A Look at Model Uncertainty in the Evaluation of Commodity Contingent Claims: A Practitioner’s Guide

by

Jovan Lukovich

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Graduate Department of Mechanical & Industrial Engineering
University of Toronto

© Copyright 2013 by Jovan Lukovich
Abstract

A Look at Model Uncertainty in the Evaluation of Commodity Contingent Claims: A Practitioner’s Guide

Jovan Lukovich
Master of Applied Science
Graduate Department of Mechanical & Industrial Engineering
University of Toronto
2013

Model uncertainty in financial markets is prevalent by the very nature of how models are constructed and used by financial practitioners. As such, a proper characterization of model uncertainty should be paramount in the eyes of every practitioner, and furthermore, a proper framework for implementing such a characterization towards financial activities should be implicit. While model uncertainty is acknowledged by practitioners, a cohesive and robust framework for determining a model uncertainty risk measure that is broadly accepted by practitioners is missing. We acknowledge this deficiency and provide a practitioner’s guide for evaluating a modern characterization of model uncertainty, specifically that of Li and Kwon, as applied to a subset of derivative related calculations, with the goal of promoting its implementation by practitioners. We promote its implementation by demonstrating the utility and flexibility of such a characterization relative to another modern model uncertainty risk measure, specifically that of Cont.
Acknowledgements

I want to first and foremost acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC) and Integrated Managed Futures Corp. (IMFC) for jointly funding my Industrial Postgraduate Scholarship (IPS1). Second, I would also like to acknowledge the Department of Mechanical & Industrial Engineering (MIE) at the University of Toronto for awarding me a MIE fellowship.

I want to thank my thesis adviser, Professor Kwon, for his advice and support throughout my degree. His insight and knowledge of modern constrained optimization techniques was enlightening, and was invaluable. I also want to thank Professor Lawryshyn of the Department of Chemical Engineering (Centre for Management of Technology and Entrepreneurship). His experience with commodity modeling and the utilization of the Kalman filter was very instrumental.

I want to thank my colleagues and superiors at IMFC, specifically Robert Koloshuk and Roland Austrup. Their insight into commodity futures and practical knowledge of the industry was invaluable and instrumental in how the thesis was framed for practitioners.

I want to thank the members of Professor Kwon’s lab, specifically Dexter Wu and Dr. Jonathan Li.

I want to thank all my family and friends for their continued support. In particular, I would like to thank my parents, Sophia and Milovan, for their undeniable support towards pursuing higher education.

Lastly, and most importantly, I want to thank my fiancée, Milka. She has been there from the beginning and will be there at the end. Without her continued support, it is very unlikely that this thesis would exist, I appreciate her every day.
# Contents

1 Introduction ........................................ 1
   1.1 Purpose & Motivation .............................. 1
   1.2 Structure ......................................... 4

2 Model Uncertainty .................................. 6
   2.1 Understanding Model Uncertainty .................. 6
   2.2 Risk, Uncertainty & Incomplete Markets .......... 8
   2.3 Risk Measures ..................................... 10
   2.4 Cont’s Measures ................................... 13
   2.5 Li & Kwon Extension ............................... 15

3 Commodity Modeling ................................ 22
   3.1 Empirical Studies ................................. 22
      3.1.1 Theory of Normal Backwardation .............. 23
      3.1.2 Theory of Storage ............................. 24
      3.1.3 Samuelson Hypothesis ........................ 26
      3.1.4 Inducing Mean Reversion ..................... 27
   3.2 Theoretical Studies ............................... 28
      3.2.1 Reduced Form Models .......................... 29
      3.2.2 Model 1: S1997-1 .............................. 34
      3.2.3 Model 2: S1997-2 .............................. 35
      3.2.4 Model 3: SS2000 ............................... 36
      3.2.5 Model 4: CS2003 ............................... 38

4 Results & Discussion ............................... 41
   4.1 Data Collection .................................... 41
   4.2 Calibration ......................................... 43
   4.3 Calculation ......................................... 49
      4.3.1 Cont’s Convex Measure ......................... 49
      4.3.2 Li-Kwon Convex Measure ....................... 51

5 Conclusions & Future Work ......................... 61
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Benchmark call options written on WTI Crude Oil futures for March 2013 delivery (CLH3C), extracted from Bloomberg and are of American type.</td>
</tr>
<tr>
<td>4.2</td>
<td>Calculated empirical moments estimated over different look back windows.</td>
</tr>
<tr>
<td>4.3</td>
<td>Estimated parameters from S1997-1 model.</td>
</tr>
<tr>
<td>4.4</td>
<td>Estimated parameters from S1997-2 model.</td>
</tr>
<tr>
<td>4.5</td>
<td>Estimated parameters from SS2000 model.</td>
</tr>
<tr>
<td>4.6</td>
<td>Estimated parameters from CS2003 model.</td>
</tr>
<tr>
<td>4.7</td>
<td>Estimated option prices for each model. The strikes that have * indicate those that were used as benchmark options.</td>
</tr>
<tr>
<td>4.8</td>
<td>Cont’s convex risk measure results from the four models. The derivative prices obtained from the CS2003 model were the reference option prices used in calculating the displayed results. CLH3C is short-form notation for a call option written on a March 2013 crude oil futures contract.</td>
</tr>
<tr>
<td>4.9</td>
<td>Adjusted empirical moments using $\epsilon = 0.0045$ from equation 4.3. All moments denoted with * are those that were changed by the moment boundary condition.</td>
</tr>
<tr>
<td>4.10</td>
<td>The percentage change in each moment upon the adjustment that allows the problem to be feasible. With regards to the notation, $m_{i,t}$, is where the first subscript indicates the moment $i \in 1, 2, 3$, and the second subscript indicates the look back window $t \in 0.25, 0.50, 1.50, 3.00$.</td>
</tr>
<tr>
<td>4.11</td>
<td>Li-Kwon’s generalization of Con’ts convex risk measure results using one benchmark option (CLH3C 100). The first half of results calculates the measure using two moments ($n = 2$), whereas the later half uses three moments ($n = 3$). CLH3C is short-form notation for a call option written on a March 2013 crude oil futures contract.</td>
</tr>
<tr>
<td>4.12</td>
<td>Li-Kwon’s generalization of Con’ts convex risk measure results using two benchmark option (CLH3C 100, and CLH3C 105). The first half of results calculates the measure using two moments ($n = 2$), whereas the later half uses three moments ($n = 3$). CLH3C is short-form notation for a call option written on a March 2013 crude oil futures contract.</td>
</tr>
</tbody>
</table>
List of Figures

4.1 Flowchart of the methodology for calculating the continuous case of the convex risk measure of model uncertainty. ............................ 52
List of Appendices

Appendix A  Background Knowledge .................................................. 71
Appendix A.1  Futures Basics ................................................................. 71
Appendix A.2  Forward VS Futures Contracts ....................................... 75
Appendix A.3  Hedging & Basis Risk ..................................................... 77
Appendix A.4  Economic Basics .............................................................. 78
Appendix A.5  Option Basics ................................................................. 83

Appendix B  Kalman Filter ................................................................. 85

Appendix C  Delta-Hedging Basics ....................................................... 88

Appendix D  MATLAB Code ............................................................... 92
Appendix D.1  Sample Code 1 ............................................................... 92
Appendix D.2  Sample Code 2 ............................................................... 94
Chapter 1

Introduction

1.1 Purpose & Motivation

Financial modeling is as much art as it is science [22]. The inherent stochasticity of market returns does not conform to any parametric form because markets are not governed by physical laws. Markets are more behavioural than many rational economists and mathematicians imply by their model assumptions. It becomes clear that identifying a fundamental theorem of asset returns is a difficult, if not impossible, task, and highlights the first important aspect of Derman’s taxonomy on model uncertainty, the inapplicability of modeling [22].

The difficulty of assessing market behaviour, and possibility of not being able to model its behaviour, has not deterred financial engineering practitioners. In fact, quite the contrary effect has incurred as there exist various models of varying complexities that attempt to model the same market dynamics. As an example, there exists multi-factor, jump-diffusion, and stochastic volatility (many more exist) models that are constructed to characterize asset price dynamics to use in the pricing of options and other complex derivatives. The obvious question therefore becomes what model is most appropriate for a given situation, and more importantly, how do we know the level of accuracy we should expect while using a model? It becomes imprudent to assume that only one model should be used at a time when we are uncertain whether or not that particular model will be able to characterize the probability of the events about to transpire. This highlights the second important aspect of Derman’s model uncertainty taxonomy, picking the incorrect model [22].

Despite the apparent deficiencies, financial models serve an important utility in financial markets and have been embraced by today’s financial community. Financial institutions now rely upon models for important every day tasks such as pricing complex securities, assessing risks,
Chapter 1. Introduction

market making, and ultimately decision making. The fundamental reason for adopting financial models is very logical, models are believed to internalize the intuition of traders and capture market risks. This belief becomes a self-fulfilling prophecy because the construction of exotic derivatives that use other simpler derivatives as their building blocks for pricing are driven by the belief that the simpler models are able to capture the inherent market risks. This game of constructing derivatives upon derivatives to diversify risks has yielded a dangerous dependence upon modeling. Practitioners begin to use models to learn how they should interpret more complex models that were constructed of the simpler models, and will ultimately use the simpler model to assist them in trading the more complex instrument. Such a cascading affect may lead to dangerous repercussions because practitioners will continue to distance themselves from the real world of random markets by travelling towards the alluring illusive world of theoretical mathematical models.

The distance between the real market world and that of financial models instantly evaporates upon the onset of a crisis. It is at this time when financial engineers (probably the largest benefactor of the advent of financial models) retrace their steps and make a very clear statement that their model is only supposed to be used in accordance to the set of predefined assumptions. The crash is often due to an unanticipated, and more importantly, unmodeled market correction. Unfortunately, it is often at this time when practitioners attempt to reconcile their model, and it has taken a number of large and public model malfunctions in the financial community before the concept of model uncertainty became prevalent in academic and practitioner circles. A recent example is the 2007 collapse in subprime mortgages and the subsequent contagion to other seemingly uncorrelated markets. No model was robust enough to characterize the subsequent shift in market regimes.

The events that followed the subprime mortgage crisis prompted Emanuel Derman and Paul Wilmott to write an oath (more of quip on the relationship and responsibilities of financial modeler’s) aptly named the *The Modelers’ Hippocratic Oath* [23] (Derman also later wrote a book). We list them here:

i: I will remember that I didn’t make the world, and it doesn’t satisfy my equations.

ii: Though I will use models boldly to estimate value, I will not be overly impressed by mathematics.

iii: I will never sacrifice reality for elegance without explaining why I have done so.

iv: Nor will I give the people who use my model false comfort about its accuracy. Instead, I will make explicit its assumptions and oversights.

---

1 Some attribute this to physics envy, i.e., financial modeler’s envy the elegance and practicality of model’s generated by physicists. Andrew Lo indicates that physics envy may have created a false sense of mathematical precision in some cases [47].
v: I understand that my work may have enormous effects on society and the economy, many of them beyond my comprehension.

What Derman and Wilmott imply with the five items of the *The Modelers’ Hipperatic Oath* is that model uncertainty is prevalent by the very nature of how models are constructed and used by financial practitioners. It becomes clear that a proper characterization of model uncertainty should be paramount in the eyes of every practitioner, and furthermore, a proper framework for implementing such a characterization towards financial activities should be implicit. Unfortunately however, to our knowledge, there remains no commonly accepted risk measure for tackling model uncertainty that is broadly accepted by practitioners.

It is in this context, specifically the lack of an accepted risk measure for the implementation and evaluation of model uncertainty, and its hopeful adoption by practitioners, that lies the focus of this thesis. As such, we aim to provide a practitioner’s guide of a modern framework for evaluating the model uncertainty [of a subset of derivative related calculations] that we believe is a useful and informative risk measure for understanding the potential impact of model uncertainty.

In order to accomplish this goal, we demonstrate the implementation of two modern approaches to assessing the model uncertainty of a class of derivatives in order to compare and contrast their utility to a practitioner. We evaluate the efficacy of incorporating both of these alternative model uncertainty risk measure frameworks in the evaluation of a commodity contingent claim. We also thoroughly outline the methodologies for implementing both model uncertainty frameworks using real-data and highlight their deficiencies as well as their proficiencies.

As explained, model uncertainty will be evaluated using two alternative methodologies. These methodologies can be interpreted as assuming a discrete or a continuous set of measures in the formulation of the solution, with the former developed by Cont [15], and the latter developed by Li and Kwon [45]. We conclude that the continuous setting is more informative than the discrete setting. In our study, we implement each model risk framework on an options contract written on a commodity futures contract. We also provide a review of their theoretical frameworks.

We have specifically chosen commodity contingent claims because the properties of commodity dynamics are complex and difficult to effectively capture, thus model uncertainty is prevalent in their characterization. We undertake a thorough investigation of empirical and theoretical properties of commodity dynamics in this thesis.
1.2 Structure

As outlined in the motivation section, the high-level purpose of this thesis is to motivate the implementation of a model uncertainty framework to a practitioner. We attempt to do this by outlining the methodology of two modern model uncertainty frameworks, and motivate their utilization in practice. En route, we explain model uncertainty in the context of derivatives, we explain why commodity derivatives were chosen as our example asset class and present an overview of commodity dynamics, as well as provide a calculated example of both methodologies. We now detail each chapter in what follows.

Before we begin, we note that if the reader is not well versed in futures and option pricing, or in basic economic theory of commodity dynamics, we recommend reading the Background Knowledge section of the appendix A. The purpose of this appendix section is to elicit some key features of commodity futures, and a few basic principles of vanilla options. We attempt to give an intuitive description of the market forces that govern commodity price dynamics from an economic perspective. Furthermore, key terms such as backwardation and contango, basis and hedging are discussed.

The second chapter discusses model uncertainty in a general context and explains the evolution of model uncertainty towards its current modern representations as it relates to derivative contracts. We begin by defining our working definition of model uncertainty and how it can be interpreted differently when evaluating derivatives. This discussion explains how modern risk measures have evolved to encapsulate model uncertainty. We formally describe coherent and convex risk measures, and describe Cont’s work on extending such measures towards derivative contracts [15]. We close with a thorough description of a generalization of Cont’s measures using a modern semidefinite reformulation as developed by Li and Kwon [45]. Such a generalization is the foundation of this thesis as we attempt to describe why using it in a practical setting is pertinent.

The third chapter introduces empirical and theoretical studies of commodity dynamics that are most pertinent to this thesis. The purpose of such a detailed background is to encapsulate the complexities of commodity dynamics, and to demonstrate why a modern approach to model uncertainty may be relevant.

The empirical studies that we present focus on the dynamics of commodities, specifically developing an understanding of the sources of return and risks that are embedded in commodity futures. We discuss the theory of normal backwardation, the theory of storage, and the Samuelson hypothesis, and explain how these theorems have motivated theoretical studies.

The theoretical studies that we present focus on modeling the resulting price dynamics of spot commodities using statistical models. Such studies fall into two regimes, either equilibrium structural models or reduced form models. Equilibrium structural models attempt to
endogenously induce relevant parameters of commodity price dynamics by modeling external exogenous factors. Reduced form models replicate observed dynamics exogenously by specifying a functional form. We focus on the later and devote a section to common reduced form models, specifically those that are Gaussian-affine.

We conclude this chapter with a detailed description of the four commodity pricing models that will be used to calculate Cont’s risk measure. We present the statistical description of each model along with the corresponding state-space representation. The state-space representation allows utilizing the Kalman filter for parameter estimation. We provide an outline of the Kalman filter algorithm in the second section of the appendix B and our state-space representation utilizes the notation of this appendix. As such, we recommend that the reader take a cursory glance in order to better follow the model descriptions that we are presenting in this chapter.

The fourth chapter presents the results of an example, followed by a discussion, of how a practitioner can utilize the framework that we have described in practice. We contrast the results of the new framework with that of Cont, and explain the benefits of the new framework. In our discussion, we detail the data that was collected, the calibration methodology that was used, and present the results obtained. When presenting the calculated results, we provide a descriptive set of instructions/guidelines to utilize the modern framework for calculating model uncertainty. This guideline also outlines how to troubleshoot the mathematical description of the model. To encourage readability, we also explicitly follow an example that demonstrates how the methodology can be implemented in practice. These instructions dissect the technical framework, and is written for a practitioner who is interested in using the framework.

The final chapter reports our conclusions and outlines how the framework can be applied to other problems. Specifically, we outline how to use a similar representation to identify the possible model uncertainty when delta hedging.
Chapter 2

Model Uncertainty

This chapter provides the framework for incorporating model uncertainty into the pricing of contingent claims in their basic form. We progress through the evolution of model uncertainty towards the penultimate result of Li & Kwon’s distribution free quantification of Cont’s convex risk measure of derivative securities.

2.1 Understanding Model Uncertainty

The impact of model uncertainty has become increasingly more relevant to financial institutions due to their reliance on theoretical models for important every day tasks such as pricing complex securities, assessing risks, and ultimately decision making. One of the first academic/practitioners to bring attention to the topic was Emanuel Derman [22], who outlined different types and sources of model uncertainty in an internal technical document during his time at Goldman Sachs. He became part of a growing trend of practitioners and academics (some of their work to follow in subsequent subsections) that began actively looking at quantitatively measuring model uncertainty. In what follows, we outline Derman’s [22] description of the sources of model uncertainty in order to build a level of intuition towards how a practitioner may tackle the problem of model uncertainty.

We begin with the most fundamental of all model uncertainties, the inapplicability of modeling. Derman explains that there are some relationships in financial markets that just can not be expressed mathematically [22]. He continues to explain that it is important to understand how a system interacts with its environment to develop plausible independent variables that influence the system you are modeling. It is important to note that implicit in the description of model uncertainty in the context of the potential inapplicability of financial modeling is that the inherent stochasticity of market returns do not conform to any parametric form because
markets are not governed by physical laws. Many dynamical features of markets are a result of the behavioural tendencies of human beings because markets are man-made.

The second type of model uncertainty is when you have the incorrect model [22]. This may occur when some of the necessary factors that influence the valuation of the security being modelled are missing. Alternatively, model factors can be improperly specified as either deterministic or stochastic, the dynamics of a factor may be incorrect, inter-factor relationships may be dependent on a particular market environment or time horizon (short-term or long-term dynamics), or a factor may be improperly estimated. Worse still, a model may be correct in principle but exogenous factors (crisis or some other shock) allow the model to deviate from observed prices.

Model uncertainty sources 3-5 relate to human error in the development (correct model, incorrect solution & badly approximated solution) and usage (correct model, incorrect use) [22]. It is easy to make a technical mistake in developing an analytical solution or time marching scheme to implement into a piece of software. Proper sensitivity analysis must be made to each model’s parameter to avoid improper estimation of a security’s value. Furthermore, even if the model is properly specified in accordance to its assumptions, human beings have a tendency to plea ignorance and use a model outside of its specification and ignore the risks and limitations that it may have. A very good example is the naive usage of the Gaussian copula model developed by David Li for the pricing of complex credit derivatives. The lack of understanding of Li’s model’s limitations by traders in the United States became the recipe for disaster during the 2008 credit crunch [39].

Model uncertainty sources 6-7 are more mechanical in nature, and relate to software and hardware bugs, and unstable data [22]. Many financial models are computer programs and must thread the needle between speed and accuracy. Furthermore, accuracy before deploying a model into the market for live trading to avoid being subject to a rogue program that behaves irrationally. Such an incident occurred in the summer of 2012 when Knight capital’s high frequency trading algorithms entered into erroneous trades that resulted in a pre-tax loss of over $450M USD once the positions were unwound [43].

The later five sources of model uncertainty (sources 3-7) may seem unwarranted in our discussion of developing a quantitative risk measure because they can be avoided with attention to detail and diligence on behalf of the practitioner. This may be true, however, it is important to note that their existence has resulted in the occasionally erratic behaviour of markets and continues to pave the path towards market efficiency and completeness. As such, understanding them will aid in possible inclusion of such dynamics towards proper modeling of the underlying in a derivative contract (i.e. whether jump processes are warranted, etc.). That being said, developing a quantitative risk measure for model uncertainty is more concerned with the applicability or the correctness of a model because these sources of uncertainty will propagate into incorrect
decision making even if there are no errors in the implementation of the model.

Now that we have a sufficient understanding of what model uncertainty is in reference to in the context of financial modeling, the next few sections build on methods to quantify such uncertainty in the development of a risk measure. We first clarify our working definition of risk, uncertainty and their relations to incomplete markets, and associate their relevance to the development of a risk measure. We then describe the underlying fabric of the coherent and convex risk measures and how they emerged. Lastly we explain how these have been extended to include the subtleties of derivative pricing both in a discrete and continuous context (to be explained).

2.2 Risk, Uncertainty & Incomplete Markets

The differences between risk and uncertainty was first highlighted by Knight [44]. In it’s basic form, risk is the lack of knowledge of an outcome from which the probability distribution is known, whereas uncertainty implies a lack of knowledge of the probability distribution that generated the outcomes. More formally and in the context of asset pricing as presented by Cont [15], let us assume that an asset $S$ follows a probability triplet $(\Omega, \mathbb{P}, \mathcal{F})$, where $\Omega$ is the set of potential market scenarios, $\mathcal{F}$ is the subset of potential paths followed or filtration of $\Omega$, and $\mathbb{P}$ is a probability measure on $\mathcal{F}$. Risk is not knowing the resulting price of the asset $S(\omega)$ until the future scenario $\omega \in \Omega$ is realized, and uncertainty is not knowing the probability measure $\mathbb{P} \in \mathcal{P}$ that is generating the future result (assuming a subset $\mathcal{P}$ of probability measures that can describe the space exists). Again, this is saying risk implies not knowing the future state of the market but knowing the probability distribution of market prices and uncertainty is not knowing the probability distribution that is generating the market prices. In literature, our definition of risk as developed in the context of asset pricing, is more commonly known as market risk and uncertainty is actually referring to model uncertainty or model risk (the two terms are inter-changeable, however, model uncertainty is more correct, and we will use it going forward).

In its simplest form, market incompleteness can be referred to as the inability to describe all traded assets using a single probability measure. To make this more clear, we note an informal definition of a derivative security as an instrument that derives its value from another asset (i.e. a contingent claim on an uncertain payoff stream of an asset). Modern day derivative pricing uses arbitrage arguments to construct these derivatives from simpler instruments that are actively traded and thus have liquid prices. Simple examples include commodity futures which are derivatives of the spot commodity (more on that in the next chapter - 3, as well as in the appendix - A), a bond can be viewed as a derivative of the term structure of interest rates (it is not usually defined in this way), or a collateralized mortgage obligation is a derivative of
interest rates and mortgage principle prepayments [22]. Looking at it this way makes the main idea of derivative pricing behave in similar fashion to a puzzle, i.e. the goal is to replicate or re-construct the payoff puzzle of a derivative under all market scenarios using a subset of other tradable instruments as puzzle pieces. If some of the puzzle pieces are missing, the derivative cannot be replicated, and we have an incomplete market (in a mathematical context). More formally, in a complete market, a market participant is expected to know all possible future states of all instruments and is able to eliminate undesirable states with a contingency claim. An incomplete market is one where there are more possible market outcomes that influence a security then there are contingent claims to hedge these possible states.\footnote{Implicitly, allowing for transaction costs is a source of market incompleteness. Nonetheless, assuming market completeness conveniently allows the math to work because you can replicate any contingent claim in real time using no-arbitrage arguments. As such, hedging such a claim with the set of assets that replicate the payoff involves the buying and selling specific amounts of these assets at every possible moment in time to yield a perfect payoff replication (continuous-time models rely on such assumptions). Such an infinite number of transactions would in theory require a payment of an infinite amount of money if transactions costs are not ignored. However, you can show that hedging your position at discrete times over the life of the asset is often sufficient.}

The aforementioned arguments clearly differentiate market uncertainty from an incomplete market, however, they can influence each other in the context of derivative pricing. We recall from elementary derivative pricing the necessity in obtaining the risk neutral measure or pricing measure (also known as the equivalent martingale measure) to price the security. The pricing measure $\mathbb{Q}$ is related to but different from the real or economic measure $\mathbb{P}$ (same $\mathbb{P}$ as in preceding paragraph) in a very fundamental way: the $\mathbb{P}$ measure describes the probabilities of the market prices whereas the $\mathbb{Q}$ measure describes the probabilities of the derivative’s payoff. (Cont differentiates these measures as an econometric measure ($\mathbb{P}$) and a pricing measure ($\mathbb{Q}$) \cite{Cont}). It is important to note that $\mathbb{Q}$ and $\mathbb{P}$ agree on the set of possible events (i.e. they have the same $\mathcal{F}$) and that the risk neutral measure is a change of measure that simply encompasses all investors risk premiums of the real measure in order to simplify the calculation of a derivative price to the expectation of the derivatives payoff under the pricing measure discounted by the risk free rate\footnote{Martingale pricing theory indicates that any tradable asset can be used as the numeraire, and that there exist a unique measure (often denoted by $\mathbb{Q}$) for every numeraire. As such, the risk neutral probability measure is denoted as such because it is the measure resulting from choosing a cash instrument paying the risk-free rate to be the numeraire.}. Emanating from our discussion on model uncertainty and market incompleteness, it is clear that if the $\mathbb{P}$ measure describes a complete market, the $\mathbb{Q}$ measure is unique and model uncertainty derives from the $\mathbb{P}$ measure alone. This result is evident because $\mathbb{P}$ and $\mathbb{Q}$ are defined over the same filtration $\mathcal{F}$, therefore, model uncertainty for a derivative whose pricing measure $\mathbb{Q}$ follows from a complete $\mathbb{P}$ resorts to model uncertainty in a traditional Knightian sense. In an incomplete market, the $\mathbb{Q}$ measure is no longer uniquely defined by $\mathbb{P}$. This is because there exists derivative instruments that cannot be uniquely replicated from existing assets resulting in an infinite number of possible prices and pricing measures. Furthermore, even if the $\mathbb{P}$ measure accurately describes all liquid instruments and is deemed correct, you may still have model uncertainty resulting from the $\mathbb{Q}$ measure because the market is incomplete. Cont’s \cite{Cont}
development of his convex risk measure for derivative securities utilizes the information inherent in the non-unique pricing measure $Q$. Cont’s formulation uses traded instruments to infer possible risk-neutral measures in his convex risk measure. Such a formulation is fundamental in how we will be treating model uncertainty in our study of commodity contingent claims.

### 2.3 Risk Measures

In what follows, we explain the development of the coherent and convex risk measure in the context of risk management practitioners. We begin with an informal definition of a market risk measure and explain how they are used in a risk management context. We follow with a brief dialogue about how developing a useful risk measure for model uncertainty and market risk are different. We conclude this section by linking the discussion of developing a useful risk measure for evaluating model uncertainty to the development of both a coherent and a convex risk measure.

