marianthi G, Ierapetritou A, Zhenya Jia, "Short-term scheduling of chemical process including uncertainty", *Control Engineering Practice* 15, (2007), *pp. 1207–1221*

Markowitz, Harry, "Portfolio Selection". *The Journal of Finance,* Vol. 7, No. 1 (Mar., 1952), *pp. 77-91.*

Margrabe, W. "The value of an option to exchange one asset for another." The Journal of Finance, (1978), Vol. 33, pp. 177-186.

Moro, J.M. Pinto, "Planning and scheduling for petroleum refineries using mathematical programming", Computers & Chemical Engineering, Vol. 24, Issues 9–10, 1 October (2000), pp. 2259–2276.

Nan, and Marc, Refinery Planning and Scheduling, (2013), Petroleum Refining and Natural Gas Processing, ASTM standards publication; http://www.astm.org/.

Neiro, S.M.S., Pinto, J.M. "Supply Chain Optimization of Petroleum Refinery Complexes", *Proceedings of the fourth international conference,* (2003)

Parsons, Cliff. , "Quantifying natural gas storage optionality: a two-factor tree model"
The Journal of Energy Markets, vol 6.1 (Spring 2013): pp. 95-VI.

Pindyck , Robert S. "Irreversibilities and the timing of environmental policy", Resource and Energy Economics 22 (2000), pp. 233–259.

Poitras, Geoffrey. "SPREAD OPTIONS, EXCHANGE OPTIONS, AND ARITHMETIC BROWNIAN MOTION:INTRODUCTION" The Journal of Futures Markets; Aug (1998); 18, 5.

Pongsadki et al, "Financial risk management in the planning of refinery operations", *International Journal of Production Economics*, 3rd August (2005)

Rollins and Insely, "Real options in harvesting decisions on publicly owned forest lands", (2003) ideas.repec.org

Schwartz, E., “The stochastic behaviour of commodity prices: Implications for valuation and hedging”. *The Journal of Finance,* (1997), 52 (3), *pp. 923-973*

Schwartz, E., J. Smith. (2000). Short-term variations and long-term dynamics in commodity prices. Management Sci. 46 893–911. J. Appl. Corporate Finance 14 8–24.

Sen and Higle, "An Introductory Tutorial on Stochastic Linear Programming Models", *Interfaces*, Vol 29 Issue 2, April (1999), *pp. 33-61*

Shimko, David C (Apr 1994). OPTIONS ON FUTURES SPREADS: HEDGING, SPECULATION, AND VALUATION: INTRODUCTION. *The Journal of Futures Markets (1986-1998), Vol.* 14, No. (2), *pp. 183*

Book Chapter

The New Investment Theory of Real Options and its Implication for Telecommunications Economics; Volume 34 of the series Topics in Regulatory Economics and Policy pp 3-33

V. Ravi, P.J. Reddy, "Fuzzy Linear Fractional Goal Programming applied to refinery operations planning", *Journal of Fuzzy Sets and Systems*, August (1998)

Wilcox, D. (1990): Energy Futures and Options: Spread Options in Energy markets. New York: Goldman Sachs & Co.

Yin, R., & Newman, D. H. (1996). The effect of catastrophic risk on forest investment decisions. *Journal of Environmental Economics and Management*, *31*(2), 186-197.

Zamora, Juan M., and Ignacio E. Grossmann. "A global MINLP optimization algorithm for the synthesis of heat exchanger networks with no stream splits." Computers & Chemical Engineering 22.3 (1998): 367-384.