Suppose we have a set of market scenarios $\omega \in \Omega$. Let $\mathcal{X}$ be a set of real valued payoff functions $X$ on $\Omega$ and let $\mathcal{P}$ be a probability measure on $\Omega$. A market risk measure is a mapping from the set $\mathcal{X}$ into $\mathbb{R}$. $\mathbb{R}$ is most often represented as a monetary value, thus, a market risk measure can be interpreted as translating the risk or uncertainty in the payoff of a financial security into a monetary value. Such a monetary value often represents the minimum amount of capital required to finance the potential losses of investing in such a security. As such, a risk measure attempts to evaluate the probability of a series of detrimental random events (random payoff function given a market scenario) that affects a security or a portfolio, and maps these random variables to the equivalent value of an asset (usually cash). In this context, risk measures serve as a useful decision making tool for financial institutions and regulators as they attempt to set the boundaries on the level of risk taking activity of a financial institution (i.e. limit or control its blow-up/left-tail risk).

As explained, model uncertainty is different from market risk. Market risk describes risk in a classical sense, i.e uncertainty in which market scenario $\omega \in \Omega$ will be realized while having a complete knowledge of the distribution $\mathcal{P}$ and possible paths $\mathcal{F}$ the market can take. This type of risk can be hedged if we assume complete markets. Model uncertainty is uncertainty in the distribution generating the market values, i.e. uncertainty of the distribution $\mathcal{P}$ among the set of possible distributions $\mathcal{P}$. Due to these differences, isolating model uncertainty has evolved into an aversion to the worst possible event a series of models produces because you do not know what $\mathcal{P}$ will be generating the actual result. This is in contrast to an aversion to a likely scenario as is common when attempting to minimize market risk and highlights the difference between how practitioners treat market risk from model uncertainty. Market risk looks at the result of averaging over market scenarios and is a natural candidate for valuation
in a Bayesian context. Model uncertainty garners an approach that maximizes the utility of the worst possible payoff resulting from a set of models. Such a representation was developed by Gilboa and Schmeidler [28], they propose the following:

$$\max_{X \in A} \min_{P \in \mathcal{P}} E^P[U(X)]$$

(2.1)

Here it is assumed that the agent has the ability to choose among a set of $A$ possible alternatives. The agents aversion to uncertainty is implicit in finding the minimum (or worst-case) among the set of models $\mathcal{P}$. The agents aversion to risk is represented by the utility function calculated with respect to a particular alternative $U(X)$.

A coherent risk measure naturally follows the idea of averting the worst case scenario [4] and is similar to the worst-case expected utility function as represented by Gilboa and Schmeidler [28]. In fact, Artzner et. al. [4] demonstrate that any coherent risk measure can be represented as the solution to finding the maximum of a linear utility function of the derivatives payoff. They propose the following form:

$$\rho(X) = \sup_{P \in \mathcal{P}} E^P[-X]$$

(2.2)

This implies that any coherent risk measure may be interpreted as the largest possible loss on a derivative contract among a set of possible models, and we note the similar construction to Gilboa and Schmeidler’s [28] representation 2.1.

Artzner et. al. [4] outline a series of properties that are necessary for a risk measure $\rho : E \to \mathcal{R}$ to be coherent. We enumerate the properties of a coherent risk measure below by following the nomenclature as utilized by Cont [15] in order to be consistent when we describe Cont’s risk measure extension [to derivatives]. As such, we will denote $X$ as a possible payoff function of a derivative among the set $E$. The set of payoff functions are applied to the set $\Omega$ of market scenarios to yield a real value $X : \Omega \to \mathcal{R}$ (a particular market scenario will result in a realization of the value of the derivative contract). A risk measure in this context is therefore a mapping from $E \to \mathcal{R}$.

1. **Monotonicity**: If a random market path $\omega \in \Omega$ applied to the payoff functions $X$ and $Y$ is more likely to be profitable for function $X$, $X$ as an instrument or portfolio has less risk than $Y$.

   $$X \geq Y \implies \rho(X) \leq \rho(Y)$$

(2.3)

2. **Risk is measured in monetary units**: If a real riskless monetary value $a \in \mathcal{R}$ was added to a payoff $X$, the model risk of $X$ would be reduced by the value of $a$.

   $$\rho(X + a) = \rho(X) - a$$

(2.4)
3. **Subadditivity**: Diversification of instruments likely reduces risk, i.e. the risk of two individual payoffs will likely offset one another since the payoffs will not yield that same result in every possible market scenario.

\[ \rho(X + Y) \leq \rho(X) + \rho(Y) \] (2.5)

4. **Positive homogeneity**: Assuming \( X \) is a payoff of an option in your portfolio, and you have [a nonnegative quantity] \( \lambda \) of them, the risk you are exposed to is \( \lambda \) times the original risk.

\[ \forall \lambda > 0, \ \rho(\lambda X) = \lambda \rho(X) \] (2.6)

Föllmer and Schnied [26] generalized coherent risk measures by relaxing the subadditivity (2.5) and positive homogeneity (2.6) conditions and replacing them with the following condition:

\[ \forall \lambda \in [0, 1], \ \rho(\lambda X + (1 - \lambda Y)) \leq \lambda \rho(X) + (1 - \lambda) \rho(Y) \] (2.7)

The resultant generalization of a coherent risk measure is referred to as a convex risk measure. We note that such a generalization allows for the inclusion of liquidity risk. Assuming a \( \lambda > 0 \) in the positive homogeneity condition of a coherent risk measure 2.6 indicates a linear relationship between the number of contracts and the risk of the resulting position. Such a linear relationship assumes no incremental risk is incurred by acquiring additional securities which is only possible under the condition of infinite liquidity. A convex risk measure does not rely on such an unreasonable assumption.

Additionally, F’öllmer and Schnied [26] note that if we assume the continuity condition, a convex risk measure can be represented as follows:

\[ \rho(X) = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^\mathbb{P}[-X] - \alpha(\mathbb{P}) \] (2.8)

We notice the similarity that this representation has to that of a coherent risk measure (2.2) represented by Artzner et. al.. The difference is the concave penalty function \( \alpha \), which penalizes a model in the set based on a set of undesirable characteristics (we note that this allows for non-linearities when constructing informative risk measures). We discuss this in more detail when we introduce Cont’s convex risk measure in the next section.
2.4 Cont’s Measures

This section highlights the key developments of Cont’s seminal paper on model uncertainty risk measures. Cont [15] introduces both a coherent and a convex risk measure for evaluating the model uncertainty of a derivative contract, however, we will demonstrate why the convex risk measure is more useful in practice. This subsection begins by outlining a series of assumptions that will properly frame the problems a risk management practitioner may have when attempting to quantify model uncertainty. With this setting in mind, we outline Cont’s description of the necessary tools required for constructing a measure of model uncertainty. Such a description naturally leads to development of both Cont’s coherent and convex risk measures, which will conclude this subsection.

Let an asset $S$ follow a probability triplet $(\Omega, \mathbb{P}, \mathcal{F})$. We assume that a risk management practitioner is interested in pricing an illiquid derivative with terminal random payoff $X \in \mathcal{F}$ that will be revealed at time $T$. Such a practitioner observes a set of liquidly traded derivatives, with payoff $(H_i)_{i \in I}$ and price $(C_i)_{i \in I}$ (note that it is more correct to assume that $C_i$ falls within the bid-ask spread), that are similar to the desired illiquid payoff $X$. We refer to this set as the set of benchmark derivatives and we assume that they are informative in the calibration of a pricing model, i.e. they provide important information about the correct risk-neutral pricing measure. Furthermore, we assume that the practitioner can choose from a set of arbitrage-free pricing models $Q \in \mathcal{Q}$ in order to compute a price for the illiquid asset. For pricing model $Q$ to be valid, the calculated price of each observable benchmark instrument must be within the bid-ask spread\(^3\). This constraint often makes the calibration of a pricing model difficult (omitting this constraint highlights the benefit of Cont’s convex risk measure). Next, we denote a set of derivatives $\mathcal{C}$ (more generally, a set of contingent claims) whose price is finite when calculated using the set of pricing models $\mathcal{Q}\(^4\)$. Lastly, we assume that we have a set of self-financing trading strategies $\phi \in \mathcal{S}$ such that the integral, $\int_t^T \phi_u dS_u$, is tractable for every $Q \in \mathcal{Q}$ (refer to Cont [15] for a detailed description). Such a self-financing strategy is attempting to hedge the contingent claim that the practitioner is trying to price.

Using the aforementioned information, Cont [15] develops a set of conditions/tools that are necessary for constructing a measure of model uncertainty (denoted by $\mu$). They are as follows:

1. Model uncertainty reduces to the bid-ask spread for liquidly traded instruments.

$$\forall i \in I, \mu(H_i) \leq |C_i^{ask} - C_i^{bid}|$$

(2.9)

2. Hedging the position with the underlying can influence the model risk of the derivative.

\(^3\)\(\forall Q \in \mathcal{Q}, \forall i \in I, \mathbb{E}^Q[H_i] \in [C_i^{bid}, C_i^{ask}]\)

\(^4\)\(\mathcal{C} = \{H \in \mathcal{F}_T, \sup_{Q \in \mathcal{Q}} \mathbb{E}^Q[|H|] < \infty\}\)
Recall that the hedging strategy is guaranteed and should have no incremental model risk to the position.

\[ \forall \phi \in S, \mu \left( X + \int_0^T \phi_t \, dS_t \right) = \mu(X) \]  

(2.10)

However, if the derivative can be perfectly replicated in a model-free way, hedging should offset the model risk.

\[ \exists x_0 \in \mathbb{R}, \exists \phi \in S, \forall Q \in \mathcal{Q}, \ Q \left( X = x_0 + \int_0^T \phi_t \, dS_t \right) = 1 \Rightarrow \mu(X) = 0 \]  

(2.11)

3. Model risk should decrease through diversification thus the convexity property is implicit.

\[ \forall X_1, X_2 \in \mathcal{C}, \forall \lambda \in [0, 1], \mu(\lambda X_1 + (1 - \lambda)X_2) \leq \lambda \mu(X_1) + (1 - \lambda)\mu(X_2) \]  

(2.12)

4. Static hedging with other traded derivatives can influence the model risk of the derivative. In similar fashion to 2.10, the hedging strategy is guaranteed but each option has model risk of 2.9, i.e., the bid-ask spread.

\[ \forall X \in \mathcal{C}, \forall u \in \mathbb{R}^K, \mu \left( X + \sum_{i=1}^K u_i H_i \right) \leq \mu(X) + \sum_{i=1}^K |u_i| (C_i^{ask} - C_i^{bid}) \]  

(2.13)

However, if the derivative’s payoff can be perfectly replicated using the traded options, hedging should offset the model risk to the bid-ask spread of the traded options.

\[ \exists u \in \mathbb{R}^K, X = \sum_{i=1}^K u_i H_i \Rightarrow \mu(X) \leq \sum_{i=1}^K |u_i| (C_i^{ask} - C_i^{bid}) \]  

(2.14)

Now by including the framework of equations 2.9-2.14 we define Cont’s coherent measure of model uncertainty for a well defined payoff \( X \in \mathcal{C} \) that can be priced by a set of models \( \mathcal{Q} \in \mathcal{Q} \) as:

\[ \mu_Q = \pi(X) - \pi(X) \]  

(2.15)

where we define a lower and upper price bounds as:

\[ \pi(X) = \sup_{Q \in \mathcal{Q}} \mathbb{E}^Q [X] \quad \pi(X) = \inf_{Q \in \mathcal{Q}} \mathbb{E}^Q [X] \]  

(2.16)

For a set of well-behaved pricing models \( \mathcal{Q} \) (recall that a valid pricing model \( Q \in \mathcal{Q} \) calculates a price of each observable benchmark instrument within the bid-ask spread) and adheres to a market where equations 2.9-2.14 are valid, Cont’s coherent model risk is simply the difference between the upper and lower bound prices from the set of models (note that the bounds comply with the representation of a coherent risk measure 2.2 as described by Artzner et. al. [4]).
Choosing a set of well-behaved pricing models is often difficult because calibrating their outputs to a set of observed benchmark securities may be challenging and thus limits the utility of the risk measure. Cont’s convex risk measure relaxes such a constraint by introducing a penalty function in the objective function that penalizes a particular pricing measure if it is unable to produce a price for a benchmark option that is within the bid-ask spread.

\[
\pi^*(X) = \sup_{Q \in \mathcal{Q}} E^Q[X] - ||C^* - E^Q[H]|| \\
\pi_*(X) = \inf_{Q \in \mathcal{Q}} E^Q[X] + ||C^* - E^Q[H]||
\] (2.17)

The vector norm in equation 2.17 is calculating a penalty for any degree of imprecision in the calculation of the observed price, where \(C^* \in [C^{ask}, C^{bid}]\) and \(H \in \mathcal{F}\) are the vectors of the \(i \in I\) observed prices and payoffs respectively (note any vector norm will suffice in such a representation). We note that the vector norm is equivalent to the \(\alpha(\mathbb{P})\) penalty function of Föllmer and Schied’s \([26]\) convex risk measure representation of equation 2.8. Cont’s convex risk measure of model uncertainty can now be written in a similar manner to the coherent measure of equation 2.15.

\[
\mu_* = \pi^*(X) - \pi_*(X)
\] (2.18)

There is one caveat when implementing Cont’s convex measure of model uncertainty, one of the chosen models \(Q \in \mathcal{Q}\) must oblige to the framework of a coherent risk measure, i.e. specifically equations 2.9-2.12\(^5\). This means that one of the pricing models must replicate the observed benchmark prices within the bid-ask spread. This ensures tractability of the bound and ensures a non-trivial result such as a negative bound for an asset (Refer to Cont’s paper \([15]\) for the proof and further arguments.).

### 2.5 Li & Kwon Extension

Li and Kwon \([45]\) propose a new framework to calculate Cont’s convex measure of model uncertainty, equation 2.18, by developing a methodology that uses an infinite set of pricing measures (semi-parametric, specified through a set of moment conditions) as opposed to the discrete set described by Cont \([15]\). Li and Kwon \([45]\) argue that the result obtained when using the infinite set of pricing measures is a more informative measure of model uncertainty because it allows for a much wider class of distributions. In this section we discuss the theoretical foundations of Li and Kwon’s formulation as well as the techniques required for developing the solution. The work by Li and Kwon is applied to European options because the payoff structure of these derivatives are well-behaved. We begin this section by briefly outlining the context of the theoretical formulations that will follow. The theoretical foundation is a series

---

\(^5\)Equations 2.13 and 2.14 have similar representation in Cont’s convex measure take into account the subtleties of incorporating the penalty function when calculating the effect of static hedging the derivative.
of convex optimization techniques used in solving the problem of determining the bounds of European option prices when the moments of the underlying distribution are known. This work was developed by Bertsimas and Popescu [5] and subsequently generalized by Gotoh and Konno [31], we explain their respective contributions. The framework of calculating the bounds of an option price using convex optimization techniques is incorporated by Li and Kwon in their work on reformulating Cont’s convex measure for model uncertainty. We conclude this section by describing this framework and reformulation.

A common goal for a risk management practitioner is to calculate the fair price of a derivative security. In order for them to accomplish this, they often look to the theoretical results of derivative pricing, such as the famous Black-Scholes formula, many of which make the assumption that the underlying asset follows a functional form. Assuming the asset follows a functional form may lead to an analytical result for calculating the price of the derivative once standard no-arbitrage assumptions are applied, however, relaxing this assumption is more useful to a practitioner concerned with model uncertainty. The framework for making no assumptions on the underlying price dynamics was developed by Bertsimas and Popescu [5]. Under the no arbitrage assumption, Bertsimas and Popescu posit the question of determining the best possible bounds on the price of a derivative security with a well-behaved payoff function if we assume the \( k \) moments of the underlying distribution are known (this implicitly assumes that a practitioner is able to empirically obtain the moments of an assets distribution). The solution to this question implies that you are attempting to infer the possible derivative prices as calculated by the infinite set of non-parametric models, thus is implicitly linked to model uncertainty.

Bertsimas and Popescu [5] represent the aforementioned problem using a European call option. A European call option was used because its payoff function is path independent, i.e., the payoff must only be realized at the expiration of the derivative contract. The optimization framework requires such path independency in order to develop tractable and stable bounds. We now present the upper bound of a European call option with strike price \( k \) as a constraint optimization:

\[
\begin{align*}
\text{maximize} & \quad \mathbb{E}^Q[\max(0, X - k)] = \int_0^\infty \max(0, x - k)Q(x)dx \\
\text{subject to} & \quad \mathbb{E}^Q[X^i] = \int_0^\infty x^iQ(x)dx = q_i, \quad i = 0, 1, \ldots, n \\
& \quad Q(x) \geq 0,
\end{align*}
\]

This representation is difficult to solve, however, using standard duality theory, the dual problem
can be represented as:

\[
\minimize \sum_{r=0}^{n} y_{r} q_{r}
\]

subject to \( \sum_{r=0}^{n} y_{r} x^{r} \geq \max(0, x - k) \quad \forall x \in \mathbb{R}_{+} \) \hfill (2.20)

Bertsimas and Popescu [5] imply that strong duality holds between equations 2.19 and 2.20, and demonstrate how the dual problem can be solved efficiently. In order for the dual to be tractable, the polynomials in the constraints have to be contained. As such, Bertsimas and Popescu [5] outline a framework for the bounds of a subset of polynomial functions that can be applied to the dual constraints. This framework is succinctly described in proposition 1 of their paper [5]. We note that Gotoh and Konno [31] make a slight modification to the proposition by stating that Bertsimas and Popescu’s work is only true under the additional condition that the variable \( x \) is strictly positive. Gotoh and Konnon [31] relax this condition to allow for negative values. Since we are using European options, this additional condition is not relevant, and we enumerate the results of Berstimas and Popescu below:

(a) The polynomial \( g(x) = \sum_{r=0}^{2k} y_{r} x^{r} \) satisfies \( g(x) \geq 0 \) if and only if there exists a positive semidefinite matrix \( X = [x_{i,j}]_{i,j=0,...,k} \), such that

\[
y_{r} = \sum_{i,j: i+j=r} x_{i,j}, \quad r = 0, \ldots, 2k, \quad X \succeq 0 \quad (2.21)
\]

(b) The polynomial \( g(x) = \sum_{r=0}^{n} y_{r} x^{r} \) satisfies \( g(x) \geq 0 \quad \forall x \in [0, a] \) if and only if there exists a positive semidefinite matrix \( X = [x_{i,j}]_{i,j=0,...,n} \), such that

\[
0 = \sum_{i,j: i+j=2l-1} x_{i,j}, \quad l = 1, \ldots, n, \\
\sum_{r=0}^{l} y_{r} \binom{n-r}{l-r} a^{r} = \sum_{i,j: i+j=2l} x_{i,j}, \quad l = 0, \ldots, n, \\
X \succeq 0 \quad (2.22)
\]

(c) The polynomial \( g(x) = \sum_{r=0}^{n} y_{r} x^{r} \) satisfies \( g(x) \geq 0 \quad \forall x \in [0, \infty] \) if and only if there exists

\[X \succeq 0 \] implies that the matrix \( X \) is among the set of semi-definite matrices.
a positive semidefinite matrix \( X = [x_{i,j}]_{i,j=0,...,n}, \) such that

\[
0 = \sum_{i, j: \ i + j = 2l - 1} x_{i,j}, \quad l = 1, \ldots, n,
\]

\[
0 = \sum_{i, j: \ i + j = 2l - 1} x_{i,j}, \quad l = 0, \ldots, n,
\]

\[
X \succeq 0 \tag{2.23}
\]

(d) The polynomial \( g(x) = \sum_{r=0}^{k} y_r x^r \) satisfies \( g(x) \geq 0 \ \forall x \in [a, b] \) if and only if there exists a positive semidefinite matrix \( X = [x_{i,j}]_{i,j=0,...,n}, \) such that

\[
0 = \sum_{i, j: \ i + j = 2l - 1} x_{i,j}, \quad l = 1, \ldots, n,
\]

\[
\sum_{m=0}^{l} \sum_{r=m}^{n+m-l} y_r \binom{r}{m} \binom{n-r}{l-m} a^{r-m} b^m = \sum_{i, j: \ i + j = 2l} x_{i,j}, \quad l = 0, \ldots, n,
\]

\[
X \succeq 0 \tag{2.24}
\]

There are two feasible regions of the dual problem 2.20, i.e. whether or not the European call option expires in the money. If the asset finishes in the money \((x > k)\), the polynomial \( x - k \) is brought to the left hand side of the constraint, and proposition (c) can be applied to the resulting polynomial. Similarly, if the asset does not finish in the money, the constraint does not change, and proposition (b) can be applied. Using proposition (b) and (c), the dual problem is reformulated as a semi-definite program, and can be efficiently solved [5]. Intuitively, it is possible to separate the constraints into \( d \) discrete regions that represent the range of values the underlying asset may have at expiration. The \( d \) regions should be chosen in order to represent the different payoff function regimes of the derivative in question. In this case, the strike price is the price that determines whether a European option ends up in or out of the money and is therefore the logical point of separation. More formally, if we assume a well-behaved payoff function \( \phi(x) \) can be separated into \( d \) regions, we have the following:

\[
\phi(x) = \begin{cases} 
\phi_0(x), & x \in [0, k_1] \\
\phi_1(x), & x \in [k_1, k_2] \\
\vdots & \vdots \\
\phi_{d-1}(x), & x \in [k_{d-1}, k_d] \\
\phi_d(x), & x \in [k_d, \infty]
\end{cases} \tag{2.25}
\]

Now, prepositions (a)-(d) can be applied to each of the \( d \) regions, and any problem with a well-behaved derivative payoff can be reformulated as a semi-definite program.

As mentioned in the preamble of this section, Li and Kwon reformulate Cont’s convex measure
of model uncertainty by considering an infinite set of measures specified by the first $l$ moments of the underlying distribution, or as represented by Li and Kwon [45]:

$$Q := \{ Q : \mathbb{E}^Q[M_i(\omega)] = m_i, \quad i = i, \ldots, l \}$$  \hspace{1cm} (2.26)

We are now equipped with the tools to begin reformulating Cont’s convex measure of model uncertainty for derivatives using the infinite set of measures described by the set 2.26. We begin by re-stating the calculation of the upper bound problem, $\pi^*$, as the following constrained optimization\(^7\).

$$\pi^* := \sup_{Q \in \mathbb{Q}, (q_j)_{j=1,\ldots,k} \in \mathbb{Q}(q_j)_{j=1,\ldots,k}} \mathbb{E}^Q[X(\omega)] - ||h^* - \text{vec}((q_j)_{j=1,\ldots,k})||$$

subject to

$$\mathbb{E}^Q[M_i(\omega)] = m_i, \quad i = 0, 1, \ldots, l$$

$$\mathbb{E}^Q[H_j(\omega)] = q_j, \quad j = 0, 1, \ldots, k$$  \hspace{1cm} (2.27)

Here, we denote $\omega \in \Omega$ as a possible risk-neutral sample path driven by an instance of the probability triplet $(\mathbb{Q}, \mathcal{F}, \Omega)$. $X(\omega)$ is the targeted payoff of the derivative that we are calculating the model uncertainty of. $H_j(\omega)_{j=0,1,\ldots,k}$ and $h^* = h_j(\omega)_{j=0,1,\ldots,k}$ are the set of benchmark derivative payoffs and market prices respectively of observed liquid derivatives that are similar to the target derivative. $M_i(\omega)_{i=0,1,\ldots,l}$ is the set of moment functions, $(\omega^l)_{i=0,1,\ldots,l}$, as applied to the underlying distribution. The set $(m_l)_{i=0,1,\ldots,l}$ is the set of empirical risk-neutral moments of the underlying distribution as calculated by the practitioner. The set $q_j(\omega)_{j=0,1,\ldots,k}$ are slack variables that represent the prices of the observed benchmark derivatives calculated under each measure in the set $(Q \in \mathbb{Q})$.

Li and Kwon introduce the following primal-dual relationships that they use to represent the reformulation of Cont’s convex measure of model uncertainty into a more tractable form (we use the notation from the paper [45]). Consider the problems

$$p_{\text{sup}} = \sup_Q \int_S \psi(\zeta) d\mathbb{Q}(\zeta) : \int_S \varepsilon(\zeta) d\mathbb{Q}(\zeta) = \varepsilon_0, \quad \int_S d\mathbb{Q}(\zeta) = 1,$$  \hspace{1cm} (2.28)

$$p_{\text{inf}} = \inf_Q \int_S \psi(\zeta) d\mathbb{Q}(\zeta) : \int_S \varepsilon(\zeta) d\mathbb{Q}(\zeta) = \varepsilon_0, \quad \int_S d\mathbb{Q}(\zeta) = 1,$$  \hspace{1cm} (2.29)

where $\mathbb{Q}$ is a non-negative probability measure, $\psi$ and $\varepsilon$ are both continuous functions that map $\mathbb{R}^n \to \mathbb{R}$ and $\mathbb{R}^n \to \mathbb{R}^m$ respectively, while $\varepsilon_0 \in \mathbb{R}^m$. The dual problems can be represented as:

$$d_{\text{sup}} = \inf_{\lambda_0, \lambda_\varepsilon} \lambda_0 + \Lambda_\varepsilon \cdot \varepsilon_0 : \lambda_0 + \Lambda_\varepsilon \cdot \varepsilon(\zeta) \geq \psi(\zeta), \quad \forall \zeta \in S$$

$$= \inf_{\lambda_\varepsilon} \sup_{\varepsilon \in S} \psi(\zeta) - \lambda_\varepsilon \cdot \varepsilon(\zeta) + \Lambda_\varepsilon \cdot \varepsilon_0$$  \hspace{1cm} (2.30)

\(^7\)We note that the lower bound problem can be represented in almost an identical form, the objective becomes $\pi_* := \inf_{Q \in \mathbb{Q}, (q_j)_{j=1,\ldots,k}} \mathbb{E}^Q[X(\omega)] + ||h^* - \text{vec}((q_j)_{j=1,\ldots,k})||$, and the constraints are the same.
\[ d_{inf} = \sup_{\lambda_0, \Lambda_e} \lambda_0 + \Lambda_e \cdot \varepsilon_0 : \lambda_0 + \Lambda_e \cdot \varepsilon(\zeta) \leq \psi(\zeta), \quad \forall \zeta \in S \]
\[ = \sup_{\lambda_0} \inf_{\Lambda_e} \psi(\zeta) - \Lambda_e \cdot \varepsilon(\zeta) + \Lambda_e \cdot \varepsilon_0 \] (2.31)

where \( \lambda_0 \in \mathbb{R} \) and \( \Lambda_e \in \mathbb{R}^m \). Li and Kwon demonstrate that strong duality holds, i.e. \( p_{inf} = d_{inf} \) and \( p_{sup} = d_{sup} \).

Using the results of equations 2.30 and 2.31, Li and Kwon convert the upper bound constrained optimization problem 2.27 into its more tractable dual representation:

\[ \pi^* := \inf_{\lambda_m, \lambda_h, s, t} s + t \]
subject to \( X(\omega) - \sum_{i=l}^{l} \lambda_m i (M_i(\omega) - m_i) - \sum_{j=l}^{k} \lambda_h j H_j(\omega) \leq s \quad \forall \omega \geq 0 \) (2.32)

The reformulation of the lower bound \( \pi_* \) is very similar to equation 2.328.

Our discussion on the work presented by Li and Kwon [45] thus far has been generalized for any derivative. We now present the formulation with our payoff of interest, the European call option with strike price \( k_0 \), i.e. \( X(\omega) = \max(0, \omega - k_0) \). Furthermore, as our benchmark derivatives, we will use a set of \( k \) European options that are written on the same underlying asset, have the same time expiration, but have different strike prices. The set of \( k \) benchmark options will include \( o \) call options with payoffs \( \max(0, \omega - k_j) \) and \( k - o \) put options with payoffs \( \max(0, k_j - \omega) \). Using this information, the first constraint of equation 2.32 can be rewritten as follows:

\[ \max(0, \omega - k_0) - \sum_{i=l}^{l} \lambda_m i (M_i(\omega) - m_i) - \sum_{j=l}^{o} \lambda_h j \max(0, \omega - k_j) - \sum_{j=o+1}^{k} \lambda_h j \max(0, k_j - \omega) \leq s \quad \forall \omega \geq 0 \]

Written this way, the constraint can be separated into a series of regions as determined by the total number of different option strike prices in the problem formulation, i.e. it can be separated as indicated in equation 2.25. Each region will have a stable polynomial function and can use Berstimas and Popescu’s propositions (a)-(d) (2.21-2.24) to be solved as a semi-definite program [5]. We note that \( \forall \omega \geq 0 \) is implicit in Berstimas and Popescu’s propositions.

In order to tackle the second constraint of equation 2.32, we make a few adjustments. First, we notice that if we remove the \( \leq t \) from the problem, we have a separate linear program (LP). In the new LP problem, we replace the norm function from the objective with the 1-norm and introduce the slack variables \( vec((z_j)_{j=1,\ldots,k}) \) to represent the individual absolute value

---

8The objective changes to \( \pi_* := \sup_{\lambda_m, \lambda_h, s, t} s + t \), while the second constraint changes slightly to
\[ \sup_{(q_j)_{j=1,\ldots,k}} \sum_{j=1}^{k} \lambda_h j q_j + ||h^* - vec((q_j)_{j=1,\ldots,k})|| \leq t \]
calculations, i.e. \( z = \| h^* - vec((q_j)_{j=1,\ldots,k}) \|_1 = \| h^* - vec((q_j)_{j=1,\ldots,k}) \|_1 \), where \( z_j = | h^*_j - q_j | \).

Introducing the slack variables \( z \) allows us to add the \( k \) 1-norms as additional constraints in the new problem. The objective variables for this linear program are vectors \( z = vec((z_j)_{j=1,\ldots,k}) \) and \( q = vec((q_j)_{j=1,\ldots,k}) \), which we denote as \( z \) and \( q \) to simplify the notation. By rearranging the constraints such that the objective variables are on the LHS, the linear program can be represented in the standard primal form. Using dual vector \( y = vec((y_j)_{j=1,\ldots,2k}) \), the dual problem can be represented as follows:

\[
\begin{align*}
\inf_{y=vec((y_j)_{j=1,\ldots,2k})} & \quad h^*_1 y_1 + \ldots + h^*_k y_k - h^*_1 y_{k+1} \ldots - h^*_k y_{2k} \\
\text{subject to} & \quad y_1 - y_{k+1} \geq \lambda_{h_1} \\
& \quad \vdots \\
& \quad y_k - y_{2k} \geq \lambda_{h_1} \\
& \quad -y_1 - y_{k+1} \geq -1 \\
& \quad \vdots \\
& \quad -y_k - y_{2k} \geq -1
\end{align*}
\]

We recall that the original problem of constructing the upper bound of Cont’s convex risk measure was reformulated into its dual representation of equation 2.32. We have reconstructed the second constraint of this reformulation into the above equation after dropping the \( \leq t \) constraint. Notice that now both the dual representation of the initial problem and the reformulation of the second constraint into its dual are both minimization problems (\( \inf \)). As such, the second constraint can drop the \( \inf \) and add the vector \( y = vec((y_j)_{j=1,\ldots,2k}) \), into the objective. The final problem can now we represented in the following form:

\[
\pi^* := \inf_{\lambda_m,\lambda_h,s,t,y} \quad s + t \\
\text{subject to} \quad \max(0, \omega - k_0) - \sum_{i=1}^{l} \lambda_{m_i} (\omega^i - m_i) - \\
\sum_{j=l}^{o} \lambda_{h_j} \max(0, \omega - k_j) - \sum_{j=o+1}^{k} \lambda_{h_j} \max(0, k_j - \omega) \leq s \quad \forall \omega \geq 0
\]

\[
\begin{align*}
& \quad h^*_1 y_1 + \ldots + h^*_k y_k - h^*_1 y_{k+1} \ldots - h^*_k y_{2k} \leq t \\
& \quad y_1 - y_{k+1} \geq \lambda_{h_1} \\
& \quad \vdots \\
& \quad y_k - y_{2k} \geq \lambda_{h_1} \\
& \quad -y_1 - y_{k+1} \geq -1 \\
& \quad \vdots \\
& \quad -y_k - y_{2k} \geq -1
\end{align*}
\]  

(2.33)
Chapter 3

Commodity Modeling

This chapter is divided into two sections, first we discuss the empirical studies that serve as the foundation for determining the empirical properties in commodities (primarily sources of mean-reversion), and second, we outline the theoretical studies developed to model these properties.

As outlined, the first section explains the sources of mean-reversion in commodity dynamics. We explain the relationship between the convenience yield and spot price as dictated by the Theory of Storage, the relationship that spot price has to time varying risk premia as hypothesized by the Theory of Normal Backwardation, and the relationship of commodity price volatility to the maturity of the contract, deemed the Samuelson Effect. We conclude this section with a brief overview on how these theories imply mean-reversion in commodity prices and we foreshadow our subsequent discussion of theoretical models by explaining the implications of these studies.

The second section summarizes the literature on theoretical commodity price dynamics models. We briefly introduce the approaches academics have taken towards developing theoretical models before dedicating a subsection on reduced form models. We conclude with a few subsections dedicated to the reduced form models chosen to evaluate Cont’s convex risk measure.

3.1 Empirical Studies

As mentioned, this section explains the theories behind the sources of commodity spot price mean-reversion, primarily the Theory of Normal Backwardation, the Theory of Storage, and the Samuelson Effect, and we explain how commodity price mean-reversion is a natural consequence of the aforementioned theories. Each of these theories attempt to empirically explain the dynamics of the time-varying term structure of futures. It is the result of these empirical studies
that have underpinned the understanding of commodity term structure, and are the basis for the theoretical development of modeling the continuous time dynamics of commodities that we discuss in the subsequent section. We touch upon the link between empirical and theoretical studies to conclude this section.

3.1.1 Theory of Normal Backwardation

The Theory of Normal Backwardation begins with the works of Keynes [42] and was complemented by the work of Hicks [36]. We briefly outline how their studies have contributed to the foundation and principles associated with this theory. We conclude with a brief discussion of key empirical studies that extend the more general framework of identifying a time varying risk premia.

In 1930, Keynes postulation that speculators should be rewarded for participating in short-term futures market is the foundation of the existence of a time varying risk premia in futures term structure [42]. He theorized that a producer is willing to sell commodity futures at a discount to the expected spot commodity in order to transfer the price risk to speculators. He believed that this behaviour of risk transfer between a hedger and a speculator constitutes the normal expected behaviour between these participants, and thus his theory is aptly named the theory of normal backwardation. The transfer of risk from hedgers to speculators has a natural downward pressure on futures prices because the speculator will only accept the risk if they are given a sufficient risk premium to bear this risk. Intuitively, the existence of a risk premium forces the futures price to be lower than the expected future spot price, \( E[S_{t,T}] > F_{t,T} \), and this difference is the risk premium paid to the speculator in compensation for bearing the spot risk, \( E[S_{t,T}] - F_{t,T} = \pi_{t,T} \), where \( \pi_{t,T} \) is time-\( t \) risk premia for a futures contract maturing at time \( T \). The size of the risk premium is proportional to the level of expected price risk of hedgers. It is important to note that this theory does not imply a backwardated term structure but rather associates the state of normal backwardation to the relationship between the expected future spot price and the futures price.

In 1939, Hicks furthered the theory of normal backwardation by expanding on why the theory postulates a downward price bias associated with active short hedgers while does not consider the opposing set of circumstances to have much influence, namely an upward price bias associated with active long hedgers [36]. He deduced that producers are more vulnerable to short-term spot price volatility than are consumers and are therefore more active short-hedging participants than consumers are long-hedging participants. His deduction relies on the business risk of a producer needing to fix a selling price of their product relative to a consumer wanting to fix an input cost. Selling the product influences nearly the entire profit margin of the producer, whereas the variable input cost is a smaller fraction of the profit of the consumer (we defined consumer in this context as a corporation that is utilizing the commodity, such as
oil, as an input cost in their own downstream production of a refined good). The additional short hedging pressure leads to a natural weakness on the demand side of the market that a speculator is happy to participate in if a sufficient risk premium is available for this service. Long-hedging pressure inducing supply side weakness in the market is not as common due to the aforementioned arguments.

The empirical research performed to determine the validity of the theory of normal backwardation focussed on quantifying the risk premia. In 1960 and 1967, working with wheat futures, Cootner supported the existence of risk premiums as described by the theory of normal backwardation. Furthermore, Cootner argues in favour of hedging pressure from both short and long hedgers (also known as net hedging pressure hypothesis [19], [20]) and claims that a speculator can earn profit by entering into a long position at the peak of short hedging pressure, and vise versa. Tesler questioned the existence of risk premiums [62]. Working postulated that the downward bias may be too small to notice empirically in a daily time series of commodity prices [67], which can be attributed to the low signal-to-noise ratio of risk premiums relative to the volatility of the daily returns\footnote{Both Working and Tesler were influential towards the theory of storage.}. To compensate for the noise, both Fama and French [24] as well as Gorton and Rouwenhorst [29] identified a positive risk premium in a diversified portfolio of commodities (the diversification dampens the noise of the signal).

Other research extending the theory of normal backwardation has attempted to link the positions of speculators and hedgers to risk premiums. Such studies evaluate the profitability of trading strategies that are driven on the positions of speculators and hedgers. Chang [13] was able to determine that prices rose more often in months where speculators had a net long position and fell more often in months where hedgers had a net long position. This realization led Chang to believe that the net activity of participants are an important determinant of risk premiums. Gorton, Hayashi, and Rouwenhorst [30] reject the role that market participants have in determining risk premiums and attribute it to changes in inventories.

3.1.2 Theory of Storage

The Theory of Storage as is it known today traces its origins to the collective work of Kaldor [40], Working [66, 67], Brennan [9] and Telser [62]. In contrast to Keynes and Hicks who focus on the idea that the supply of speculative services drive commodity dynamics, Kaldor, Working, Brennan and Telser focus on how the supply of storage influences commodity dynamics. The theory of storage began with the insights of Kaldor, and was expanded and verified by the empirical works of Working, Brennan, and Telser. We discuss their studies and contributions in what follows.

In his 1939 study [40], Kaldor postulated that the yield for commodity inventory holders is
split up into two components. The first is the cost of financing and the cost of storing the physical commodity. The second is the benefit of being able to use the inventory the moment it becomes commercially viable to do so (termed convenience yield). Kaldor postulated that the components of the yield are observable in the difference between the spot commodity ($S$) and futures ($F$) with the following representation: $F - S = storage costs + interest costs - convenience yield$.

In 1948 and 1949 [66, 67], Working validated the work of Kaldor with empirical evidence. He used physical futures data (wheat futures) and incorporated as much inventory information as was available towards the study of the empirical relationship between the basis and the prevailing level of inventory. As expected, he found the relationship to be non-linear and inversely related, thus confirming Kaldor’s theory\(^2\), and his empirical study resulted in two concluding remarks. The first being that backwardation is a reliable indicator of the scarcity of the commodity, and the second is that in times of scarcity, Kaldor’s convenience yield was large enough to induce a negative price of storage, i.e. the owner of inventory is being compensated more than the physical costs. In 1953 [68], Working argued that short-hedgers are not solely motivated by risk avoidance activities. He proposes an alternative view that the hedging activity of producers are an expected operational activity of merchandizing and that owning a good is a risk taking activity. He explains that short-hedging a commodity futures contract is often attributed to basis speculation because the producer is holding inventory (in this case, the expectation of increasing basis). Furthermore, the consumer of the commodity during periods of high prices may buy the commodity futures and sell the output product forward when the processing margin is large enough (the example is the flour miller buying wheat futures and selling forward flour to lock in the margin). The motivating factor of the producer in this action is clearly risk taking, while the motivation of a consumer is clearly risk avoidance (no speculator interactions were discussed). These principles are in contrast to the risk avoidance principles that holders of inventory are expected to have in the theory of normal backwardation.

In 1958 [9], Brennan studied five commodities including those that did not have futures contracts (eggs, butter, cheese, wheat and oats). Brennan studied the relationship between inventory and basis and was able to confirm the same inverse relationship as Working using a larger subset of commodities. More importantly however, Brennan expanded the definition of convenience yield by including the unobserved benefits of holding a commodity from the perspective of both the producer and a wholesaler. He argues that from the perspective of a producer, being able to take advantage of a rise in demand and prices without having to adjust their production schedule has value. From the perspective of a wholesaler in a similar environment, being able to respond to an increase in the flow of orders has value. The summation of these benefits is incorporated into the value of the convenience yield.

\(^2\)Recall that a highly backwardated futures term structure implies negative basis and according to the theory of storage, a low level of inventory.
In 1958 [62], similar to Working and Brennan, Tesler also studied the relationship between basis and inventory using cotton and wheat. He was also able to confirm the same relationship between inventories and basis as those before him, however, he identified the relationship how seasonal patterns of a commodities harvest affect its convenience yield. He noted that the patterns of inventory build-up and decline follow its harvest schedule, and that the convenience yield tends to be higher before a harvest when inventory has not yet been replenished.

### 3.1.3 Samuelson Hypothesis

The Samuelson Hypothesis was an inadvertently generalized proclamation by Samuelson in his 1965 study [58]. Samuelson sought out to provide a theorem (generalized stochastic model with unspecific distribution) that the next period change in futures price is uncorrelated with any previous price changes and is therefore unpredictable. Samuelson models the behaviour of futures prices (uses wheat futures as his example), and in the development of his models arguments, he proclaimed that *it is a well known rule of thumb that nearness to expiration date involves greater variability or riskiness per hour or per day or per month than does farness.* This proclamation implies that longer dated maturity futures have less variance than nearer dated maturity futures, and became what is know known as the Samuelson hypothesis.

Samuelson does not provide any formal proof of Samuelson hypothesis but rather develops a representative mathematical example. Samuelson’s example relies on a mathematical description of futures prices where he assumes that a futures price is the expectation of the future spot price and the spot price follows a first order autoregressive process (AR(I))with a coefficient equal to 0.5. With this construction, Samuelson argues that a larger fraction of shared potential sources of variation exist among longer dated maturity futures contracts than among shorter termed maturity futures contracts. He claims that sharing more elements of potential variation\(^3\) stabilizes the return stream because the longer dated maturity futures prices may approach an ergodic state and have lower variance (he notes that approaching an ergodic state is not necessary in his model construction). The specification of Samuelson’s example (AR(I) model with coefficient of 0.5), implies that the spot price is stationary and reverts to a long term mean of zero. This is important because his model description implies mean reversion in the spot price by construction, which although never explicitly stated, can be inferred as a consequence of the Samuelson hypothesis.

Empirical studies of the Samuelson hypothesis by Anderson in 1985 as well as by Milonas in 1986 have verified the theory for agriculture markets [3, 50]. However, the hypothesis was not nearly as prominent in metals (both precious and industrial) as indicated by Anderson and by

---

\(^3\)Sharing more elements of variation means that the noise component in the AR(I) process of each futures price maturing at time \(t : t < T\) is contained in all futures prices maturing at time \(z : z > T\). Therefore, all longer dated futures prices (when looking at the term structure of futures) encompasses the variation of all shorter dated futures thus each incremental maturing futures has less variability relative to a shorter maturity.
Milonas [3, 50], and no evidence in fixed income and currency futures as researched by Milonas as well as by Grammatikos and Saunders in 1986 [50, 32]. There have been attempts to describe the economic reasoning for the existence of the Samuelson effect as well as identify why it exists in some markets but does not in others.

In 1983, Anderson and Duthie [2] argue that the Samuelson hypothesis is a consequence of the rate of information flow for a maturing contract. The information that they refer to are related to the factors that may affect the price of the underlying commodity of a futures contract. These factors are priced into the futures contract at an increasing rate because the information becomes more readily available as the maturity of the contract nears. A simple example is how the prevailing weather conditions affect agriculture futures. Information about the weather conditions become readily available as the contract nears maturity, thus the *informational flow* of the weather and its possible effects become known and are priced into the market. Anderson and Duthie argue that the fluctuations in price for a commodity become large as the rate of information about the commodity becomes clear. This implies that the uncertainties of a commodity become realized as the contract matures and are priced-in at an increasing rate (resulting in a higher volatility). Similarly, a longer maturing contract has little informational content with regards to realizing the implicit uncertainties.

In 1996, Bessembinder, Coughenour, Seguin and Smeller reject Anderson and Duthie’s assertion that the Samuelson effect is induced by the increased informational flow of a maturing contract [6]. They argue that they were unable to identify any reasons for information to cluster around futures delivery dates and assert that it is not a necessary condition to induce the Samuelson effect. Bessembinder *et al.* predict that the Samuelson hypothesis will only hold in those markets where there is an expectation of reversing spot price changes in the near future (essentially mean reversion). Furthermore, they argue that an investor will only hold the spot commodity if they are sufficiently compensated to do so in order to offset the imminent spot losses. Such compensation is reflected in the slope of the futures term structure, and Bessembinder *et al.* were able to predict that the Samuelson hypothesis is observable in markets that have a negative covariation between the changes in spot price and slope of the futures term structure. They conclude that such markets must have a positive covariation between convenience yields and spot prices which leads to mean reversion in spot price (as predicted by the Samuelson hypothesis).

### 3.1.4 Inducing Mean Reversion

We begin by summarizing the key empirical findings of the three pillars of commodity investment:

(a) Keynes and Hicks predict that long investors in commodity futures have historically earned
a risk premium because the futures price is a downward bias estimate of the expected spot price.

(b) Commodity futures risk premiums vary across commodities and over time depending on the level of physical inventories, as predicted by the theory of storage (negatively associated, i.e., low inventories induce high risk premiums).

(c) Convenience yield is a decreasing, non-linear relationship of inventories.

(d) A large convenience yield induces a backwardated futures term structure.

(e) Price measures, such as the futures basis, prior futures returns, and spot returns reflect the state of inventories and are informative about commodity futures risk premiums.

(f) Price volatility is a decreasing function of futures maturity as predicted by the Samuelson Hypothesis.

Each and every empirical finding is an explanation of why the spot and futures price of commodities mean revert. It is clear that the primary drivers of mean reversion are time variation in the convenience yield and risk premiums. However, what is unknown is the exact non-linear relationship that the convenience yield has to storage and variation in risk premiums. This highlights the goal and purpose of theoretical models. They are constructed by identifying the number of factors responsible for driving the convenience yield variation that is inducing mean reversion in the spot and futures price, and either explicitly fitting the factors to the observed data or exogenously specifying them. Modeling the convenience yield is often represented by or included in the cost of carry, i.e. the cost of carry is the basis and encompasses both the cost of storage and the convenience yield.

3.2 Theoretical Studies

It has been documented that the stochastic behaviour of commodities is vastly different from conventional assets in a manner that makes modeling their price dynamics more difficult [29]. More specifically, as mentioned in our discussion on empirical studies in section 3.1, empirical testing has demonstrated that commodity futures are often in backwardation, spot and futures prices are often mean-reverting, price volatility is correlated with the degree of backwardation, term structure of commodity forward price volatility typically decreases with longer-dated contracts, and many commodities exhibit seasonality in price and volatility. These behaviours play an important role in developing models for valuing financial contingent claims on commodities and clearly must be properly understood before they can be included into any model.

Modeling commodity price dynamics can be categorized in structural equilibrium models and reduced form models. Structural equilibrium models attempt to endogenously replicate the
explicit economic relationships that exist for storable commodities by using exogenous factors that influence the price. Such exogenous factors can be consumption, production, weather, technical development or simply inventory levels. They assume a stochastic form for each of the exogenous variables, apply equilibrium and no-arbitrage arguments and obtain the desired outputs endogenously. Examples of outputs are spot prices or convenience yields. Reduced form modeling on the other hand explicitly assumes stochastic processes for the input variables and are thus specifying them exogenously. Reduced form models therefore assume spot price and convenience yield are inputs into the model as opposed to being outputs as in structural equilibrium models. It becomes clear that one approach attempts to induce mean reversion of the underlying commodity by modeling the relationship that economic drivers have towards storage and price while the other explicitly models mean-reversion in the price process itself.

Reduced form models dominate the current literature because of their flexibility and their ability to explain futures and options prices well. Structural equilibrium models may be more economically sound, however, they often require a large amount of data and parameters. Due to these reasons, reduced form models will be used in this thesis when calculating Cont’s convex measure, and we provide an overview of Gaussian-affine reduced form models in the subsequent section. Before we continue however, we illustrate a popular structural equilibrium model to serve as a comparison to the reduced form models that we present later.

In 2000, Routledge, Seppi, and Spatt develop a structural equilibrium model that attempts to capture the dynamics between the levels of inventory of a physical commodity to the commodity’s time-varying convenience yield [57]. They model the level of inventories endogenously in discrete time subject to exogenous supply shocks. Furthermore, inventory is constricted to being a nonnegative quantity. Routledge, Seppi, and Spatt claim that once commodity levels, that are supposed to be in excess of those already committed to the production process, are driven to zero, the commodity has a high immediate usage consumption value resulting in backwardation and positive convenience yields. This is the fundamental assumption of the Routledge, Seppi, and Spatt model as they attempt to incorporate the timing option into determining the value of the commodity. Their analysis also assumes that the immediate-use consumption value is driven by a mean-reverting Markov process, and they solve for the equilibrium inventory of competitive risk-neutral agents. The resultant price process is regime-shifting, including one regime with positive inventories and one with zero inventories, and the model output is the convenience yield.

### 3.2.1 Reduced Form Models

In 1976, Black provides one of the original spot commodity pricing models by introducing a one-factor model [8]. His paper made the connection between his work on equity options with dividends to pricing futures and forwards by representing the convenience yield of a storable
commodity as the dividend yield from his original work. This model assumed that the spot price followed a geometric Brownian motion (GBM) with a constant interest rate and convenience yield. Similar assumptions were made by Brennan and Schwartz in their 1985 model [10]. They were valuing natural resource investments using a real options approach. While the assumptions of the model was similar to that of Black, Brennan and Schwartz were valuing long-dated contracts. Both of these models were useful in that they were tractable and easy to implement. However, assuming constant interest rates and convenience yield implies a constant cost of carry and a constant term structure of volatility. Empirically, we know from the Samuelson hypothesis that this is incorrect, and furthermore it does not induce mean reversion in the spot price, which empirically has been proven to exist.

A number of one-factor models were introduced to address the lack of mean reversion in the spot price that was limiting the one-factor GBM model. While many of these came after the introduction of a two-factor model, in the interests of logical model progression, a subset of one-factor models will be introduced first. In 1997 Ross developed a one-factor model that induced mean reversion in the spot price explicitly by assuming the spot price followed a mean-reverting process. Also in 1997, the first model of Schwartz paper is a one-factor model, that assumes the log of the spot price follows a mean-reverting process.

There are a few one-factor models that follow the Heath, Jarrow, and Morton (HJM) no-arbitrage approach to modeling the entire futures term structure evolution [34]. Such a structure uses the initial state of the term structure as given and models the drift of the risk-neutral process for term structure evolution in a way that is consistent with no arbitrage. In 1994, Cortazar and Schwartz follow such an approach [16]. Amin, Ng, and Pirrong in 1995 [1] develop a similar model.

The primary issue with 1-factor reduced form models is that, by construction, they assume that all futures maturities are perfectly correlated. As we have discussed previously in our empirical discussions on commodity price dynamics, perfect correlation among futures maturities does not exists, and is in fact also another violation of the Samuelson’s hypothesis. Adding more factors in modeling the spot price process and the cost of carry often yields a better fit to the observed data. As such, the natural progression of model development was in fact to add more factors to better describe the term structure of futures data.

It was the seminal paper of Gibson and Schwartz in 1990 that first introduced a two-factor model that explicitly induced mean reversion in the spot price process [27]. The two factors are the spot price and the convenience yield. The structure of their model assumes that the spot price of oil has a lognormal stationary distribution following Geometric Brownian Motion (GBM) and the convenience yield follows a stochastic mean reversion Ornstein-Uhlenbeck (OU) process. Furthermore, the spot price and convenience yield follow a joint diffusion process that assumes a constant correlation between the convenience yield and spot price. At the time of
its introduction, this model did not have a closed form solution for a futures price but rather required numerical treatment. This model was the first to model an inverse volatility term structure. The second model of Schwartz's 1997 paper develops a closed form solution of the Gibson and Schwartz two-factor model and explained how the Kalman filter can be used to estimate its parameters [59].

In 2000, Smith and Schwartz present another two-factor model [60]. The two factors were the long-term mean of the underlying spot process and the short term deviations, thus the addition of the two variables yielded the spot process. The long-term mean followed GBM whereas the short-term deviations were mean reverting to a long term mean of zero. Smith and Schwartz were able to demonstrate that even though the definition of their factors differed from those of the Gibson and Schwartz model, the model can be considered equivalent. However, they allow risk premium to be defined as a function of the short term deviation from the long-term mean, which may be a more natural way to associate risk premium to spot price fluctuations.

In 2004, Nielson and Schwartz develop a two-factor model that extends that of Gibson and Schwartz in a way that the volatility of both factors is a function of the convenience yield. They were able to demonstrate that this had a positive influence when modeling options [52]. In 2004, Ribeiro and Hodges present a similar model to that of Nielson and Schwartz. However, they model the convenience yield as a Cox-Ingersoll-Ross process to ensure that it is arbitrage free. Second, they introduce a time-varying volatility factor that is proportional to the square root of the convenience yield [54].

The third model of Schwartz's 1997 paper introduces a three-factor model [59]. The first two factors are identical to those of the two-factor model from the same paper in specification and functional form (spot price and convenience yield), while the third factor is interest rates. The model assumes that interest rates followed a Vasiček model/process [64]. The model suggests estimating the parameters for the interest rate process separately from the other two factors before integrating it into the calculation of the spot process.

In 1998, Milterson and Schwartz developed a three-factor model that is an extension of the third Schwartz model. It is developed under the HJM framework. They differentiate between forwards and futures in their analysis due to the stochasticity of interest rates (refer to appendix A.2 for a discussion on the differences between a futures contract and a forward). Milterson and Schwartz derived a closed-form solution for the price of options on commodity futures, which they state as a generalization of the Black-Scholes/Merton framework.

In 1998, Hilliard and Reis extended the work of Schwartz by replacing the Vasiček process to the Hull and White [38] interest rate model [37]. The also introduce jumps in the spot process. They notice that the inclusion of stochastic interest rates has an affect on futures prices while does not affect forward prices. Furthermore, they noticed that including jumps in the spot process did not affect the futures price whereas differences existed when calculating option
prices.

In 2002, Sørensen introduced a model that includes seasonality into the model description [61]. Seasonality is defined as seasonal effects that affect the convenience yield of a commodity in a deterministic way, i.e., in the case of agricultural commodities, convenience yield is high before the harvest because of depleted stocks. Sørensen’s model uses deterministic weighted function of sine and cosine functions as the primary factor, and adds two additional factors. Richter and Sørensen construct a similar model in 2002, but add the seasonality to a three-factor model [55]. The three factors are spot price, convenience yield and spot price volatility, and they add a separate deterministic function to the spot volatility and convenience yield volatility. Richter and Sørensen choose to keep the seasonality component in the volatility of the spot price and that of the convenience yield the same to simplify their numerical recipe. While seasonality has been exhibited in agricultural commodities, such behaviour is not prevalent in crude oil futures (the data that we will be using in this thesis).

In 2003, Cortazar and Schwartz develop a three factor model that is an extension of the Gibson and Schwartz formulation [17]. They restate the Gibson and Schwartz model in a more parsimonious way by reformulating the model using the log spot price and the *de-meaned* instantaneous convenience yield as their two factors. In their parsimonious two-factor description, they assume that the long-term price return is constant. The three-factor model makes the long-term price return stochastic, and is the third factor. They argue that their representation is parsimonious and easy to understand, and should become a standard for practitioners (They demonstrate how a spreadsheet model can be developed for pricing contingent claims to appease practitioners).

In 2005, Cassius and Collin-Dufresne develop a three-factor model that they claim is maximal, as it encompasses many of the popular two and three factor models that we have discussed into a general three-factor setting. In their model, they allow the convenience to depend on both the interest rates and spot prices. Furthermore, they model risk premia as linear functions of the state variables. In 2006, Cortazar and Noranjo develop a n-factor Gaussian model that generalizes most Gaussian-affine models [18]. They demonstrate how their representation can be converted into many of the popular two and three-factor models (many of which are mentioned in our discussion).

In 2009, Trolle and Schwartz tackle the problem of unspanned volatility, i.e. the problem that volatility risk can not be hedged by the underlying commodity [63]. Such a problem implies that volatility dynamics are unspanned by the underlying and justifies the existence of an options market. They claim that futures prices are driven by three factors, and propose a HJM type three factor model to describe it. The factors that they implement are spot price and two others relating to the forward cost of carry curve. Additionally, they claim that option prices are driven by two additional volatility factors.
Most of the models we have presented are Gaussian-affine. As demonstrated, the general structure of these models has been to add additional factors to represent variations of the spot price and the cost of carry in order to improve the goodness of fit to the observed futures and option of futures prices. We have also briefly discussed models that have included stochastic volatility models (that are Heston-like [35]), included stochastic equilibrium level models, added discontinuous jumps into the process, and added deterministic seasonality effects into the model description. These models were mentioned but were not the focus of the literature search. We have presented a subset of the Gaussian-affine family of reduced form models because they are representative of the models that we have chosen to focus on for this thesis due to their general tractability and relative ease of parameter estimation using the Kalman filter.

The Gaussian-affine models that we have chosen to calculate Cont’s convex risk measure are the one-factor Schwartz 1997 model (S1997-1) [59], the two-factor Schwartz 1997 model (S1997-2) [59], the two-factor Schwartz and Smith model (SS2000) [60], and the three-factor Cortazar and Schwartz model (CS2003) [17]. As explained, Gaussian-affine models naturally allow for parameter estimation using the Kalman filter. Furthermore, these 4 specific models were chosen because they are examples of three generations of commodity model development, the one-, two-, and three-factor generations. They encompass a wide array of models and are representative of what many practitioners have been using when evaluating commodity contingent claims. While not all encompassing of some of the more modern valuation models ([63], [12]), they will suffice in the context of this thesis.

In the next few sections, we provide a thorough description of the models chosen to evaluate Cont’s convex risk measure. Each model description will explain the intuition behind the stochastic representation of each factor followed by outlining the real world dynamics ($P$), the risk-neutral world dynamics ($Q$), the expression for calculating the futures price and the state space model that is ultimately used for parameter estimation.

The estimation techniques used to evaluate the parameters of each model are often as difficult as developing the description of the model dynamics itself. Each of the four models that we have chosen is Gaussian-affine, and have a closed form solution of the futures price that exemplifies linear relationship between the logarithm of futures prices and the models underlying factors. As such, implementing the Kalman filter for parameter estimation was a natural choice given the structure of these models. We turn the reader to the appendices B where we build the intuition behind using the Kalman filter along with a mathematical description of the filter’s algorithm.

---

*[4]There also exists regime switching models. Such models assume a commodity behaves differently when in different regimes, and applies different pricing rules when it is calculated to be in one. A simple example is to change valuation methodologies based on whether or not the convenience yield is high. If it is high, the spot price is expensive that may have been caused by an exogenous shock, and mean reversion is assumed to be imminent. If spot is low, a simple cost of carry model is used. This is the framework utilized Bühler, Korn, and Schöbel in 2004 [11].*
With regards to notation, the real world dynamics ($\mathbb{P}$), the risk-neutral world dynamics ($\mathbb{Q}$), and the expression for calculating the futures price will be described using the notation of the article that the model was presented in. The notation of our state space model follows the notation contained in the appendix on the Kalman filter $B$.

### 3.2.2 Model 1: S1997-1

The S1997-1 model was developed by Schwartz in his 1997 seminal paper [59]. It is a one-factor model that assumes the log of the commodities spot price follows an Ornstein Uhlenbeck process (mean-reversion). The real-world dynamics are as follows:

$$dS = \kappa(\mu - \ln S)S dt + \sigma S dz$$  \hspace{1cm} (3.1)

Applying Ito’s Lemma and setting $X = \ln S$, we get:

$$dX = \kappa(\alpha - X) dt + \sigma dz$$  \hspace{1cm} (3.2)

In the above equations 3.2, $\kappa$ is the rate of mean reversion, $\alpha$ is the long run mean log spot price, $\sigma$ is the volatility of the process and $dz$ is an increment to a standard Brownian motion under the real measure. Re-writing 3.2 under the risk-neutral measure yields the following:

$$dX = \kappa(\alpha^* - X) dt + \sigma dz^*$$  \hspace{1cm} (3.3)

$\alpha^* = \alpha - \lambda$ where $\lambda$ is the market price of risk and is assumed to be constant. $dz^*$ is an increment to a standard Brownian motion under the risk neutral measure.

To calculate the futures price at time $T$, we note that the conditional distribution of $X$ at time $T$ under the risk-neutral measure is normally distributed, i.e., $F(S,T) = \mathbb{E}[S(T)|\mathcal{F}_T]$. By evaluating this equation, we can solve for the futures price using the parameters of the preceding processes:

$$\ln F(S,T) = e^{-\kappa T}X + (1 - e^{-\kappa T})\alpha^* + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa T})$$  \hspace{1cm} (3.4)

Representing the S1997-1 model in state space form is trivial because it is a one-factor model. We make a small alteration to the notation of the observed process $z_k$ by introducing a constant matrix $D$. The addition of this matrix makes a slight change to the a posteriori estimate of the corrector equations 14 that we introduced in the appendix\(^5\). We also note that the values in $A$ and $B$ differ from the derivation in Schwartz [59]. Schwartz makes the assumption of a

\(^5\)The equation becomes $\hat{x}_k = \hat{x}_k^- + K_k(z_k - [H\hat{x}_k^- + D])$
small time step in the Kalman Kilter ($\Delta t$) whereas we relax that assumption and present the analytical solution of the mean-reverting process. Lastly, $T_i$, where $i \in \{1 \cdots N\}$, is the time to expiration of the particular futures contract.

$$x_t = \begin{bmatrix} X_t \end{bmatrix} \quad A = \begin{bmatrix} e^{-\kappa \Delta t} \end{bmatrix} \quad B = \begin{bmatrix} (1 - e^{-\kappa \Delta t})\alpha^* \end{bmatrix} \quad Q = \begin{bmatrix} \sigma^2 \end{bmatrix}$$

$$z_t = \begin{bmatrix} \ln F_t(S, T_1) \\ \vdots \\ \ln F_t(S, T_N) \end{bmatrix} \quad H = \begin{bmatrix} e^{-\kappa T_1} \\ \vdots \\ e^{-\kappa T_N} \end{bmatrix} \quad D = \begin{bmatrix} (1 - e^{-\kappa T_1})\alpha^* + \frac{\sigma^2}{4\kappa} (1 - e^{-2\kappa T_1}) \\ \vdots \\ (1 - e^{-\kappa T_N})\alpha^* + \frac{\sigma^2}{4\kappa} (1 - e^{-2\kappa T_N}) \end{bmatrix}$$

### 3.2.3 Model 2: S1997-2

The S1997-2 model was also developed by Schwartz and introduced in his 1997 paper [59]. It is a two-factor model that assumes both the instantaneous convenience yield and the log of the commodities spot price evolve in accordance to the following joint stochastic process (real-world joint-dynamics):

$$dS = (\mu - \delta)Sdt + \sigma_1 Sdz_1$$
$$d\delta = \kappa(\alpha - \delta)dt + \sigma_2 dz_2$$

(3.5)

where the increments to the Brownian motion are correlated with $dz_1dz_2 = \rho dt$. The parameter $\mu$ represents the long-term total return on the commodity, which encompasses both the price appreciation of the asset along with the long-term convenience yield, the parameter $\alpha$ represents the long-term convenience yield, and the state variables $S$ and $\delta$ represent the spot price and the instantaneous convenience yield respectively.

We again define $X = \ln S$ and apply Ito’s Lemma to get the following:

$$dX = (\mu - \delta - \frac{1}{2}\sigma_1^2)dt + \sigma_1 dz_1$$

(3.6)

When represented in this form, the log spot price follows a GBM process while the convenience yield follows a mean-reverting (OU) process.

By construction, the model assumes that owning the spot commodity entitles the holder to a convenience yield $\delta$ (treated like a dividend). This implies the numeraire of the risk neutral process must be reduced by this amount. Furthermore, the convenience yield cannot be spanned by tradable assets and it is therefore not possible to hedge. As such, the convenience yield
Chapter 3. Commodity Modeling

process requires a market price of risk \( \lambda \). We now present the risk-neutral dynamics:

\[
\begin{align*}
    dS &= (r - \delta)Sdt + \sigma_1 Sdz_1^* \\
    d\delta &= \kappa [(\alpha - \delta) - \lambda]dt + \sigma_2 dz_2^* \\
    dz_1^* dz_2^* &= \rho dt
\end{align*}
\]

(3.7)

Schwartz [59] derive the partial differential equation that the aforementioned risk-neutral dynamics 3.7 must satisfy and present the following solution for the log futures price:

\[
\begin{align*}
    \ln F(S, \delta, T) &= X - \delta \left(1 - e^{-\kappa T}\right) + A(T) \\
    A(T) &= (r - \hat{\alpha} + \frac{1}{2}\frac{\sigma_2^2}{\kappa^2} - \frac{\sigma_1\sigma_2\rho}{\kappa})T + \frac{1}{2}\frac{\sigma_2^2}{\kappa^3} + (\hat{\alpha}\kappa + \sigma_1\sigma_2\rho - \frac{\sigma_2^2}{\kappa}) \left(1 - e^{-2\kappa T}\right) \kappa^2
\end{align*}
\]

(3.8)

As mentioned, the S1997-2 model is a two factor model that assumes the spot price and convenience yield follow mean-reverting processes. Neither of these processes are observable, and are the state variables of S1997-2’s state-space form. We present the derivation of parameter estimation that is to be used in the Kalman filter as presented by Schwartz [59].

\[
\begin{align*}
    x_t &= \begin{bmatrix} X_t \\ \delta_t \end{bmatrix} \\
    A &= \begin{bmatrix} 1 & -\Delta t \\ 0 & 1 - \kappa\Delta t \end{bmatrix} \\
    B &= \begin{bmatrix} (\mu - \frac{1}{2}\sigma_1^2)\Delta t \\ \kappa\alpha\Delta t \end{bmatrix} \\
    Q &= \begin{bmatrix} \sigma_1^2\Delta t & \rho\sigma_1\sigma_2\Delta t \\ \rho\sigma_1\sigma_2\Delta t & \sigma_2^2\Delta t \end{bmatrix} \\
    z_t &= \begin{bmatrix} \ln F_t(S, \delta, T_1) \\ \vdots \\ \ln F_t(S, \delta, T_N) \end{bmatrix} \\
    H &= \begin{bmatrix} 1 & -1 - e^{-\kappa T_1}\kappa \\ \vdots & \vdots & \vdots \\ 1 & -1 - e^{-\kappa T_N}\kappa \end{bmatrix} \\
    D &= \begin{bmatrix} A(T_1) \\ \vdots \\ A(T_N) \end{bmatrix}
\end{align*}
\]

3.2.4 Model 3: SS2000

The SS2000 model was developed by Schwartz and Smith, in their 2000 paper [60]. They develop a two-factor model that assumes the price of a commodity is a function of the long term equilibrium, \( \xi \), and short-term deviations, \( \chi \). Economically, they explain that short term deviations are due to disruptions in supply and demand and are not expected to persist. The long-term equilibrium level is the true price of the commodity and represents the resulting expected price once long term expectations of supply and demand are considered. Schwartz and Smith argue that the differences between short-term and long-term futures contracts give information regarding estimating the value of these two factors. The difference between short-term and long-term contracts provide insights towards the possible short-term deviations, and long-term futures contracts are informative toward the long-term equilibrium level because of
the assumption that long-term prices elicit the expectation of whether short-term disruptions will dissipate or change price levels permanently.

Schwartz and Smith justify not modeling the spot and convenience yield process explicitly because the difference in price between long-term and short-term contracts is a way to evaluate how basis is driving commodity prices (also note that convenience yield is an important aspect of basis). They continue this justification by stating the equivalence of the SS2000 model to the Gibson and Schwartz two-factor model \[27\]^6 and derive a set of equations to represent one model in terms of the other. The main difference is that by construction, the SS2000 model does not estimate the risk-free rate and it is implicit in the solution to the futures price (described below). This difference can be dismissed by setting the risk free rate as the yield on the US T-Bill, thus making the two models equivalent (stochastically). The main reason for using the long-term equilibrium price and short-term deviations as the two factors driving commodity prices is that they believe it to be more intuitive.

The short term deviations, $\chi$, are stochastic and are expected to revert to zero and follow an Ornstein Uhlenbeck process. The long term equilibrium level, $\xi$, is also stochastic and is assumed to follow geometric Brownian motion. The joint-diffusion of the real-world process is as follows (we use the notation of the original paper \[60\]):

\[
\begin{align*}
d\chi &= -\kappa \chi dt + \sigma_{\chi} dz_{\chi} \\
d\xi &= \mu_{\xi} dt + \sigma_{\xi} dz_{\xi}
\end{align*}
\]

(3.9)

where the increments to the Brownian motion are correlated with $dz_{\chi} dz_{\xi} = \rho_{\chi \xi} dt$. Once the state variables are determined for any time $t$, we can calculate the spot price at time $t$ with the following relationship: $\ln S_t = \xi_t + \chi_t$.

Neither of the factors in this model are hedgeable and a risk adjustment must be made. We introduce additional parameters $\lambda_{\xi}$ and $\lambda_{\chi}(\chi_t)$ that specify reductions in drift for each process. Incorporating these risk adjustments, we obtain the following risk neutral representation of the SS2000 model dynamics:

\[
\begin{align*}
d\chi &= (-\kappa \chi - \lambda_{\chi}(\chi)) dt + \sigma_{\chi} dz^*_{\chi} \\
d\xi &= (\mu_{\xi} - \lambda_{\xi}) dt + \sigma_{\xi} dz^*_{\xi} \\
dz^*_{\chi} dz^*_{\xi} &= \rho_{\chi \xi} dt
\end{align*}
\]

(3.10)

Schwartz and Smith \[60\] offer two representations for the values of $\lambda_{\chi}(\chi_t)$. The first assumes a constant reduction in drift ($\lambda_{\chi}(\chi_t) = \lambda_{\chi}$), while the second assumes a time varying risk adjustment based on the current short term deviation ($\lambda_{\chi}(\chi_t) = \beta \chi_t + \alpha$). We implement the later because it offers a different perspective to the estimation process and differentiates it from the SS1997-2 model. Adding the aforementioned time varying risk premium changes the

---

^6The two factors are spot price and convenience yield, and is very similar to S1997-2
risk neutral dynamics of the short term variation factor to \( d\chi = (-\kappa^* \chi - \alpha)dt + \sigma_\chi dz_\chi^* \) where \( \kappa^* = \kappa + \beta \). We can now write the futures price as

\[
ln F(\xi, \chi, T) = e^{-\kappa^*T} \chi + \xi + A(T)
\]

\[
A(T) = \mu^* \xi T - \frac{1}{2} \left(1 - e^{-2\kappa^*T}\right) \frac{\sigma^2_\chi}{\kappa^*} + 2 \left(1 - e^{-\kappa^*T}\right) \frac{\rho_\chi \xi \xi \sigma_\chi \sigma_\xi}{\kappa^*} \right) \tag{3.11}
\]

\[
\mu^* = \mu \xi - \lambda \xi
\]

We conclude the discussion on the SS2000 model by presenting the model’s state space representation. The parameters used in the state space estimation that we present in what follows are those derived in the paper [60]. As we’ve already discussed, the two factors of the SS2000 model are unobservable and will serve as the components of our state vector (again, these are the short-term variations \( \chi \) and the long term equilibrium \( \xi \)).

\[
x_t = \begin{bmatrix} \chi_t \\ \xi_t \end{bmatrix} \quad A = \begin{bmatrix} e^{-\kappa \Delta t} & 0 \\ 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ \mu \Delta t \end{bmatrix} \quad Q = \begin{bmatrix} (1 - e^{-2\kappa \Delta t}) \frac{\sigma^2_\chi}{2\kappa} & (1 - e^{-\kappa \Delta t}) \frac{\rho_\chi \xi \sigma_\chi \sigma_\xi}{\kappa} \\ (1 - e^{-2\kappa \Delta t}) \frac{\sigma^2_\xi}{\kappa} & (1 - e^{-\kappa \Delta t}) \frac{\rho_\chi \xi \sigma_\chi \sigma_\xi}{\kappa} \end{bmatrix}
\]

\[
z_t = \begin{bmatrix} ln F_t(\chi_t, \xi_t, T_1) \\ \vdots \\ ln F_t(\chi_t, \xi_t, T_N) \end{bmatrix} \quad H = \begin{bmatrix} e^{-\kappa T_1} & 1 \\ \vdots & \vdots \\ e^{-\kappa T_N} & 1 \end{bmatrix} \quad D = \begin{bmatrix} A(T_1) \\ \vdots \\ A(T_N) \end{bmatrix}
\]

### 3.2.5 Model 4: CS2003

The CS2003 is a three-factor stochastic model, and was developed by Cortazar and Schwartz in their 2003 paper [17]. The CS2003 model extends the S1997-2 model that we’ve already introduced. Cortazar and Schwartz begin the description of the CS2003 model with a reformulation of the parameters that govern the dynamics of the S1997-2 model. This reformulation naturally presents the third factor in an intuitive way. We will summarize the logic as presented in the paper [17].

We first recall from equation 3.5 the following parameter and state variable descriptions: the parameter \( \mu \) represents the long-term total return on the commodity, which includes the price appreciation of the asset along with the long-term convenience yield, the parameter \( \alpha \) represents the long-term convenience yield, and the state variable \( \delta \) represents the instantaneous convenience yield. We define a new state variable \( y \) as the short-term variation in the instantaneous convenience yield relative to its long-term mean. More formally, \( y = \delta - \alpha \), and replaces the instantaneous convenience yield state variable in the model formulation (referred to as the de-meaned instantaneous convenience yield in the paper). Furthermore, we introduce the parameter \( \nu \) as the long-term price return on the commodity attributed strictly to price
appreciation. As we’ve explained, the parameter \( \mu \) encompasses both the price appreciation and the long-term convenience yield of the asset, therefore, \( \nu \) can be expressed as \( \nu = \mu - \alpha \). Using these newly defined variables, we get the following dynamics for the S1997-2 model:

\[
dS = (\nu - y)Sdt + \sigma_1 Sdz_1
\]
\[
dy = -\kappa y dt + \sigma_2 dz_2
\]
\[
d\nu = a(\bar{\nu} - \nu) dt + \sigma_3 dz_3
\]

(3.12)

This representation is more succinct than it was in its original form, however, Cortazar and Schwartz argue that this representation encompasses the same explanatory power of commodity dynamics and is therefore more parsimonious than the original. They also argue that this representation is a more natural way to look at the fundamental drivers of spot commodity dynamics. These fundamental drivers, as implied by Equation 3.12, imply that the spot commodity dynamics are influenced by long-term changes to macro-economic factors such as technology, inflation or demand (factors that are attributed to the long-term price appreciation of the commodity) and short-term inventory level fluctuations due to supply shocks (factors that are attributed to the short-term perturbations in the convenience yield). It is intuitive to think that the long-term price appreciation of the commodity is also stochastic, and this is in fact the third factor in the CS2003 model.

The spot price, \( S \), the de-meaned instantaneous convenience yield, \( y \), and the long-term price return, \( \nu \), are all stochastic and the joint-diffusion of the real-world process is as follows (we use the notation of the original paper [17]):

\[
dS = (\nu - y)Sdt + \sigma_1 Sdz_1
\]
\[
dy = -\kappa y dt + \sigma_2 dz_2
\]
\[
d\nu = a(\bar{\nu} - \nu) dt + \sigma_3 dz_3
\]

(3.13)

where the increments to the Brownian motion are correlated with \( dz_1 dz_2 = \rho_{12} dt, dz_1 dz_3 = \rho_{13} dt, dz_2 dz_3 = \rho_{23} dt \). All parameters in equation 3.13 have been introduced except for \( \bar{\nu} \) and \( a \). Since \( \nu \) is the long-term price appreciation of the asset, \( \bar{\nu} \) is the long-term average that it converges to with a mean-reversion shaping parameter \( a \). None of the three factors in this model are hedgeable and a risk adjustment must be made. We introduce additional parameters \( \lambda_i, i \in \{1 \cdots 3\} \) that specify reductions in drift for each process. Incorporating these risk adjustments, we obtain the following risk neutral representation of the CS2003 model dynamics:

\[
dS = (\nu - y - \lambda_1)Sdt + \sigma_1 Sdz_1^*
\]
\[
dy = (-\kappa y - \lambda_2) dt + \sigma_2 dz_2^*
\]
\[
d\nu = a((\bar{\nu} - \nu)) - \lambda_3) dt + \sigma_3 dz_3^*
\]
\[
dz_1^* dz_2^* = \rho_{12} dt, dz_1^* dz_3^* = \rho_{13} dt, dz_2^* dz_3^* = \rho_{23} dt
\]

(3.14)
Cortazar and Schwartz [17] derive the partial differential equation that the aforementioned risk-neutral dynamics 3.14 must satisfy and present the following solution for the log futures price (we again denote $X = \ln S$):

$$\ln F(S, y, \nu, T) = X - y \frac{1 - e^{-\kappa T}}{\kappa} + \nu \frac{1 - e^{-aT}}{\kappa} + A(T)$$

$$A(T) = -\lambda_1 T + \frac{\lambda_2 - \sigma_1 \sigma_2 \rho_{12}}{\kappa^2} (\kappa T + e^{-\kappa T} - 1) + \frac{\sigma_2^2}{4\kappa^3} (e^{-2\kappa T} + 4e^{-\kappa T} + 2\kappa T - 3)$$

$$+ \frac{a\bar{\nu} - \lambda_3 + \sigma_1 \sigma_3 \rho_{13}}{a^2} (aT + e^{-aT} - 1) - \frac{\sigma_3^2}{4a^3} (e^{-2aT} + 4e^{-aT} + 2aT - 3) - \frac{\sigma_2 \sigma_3 \rho_{23}}{\kappa^2 a^2 (a + \kappa)}$$

$$+ (\kappa^2 e^{-at} + \kappa ae^{-aT} + \kappa a^2 T + \kappa ae^{-\kappa T} + a^2 e^{-\kappa T} - \kappa ae^{(\kappa + a)T} - \kappa^2 - \kappa a - a^2 + \kappa^2 a T)$$

(3.15)

We conclude the discussion on the CS2003 model by presenting the model’s state space representation. The parameters used in the state space estimation that we present in what follows are those inferred from the derivations in the paper [17]. All three factors of the CS2003 model are unobservable and will serve as the components of our state vector (again, these are the log spot price, $X$, the de-meaned instantaneous convenience yield, $y$, and the long-term price return, $\nu$).

$$x_t = \begin{bmatrix} X_t \\ y_t \\ \nu_t \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & -\Delta t & \Delta t \\ 0 & -\kappa \Delta t & 0 \\ 0 & 0 & -a \Delta t \end{bmatrix}$$

$$B = \begin{bmatrix} -\frac{1}{2} \sigma_1^2 \Delta t \\ 0 \\ a\bar{\nu} \Delta t \end{bmatrix}$$

$$Q = \begin{bmatrix} \sigma_1^2 & (1-e^{-\kappa \Delta t}) \sigma_1^2 \sigma_2 \rho_{12} & (1-e^{-a \Delta t}) \sigma_2 \rho_{13} \\ (1-e^{-\kappa \Delta t}) \kappa \sigma_1^2 \sigma_2 \rho_{12} & \sigma_2^2 & (1-e^{-a \Delta t}) \sigma_3 \rho_{23} \\ (1-e^{-a \Delta t}) \sigma_1^2 \sigma_3 \rho_{13} & (1-e^{-a \Delta t}) \sigma_2 \rho_{13} & \sigma_3^2 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & -e^{-\kappa T_1} & -e^{-a T_1} \\ \vdots & \vdots & \vdots \\ 1 & -e^{-\kappa T_N} & -e^{-a T_N} \end{bmatrix}$$

$$D = \begin{bmatrix} A(T_1) \\ \vdots \\ A(T_N) \end{bmatrix}$$
Chapter 4

Results & Discussion

This chapter presents the results that we obtained from our analysis of model uncertainty, and comments on some of the observations that we made. Furthermore, we provide a thorough overview of the methodology applied in calculating the convex risk measure of Li-Kwon. The goal of this overview is to serve as a manual for implementing the strategy using real-data. We discuss some of the challenges that were faced in such an implementation. We do not categorically separate results from discussion but rather integrate our discussion points as we present the various aspects of the results.

We will categorize this chapter into three components: data collection, calibration, and calculation. In the data collection subsection, we detail the data that was used in this study and any assumptions that were made in their retrieval. In calibration, we outline the optimal parameters that were obtained using the Kalman filter algorithm along with the resulting option prices obtained. We also highlight some features of our implementation of the Kalman filter to obtain the set of optimal parameters for each model. Lastly, in calculation, we show the results of both Cont’s convex measure and that of Li and Kwon. Before presenting the results obtained from Li and Kwon’s measure, we provide a detailed description of the steps involved in its calculation.

4.1 Data Collection

All derivative data collected was related to the WTI Crude Oil contract traded on the Chicago Mercantile Exchange (CME).\footnote{It actually trades on the New York Mercantile Exchange (NYMEX), but the NYMEX is owned by the CME.} All data collection used December 3, 2012 as the valuation date. All options and futures were for delivery in March 2013.

We collected data for three purposes. First, we collected WTI Crude Oil futures term structure...
data for the purpose of calibrating the four commodity spot models. Second, we collected
benchmark option prices for calculating the two model uncertainty measures. Third, empirical
moments were calculated for the Li-Kwon extension.

The Kalman filter algorithm that we implemented to estimate the parameters of the four models
requires a time series of five separate contract maturities, specifically the 1, 5, 9, 13, and 17
nearest contracts. Since we are calculating the expectation of the spot price in March 2013
starting in December 2012, the spot price month was December 2012, the first nearest contract
expiration was January 2013, etc. These months provide information on the unobserved spot
process, and upon calibration, were compared to the December futures contract (used as proxy
for spot market).

The term structure data was collected for four different look back windows, 0.25, 0.5, 1.5 and
3.0 years respectively. The reason for this is because we wanted to observe optimal parameter
estimation values when different amounts of historical data was present. We later calculate
Cont’s model uncertainty under the different look back windows.

As mentioned, the option data that was obtained was for March delivery. Table 4.1 presents the
options that were used as benchmark derivatives when calculating the risk measures.

Table 4.1: Benchmark call options written on WTI Crude Oil futures for March 2013 delivery
(CLH3C), extracted from Bloomberg and are of American type.

<table>
<thead>
<tr>
<th>Date</th>
<th>CLH3C 100.00</th>
<th>CLH3C 105.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012-12-03</td>
<td>1.18</td>
<td>0.56</td>
</tr>
</tbody>
</table>

European options on crude oil futures are illiquid. We have therefore chosen American type
options\(^2\) as benchmark options when calculating model uncertainty even though they are path-
dependent options and may differ significantly from their European counterparts. We note
however that an American options value will converge to the value of the European option when
it is deep out of the money. This is because the time value of early exercise of an American option
written on a futures contract is proportional to the likelihood of a large beneficial price spike
before expiration that is expected to dissipate before the contract matures\(^3\). This likelihood
is strongly diminished the further out of the money the American option is. As such, we
determined that value of an American option that is a minimum of 10% out of the money is
sufficiently close to that of its European counterpart and was used for this study.

\(^2\)Recall that the only difference between an American and European option with strike \(k\) and expiration \(T\)
written on the same underlying is that an American option gives you the right to an early exercise whereas a
European option must be held until expiration.

\(^3\)We note that an American option written on an equity may be a candidate for early exercise in the case when
the underlying stock goes ex-dividend. However, we discuss in the appendices A that the dividend of a future is
the convenience yield which only accrues to the holder of the spot commodity. As such, the time-value of early
exercise of an American option written on a commodity futures contract is not affected by the underlying going
ex-dividend because such an equivalent does not exists in futures.
The empirical moments were calculated using four different look back windows, 0.25, 0.5, 1.5 and 3.0 years respectively. The moments were calculated using a time series of the daily closing price data of the March crude oil futures contract.

Table 4.2: Calculated empirical moments estimated over different look back windows.

<table>
<thead>
<tr>
<th></th>
<th>0.25 year</th>
<th>0.50 year</th>
<th>1.50 year</th>
<th>3.00 year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>92.88</td>
<td>91.89</td>
<td>95.93</td>
<td>93.24</td>
</tr>
<tr>
<td>2</td>
<td>8644.49</td>
<td>8467.86</td>
<td>9248.06</td>
<td>8748.07</td>
</tr>
<tr>
<td>3</td>
<td>806025.83</td>
<td>782591.09</td>
<td>895961.71</td>
<td>826038.63</td>
</tr>
</tbody>
</table>

Recall from the description of the Li-Kwon methodology (2.5) that the moments used when calculating the model uncertainty are those of the risk-neutral measure. As such, implicit to our empirical calculations of the moments is that we are assuming that the empirical moments is the most-likely estimate for the risk-neutral moments. Such an assumption is not as unlikely as it sounds. We note that, according to standard no arbitrage theory, there exists a risk-neutral pricing measure that incorporates all risk premiums of all investors and allows a practitioner to value a derivative as the expectation under this measure. Furthermore, in a complete market, the practitioner is guaranteed that the expectation under this measure is actual value of the asset. Accordingly, if we assume some form of market pseudo-completeness\(^4\), both futures and options are derivatives that are priced as the expectation of the same risk-neutral measure. As such, since the underlying of the options contracts that we are valuing are those written on futures, the empirical moments that we calculate are actually a good approximation of risk-neutral moments.

4.2 Calibration

In what follows, we describe the methodology that we used for estimating the parameters of the chosen models. We first discuss the estimation procedures that we implemented to identify the optimal set of parameters for each model. We present the parameters, show how well they fit to the data, and present the calculated option prices. We conclude with a quick discussion of the observed differences among the calculated option prices and parameters over different look back periods.

The Kalman filter was used to calibrate our model parameters to observed data in order to obtain the unobserved spot price process (the components of the unobserved state variables yield the spot price)\(^5\). Before we continue the remainder of our calibration, we interlude to

\(^4\)We do not formally assume complete markets because the framework that we are working with is to identify model uncertainty on markets that are illiquid thus by construction implies an incomplete market.

\(^5\)Kalman filter algorithm basics described in appendix B
discuss how we managed the sensitivity of the Kalman filter’s gain matrix to the measurement error related to the estimation of the state variables.

The measurement equations that were derived for each commodity price model are those that relate the state variables to the observable [liquid] futures prices quoted on futures exchanges. It is easy to make the assumption that the exchange has perfect information relating to the activity of their marketplace and therefore are able to set a perfect settlement price at market close for each maturing futures contract. This assumption would imply that the covariance of such measurements is zero (i.e. setting $R = 0$ in the gain matrix)\(^6\). However, as utilized in Smith and Schwartz [60], we assume that the measurement covariance is non-zero due to the combination of possible asynchronous reporting of pricing and possible error in the model’s fit to futures prices. Furthermore, we assume that each measurement error in one futures price is independent of all other futures prices of different maturity, which results in a diagonal covariance matrix\(^7\). These assumptions not only allow us to relax the assumption of perfect information of exchanges but it also allows for better flexibility in the Kalman filter’s ability to evaluate better state variable estimates by incorporating more of the previously computed state variables in the new result. This makes the system less reliant on each futures price and can allow for other dependencies that relate to the a priori estimate.

When calibrating, we maintained consistency among the models by using the same futures term structure data during parameter estimation. In order to identify the set of optimal parameters conditional to the observed data, we implemented the Kalman filter in a way that each iteration calculates the likelihood of the unobserved variables given the observed futures term structure. Furthermore, we used the first nearest futures contract to proxy as the spot process\(^8\). As such, in order to identify the optimal set of parameters, we attempted to maximize the likelihood of the vector of unobserved variables penalized by the norm of how well the result fit to the first nearest contract. Our objective function was as follows:

$$\max \ E[\bar{x}_L] - \|\bar{x} - S_0\|_2$$

Where $\bar{x}_L$ is a vector representing the log-likelihood of each individual unobserved state vector indexed by time, $\bar{x}$ is a vector representing the estimated spot process from the Kalman filter algorithm, and $S_0$ is a time series of the first nearest contract represented as a vector. We penalize the mean of the log-likelihood vector by the 2-norm of the variation in the [proxy for] observed spot price relative to the estimated spot price in order to minimize the likelihood of

\(^6\)This would also elude the possibility of imperfect markets yielding imperfect prices.

\(^7\)This assumption considerably simplifies the parameter estimation of the $R$ matrix because only $n$ variables have to be estimated as opposed to $n \times n$

\(^8\)It is important to note that the December futures contract is not the actual spot price but provides a good approximation. Spot prices are rarely observed for physicals due to the difficulty associated with striking a price for immediate delivery. This difficulty implies that an active spot market is nearly impossible and price discovery is often transferred to futures.
any large individual data deviations from the observed spot price (again, this is approximated by the December first nearest contract).

Recall that Cont’s convex risk measure expects that at least one model is calibrated to observed derivatives prices, which in our case were European options. In order to incorporate this into one of the models, we added another 2-norm penalty to our objective function as follows:

$$\max \ E[\bar{x}_L] - \|\bar{x}_L - \bar{S}_0\|_2 - \lambda \|\bar{C} - C_0\|_2$$

where $\bar{C}$ and $C_0$ are vectors of the estimated and observed call option vectors respectively, and $\lambda$ is a shaping parameter used to adjust the influence of the call option price estimation error in the objective function. This was done because the number of available futures observations outweighed the total number of call option observations, and in order to ensure that they were equally weighted in the objective function, the shaping parameter was set to equal the ratio of number of futures observations relative to the number of option observations.

When estimating the parameters for each of the chosen models, we note that due to the high-degree of non-linearity associated with the calculation of the optimal set of parameters, implementing an algorithm to determine the global maximum or minimum of a large parameter set is non-trivial. However, implementing/developing an algorithm to calculate the global maximum or minimum is not the focus of this thesis, and we therefore implemented [what we deem as] a constrained brute force search algorithm to estimate all parameters. In the implementation of a constrained brute force search algorithm, we used the results of the academic paper that the model, whose parameters we were estimating, originated from to estimate the bounds of each parameter. We assumed that the estimated bounds of each parameter represented the upper and lower bound of a uniform distribution, and randomly selected a value for each parameter during each iteration of the estimation algorithm. During each iteration, we calculated an instance of the objective function and compared it to the maximum result that we had obtained (the maximum value was set to an arbitrarily large value to begin the algorithm). If the calculated objective function of the iteration proved to be larger than the prevailing maximum, it was stored along with the parameters that were used to generate the result. In order to ensure the approach was as exhaustive, we used a few hundred thousand iterations. Such an approach was chosen because calculation speed was not crucial to this thesis, and these models were implemented for the sole purpose of calculating Cont’s convex model uncertainty measure.

Before we continue, we want to highlight a couple differences between the methodology utilized by Li and Kwon [45] when calculating Cont’s measure to the one that we utilized in this thesis. Li and Kwon used the example as presented by Lo [46] for their computational experiment on equity derivatives. In Lo’s example, two classes of price processes, specifically lognormal diffusions and mixed diffusion-jump processes, were chosen as the possible drivers of the un-
deriving stock price. These two processes were chosen because Lo demonstrated that given any dataset, the risk-neutral variances of the two processes were numerically identical [46]. As such, when Li and Kwon set up the semi-parametric methodology, they fixed the moments of their semi-infinite set of pricing measures to those computed by the processes chosen in the Lo paper. This implies that both the discrete set of models chosen to compute Cont’s measure as well as the continuous set of models chosen to compute Cont’s measure using the Li-Kwon methodology had identical moments. Furthermore, in the computational experiment of Li and Kwon, they did not estimate the parameters from observed price data, but rather chose parameter values. We argue that such a framework, specifically the calibration of each model’s resulting distribution to the identical set of moments, is difficult when using commodity derivatives due to the additional complexities the empirical properties of commodity dynamics relative to those of equities (as discussed in detail in a previous section - 3.1). Furthermore, additional complexities arise when calibrating to real price data (mentioned in aforementioned paragraph). It is because of these practical issues that we relax the condition of matching moments, and justify such a relaxation by relying on the fact that each model is calibrated to the same futures and option data.

We now present the optimal parameters calculated when using the Kalman filter on the rolling term structure of futures data. The estimated parameters for the S1997-1 can be found in table 4.3, while those estimated for models S1997-2, SS2000, and CS2003 can be found in tables 4.4, 4.5, 4.6 respectively.

<table>
<thead>
<tr>
<th>Look-back (yrs)</th>
<th>κ</th>
<th>σ</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.6196</td>
<td>0.88</td>
<td>4.074</td>
</tr>
<tr>
<td>0.50</td>
<td>0.2467</td>
<td>0.8707</td>
<td>3.222</td>
</tr>
<tr>
<td>1.50</td>
<td>0.2003</td>
<td>0.8489</td>
<td>2.9323</td>
</tr>
<tr>
<td>3.00</td>
<td>0.269</td>
<td>0.8637</td>
<td>3.3129</td>
</tr>
</tbody>
</table>

Table 4.4: Estimated parameters from S1997-2 model.

<table>
<thead>
<tr>
<th>Look-back (yrs)</th>
<th>ρ</th>
<th>σ2</th>
<th>λ</th>
<th>α</th>
<th>κ</th>
<th>σ1</th>
<th>μ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.5791</td>
<td>0.8375</td>
<td>0.1892</td>
<td>0.0156</td>
<td>2.1021</td>
<td>0.4965</td>
<td>0.0161</td>
</tr>
<tr>
<td>0.50</td>
<td>0.1476</td>
<td>0.634</td>
<td>0.0444</td>
<td>0.1565</td>
<td>1.0338</td>
<td>0.482</td>
<td>0.0141</td>
</tr>
<tr>
<td>1.50</td>
<td>0.1666</td>
<td>0.705</td>
<td>0.1806</td>
<td>0.1689</td>
<td>1.5331</td>
<td>0.4802</td>
<td>0.000907</td>
</tr>
<tr>
<td>3.00</td>
<td>0.2794</td>
<td>0.7629</td>
<td>0.4003</td>
<td>0.252</td>
<td>1.6897</td>
<td>0.459</td>
<td>0.0214</td>
</tr>
</tbody>
</table>

Referring to the tables of the calculated parameters for each model (4.3, 4.4, 4.5, and 4.6), each look back window contains a slightly different optimal set of parameters. Using Derman’s nomenclature from our discussion on model uncertainty 2.1 [22], such a variation in parameter estimation can result from choosing an incorrect model, having a badly approximated solution, or even the inapplicability of modeling. We briefly discuss the aforementioned possible sources
of our observed parameter variation, however we want to bring attention to the fact that model uncertainty is implicit in the existence of such variation.

If we have chosen an incorrect model, the estimated parameters from an arbitrary sample of price history would be sensitive to any changes in the sample. Although this is not specifically illustrated in the aforementioned tables, the changes of the parameter estimates using different look back windows may be indicative of choosing an incorrect model. Alternatively, we may have decided on an incorrect look back window of historical data to properly estimate the data.

A badly approximated solution is also quite likely given the non-linearities of optimizing the set of parameters of the models that we have chosen. However, even though our constrained brute force approach is rather elementary, it is exhaustive over a constrained space of possible parameter values. While it is unlikely to achieve the globally optimal solution, it is less likely to get stuck in a locally optimal solution. This implies that we may have found the optimal solution space neighbourhood, but due to the complex structure and poor behavioural characteristics of log-likelihood surfaces, it is difficult to know for sure. That being said, the constrained brute force approach did yield a [optically] larger variation among the optimized parameter set over different look back periods. This highlights the fact that a larger parameter set results in a more complex log-likelihood surface, and it is more likely that each optimal solution is in fact simply a local maximum.

The inapplicability of modeling the futures term structure and option price data using an affine-
Gaussian model is also possible. Understanding how many factors are driving a commodity price process is challenging. It is documented that using 3 factors often aptly captures the futures term structure while an additional two are necessary to encompass most of the remaining additional orthogonal drivers of option prices [63]. While this may be true, what if non-Gaussian affine models are more appropriate?

While we did not do out-of-sample testing of our estimated parameter sets, we applied each optimal parameter set to the data of each of the different look back windows. What we noticed was that the look back window that the parameters were optimized for exhibited, visually, a better fit, however, they did not deviate significantly when applied to the obtained data from different look back windows. This was more apparent the more parameters that we had, i.e. each of the CS2003 model’s optimal parameter sets appeared to fit reasonably well to futures data of different look back windows. That is interesting in itself because arguably the CS2003 parameter set variation is the highest among the chosen models. This further illustrates the complexities of the log-likelihood surface and highlights a potential high level of susceptibility to model uncertainty. While finding optimal parameters for a model is an interesting research topic in itself, we do not provide any further detail on this topic because it is not paramount to the thesis. We provided the brief discussion above in order to demonstrate actual model uncertainties that are present in our calculations. This should motivate the semi-parametric form for calculating model uncertainty as being more informative because identifying all potential sources of model uncertainty in all possible models is much more tedious than the semi-parametric form we present in the next subsection.

Before proceeding to presenting the calculated options prices, we note that the contract for March delivery expires on February 20 (third business day prior to the twenty-fifth calendar day of the month preceding the delivery month), however, delivery of the spot commodity must actually occur in the delivery month (March in this case). As such, even though the contract does not exist in March, the fact that its value is derived from accepting or taking delivery of the spot commodity in March, we assume that modeling the option price remains a derivative of the spot commodity in March. Due to the variation in delivery during the month of March, we simulated approximately 75 days of data when calculating derivative prices (beginning on December 2, 2012, and ending mid-month in March). We refer you to table 4.7 for the derivative prices that were calculated using the different look back periods. Each option price was calculated using a Monte Carlo simulation.

It is clear that the S1997-1 model was unable to model option prices effectively. This is to be expected of a one-factor model because they are not able to incorporate the volatility of a futures term structure into their model description as well as higher factor models do. Both two factor models, SS2000 and S1997-2, calculated prices that were closer to those that were actually observed (4.7). This may imply that most of the uncertainty in option price data was captured using two stochastic factors. The CS2003 model was used as the benchmark model
Table 4.7: Estimated option prices for each model. The strikes that have * indicate those that were used as benchmark options.

<table>
<thead>
<tr>
<th>Lbk(yrs)</th>
<th>100*</th>
<th>105*</th>
<th>107.5</th>
<th>110</th>
<th>112.5</th>
<th>115</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5</td>
<td>12.6296</td>
<td>11.1081</td>
<td>10.4162</td>
<td>9.762</td>
<td>9.1449</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>12.3757</td>
<td>10.8496</td>
<td>10.1537</td>
<td>9.5002</td>
<td>8.8844</td>
</tr>
<tr>
<td>S1997-2</td>
<td>0.25</td>
<td>0.9881</td>
<td>0.3895</td>
<td>0.2332</td>
<td>0.1358</td>
<td>0.0763</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.1708</td>
<td>0.5247</td>
<td>0.3468</td>
<td>0.2265</td>
<td>0.147</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>0.9063</td>
<td>0.352</td>
<td>0.2104</td>
<td>0.1218</td>
<td>0.0688</td>
</tr>
<tr>
<td></td>
<td>3.0</td>
<td>0.6522</td>
<td>0.2152</td>
<td>0.117</td>
<td>0.0611</td>
<td>0.0306</td>
</tr>
<tr>
<td>SS2000</td>
<td>0.25</td>
<td>2.5293</td>
<td>1.8102</td>
<td>1.5253</td>
<td>1.2824</td>
<td>1.0761</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.135</td>
<td>0.6823</td>
<td>0.5237</td>
<td>0.4008</td>
<td>0.3054</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>1.149</td>
<td>0.7059</td>
<td>0.5493</td>
<td>0.4261</td>
<td>0.3289</td>
</tr>
<tr>
<td></td>
<td>3.0</td>
<td>1.6025</td>
<td>1.0331</td>
<td>0.8242</td>
<td>0.655</td>
<td>0.5181</td>
</tr>
<tr>
<td>CS2003</td>
<td>0.25</td>
<td>1.1978</td>
<td>0.5854</td>
<td>0.3991</td>
<td>0.2683</td>
<td>0.178</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.2991</td>
<td>0.564</td>
<td>0.3583</td>
<td>0.2217</td>
<td>0.1348</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>1.2235</td>
<td>0.5124</td>
<td>0.3172</td>
<td>0.1907</td>
<td>0.1112</td>
</tr>
<tr>
<td></td>
<td>3.0</td>
<td>1.2354</td>
<td>0.5776</td>
<td>0.4337</td>
<td>0.3009</td>
<td>0.2061</td>
</tr>
</tbody>
</table>

in calculating Cont’s convex measure of model uncertainty and was therefore calibrated to the benchmark option prices. As mentioned, the CS2003 model was the one chosen to perfectly calibrate to the observed market prices. However, our parameter estimation methodology was unable to identify a set of parameters that perfectly replicated the two observed option prices when calibrating the CS2003 model. This does not limit us in the calculation of Cont’s convex risk measure to be described below.

### 4.3 Calculation

As explained, we present the final calculations of both Cont’s and Li-Kwon’s convex risk measure. We provide a detailed description of the Li-Kwon methodology before presenting the results. The detailed description should be viewed as a practitioner’s guide to the Li-Kwon methodology.

#### 4.3.1 Cont’s Convex Measure

We have calculated Cont’s measure using three different penalty norms, specifically the $\| \cdot \|_1$, $\| \cdot \|_2$, and the $\| \cdot \|_{\infty}$. Moreover, we calculated a separate measure using each of the four different look back windows. Each look back window’s model uncertainty was calculated using the
calculated option prices from table 4.7 and the calculated prices from the CS2003 model were assumed to be perfectly calibrated to the market thus the calculated prices of the benchmark derivatives serve as the benchmark prices (as described in the previous subsection). The average model uncertainty for each derivative was calculated and displayed in order to present a representative result for the derivative.

Table 4.8: Cont’s convex risk measure results from the four models. The derivative prices obtained from the CS2003 model were the reference option prices used in calculating the displayed results. CLH3C is short-form notation for a call option written on a March 2013 crude oil futures contract.

<table>
<thead>
<tr>
<th>Lbk(yrs)</th>
<th>CLH3C 107.5</th>
<th>CLH3C 110</th>
<th>CLH3C 112.5</th>
<th>CLH3C 115</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-norm</td>
<td>0.25</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>1.50</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>AVG</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>2-norm</td>
<td>0.25</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>1.50</td>
<td>0.025</td>
<td>0.028</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>AVG</td>
<td>0.006</td>
<td>0.007</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>∞-norm</td>
<td>0.25</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>0.001</td>
<td>0.015</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>1.50</td>
<td>0.039</td>
<td>0.042</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>AVG</td>
<td>0.010</td>
<td>0.014</td>
<td>0.008</td>
<td>0.000</td>
</tr>
</tbody>
</table>

The CS2003 model was calibrated to the benchmark options. As mentioned in the previous section, the estimated parameter set obtained for this model did not yield the exact benchmark prices of table 4.1. We decided to use the output prices of the CS003 model as the new benchmark prices when calculating Cont’s model uncertainty and we claim that this does not diminish the obtained results. The results are not diminished because the calculated parameter set is the best possible result given our methodology. Furthermore, we calibrated the option prices using a European option pricing model whereas the option prices were those from options of American type. We have stated that the American option price should converge to the European option price the further out of the money the option is. However, in our example, the option expires in four months time and the time premium of an option that is ≥ 10% out of the money likely has a non-zero early exercise premium. However, our calibration is sufficiently close to the observed American type option prices to be considered useful for a practitioner using real data. Cont’s measure is a tool attempting to quantify variation in the results in computing the same derivative, and our methodology does not diminish this in the context of a practitioner garnering information about potential variation in the computed derivative.
prices.

Referring to the results of table 4.8, the model uncertainty of many derivatives is zero, thus implying that there is no model uncertainty in calculating the prices of these illiquid securities (similar result was obtained by Li and Kwon [45]). The reason why this occurs is that the derivative prices that are not calculated by the benchmark model are penalized by the deviation that the model output has to the benchmark prices. Since the benchmark model is not penalized, it often becomes both the upper and lower bound of the measure. This result highlights the limitation of calculating the aforementioned convex risk measure using a discrete set of models.

4.3.2 Li-Kwon Convex Measure

Before we present the calculated results of the Li-Kwon convex measure, we outline the steps involved in setting up the SDP formulation and ultimately its solution. We will follow a common example throughout the description in order to make each step in the procedure more explicit. The example will be developed as we continue with the description. Furthermore, we have outlined the procedure in a flowchart (4.1), and we use it as a reference in what follows.

1. **Identify derivative**

   We first must identify the derivative that we want to calculate the model uncertainty of. It is important to note that when using the Li-Kwon framework, we are only capable of calculating the model uncertainty of derivatives that are not path dependent, i.e, a derivative whose value is dependent on the ending price of the underlying as opposed to the path it took to get there\(^9\). One class of derivatives that are not-path dependent are European call and put options. Furthermore, we want to identify the model uncertainty of illiquid options (liquid options restrict model uncertainty to the bid-ask spread).

   The working example that we will be using is a thinly traded, or nearly illiquid, European call option with strike price \(k_0\).

2. **Identify candidate benchmark derivatives**

   A benchmark derivative must be one that has a similar payoff structure to that of the derivative whose model uncertainty we are calculating. In our example, we are calculating the model uncertainty of an illiquid European option with strike price \(k_0\), and a time to maturity of \(T\). Any liquid European call or put option with a different strike price is a possible candidate benchmark derivative. We note that choosing any number of benchmark derivatives is possible, however, we have decided to simplify this example by choosing only one. As such, we have chosen a call option with a strike price of \(k_1\), and

---

\(^9\)Refer to 2.5 for more details.
identify derivative practitioner
identify candidate benchmark derivatives
calculate moments of underlying distribution
partition space into breakpoints of derivative payoffs
convert each partition into its corresponding SDP formulation
choose norm penalty function
re-write norm in dual representation
adjust moments
calculate SDP problem
feasible and bounded solution?
stop

Figure 4.1: Flowchart of the methodology for calculating the continuous case of the convex risk measure of model uncertainty.
3. Calculate moments of underlying distribution

In order to calculate the empirical moments of the underlying distribution, we need a representative sample of the underlying future distribution. This sample can be generated in many different ways. Arguably, the simplest method is to take the most recent end-of-day prices for a short period of time prior to the valuation date. Determining the optimal historical data window for retrieving end-of-day prices that will ultimately be reflective of the underlying’s future distribution is very subjective. Practitioners differ on, to name a few, determining the historical data window, on the level of noise in price data, and on whether intra-day prices are better than end-of-day settlement prices. We recommend using multiple look back windows to determine the sensitivity of model uncertainty to such a parameter. Furthermore, we must determine the number of moments we want to use in our calculation. In this case, theoretically, the law of diminishing returns applies, i.e., each additional moment does not add much additional benefit to the size of the bound after roughly 3-4 moments. Additionally, adding each moment complicates the computation and makes the SDP formulation brittle [5], [31]. As such, using three or four moments is often sufficient in practice and is what we recommend. However, for the purpose of our example, we have decided that the first two moments are sufficient for explaining the SDP formulation.

4. Partition space into breakpoints of derivative payoffs

The idea behind partitioning the space into the breakpoints of the derivative payoffs is to reformulate the constraints of the Li-Kwon formulation into SDP constraints (will be explicitly shown in the next step). The space that is being partitioned is the space of possible values that the underlying asset can have, i.e., the set of positive real numbers \( \mathbb{R}_+ \). As such, partitioning the space is simply identifying regions where the derivative’s payoff function is continuous and can be explicitly evaluated. We add depth to this explanation through the following example.

Recall the first constraint of the Li-Kwon upper bound formulation 2.33 (note that the lower bound formulation is very similar to what we describe below, simply change the \( \leq \) to a \( \geq \) and re-evaluate),

\[
\max(0, \omega-k_0) - \sum_{i=l}^l \lambda_{m_i}(\omega^i-m^i) - \sum_{j=l}^o \lambda_{h_j}\max(0, \omega-k_j) - \sum_{j=o+1}^k \lambda_{h_j}\max(0, k_j-\omega) \leq s \ \forall \omega \geq 0
\]

In our simplified example, we are using one benchmark call option with strike price \( k_1 \),
and we are assuming that we have computed the first two moments of the underlying. We can now re-write this equation as:

\[
\max(0, \omega - k_0) - \sum_{i=1}^{2} \lambda_{m_i}(\omega^i - m_i) - \lambda_{h_1} \max(0, \omega - k_1) - s \leq 0 \quad \forall \omega \geq 0 \quad (4.1)
\]

In the context of progressing towards the SDP formulation, we have defined a breakpoint as a partitioning of the positive real line where the above function/constraint evaluates to different values. In our example, the two breakpoints are the strike price of the benchmark option, \(k_0\), and the strike price of the illiquid option \(k_1\). The partitioning of the space is as follows, assuming \(k_0 > k_1\) and denoting equation 4.1 as \(\phi(\omega)\):

\[
\phi(\omega) = \begin{cases} 
\phi_0(\omega), & \omega \in [0, k_1] \\
\phi_1(\omega), & \omega \in [k_1, k_0] \\
\phi_2(\omega), & \omega \in [k_0, \infty) 
\end{cases}
\quad (4.2)
\]

In every partition, the equation (4.1) will alter slightly, in the following way:

(a) \(\omega \in [0, k_1]\) Neither call option are in the money in this partition, the equation can therefore be re-written as:

\[
\lambda_{m_1}(\omega - m_1) - \lambda_{m_2}(\omega^2 - m_2) - s \leq 0
\]

(b) \(\omega \in [k_1, k_0]\) Only the benchmark call option is in the money over this partition, the equation can therefore be re-written as:

\[
\lambda_{m_1}(\omega - m_1) - \lambda_{m_2}(\omega^2 - m_2) - \lambda_{h_2}(\omega - k_1) - s \leq 0
\]

(c) \(\omega \in [k_0, \infty]\) Both the benchmark and illiquid call options are in the money over this partition, the equation can therefore be re-written as:

\[
(\omega - k_0) - \lambda_{m_1}(\omega - m_1) - \lambda_{m_2}(\omega^2 - m_2) - \lambda_{h_2}(\omega - k_1) - s \leq 0
\]

5. **Convert partition into its corresponding SDP formulation**

In our discussion of the Li-Kwon methodology for solving Cont’s risk measure (2.5), we discussed the conversion of a strictly positive polynomial, \(\sum_{r=0}^{n} y_r \omega^r \quad \omega \geq 0\), to a list of constraints for a semidefinite matrix \(X\) [5]. Each partition is a candidate for such a conversion once it is represented in the appropriate polynomial form because each partition
can each be represented as being strictly nonnegative. We demonstrate how each partition is converted into the appropriate polynomial representation, then we present how each polynomial can be converted into a list of constraints that the semidefinite matrix $X$ must follow. Additionally, we note for clarity that each partition utilizes its own matrix, and in our example, each partition will yield a separate $3 \times 3$ matrix, $(x_{ij})_{i,j=0,\ldots,2}$. We only compute the constraints of the upper bound, but the lower bound formulation is very similar and is omitted.

Before evaluating each partition individually, each partition has a few constraints in common. The semidefinite matrix $X \succeq 0$ satisfies $\sum_{i,j: i+j=2l-1} x_{i,j} = 0$, $l = 1, \ldots, n$. As such, the matrix $X$ corresponding to each partition in our example will have the following constraints:

$$x_{0,1} + x_{1,0} = 0 \quad x_{1,2} + x_{2,1} = 0$$

(a) $\omega \in [0, k_1]$ We re-write equation 4.1 over the partition as:

$$\omega^0 \left( s - \lambda_{m_1} m_1 - \lambda_{m_2} m_2 \right)_{y_0} + \omega^1 \left( \lambda_{m_1} + \lambda_{m_2} \right)_{y_1} + \omega^2 \left( \lambda_{m_2} \right)_{y_2} \geq 0$$

Under such a representation, we can use the following [5]:

$$\sum_{r=0}^{l} y_r \binom{n - r}{l - r} a^r = \sum_{i,j: i+j=2l} x_{i,j}, \quad l = 0, \ldots, n,$$

Which yields the following SDP constraints:

$$x_{0,0} = y_0$$
$$x_{2,0} + x_{1,1} + x_{0,2} = 2y_0 + y_1 k_1$$
$$x_{2,2} = y_0 + k_1 y_1 + k_1^2 y_2$$

(b) $\omega \in [k_1, k_0]$ We re-write equation 4.1 over the partition as:

$$\omega^0 \left( s - \lambda_{m_1} m_1 - \lambda_{m_2} m_2 - \lambda_{h_1} k_1 \right)_{y_0} + \omega^1 \left( \lambda_{m_1} + \lambda_{h_1} \right)_{y_1} + \omega^2 \left( \lambda_{m_2} \right)_{y_2} \geq 0$$

Under such a representation, we can use the following [5]:

$$\sum_{m=0}^{l} \sum_{r=m}^{k+m-l} y_r \binom{r}{m} \binom{n - r}{l - m} a^{r-m} b^m = \sum_{i,j: i+j=2l} x_{i,j}, \quad l = 0, \ldots, n.$$
Which yields the following SDP constraints:

\[ x_{0,0} = y_0 + k_1 y_1 + k_1^2 y_2 \]
\[ x_{2,0} + x_{1,1} + x_{0,2} = 2y_0 + y_1(k_1 + k_0) + 2y_2k_1k_0 \]
\[ x_{2,2} = y_0 + k_0 y_1 + k_0^2 y_2 \]

(c) \( \omega \in [k_0, k_\infty] \) We re-write equation 4.1 over the partition as:

\[ \omega^0 \left( k_0 + s - \lambda_{m_1} m_1 - \lambda_{m_2} m_2 - \lambda_{h_1} k_1 \right) + \omega^1 \left( \lambda_{m_1} + \lambda_{h_1} - 1 \right) + \omega^2 \left( \lambda_{m_2} \right) \geq 0 \]

Under such a representation, we can use the following [5]:

\[ \sum_{r=l}^{n} y_r \binom{r}{l} a^{r-l} = \sum_{i,j: i+j=2l} x_{i,j}, \quad l = 0, \ldots, n, \]

Which yields the following SDP constraints:

\[ x_{0,0} = y_0 + k_0 y_1 + k_0^2 y_2 \]
\[ x_{2,0} + x_{1,1} + x_{0,2} = y_1 + 2y_2 k_0 \]
\[ x_{2,2} = y_2 \]

6. **Choose norm penalty function**

Recall the second constraint of the Li-Kwon upper bound 2.33:

\[ \sup_{(q_j)_{j=1}^{k}} \sum_{j=1}^{k} \lambda_{h_j} q_j - ||h^* - vec((q_j)_{j=1}^{k})|| \leq t \]

where \( h^* \) is a vector of benchmark derivative prices.

The norm function of the above equation must be semi-definite representable\(^{10}\). In our example, we only have one benchmark derivative and we can use a \( \|\cdot\| \), to represent the second constraint as follows:

\[ \sup_{q_1} \lambda_{h_1} q_1 - ||h_1^* - q_1|| \leq t \]

7. **Re-write norm in dual representation**

The goal is to isolate the contents that are to the left of \( \leq t \) and notice that it forms

\(^{10}\)A convex function \( f : \mathbb{R}^m \to \mathbb{R} \cup \{\infty\} \) is called semi-definite representable (SDr) if its epigraph \( \{(x, t) | f(x) \leq t\} \) is an SDr set, i.e. a set that can be expressed as \( A(x, t) - B \succeq 0 \), where \( A \) denotes a linear operator and \( B \) denotes a constant matrix [45].
another constrained optimization problem. The dual of this problem is a minimization problem, \( \inf \), which can be put into the objective function of the Li-Kwon upper bound 2.33. As such, it adds additional constraints to the overall problem.

Depending on the chosen normal, the dual relationship is often not easily constructed. Most common primal-dual relationships are available in many introductory optimization textbooks. In our example, we have chosen the 1-norm, and its dual representation is presented below:

\[
\inf_{y_1, y_2} h_1^* y_1 - h_2^* y_2
\]

subject to

\[
\begin{align*}
& y_1 - y_2 \geq \lambda_{h_1} \\
& -y_1 - y_2 \geq -1 \\
& y_1, y_2 \geq 0
\end{align*}
\]

\( h_1^* \) is the collected option price data and \( y \) is the dual variable.

8. **Calculate SDP problem**

There are a few free optimization solver software packages available, such as SeDuMi and SDPT3. Furthermore, there is software that provides a useful language for writing optimization problems into MATLAB, which translates the program when compiled into a format suitable for an optimization solver that it supports. In particular, YALMIP and CXV are well developed interfaces that implement both SeDuMi and SDPT3 (among others). We used CVX for the computation of this thesis [21, 33], but any optimization package that handles SDP formulations should be suitable. Once a solver has been chosen, the constraints that we have developed in the preceding steps can be transcribed into the software package of your choice, and ultimately solved.

9. **If infeasible or unbounded, adjust moments, otherwise solution is obtained**

Many SDP problems are sensitive to their input parameters, and the formulation that we have described above is no different. If the initial attempt at a calculation resulted in an infeasible solution, it is likely that the constraints were incorrectly transcribed. We recommend re-checking the primal dual relationship to ensure that it was properly represented and calculated. If the constraints were properly transcribed however, and the result is an unbounded solution, we recommend modulating the moments.

Even though the moments were empirically calculated, due to computational restrictions in the feasible region of a SDP problem, it is necessary to change the moments to a suitable level that enables a solution. The procedure that we follow is we chose a moment that we believe to be the most relevant from the set of calculated empirical moments. Once this moment has been determined, small adjustments are made to the other moments until we enter into the feasible region. While this procedure is subjective, the resulting bound remains informative. We present an example procedure that we have followed.
When choosing a representative moment from the set of empirically calculated moments, the main decision criteria should be such that the moment is most influential/related to the derivative whose model uncertainty we are calculating. Since we are calculating the model uncertainty of a European option in our example, we argue that the most relevant moment is the second moment $m_2$. We logically deduce this by acknowledging the dependence an option contract has on its implied volatility (we discuss this briefly in the appendix A). The most likely estimator of the future implied volatility is the historical volatility, which is closely related to the second moment of a historical distribution (sample). Upon choosing the second moment, we have noticed a relationship that is required between the first three moments. Assuming we set the second moment to the calculated empirical moment, the first and third follow these bounds:

\[
\begin{aligned}
    m_1 & \leq \sqrt{(1 - \epsilon)m_2} \\
    m_3 & \geq m_2^2 (1 + 2\epsilon)
\end{aligned}
\]  

(4.3)

The value $\epsilon$ is near zero (we used a value of roughly 0.0045), and is only implemented to ensure that the solution is bounded. If the calculated moments violate the above bounds, they have to be adjusted. The adjusted moments are often quite small relative to the ones originally calculated and remain informative when calculating model uncertainty. Please refer to the appendix for the MATLAB implementation of this problem (D.1).

We decided to present the model uncertainty for each look back window in order to follow the same format that we presented Cont’s measure. We separate the results based on the number of benchmark options that we used in order to demonstrate how the additional information on the risk neutral measure implicit in the option price affects the measure.

Upon first glance of tables 4.11 and 4.12, it is clear to see that the Li-Kwon methodology has determined that model uncertainty does in fact exist when evaluating every derivative. This is in stark contrast to when we evaluated Cont’s measure of model uncertainty where the trivial answer of zero uncertainty was common.

When looking at the influence of adding moments or benchmark derivatives, we notice that the additional benchmark option tightens the bounds of the model uncertainty calculation significantly more than the additional moment. This implies that the additional benchmark option allows it to capture more information of the risk-neutral measure and calculate a tighter bound. This may imply that adding an additional benchmark derivative adds more relative information about the risk-neutral distribution than adding an additional moment does to underlying distribution. We posit that adding a third benchmark option would have less effect than adding the second (law of diminishing returns).
It is interesting to note that the additional moment \((n = 3)\) does not lower the bound in all cases, and it has far less influence than the addition of a benchmark derivative. We posit two alternative interpretations of this. First, one could assume that a potential tail event is hidden in the third moment that would increase the value of an option thus increasing the bounds in certain cases. This may imply that most of the information is contained in the first two moments of a distribution during normal periods, whereas during periods of stress, the higher moments behave differently. Alternatively, each additional moment is subject to the law of diminishing returns, and does not provide as much additional information.

Table 4.9: Adjusted empirical moments using \(\epsilon = 0.0045\) from equation 4.3. All moments denoted with * are those that were changed by the moment boundary condition.

<table>
<thead>
<tr>
<th></th>
<th>0.25 year</th>
<th>0.50 year</th>
<th>1.50 year</th>
<th>3.00 year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>92.76*</td>
<td>91.81*</td>
<td>95.93</td>
<td>93.24</td>
</tr>
<tr>
<td>2</td>
<td>8644.49</td>
<td>8467.86</td>
<td>9248.06</td>
<td>8748.07</td>
</tr>
<tr>
<td>3</td>
<td>810961.49*</td>
<td>786234.14*</td>
<td>897361.03*</td>
<td>826038.6</td>
</tr>
</tbody>
</table>

Table 4.10: The percentage change in each moment upon the adjustment that allows the problem to be feasible. With regards to the notation, \(m_{i,t}\), is where the first subscript indicates the moment \(i \in 1, 2, 3\), and the second subscript indicates the look back window \(t \in 0.25, 0.50, 1.50, 3.00\).

<table>
<thead>
<tr>
<th></th>
<th>(m_{1,0.25})</th>
<th>(m_{1,0.50})</th>
<th>(m_{3,0.25})</th>
<th>(m_{3,0.50})</th>
<th>(m_{3,1.50})</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.14%</td>
<td>-0.09%</td>
<td>0.61%</td>
<td>0.47%</td>
<td>0.16%</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.11: Li-Kwon’s generalization of Con’ts convex risk measure results using one benchmark option (CLH3C 100). The first half of results calculates the measure using two moments \((n = 2)\), whereas the later half uses three moments \((n = 3)\). CLH3C is short-form notation for a call option written on a March 2013 crude oil futures contract.

<table>
<thead>
<tr>
<th>Lbk(yrs)</th>
<th>CLH3C 107.5</th>
<th>CLH3C 110</th>
<th>CLH3C 112.5</th>
<th>CLH3C 115</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n = 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>0.431</td>
<td>0.301</td>
<td>0.175</td>
<td>0.068</td>
</tr>
<tr>
<td>0.50</td>
<td>0.401</td>
<td>0.295</td>
<td>0.176</td>
<td>0.073</td>
</tr>
<tr>
<td>1.50</td>
<td>0.866</td>
<td>0.762</td>
<td>0.657</td>
<td>0.553</td>
</tr>
<tr>
<td>3.00</td>
<td>0.876</td>
<td>0.775</td>
<td>0.674</td>
<td>0.572</td>
</tr>
<tr>
<td>AVG</td>
<td>0.643</td>
<td>0.533</td>
<td>0.420</td>
<td>0.317</td>
</tr>
</tbody>
</table>

| \(n = 3\) |             |           |             |           |
| 0.25     | 0.501       | 0.360     | 0.231       | 0.122     |
| 0.50     | 0.187       | 0.086     | 0.055       | 0.011     |
| 1.50     | 0.866       | 0.762     | 0.657       | 0.553     |
| 3.00     | 0.782       | 0.649     | 0.517       | 0.384     |
| AVG      | 0.584       | 0.464     | 0.365       | 0.267     |

In the detailed list of instructions on the implementation of the Li-Kwon methodology, we made note of an empirical relationship among the set of empirical moments that was necessary to obtain a bounded solution (equation 4.3). When we formulated this relationship, we assumed
Chapter 4. Results & Discussion

Table 4.12: Li-Kwon’s generalization of Con’ts convex risk measure results using two benchmark option (CLH3C 100, and CLH3C 105). The first half of results calculates the measure using two moments \((n = 2)\), whereas the later half uses three moments \((n = 3)\). CLH3C is short-form notation for a call option written on a March 2013 crude oil futures contract.

<table>
<thead>
<tr>
<th>Lbk(yrs)</th>
<th>CLH3C 107.5</th>
<th>CLH3C 110</th>
<th>CLH3C 112.5</th>
<th>CLH3C 115</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n = 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>0.172</td>
<td>0.175</td>
<td>0.120</td>
<td>0.076</td>
</tr>
<tr>
<td>0.50</td>
<td>0.110</td>
<td>0.085</td>
<td>0.043</td>
<td>0.154</td>
</tr>
<tr>
<td>1.50</td>
<td>0.361</td>
<td>0.488</td>
<td>0.453</td>
<td>0.417</td>
</tr>
<tr>
<td>3.00</td>
<td>0.354</td>
<td>0.479</td>
<td>0.439</td>
<td>0.398</td>
</tr>
<tr>
<td>AVG</td>
<td>0.249</td>
<td>0.307</td>
<td>0.263</td>
<td>0.261</td>
</tr>
<tr>
<td>(n = 3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>0.172</td>
<td>0.240</td>
<td>0.151</td>
<td>0.076</td>
</tr>
<tr>
<td>0.50</td>
<td>0.110</td>
<td>0.085</td>
<td>0.043</td>
<td>0.154</td>
</tr>
<tr>
<td>1.50</td>
<td>0.358</td>
<td>0.488</td>
<td>0.453</td>
<td>0.417</td>
</tr>
<tr>
<td>3.00</td>
<td>0.334</td>
<td>0.441</td>
<td>0.382</td>
<td>0.322</td>
</tr>
<tr>
<td>AVG</td>
<td>0.243</td>
<td>0.314</td>
<td>0.257</td>
<td>0.242</td>
</tr>
</tbody>
</table>

that the second moment, \(m_2\), was the most relevant in the calculation of our example. We make the same assumption in the results that we obtained, and using this relationship, we present the re-calculated moments in table 4.9, where we have set \(\epsilon = 0.0045\). Five out of the calculated eight moments were outside of the boundary as set by equation 4.3) and were adjusted to the boundary. The resulting differences are in table 4.10.

While making changes to the empirical moments is an undesirable consequence of the methodology, we argue that the changes in moments, as denoted in table 4.10 are inconsequential to the informational content obtained in the calculation. To cement the point that the differences are inconsequential, recall that we chose four different look back windows because picking a representative sample of the most likely future distribution of an asset is very subjective to the practitioners view on markets. Furthermore, in our presentation of the Li-Kwon methodology, we do not identify any particular look back window as being correct because they are statistically equally likely. What this implies is that theoretically, any combination of moments from the calculated subset are likely and making the small alterations to the moments that we made lie within this subset.

Nonetheless, the stability of the algorithm is a concern, but we have introduced a methodology to circumvent the largest issues that are likely in the given context. The MATLAB code for implementing the SDP formulation can be found in the appendix (D). We present the code for the problem formulation containing two moments with one benchmark option (section D.1) as well as the code for the problem formulation of three moments with two benchmark options (section D.2).
Chapter 5

Conclusions & Future Work

As outlined in our introduction, we aimed to provide a practitioner’s guide of a modern framework for evaluating the model uncertainty [of a subset of derivative related calculations] that we believe is a useful and informative risk measure for understanding the potential impact of model uncertainty.

We demonstrated the utility and flexibility of the Li-Kwon methodology for calculating a convex measure of model uncertainty, and compared it to the more parameterized methodology as proposed by Cont. We argue, with conclusive evidence from the results of tables 4.8, 4.11, and 4.12, that the Li-Kwon methodology is more informative towards identifying the level of model uncertainty in a derivatives contract. We also provided a detailed description of the exact methodology required for implementing the Li-Kwon framework, 4.1, that simplifies the sometimes onerous task of tackling an SDP problem.

However, due to current limitations associated with solving complex SDP problems, the Li-Kwon formulation is sensitive to its input parameters, specifically, the moments of the risk-neutral distribution. We spent time describing a method to circumvent the sensitivity to the input moments in order to obtain results that remain meaningful and informative. Specifically, subject to our objective arguments that justify our assumption that the second moment provided the most relevant information for our example, we provided a method to alter the calculated first and third moments relative to the fixed second moment (refer to equation 4.3) in order to arrive at a bounded solution without sacrificing all the informational content of the problems formulation. While suboptimal in necessity, the changes in moments are small in comparison to their original calculated values and we argue are inconsequential towards obtaining the final result.

Furthermore, the presented Li-Kwon framework is only applicable to path independent options. This limits the applicability to the real world because path independent derivatives form only
a small subset of derivatives trading volume on both exchanges and OTC\textsuperscript{1}. In our example, we were able to circumvent this problem by relating the prices of deep out of the money American type options to those of their European counterparts. The logic is that these options converge to the same price the deeper out of the money the strike price is and can be informative towards the range of values expected of a European option. As a consequence to that, the utility of the Li-Kwon framework may appear diminished, however, striking the bound of a simple derivative that may be illiquid may be informative towards more complex derivatives that use the simpler derivative in its construction, such as compound options.

Nonetheless, in order to make the Li-Kwon formulation a robust methodology for calculating model uncertainty that can serve as a benchmark for model uncertainty calculations for financial practitioners, both a proper method for calculating a set of accurate risk-neutral moments and extending the framework to allow path dependent options are necessary. While extending the methodology to include path dependent options is not trivial\textsuperscript{2}, it may be possible to construct a robust methodology for calculating the risk-neutral moments in a manner that will always yield a bounded result.

As an example, Berstimas and Popescu [5] propose a method for obtaining the bounds of a set of risk-neutral moments given a set of liquidly traded options. Once these are calculated, we could use the resulting moment bounds as inputs into the Li-Kwon formulation. However, we would have to use a sufficient number of options in order to get a tight bound which involves solving a complex SDP problem.

In conclusion, we believe that we have developed an adequate practitioner’s guide that highlights the Li-Kwon formulation for calculating a convex measure of model uncertainty. Along the way, we demonstrated the formulations utility, and have logically deduced how to circumvent a couple idiosyncrasies associated with its calculation.

Before concluding the thesis, and as a demonstration of the flexibility of the Li-Kwon methodology, we present a methodology for calculating the model uncertainty in delta-hedging, and again contrast our continuous semi-parametric measure approach with that of a discrete set of parametric measure.

### 5.1 Delta-Hedging Under Model Uncertainty

We present a methodology for calculating the optimal contents of the replicating portfolio subject to model uncertainty using a similar framework to that of the work of Li and Kwon [45].

\textsuperscript{1}OTC - Over The Counter

\textsuperscript{2}It is possible that the Li-Kwon formulation may serve as a special case for path independent options that is part of a larger more general framework.
From our discussion on delta-hedging (refer to appendices for this discussion C, it becomes clear that delta-hedging is reliant on the model chosen to describe the dynamics of the asset. If the model is misspecified, understanding the potential impact on profit and loss of a hedging strategy, and attempting to find an optimal delta to minimize this risk are concerns that a risk management practitioner would want to elicit. In this section, we describe a potential framework, using both a discrete and infinite number of measures, that can address both of these related concerns using a specific class of derivatives. We demonstrate a framework that converges the two aforementioned concerns into one problem\(^3\).

A self-financing delta-hedging strategy as derived from a given model falls into two categories. If the model describes a complete market, a replication hedging strategy (described in the previous section) is appropriate. However, if the model describes an incomplete market, a risk-minimization hedging strategy is used. A risk-minimization hedging strategy is similar to that of a replication hedging strategy by construction, however, a perfect hedge does not exist.

Keeping with our example problem from section 2.5, we have decided to hedge a short position in an illiquid European call option using the underlying asset, and we will be describing a risk-minimization hedging strategy. With this in mind, a short position in an illiquid call option is hedged by a long position in the underlying asset, specifically, by buying delta units of the asset. Recall that the delta is the sensitivity of the derivatives value to fluctuations in the underlying, as such, we define a set of deltas for an arbitrary payoff \(X\) on an asset \(S\) as:

\[
\Psi := \left\{ \psi = \frac{\partial E^Q[X]}{\partial S} \mid Q \in Q \right\}
\]  

The set \(\Psi\) can be a collection of vectors or scalars. If it is decided that a dynamic hedge is the appropriate hedging methodology to be implemented, the set \(\Psi\) will represent a series of vectors that is reflective of the rebalancing methodology chosen. Alternatively, if it is decided that a static hedge is more appropriate, each element of the set will be a scalar value.

A call option will always have a delta that is strictly in the range between 0 and 1. This naturally follows when evaluating the behaviour of a call option payoff at extreme values of the underlying asset. If the call option is significantly in the money \((S >> K)\), the value of the call option will almost perfectly track the underlying value and will therefore have a delta near 1. Conversely, if the option is significantly out of the money \((S << K)\), the delta will be near zero. As such, in a discrete setting, the set of \(n\) deltas will be a subset of the open interval between 0 and 1, \((\psi_i)_{i=1,...,n} \in \Psi \subset (0,1) = \{x \in \mathbb{R} | 0 \leq x \leq 1\}\). Each element of the set is obtained by collecting the delta from a set of pricing measures \(\{Q_i \mid i = 1, \ldots, n\} \in Q\). In a continuous setting, collecting an infinite number of deltas is not feasible. However, we argue that as the number of deltas in the set approaches infinity, relaxing the set of infinite deltas to

\(^3\)Using the same arguments we presented in a previous section (2.5) on the continuous model uncertainty framework as developed by Li and Kwon [45], we claim that it is more informative to use an infinite set of measures when incorporating model uncertainty into a delta-hedging strategy.
its superset, the open interval between 0 and 1, will provide informative results in the problem formulations to follow

Cont [15] introduces a methodology to evaluate the robustness of hedging strategies to model uncertainty using his coherent measure of model uncertainty (can also be applied to the convex measure). He states that the risk measure \( \mu_Q(X - \int_0^T \phi_t dS_t) \) quantifies the potential impact that model misspecification may have on the profit and loss (PnL) of a hedging strategy, where \( \phi_t \) is a self-financing hedging strategy derived from a given model \( Q \in Q \), and \( X \) is a payoff function. Such a methodology can be interpreted as attempting to identify which hedging strategy from a set of optimal trading strategies, as calculated for a specific pricing measure, is most likely to lose the least amount of money if it is applied to the complete set of other pricing measures.

Cont’s problem formulation for identifying the variation in PnL implicitly incorporates the possibility of profiting from a hedging strategy into the calculation of its bounds. We argue that including the possibility of profit dilutes the informational content of the measure because the practitioner is only interested in the potential loss that a hedging strategy may have. By including the possibility of profit, Cont’s measure of model uncertainty will be larger than if such a possibility is omitted, however, no additional risk is actually present. It is therefore possible to conjecture that an informative measure of the model uncertainty present in the PnL of a hedging strategy should not include the possibility of profit. Removing the possibility of profit in Cont’s framework therefore sets the lower bound of the potential loss to the trivial value of zero, and the measure simplifies to the upper bound i.e. \( \pi / \pi_\ast = 0 \Rightarrow \mu_Q / \mu_\ast = \pi / \pi_\ast \). Using this information, we can now reformulate Cont’s framework, in the context of hedging an illiquid call option, as finding the optimal delta that will minimize the maximum amount of potential loss. Formally, we present this formulation with the following \( \min \max \) representation:

\[
\mu_Q = \pi = \inf_{\psi \in \Psi} \sup_{Q \in Q} E^Q[X(\omega)] - \psi E^Q[\omega] \tag{5.2}
\]

In a monetary context, the resulting output should be the smallest amount of money lost if your delta-hedging strategy is misspecified. Also, the \( \psi \in \Psi \) that minimizes the objective function represents the optimal number of units of the underlying required to purchase in order to minimize your potential loss. If we recall from the introductory paragraph when motivating

\[\text{In a continuous setting, the open interval between 0 and 1 is the superset of the set of deltas and therefore allows us to find the optimal lower bound of the problem. If we find a tight optimal upper bound, we can add some merit to the claim that the set of deltas in a continuous setting approaches the open interval between 0 and 1, } \Psi \to [0, 1] \text{ as } n \to \infty.\]

\[\text{This problem is often restated as the minimization of the the absolute difference between a perfect hedge and the resulting imperfect hedge. This framework includes the possibility of earning a profit and incorporates an absolute value in the objective. When we look at the continuous setting, including the absolute value makes the problem difficult to solve using the convex optimization methods described in subsection 2.5. Limiting the problem to minimizing the potential loss removes the absolute value from the objective and makes the problem more tractable computationally.}\]
our formulation in the context of model misspecification, identifying the potential impact on PnL of a hedging strategy is satisfied by calculating $\mu_Q$, and determining an optimal delta to minimize the impact on PnL is satisfied by identifying $\psi \in \Psi$. Therefore, we were able to show that both concerns are addressed using the reformulation of equation 5.2.

We can now use the semi-parametric set of pricing measures, $Q := \{Q : E_Q[M_i(\omega)] = m_i, \ i = i, \ldots, l\}$ to reformulate the problem in a continuous setting. We switch now to the convex representation of Cont’s measure of model uncertainty and use the results of Li and Kwon [45], specifically the equation for calculating the upper bound 2.27. We present the reformulation below:

$$
\mu^* = \pi^* := \inf_{\psi \in \Psi} \sup_{Q \in Q, (q_j)_{j=1,\ldots,k}} \mathbb{E}_Q[X(\omega)] - \psi \mathbb{E}_Q[\omega] - ||h^* - vec((q_j)_{j=1,\ldots,k})||
$$

subject to

$$
\mathbb{E}_Q[M_i(\omega)] = m_i, \ i = 0, 1, \ldots, l
$$

$$
\mathbb{E}_Q[H_j(\omega)] = q_j, \ j = 0, 1, \ldots, k
$$

(5.3)

Using the primal-dual relationships introduced by Li and Kwon (2.28 and 2.30), we can re-write the maximization portion of equation 5.3 into a more tractable form. Note that the dual of the inner maximization is a minimization, which can be joined with the outer minimization problem in the following manner:

$$
\mu^* = \pi^* := \inf_{\lambda,\lambda_h,s,t,\psi \in (0,1)} s + t
$$

subject to

$$
X(\omega) - \sum_{i=0}^{l} \lambda_{m_i}(M_i(\omega) - m_i) - \psi \omega - \sum_{j=1}^{k} \lambda_{h_j}H_j(\omega) \leq s \ \forall \omega \geq 0
$$

$$
\sup_{(q_j)_{j=1,\ldots,k}} \sum_{j=1}^{k} \lambda_{h_j}q_j - ||h^* - vec((q_j)_{j=1,\ldots,k})|| \leq t
$$

$$
0 \leq \psi \leq 1
$$

(5.4)

Equation 5.4 can be solved using the same methods that were utilized when describing the solution of equation 2.32 (Li and Kwon’s extension for solving Cont’s convex risk measure).
Bibliography


[65] G. Welch, G. Bishop An Introduction to the Kalman Filter. UNC-Chapel Hill Transcript, 2006, TR 95-041


A Background Knowledge

This appendix chapter is dedicated towards providing background knowledge for sufficient understanding of the financial instruments that are presented in this thesis. As such, we provide both a description of the mechanics of the financial tools (futures and options) and the technical aspects that serve as the focus for much of the remaining thesis content.

We begin by describing futures from the perspective of finance practitioners along with an explanation of futures market participants and basic futures terminology (A.1). We subsequently explain the difference between a futures and a forward contract using the culminating work of Cox, Ingersoll and Ross [14] as a reference (A.2). Subsequently, the concept of basis is introduced and is described in the context of hedging using a futures contract (A.3). Basis is explored further by introducing basis risk and we explain how this may result in an imperfect hedge when a practitioner uses a futures contract. Furthermore, we explain the idea behind optimal roll conditions to mitigate this risk. The next section describes the interplay of the economic forces that yield the price of a futures commodity contract (A.4). The primary focus is on the interaction between a cash market for spot commodity and a physical market for storage, and how the equilibrium forces of these two markets influence the price of a commodity futures contract. Common shapes of the term structure of commodities will be explained in this context. Lastly, we go over option basics with a focus on terminology (A.5).

A.1 Futures Basics

In general, a futures contract is a contingent claim on an underlying asset that stipulates the acceptance or delivery of a standard quantity and quality of the asset at a pre-specified date and price in the future (many maturities exist and are pre-set by the exchange). The contract is a legal obligation between two parties, the buyer, who is assigned the long position, and the seller, who is assigned the short position, that agree upon a price for future transaction. Note that this implies, by construction, the existence of a long for every short, i.e., a buyer for every seller of a futures contract.

For no-arbitrage conditions (no potential for risk-less profits, i.e., no free lunch) to be met, the
price agreed upon between the two parties is not the expectation of the future spot price but rather a price that is reflective of the current spot price offset by the amount that it costs to store the asset until maturity (slightly more complex, but this point will be elaborated upon in more detail later). The participant who has a long (short) exposure will gain (lose) if the asset appreciates before its maturity date in an equal amount as the loss (gain) observed by the participant with the short (long) position. The payoff function is linear, i.e., if the futures price at time $t$ is $F_t$, and the futures price at maturity is $F_T$ for $T > t$, the long position will have a payoff of $F_T - F_t$, while the short position will have a payoff of $F_t - F_T$. If the contract is held until maturity, the participant with the short position must deliver the asset to the participant with the long position. Delivery of the asset is not common practice however, market participants usually offset their exposure by taking the opposite position thus voiding their obligation to accept or make delivery.

There exist futures written on many asset classes and trade on many exchanges all around the world. Each exchange lists contract specifications for each futures contract that they offer to market participants. Such specifications provide important information concerning the size of the contract (varies from the number of bushels of corn, barrels of oil, pounds of pork bellies, or multiplier of an equity index), the exact quality or grade of what will be accepted or delivered at delivery, the settlement procedures for acceptance and delivery of a physical commodity (financial futures are often cash settled), and the last trade date. We will not go into detail about the different futures exchanges, however, among the largest futures exchange are the Chicago Mercantile Exchange (CME), the Chicago Board of Trade (CBOT), the New York Mercantile Exchange (NYMEX), the London Metals Exchange (LME), the London Interbank Financial Futures Exchange (LIFFE), the International Petroleum Exchange (IPE), the Tokyo Commodity Exchange (TCE), and the Sydney Futures Exchange (SFE). It is of relevance to note that futures exchanges tend to specialize and offer products of a particular sector that is often based on their location, I end this paragraph with a couple examples. The SFE is the only exchange in the world to offer futures contracts on 10-Year Australian Government Bonds, the SPI 200 equity index, and Australian 90-day bank bills. The LME offers the most liquid aluminium, zinc, and nickel forward (more on the difference between forwards and futures later (A.2). The CBOT, due to its location in the American Midwest offer the most liquidity for wheat, soybeans, and corn futures.

In order for a practitioner to participate in a futures market, they are expected to post margin to the exchange upon the initiation of the position. The amount of margin required varies for each contract and is set by the exchange (the exchange has the authority to change the margin requirement whenever they feel it is necessary). Futures exchanges also serve to eliminate counterparty risk by acting as the short for every long position and the long for every short position. This is accomplished by a division within every exchange, called a clearinghouse, who serve the function of a complex matching system that offsets all futures transactions as
they occur. Therefore, every futures market participant only has the clearinghouse as their counterparty instead of some unknown entity. This serves to legitimize futures markets and to increase participation and liquidity in them. Since the clearinghouse is now bearing the risk of counterparty default, exchanges implement a series of rules to mitigate this risk. The most important of these rules is the requirement of daily marking-to-market. Marking-to-market is the practice of settling your daily gain or loss in order to have zero net monetary obligation at the beginning of the next day. Mechanically, if the underlying market appreciates during the day, the short position must pay the long position the exact amount that they have gained and vise versa. However, since the clearinghouse is the counterparty, the short pays the clearinghouse who in turn transfers the money to the holder of the long position. If the short defaults on their payment, the long position is still paid by the clearinghouse and the clearinghouse only loses the daily variation of the contract. The transfer of money based on this practice of marking-to-market is taken out of the initial margin posted when the position was initiated. If the margin in the account drops too low (below the maintenance margin), the clearinghouse requests the participant to post more margin. Furthermore, to limit the maximum amount an exchange may lose upon default of a participant, the exchange often specifies the maximum a contract price may move on any day.

The mechanics of marking-to-market is best explained by an example. We will look at Light Sweet Crude futures that trade on the NYMEX (owned by the CME), which as of the beginning of August, 2012, has an initial margin requirement is $6,548 USD, and a maintenance margin is $4,850 USD. Assume that we have a long position in Light Sweet Crude for delivery in September, 2012 at a price of $91.50 per barrel and the contract specification indicate that the contract size is 1,000 barrels (thus a $0.01 move corresponds to $10 per contract). If the price at the end of the next trading day increases to $92.00, we will receive an increase of $500 ($0.50 increase per barrel on 1,000 barrels) in our margin account for a total of $7,048. This money was given to the clearinghouse by our offsetting short position and transferred to us. We now have cleared the daily price movements of the contract, and are effectively reset to zero, i.e., it doesn’t cost anything to enter the position since daily movements are netted out. Suppose that the next day the price of crude falls to $89.00. This means that we must pay $3,000 to the clearinghouse, who transfers that money to the short position. This will lower our margin account to $4,048, which is lower than the maintenance margin of $4,850. Since it is lower, we are required to post additional margin until we are back to the initial margin requirement of $6,548, which in our case is a posting of $2,500. Assume that we want to exit our position, and we transact a short position in the same contract the following day at $90.00. The short position offsets our initial long position (voiding our obligation to deliver 1,000 barrels of light sweet crude at delivery), and effectively sets an exit price for our initial transaction at $90.00. To complete the transaction however, we receive the $1,000, which credits our margin account to $7,548 before we close the trade and receive the initial margin we deposited back.
Appendix A. Background Knowledge

There are two relevant concepts, open interest and volume, in futures trading that can be explained in the context of the aforementioned example. Open interest and volume are two metrics for measuring liquidity of a specified futures maturity. Volume is incremented for a specific maturity date if there occurs a transaction in that maturity. In our above example, we obtained a long exposure in September, 2012 light sweet crude. This transaction increases the volume in the September 2012 light sweet crude contract by 1. Open interest on the other hand is a measure of the number of offsetting contracts of a specific maturity that must mark-to-market. In our above example, we entered a new long position in light sweet crude that was only possible because there was another participant with an offsetting short position that also entered the market at the same time. This transaction increased open interest in the September contract by 1 because a long and a short position that need to be marked-to-market every night were entered. When we transacted a short position, volume in September increased by 1. Open interest on the other hand depends on the exposure of the participant who took the long position to our offsetting short position. If the other participant was initially short and the long position was used to offset his short position, open interest would decrease because the total number of offsetting mark-to-market contracts decreased. However, if the other participant is participating in the market for the first time, their long position would only take our long position, thus not affect open interest. Practitioners use volume and open interest (more open interest) to determine what maturity offers the most liquidity if they are interested in gaining price exposure (long or short) to an asset.

Before we move on to the next section, I want to briefly touch upon the different type of futures market participants: market makers, speculators, and hedgers. Each participant provides a specific utility to the marketplace and all are important in the construction of futures markets. A market maker is responsible for providing liquidity to a contract on either the short or long side if buying or selling pressure occurs. Furthermore, they are usually compensated for this service by receiving a pre-defined fee, set by each exchange, for each transaction they make. It is important to note that these participants are not primarily interested in generating return from the underlying price movements of the contract but rather from the exchange rebates they receive for participation. Hedgers are participants whose sole purpose is to use futures contracts to offset the price risk of a risk exposure they have, thus they usually own or will receive an asset with an offsetting payoff to that of the futures contract, i.e., a decrease in value of their asset will increase the value of the futures contract. The existence of hedgers is the the primary reason for the development of sophisticated futures markets and the mechanics of hedging and the risks involved will be discussed in greater detail in a subsequent section (A.3). Speculators are participants interested in gaining price exposure (either long or short) to an asset for the purpose of earning a profit. These profits are manifested in one of two ways, proper anticipation of spot movements in the asset, or from capturing mechanical features of the term structure. Furthermore, to tie in their relation with hedgers, speculators can be seen as accepting the price risk that hedgers are offsetting for a premium (the existence of such a
hedging risk premium has been debated, empirical studies relating to a hedging risk premium were covered in 3.1.1).

A.2 Forward VS Futures Contracts

Both futures and forwards are financial instruments created for market participants that are interested in locking in a future price of an asset now. As such, the description of a futures contract, as written in the previous section (A.1), is in principle the same as a forward contract, however, differences must exist otherwise a market place would not exist for both instruments. These differences are mechanical in nature, and elicit differences in how the fair value of each contract is determined. What follows is an overview of these differences, proceeded by a description of how the fair market price is determined from a non-technical perspective.

A forward contract is a non-standardized agreement between two unknown counter-parties while a futures contract is a standardized agreement between a counter-party and the clearinghouse of an exchange. A forward contract trades in the Over The Counter (OTC) market and does not operate under the strict regulations of the National Futures Association (NFA) and the Commodity Futures Trading Commission (CFTC). Thus, the holder of a forward contract is not expected to mark-to-market the daily price fluctuations with the counter-party holding the opposite exposure, whereas this is a requirement for a futures market participant, and a forward contract is not limited to a maximum daily price variation. Not having to mark-to-market daily price variations implies that the forward contract holder must accept the risk that the counter-party may not be able to make the necessary payment at the expiration of the contract (This default risk is assumed to be small because those participating in OTC markets are often large investors with appropriate risk controls.). The practice of marking-to-market at the end of every trading day essentially eliminates counter-party default risk. Furthermore, since the forward contract is non-standardized, a forward contract may not have sufficient liquidity if the participant wishes to offset his exposure before the end date as stipulated in the contract. The standardization of contract specifications coupled with the virtual elimination of default risk has made futures contracts more appealing to some investors, however, both are actively traded.

Before we begin discussing how to obtain the fair market price of a forward and a futures contract with maturity $T$ at time $T_0$ where $T > T_0$, we want to ensure that the difference between a contracts price and its value are clear. We then use this information to construct a series of rational arguments to understand the logic behind a futures or forward contracts fair market price. The price of a forward or futures contract is the amount the underlying asset would cost today for delivery on the maturity date. The value of the forward or futures contract is the difference in price between the initial price of the contract when it was initiated in the
past and a contract with the exact same specifications initiated today. Thus, the price of the forward or futures contract will change every day based on supply and demand arguments of the underlying market whereas the value of a forward contract depends on the variation of an old contracts initiation price and its current price.

Implicit in this discussion of the difference between the price and value of a futures or forward contract is that both the forward and futures price must converge to the prevailing spot price of the asset at maturity. Furthermore, no-arbitrage conditions force the price of each contract to be the underlying spot price adjusted by the cost of carrying the asset until maturity (this can be negative). The existence of no-arbitrage forces the value of a forward or futures contract upon initiation to be zero, i.e., no cash is exchanged between counter-parties.

Now if we assume that we can determine the prevailing daily spot price, we can compute a price for the contract every day. For the sake of being pedantic, let's assume that time equals the maturity date \((t = T)\), thus the price of the contract is the spot price, and the value is zero. If we go back one day, the price of the contract is now yesterday's spot price adjusted by the cost of carry and the value of the contract is zero. Now, if we go back \(N\) days where \(N < T - T_0\), the price is the now the spot price from \(N\) days ago adjusted by the cost of carry, and the value of the contract is zero. To tie this all together, the fair market price of a futures or forward contract at \(T_0\) is the price that ensures that all remaining payment streams have an initial value of zero, or alternatively, the fair market price must vary through time in such a way that all newly created contracts will always have a value of zero when initiated [14]. In order to understand how this is related to the initial price of a forward and a futures contract, recall the practice of marking-to-market using futures. This procedure implies that if the settlement price of the asset today is different from the last settlement price, the party that has benefited from the price change must immediately be paid by the party that has not. In contrast, no cash is exchanged in a forward agreement until the maturity date. The summation of the cash flows from the daily marking-to-market of futures is equal to the final payment of the forward contract, therefore the different payoff streams result in an identical net payoff. Since we know that all remaining payment streams must have a value of zero by construction, and we know that all cash flows must be discounted at the prevailing interest rate, we conclude that unless the interest rate is constant during the lifetime of the contract, the initial price of a futures and forward contract must be different.

The implication of this is that if the underlying asset is positively correlated to rising interest rates, the futures contract price is greater than the forward contract price. To see this, assume all daily cash flows received, whether positive or negative, are re-invested at the prevailing interest rate. As such, in a rising rate environment, the average discount rate applied to the net payoff will be lower for a futures contract as opposed to a forward contract, and the opposite is true in a declining rate environment. Therefore, if you are trading an asset that is positively correlated to interest rates and you are in an increasing interest rate environment, it is more
likely that the assets price is increasing and you will be discounting these cash flows at a lower average rate, thus, making the futures price higher than the forward (Another way of looking at this particular example is that you can participate in the rising rate environment by investing your cash flow at a higher interest rate than the initial opportunity cost of capital.). Similar, but opposing, arguments can be made to assets with negative correlation to interest rates.

A.3 Hedging & Basis Risk

Hedging is the practice of trying to eliminate the price risk of an asset by taking an exposure in an instrument that will completely offset any gains or losses experienced by the asset. Such practice is common for consumers wishing to lock in a price to buy an asset (long hedge) or suppliers wishing to lock in a price to sell their asset (short hedge). The existence of futures markets was built on providing a marketplace for hedging, and they were constructed with this utility in mind.

Before we begin, the concept of basis will be introduced along with some terminology, we then continue with explaining why the idea of basis risk may lead to an imperfect hedge. Simply, the basis is the difference in price between two instruments that have the same underlying but different maturity dates. The most common is the difference between the most liquid futures contract (often the first contract) and the spot price. Theoretically, the futures price will converge to the spot price as it reaches maturity and basis will therefore equal zero at maturity. Such a situation is ideal in a hedge because you are using a derivative (futures) to hedge your spot exposure, thus you want as many spot movements to be reflected in the futures (idea of negatively correlated payoff streams).

If the futures price is greater than the spot (positive basis), this market is said to be in contango and the futures price approaches the spot price at maturity from above. Alternatively, if the futures price is lower then the spot price (negative basis), this market is said to be in backwardation and the futures price approaches the spot price at maturity from below. This type of behaviour implies that if a market is expected to have a stable term structure, a trader would profit from being short a market that is in contango or from being long a market that is in backwardation. The amount of curvature in the term structure is dynamic and is a function of current market variables. This will be discussed further in the next section (A.4).

Basis risk is the risk that the instrument you are using to hedge your spot exposure is not perfectly negatively correlated. This can result in many different scenarios, but we will focus on two such scenarios that are specific to using futures. The most common scenario is that a hedger will choose a hedging instrument that has some mechanical aspect that differs from the asset they are expected to buy or sell. Examples of differing mechanical aspects include dissimilar payoff maturities, the quality of the hedging instrument is not identical (different crop
years or grades of gasoline), or the asset payoff that we want hedged does not have a liquidly traded futures contract. In the later case, the hedger often uses some combination of other similar products that have historically provided some level of negative correlation (a common textbook example is hedging jet fuel with a combination of heating oil and crude oil).

The other scenario of basis risk that we present is arguably more interesting, and it stems from the fact that even if you are hedging using a futures contract that has the exact same characteristics as the asset you want hedged, a futures contract sometimes expires at a premium or discount to the spot price. This phenomena may only occur in physicals because cash settled futures always converge to the spot price, expiration at a discount or premium is often due to structural aspects of delivering the commodity. As an example, if the exchange mandated warehouses are all full, and you are expected to accept delivery from another spare warehouse that has different price contingencies, the futures would trade at a discount to spot to the amount of the expected additional warehousing and delivery costs. Another interesting example is when there are only a few participants remaining in the nearest futures contract and it is approaching expiration. For simplicity, let us assume a hypothetical situation where only two participants remain in a market near its maturity, a long and a corresponding short. The long knows that they can guarantee the spot rate if they wait until expiration by accepting delivery and then immediately selling it at the spot rate, therefore they may submit an offer to sell above the prevailing spot rate. The short may not want to exit their position at the elevated price but will submit a bid at a price nearer to the spot rate. If the market expires without any trade taking place, the short, if they do not already own the commodity, will purchase the asset in the spot market and delivery it to the long thus effectively converging the futures price to the spot. However, this does not happen until after the futures contract has expired thus forcing the exchange to submit a final settlement price for the futures that is above the spot price. Note that most traders with a short position in such a situation would often not wait until expiration and would accept the offer even if it is above the spot price to avoid having to make physical delivery.

A.4 Economic Basics

The purpose of this section is to elicit some economical arguments toward the dynamics of commodity markets in order to give the reader some intuition toward the reason that theoretical models are constructed the way they are. In order to do this, we construct arguments behind the motivations and actions of producers and consumers of storable commodities. We then show that the dynamics of commodity spot and futures prices is the result of the interaction between two markets, a cash market for spot purchases and sales of the commodity, and a market for storage of a commodity. This is accomplished by explaining the details of each market followed by examples of how these markets interact with one another to elicit price
movements in the underlying commodity. Lastly, we explain how this is all related to futures markets. Much of the content and notation utilized in this section is summarized from the work of Robert Pindyck [53], who provides a series of logical arguments towards describing empirical observations of commodity price dynamics, specifically those of oil and gas markets.

A producer of a commodity is often most concerned with earning a stable profit from a pre-determined production schedule. This production schedule is the sum of the expected production levels for the period in question and the level of inventories the producer has in storage. It becomes clear that any unexpected short-run price dynamics or increase in market volatility may disrupt the production schedule of a producer (increase demand for hedging). If a short-run price shock does occur, a producer can maintain scheduled deliveries and operations if sufficient inventories are available in storage (buffer effect). If a shock persists, the diminishing levels of inventory will diminish its utility as a buffer and market prices should rise (this is exactly what happens). This exogenous shock of market variables implies that the value of inventory, from the perspective of a producer, is dynamic, and an intelligent producer should change his production schedule accordingly. It is important to note that a producer can reduce their costs over time by selling out of inventory during periods of low inventory/high demand, and increase their depleted stock during periods of high inventory/low demand. Furthermore, the level of inventories is finite, and can never be negative. All of the aforementioned arguments are similar for a consumer, and we refer you to Pindyck [53].

The cash market is for purchase or sale of a commodity for immediate delivery, and transacts at the prevailing spot price. The spot price will dictate the demand (consumption) and supply (production) of a commodity according to normal market forces (i.e supply is positively correlated and is an upward sloping function with respect to price and demand is negatively correlated and a downward sloping function with respect to price). We write the demand and supply functions as:

\[ Q = Q(P; z_1, \epsilon_1) \]
\[ X = X(P; z_2, \epsilon_2) \]  \(5\)

where \( P \) denotes the spot price, \( z_1 \) and \( z_2 \) are respectively vectors of demand and supply shifting variables, and \( \epsilon_1 \) and \( \epsilon_2 \) are random shocks.

Due to the existence of dynamic inventory levels, the spot price does not equate consumption and production. The cash market is therefore characterized as the relationship between the spot price and the changes in inventory that result from the changing consumption and production schedules. Pyndick [53] refers such a change inventory, the difference between production and consumption, as net demand, or rather the demand for production in excess of consumption. By denoting \( N_t \) to be the level of inventory at time \( t \), we can denote net demand at time \( t \) to be \( \Delta N_t \). Referring to \( 5 \), \( \Delta N_t \) can be represented as:

\[ \Delta N_t = X(P_t; z_{2t}, \epsilon_{2t}) - Q(P_t; z_{1t}, \epsilon_{1t}) \]  \(6\)
Appendix A. Background Knowledge

This equation implies that the cash market is in equilibrium when \( \Delta N_t = 0 \), i.e., when total demand is equal to supply or when net demand is zero. The following inverse net demand function isolates the spot price to yield the fundamental relationship of the cash market:

\[
P_t = f(\Delta N_t; z_{1t}, z_{2t}, \epsilon_t)
\]  

The inverse net demand function is upward sloping because an increase in spot price will increase the supply of the commodity and decrease the demand, thus ultimately increasing the level of inventories.

The market for storage is a market for owning stock or inventory of a commodity. As with any market, the price of storage is the marginal value of the flow of services that accrue when holding a marginal unit of a good, which in this case is inventory. The flow of services in the market for storage is attributed to the value of being able to maintain a production or consumption schedule during supply disruptions, as such, the holders of inventory are expected to make a payment for the privilege of owning a marginal unit of storage. The value of the flow of services accruing from holding the marginal unit of inventory is termed the marginal convenience yield, denoted by \( \psi \), and the demand for storage function can be written as \( N(\psi) \). Note that the demand for storage can be thought of as the amount of inventory required given the economic environment for the commodity. The relationship between the marginal convenience yield and the demand for storage is the source of the equilibrium forces in this market, and we discuss each one in turn below.

Marginal convenience yield \( \psi \) must be paid by the holder of inventory, and it can be thought of as consisting of three components. The first is the cost associated with physical storage such as the cost of renting a warehouse to store the commodity. Second, is the foregone interest earned on the capital that has been committed to purchasing the commodity (opportunity cost of capital). These first two components are often jointly referred to as the cost of carry. The final component is the expected depreciation of the commodities spot price over the period of which the commodity is kept in inventory. Intuitively, the expected depreciation is often high during supply shocks because market participants associate a higher value with owning the commodity today (since there is less of it) than in the future. It is important to note that the holder of the inventory does not physically pay the expected depreciation of a commodity in inventory as they would a conventional cost. By not selling the inventory when its expected depreciation is high, a holder of the commodity is effectively increasing their opportunity cost of capital because they are not participating in the rising flow of services associated with holding the commodity. Usually, high expected depreciation is value that has accrued to the holder of the commodity, and they want to realize this value by selling their inventory (note expected depreciation is usually separated from the cost of carry in the calculation of marginal convenience yield, and is often called the convenience yield).
Pyndick [53] explains that the demand for storage is not only influenced by the marginal convenience yield, but also by the expected rates of consumption and production, the spot price of the commodity, and the volatility of the spot price. Consumption and production rates are subject to seasonal variation in weather or sudden supply or demand shocks, and the occurrence of these events are likely to increase the demand for storage because having inventory will buffer unexpected costs with the spike in demand. The spot price of a commodity is a factor because a holder of inventory should be willing to pay more for storage of a commodity or good that is worth more. Lastly, high volatility in the spot price of a market should increase the demand for storage because the wild swings in prices make maintaining production or consumption schedules more difficult, therefore, participants will want to buffer the effects of these wild swings in prices by increasing their inventory. These factors can be written as a demand for storage function

\[ N(\psi; \sigma, z_3, \epsilon_3) \]

but more importantly, the following inverse demand function of the demand for storage yields the fundamental relationship in the market for storage:

\[ \psi = g(N; \sigma, z_3, \epsilon_3) \]  

where \( \sigma \) refers to the volatility of the commodity spot price (note storage is in greater demand the more volatile the commodity spot price is), \( z_3 \) is a vector of demand shifting variables that include the spot price, production, and consumption, and \( \epsilon_3 \) is a random shock. The inverse demand for storage function is a downward sloping function with respect to the demand for storage, i.e., the demand for more inventory given a current level of inventory. This is intuitive because the lower the level of inventory, the lower its utility as a buffer to adverse short-term price dynamics will be. This implies a higher value associated with holding inventory when there is less of it, which is associated with a higher cost of storage (just to embed this again, cost of storage is a part of the marginal convenience yield).

To clarify the dynamics of the cash market and the market for storage, we consider the effect of cold weather increasing the demand for consumption of heating oil (example used in Pyndick [53]). We begin by outlining our two assumptions of the market environment immediately prior to the onset of the colder weather. First, we assume that the market expects the cold weather to be temporary. Second, we assume that the market is in steady-state equilibrium, which implies consumption equals production or net demand is zero \((\Delta N = 0, \text{ the immediate change in inventory is zero})\). Recall the cash market relationship from equation (7), abbreviated as \( f(\Delta N) \). The increase in demand for consumption will shift the \( f(\Delta N) \) curve upwards, thus resulting in a higher spot price. Since this upward shift is expected to be temporary, current inventory levels will be drawn thus dampening the increase in spot price and setting \( \Delta N < 0 \). Now recall the market for storage relationship from equation (8), abbreviated as \( g(N) \). The drawing down of inventories will increase the price of storage, i.e., the convenience yield, \( \psi \), will increase as we travel up the curve of \( g(N) \) as the level of inventories continues to decrease. As warmer weather returns, we expect \( f(\Delta N) \) to shift back down to its original level. However,
in order to get back to equilibrium ($\Delta N = 0$), the inventory has to re-accumulate the amount that it was drawn down. Thus, the new spot price will be slightly higher until equilibrium is reached because $\Delta N > 0$ until the level of inventory is reset to its original amount. During this time, the convenience yield will travel back down $g(N)$ to its original level as the level of inventory gets replenished to its original amount.

We now conclude by relating the futures price to our discussion on the cash market and the market for storage (notation used in Pindyck [53]). Before we accomplish this however, we first recall the components of the marginal convenience yield, they are the physical cost of storing the commodity, foregone interest on capital used to purchase commodity, and the expected depreciation in spot price. Let us denote the marginal convenience yield accrued over the time period from $t$ to $t + T$, as $\psi_{t,T}$. Furthermore, let us denote the spot price as $P_t$, the risk free interest rate earned over the period $T$ as $r_T$, the per unit cost of physical storage for the time period $T$ as $k_T$, and the current futures price for delivery at time $T$ as $F_{t,T}$. We now rewrite the three components of the convenience yield. The first two are simply $k_T$ and $r_T P_t$. However, to understand the expected depreciation in spot price, we recall the idea of futures price convergence to the spot price at maturity. As such, the futures price $F_{t,T}$ can be regarded as being a good estimate of $P_T$, or the difference between the futures and current spot price can be seen as the expected depreciation in spot price. Bringing this together, we get $\psi_{t,T} = k_T + r_T P_t + (P_t - F_{t,T})$, or re-arranged, we get the following relationship for a futures contract:

$$ F_{t,T} = (1 + r_T) P_t + k_T - \psi_{t,T} \tag{9} $$

Note the return from holding a unit of the commodity is $\psi_{t,T} - k_T + (P_{t,T} - P_t)$ (recall that we have referred the expected depreciation as a payment, however, this is not the case if the value of the inventory appreciates while it is in storage and you sell the inventory at an elevated price). The holder of a futures contract would not expect the same payoff since the stipulation of a futures contract is for future delivery, therefore all the benefit of owning the inventory go to the owner of the inventory as opposed to the owner of the futures contract. This is why the value of the convenience yield is subtracted from the the capital appreciation of the spot price in equation (9). Note that if $\psi_{t,T} - k_T$ is positive, we will have backwardation, i.e spot price higher than futures price, and if the opposite is true, we will have contango, i.e., the futures price is higher than the spot price. For a resource commodity such as oil, it is common for the it to exhibit contango in the short term and backwardation in the long term. This is because most of the value of having a reserve is associated with having the ability to begin production when it is profitable to (almost like an option premium). The sum of these convenient effects often make the convenience yield of these markets high, and therefore they are often in backwardation. However, once the crude is extracted, the convenience yield drops, thus near term crude is often in contango (not a rule, but rather a description of an observation that has occurred).
Appendix A. Background Knowledge

A.5 Option Basics

In general, an option contract is a contingent claim on an underlying asset that stipulates a future payoff that is a function of the underlying assets price at maturity or at any time along the assets price path. In its most basic form, an option is an agreement between two parties, the buyer, who is assigned the long position, and the seller, who is assigned the short position, that agree upon a price for future transaction and gives the long position the right to transact at this price if they see fit. As you may recall, a futures contract stipulates an obligation to transact at maturity whereas an option stipulates the right to transact at maturity (or different times depending on the covenants of the contract), which usually occurs if the payoff of the option is non-zero (in the money). This obligation for future transaction implies one possible futures price for a given a maturity, whereas a participant in the options market may pick from a series of future prices (called the strike price) when entering into an options contract for a given maturity and has the right to transact at this price. Option payoff functions vary tremendously, however, those that are most often traded at an exchange have been categorized as vanilla options, aptly named for their more simplistic structure. We will describe the mechanics of a subset of these vanilla options, specifically European options.

A European option gives the participant with the long position the right to transact an asset at a pre-specified price, called the strike price, only at the maturity of the contract. This right is not free, and the buyer of such a contract must pay a premium to the seller for this right. If the long participant wants the right to buy the asset in the future, he enters into a call option agreement, and if the long participant wants the right to sell the asset in the future, he enters into a put option agreement. Since these contracts only stipulate the right to transact, the long position will only exercise (buy or sell the underlying asset) his option if it is profitable. Thus, the payoff function for a European call option at maturity given a strike price \( X \) and a future asset price of \( F_T \) is \( \max(F_T - X, 0) \), similarly, for a European put option with the same conditions, the payoff function is \( \max(X - F_T, 0) \). To better understand these payoff functions, consider a participant with a long call option (the holder of the option) with strike price \( X \). The holder of the long call will only buy the asset at \( X \) at maturity if the assets spot price at the contracts maturity is greater than \( X \) because it would not be to the benefit of the option holder to buy an asset at \( X \) when the spot price is less then \( X \), thus a payoff function of \( \max(F_T - X, 0) \). Similar arguments can be made for the payoff of a put option. It is important to note that mechanically, it is more intuitive to think that an actual transaction of the asset takes place when the option expires (reaches its maturity). However, what actually happens is that the price of the option converges to the options payoff near maturity, thus, similar to futures markets, option market participants usually offset their long position with a corresponding short (with the same maturity and strike).

The cost of an option that the long position must pay the short position, called the options
premium, must oblige to no arbitrage arguments upon initiation. To understand how the market
determines the value of this premium, we will walk through a series of intuitive arguments
to discuss the relevance of different inputs that ultimately result in determining the option
premium. An European option holder will only get paid (option expires in the money) if the
price of the underlying at maturity is higher than the strike price in the case of a call option
or if the price is lower than the strike price in the case of a put option. Ultimately, this implies
that the price of an option upon initiation of the contract depends on the likelihood that the
price of the underlying asset at maturity will confine to either of the aforementioned conditions
(depending on whether it is a call or put), and the more likely it does is directly proportional
to the options premium. This implies that if it is more likely that the asset price will be in
the money at the maturity of the contract, it will cost more to lock in the price and vise versa.
This relationship (for no arbitrage conditions to hold) is such that the premium paid for the
option is the average of all possible payoffs that are a result of all possible price paths that the
underlying asset can have.

The main two factors that influence the likelihood of an asset finishing in the money at maturity
are the expected volatility of the underlying assets price and the length of time until maturity.
If the expected price volatility of an asset is high, the likelihood of an option expiring in the
money is higher, similarly, the longer the maturity of the contract, the more time the asset has to
eventually end up in the money. Thus, both price volatility and length of maturity are positively
correlated to option premium. Another factor that does not influence the option premium as
much as price volatility and length of maturity is the expected interest rate. Expected interest
rates are negatively correlated to the option premium of a put option and positively correlated
to the option premium of call option because of the opportunity cost associated with the capital
required to garner the same exposure in the underlying as in the option. Participating in an
option market is a leveraged position since the premium paid to participate is much lower
than purchasing or selling short the underlying. Thus, the monetary difference between the
underlying asset and the premium paid can be invested at the prevailing interest rate. A call
option holder can invest the difference in underlying price and premium paid and would benefit
from an increase in interest rates thus increasing the initial price of the option. In contrast,
the holder of a put option is seen as equivalent to selling the underlying short. Since selling
the underlying short gives you capital that you can invest that the put option holder does not
have, an increase in rates is to the detriment of the option holder because they do not have
the capital to participate in its appreciation. It is important to note that most option pricing
models assume constant interest rates over the maturity of the contract for short term options
because the stochasticity of interest rates is low enough to have little influence in the option
premium.
B Kalman Filter

Each of the chosen models we describe attempt to explain the evolution of futures prices by extracting the unobserved factors that are driving the price process. Each of these factors has a degree of randomness that is assumed to be Gaussian. By construction, these factors are implicit in the models solution of the futures price and can be represented in linear state-space form via a discretization process of the joint-diffusion stochastic processes (one for each factor / state-space variable). Such a state-space representation of Gaussian-affine models allows for the utilization of the Kalman filter for extracting the unobserved factors driving the futures prices. We explain the Kalman filter in what follows using the notation found in Welch and Bishop’s introductory notes on the subject [65] (for an in depth look at the Kalman filter, we refer you to Kalman’s seminal paper [41] and Maybeck’s often cited textbook [48]).

The Kalman filter is a predictor-corrector algorithm. At each time step k, the most likely value of the state variables is predicated contingent on all previous estimated state values. This prediction is subsequently corrected by taking a measurement of a process that has a known relationship to the unobserved state variables. The correction of the state variables is calculated in a way to maximize the likelihood of the conditional distribution (which coincides with minimizing the variance of the corrected state variable).

Before we continue with the computational aspects of the filter, we assume that a system can be described by the following linear stochastic difference equations

\[ x_k = Ax_{k-1} + Bu_k + w_{k-1} \]
\[ z_k = H x_k + v_k \]  

(10)

- \( x_k \) are the state variables governed by an unobserved process
- \( z_k \) are the observed measurements of a process that have a dependency on / are driven by the state variables \( x_k \)
- \( A \) is parameterization of unobserved variables with respect to all previous unobserved variables

\(^6\)For our discussion, we will assume that matrices \( A, B, \) and \( H \) are constant, however, this is not a requirement.
- $u_k$ is optional control input\footnote{There is no control input in any of the models that we will specify in subsequent sections, however, there are constants that are carried through in the algorithm embedded in the $B$ matrix by setting $u_k$ to a $n \times n$ identify matrix.}
- $H$ is parameterization of observed variables with respect to the unobserved variables
- $w_{k-1}$ and $v_k$ are process and measurement noise respectively and are assumed to be normally distributed with mean 0 and covariance $Q$ and $R$ ($P(w) \sim N(0, Q), P(v) \sim N(0, R)$)

As explained, the Kalman filter is a predictor-corrector algorithm that attempts to calculate the most statistically likely estimate of an unobservable process. The algorithm accomplishes this by calculating an initial \textit{a priori} estimate at every time step $k$, denoted $\hat{x}_{k-1}$, and subsequently updates this prediction by making an \textit{a posteriori} estimate, denoted $\hat{x}_k$, upon making a measurement of observed process at time step $k$. We can now define our \textit{a priori} and \textit{a posteriori} estimation errors at each time step of the algorithm as $e_{k-1} = x_k - \hat{x}_{k-1}$ and $e_k = x_k - \hat{x}_k$ respectively, where, again, $x_k$ is the actual value of the state variable. Furthermore, we can also define our \textit{a priori} and \textit{a posteriori} error covariance at each time step of the algorithm as $P_{k-1} = \mathbb{E}[e_{k-1} e_{k-1}^T]$ and $P_k = \mathbb{E}[e_k e_k^T]$.

Based on the above notation, it is logical to assume that $\hat{x}_{k-1}$ and $\hat{x}_k$ are related since they are both estimates of the same process and only differ in that they are calculated subject to the available information at that time. Furthermore, it can be deduced that $\hat{x}_k = f(\hat{x}_{k-1})$ by realizing that $\hat{x}_{k-1}$ is the first estimation of the unobserved process $x_k$, which is then updated to the final estimate $\hat{x}_k$ after observing $z_k$. Now if we assume that the measurement process is completely deterministic, we have $z_k = Hx_k$. Continuing with this assumption, we note that our best estimate of $z_k$ prior to actually observing it is $\hat{z}_k = H\hat{x}_{k-1}$. The difference between $z_k$ and $\hat{z}_k$ is called the measurement innovation or residual. We can now express the \textit{a posteriori} estimate as a linear combination of the \textit{a priori} estimate and the measurement residual:

$$\hat{x}_k = \hat{x}_{k-1} + K_k(z_k - H\hat{x}_{k-1})$$ \hspace{1cm} (11)

The matrix $K_k$ is called the \textit{gain} matrix. To make the Kalman filter algorithm efficient towards identifying the most likely value of the state process, the value $K_k$ must be chosen to minimize the \textit{a posteriori} error covariance, $P_k$. We will not derive this value\footnote{Refer to [48] for complete derivation of gain matrix.}, however, we present the final result:

$$K_k = \frac{P_{k-1} H^T}{HP_{k-1} H^T + R}$$ \hspace{1cm} (12)

It is interesting to evaluate the limits of the \textit{gain} matrix subject to changes in the measurement procedure. If we assume that we are completely confident in the accuracy of the measurement procedure, we can intuitively state that the \textit{residual} is more important in calculating the \textit{a posteriori} estimate of the state process. Mathematically, assuming complete confidence in the
Appendix B. Kalman Filter

accuracy of the measurement procedure is saying that the covariance of the measurement noise, \( R \), is zero. By calculating the limit of the gain matrix as \( R \) approaches zero, we see that \( K_k \) approaches \( \frac{1}{H} \) and, as expected, gives more weighting to the residual. Conversely, if we have absolutely no confidence in the accuracy of the measurement procedure, we can intuitively state that the \textit{a priori} estimate is more important in calculating the \textit{a posteriori} estimate of the state process. By calculating the limit of the gain matrix as \( R \) approaches infinity, we see that \( K_k \) approaches zero and, as expected, doesn’t add any weight to the information in the residual (Note that we get a similar result if we assume the \textit{a priori} covariance is zero, i.e., the estimate is perfect).

We now have the necessary background information and intuition to develop the Kalman filter algorithm. As explained, the filter makes a prediction of the state variable contingent to all previously estimated values of the process. This procedure is called the \textit{time update equations} and is responsible for projecting forward the state equations mean and covariance. This estimate is then corrected when a measurement of a process that is driven by the unobserved process is obtained. This procedure is called the \textit{measurement update equations} and is responsible for updating the state process estimate with the information from the measured process.

We conclude this introduction to the Kalman filter with the final resulting equations of the algorithm.

\textit{Time Update Equations (Predictor)}

\[
\begin{align*}
\hat{x}_{k-1}^- &= A\hat{x}_{k-1}^- + Bu_{k-1}^- \\
P_{k-1}^- &= AP_{k-1}^-A^T + Q \\
\end{align*}
\]  \hspace{1cm} (13)

\textit{Measurement Update Equations (Corrector)}

\[
\begin{align*}
K_k &= P_{k}^-H^T(HP_{k}^-H^T + R)^{-1} \\
\hat{x}_k &= \hat{x}_{k-1}^- + K_k(z_k - H\hat{x}_{k}^-) \\
P_k &= (I - K_kH)P_{k}^- \\
\end{align*}
\]  \hspace{1cm} (14)
C Delta-Hedging Basics

In what follows, we will be describing the idea behind delta-hedging. In this context, the concepts of contingent claim replication, no-arbitrage condition, and self-financing strategies naturally arise. We use the Binomial model to introduce these topics due to its simplicity before moving to the Black-Scholes-Merton continuous time model [7, 49].

In order to set up the Binomial model, we assume that the market consists of three assets: a stock ($S$), a money market instrument earning interest at $r$, and a contingent claim on the stock ($X$) that expires at time $T$. For illustrative purposes, we use a European call option with strike price $k$ written on the stock as the contingent claim in our example, but the framework is applicable to any contingent claim. We only need to model the possible outcomes of the stock because the money market instrument is risk-free and is guaranteed to earn $r$ over the period, and the value of the call option is contingent to the outcomes of the stock at expiration. The Binomial model assumes the price of the stock after a time period $T$ has two possible outcomes ($\Omega = \{\omega_1, \omega_2\}$), i.e., $S(T, \omega_1) = s_1$ and $S(T, \omega_2) = s_2$. Without loss of generality, we also assume that $s_1$ is larger than the current price ($s_0$), $s_2$ is lower than $s_0$, and the strike price of the call option lies between $s_1$ and $s_0$.

We posit that there exists a portfolio that holds $a$ units of stock and $b$ units of the money market instrument that perfectly replicates the possible values of the contingent claim at expiration. Such a portfolio is considered the replication portfolio, denoted by $H$, and since all monetary outcomes of holding the portfolio or the contingent claim are identical, the cost to set up the position must be the same according to the no-arbitrage condition. This implies that if the price of the contingent claim in the market ($x_o$) is different from the cost of initiating the replicating portfolio, a market participant can sell the more expensive option to buy the cheaper alternative and earn a risk free profit when the positions unwind at time $T$. The described portfolio, (units of the claim, money market instrument and stock) is an instance of a self-financing strategy that obliges to the no-arbitrage condition. Informally, a self-financing strategy is a strategy where no additional resources can be allocated to the portfolio after the initial investment. Any changes in the quantity of the portfolio holdings must come from buying or selling other holdings within the portfolio. Now, no-arbitrage condition states that any self-financing strategy can not have a positive expectation of profit if the cost of setting up the initial portfolio holdings is free (or
If we denote the potential payoffs of the contingent claim $X$ as $X(T, \omega_1) = x_1$ and $X(T, \omega_2) = x_2$, we have the following relationship:

\[
\begin{align*}
H(T, \omega_1) &= as_1 + b(1 + r) = x_1 \\
H(T, \omega_2) &= as_2 + b(1 + r) = x_2
\end{align*}
\]

This 2X2 linear system has a unique solution for $a$ and $b$:

\[
a = \frac{x_1 - x_2}{s_1 - s_2} \quad b = (1 + r)^{-1} \left( x_1 - s_1 \frac{x_1 - x_2}{s_1 - s_2} \right)
\]

The replication portfolio, $H_0 = as_0 + b$, is also called the hedging portfolio. If we look at the amount of stock purchased, $a$, you notice that it is calculated by looking at the range of values the contingent claim can have relative to the values of the underlying stock, i.e., $a = \frac{\Delta X}{\Delta S}$. This implies that the amount of stock purchased $a$ is actually an amount that immunizes the sensitivity of the price of the claim relative to the price of the underlying. Option traders refer to this value as the options delta. Using this interpretation and assuming the Binomial model can have multiple periods, the delta of the option must change to reflect the information that is gained, i.e., we now know the new price of the instruments in the marketplace, and we must re-establish our hedge based on the new set of potential future values that the next period will have.

Before moving on to the continuous time case, we note that the probability of the stock making an up-move to $s_1$ relative to a down move $s_2$ is not relevant towards calculating the value of the contingent claim. Recall that the goal of the replicating portfolio is to replicate all possible future outcomes regardless of how probable or not they are. This implies that the price of the contingent claim is only dependent on the variation in possible prices of the underlying stock as opposed to the expected profit of the underlying stock.

In a continuous time setting, we assume the stock price follows geometric Brownian motion, $dS(t) = \alpha S(t) dt + \sigma S(t) dB(t)$ where $\alpha$ and $\sigma$ are the drift and volatility of the stock respectively, and $dB(t)$ represents an increment of a Brownian motion. We also assume that the money market instrument satisfies $dR(t) = r R(t) dt$, where $r$ is the continuously compounded rate of return. Next we let the hedging/replicating strategy contain $a^H(t)$ units of stock and $b^H(t)$ units of the money market at any time $t$. The price $P(t)$ of the claim at time $t$ is a function of the stock price. We denote this function by $P(t) = C(t, S(t))$ for a function $C(t, x)$ and note that the price of the claim $P(t)$ must equal the value of the hedging portfolio $H(t)$ at all times.

If we assume that the function $C(t, x)$ can be differentiated twice with respect to $x$ and once with respect to $t$, we can use Itô’s Formula to represent the dynamics of the value of the
Appendix C. Delta-Hedging Basics

contingent claim as

\[
dP(t) = \left( \frac{\partial C(t, S(t))}{\partial t} + \alpha \frac{\partial C(t, S(t))}{\partial x} S(t) + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 C(t, S(t))}{\partial x^2} S^2(t) \right) dt + \sigma \frac{\partial C(t, S(t))}{\partial x} S(t) dB(t)
\]

(15)

Now, by using the self-financing property of the joint portfolio containing the contingent claim
and hedge portfolio, we know that any changes in the value of the hedging portfolio must equal
the changes in the price of the claim. Stating this mathematically, we get

\[
dP(t) = dH(t) = aH(t)dS(t) + bH(t)dR(t)
\]

(16)

Now comparing equations 15 and 16 and using the relation \(C(t, S(t)) = P(t) = H(t) = aH(t)S(t) + bH(t)R(t)\), we can get the replicating strategy for the contingent claim, i.e., solve
for \(aH\) and \(bH\).

\[
aH(t) = \frac{\partial C(t, S(t))}{\partial x} \quad bH(t) = R^{-1}(t)(C(t, S(t)) - aH(t)S(t))
\]

It’s interesting to note the similarities between the continuous time results above to that of
the discrete result obtained when working with the Binomial model. Like \(a\) in the Binomial
model, \(aH(t)\) is attempting to immunize the sensitivity of the price of the claim relative to
the price of the underlying. This becomes more clear when we rewrite \(aH(t)\) into its equivalent
representation, i.e., \(\frac{\partial P(t)}{\partial S(t)}\). Now that we have found the replicating strategy for the contingent
claim, we require only to derive an equation to solve the function \(C(t, x)\), which incidentally is
the famous Black-Scholes-Merton partial differential equation. It can be derived by substituting
the values of \(aH(t)\) and \(bH(t)\) into equation 16, comparing the resulting
\(dt\) terms with that of
equation 15, and re-arrange to get:

\[
\frac{\partial C(t, x)}{\partial t} + r x \frac{\partial C(t, x)}{\partial x} + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 C(t, x)}{\partial x^2} = rC(t, x)
\]

It can be solved by noting that the terminal condition is \(C(T, x) = f(x)\), where \(f\) is the payoff
function of the contingent claim, i.e., \(X = f(S(T))\), which for a call option with strike \(k\) is
\(max(0, S(T) - k)\).

With this framework, the concept of delta hedging naturally arises. We first note that the delta
of being long an asset is 1 (and short an asset is -1), i.e., any change in the asset is perfectly
reflected in a portfolio that owns that asset. Now, if we assume that we have a short position
in any derivative, we can hedge the daily fluctuations of that position by buying delta units
of the underlying asset. Such a hedging strategy is called being \textit{delta neutral} because we are eliminating price variation of the portfolio contents. As we’ve discussed, the delta will change with time as the value of the contingent claim changes and can only be a riskless strategy if we re-adjust our hedge in continuous time. This is impossible in practice, however, we can choose to buy or sell the appropriate amount of the underlying at regular intervals, this is called dynamic hedging. Alternatively, we can set up the delta hedge once when the short derivative position is established and leave it until expiration, this is called static hedging. Since buying and selling assets have transaction costs, a trader must strike a balance between the importance of being delta neutral relative to the loss associated with a large number of transactions.

To conclude this section, we list the steps required to initiate the hedging portfolio to hedge a short position of a European call option at $t = 0$.

1. Sell a the European call and collect $C(0, S(0))$

2. Borrow $b^H(0) = (a^H(0)S(0) - C(t, S(0)))$ units of the money market instrument

3. Buy $a^H(0) = \frac{\partial C(0, S(0))}{\partial x}$ shares at price $S(0)$

At every time $t$, we adjust the amount of stock we own in order to have $a^H(t) = \frac{\partial C(t, S(t))}{\partial x}$ shares. Such a portfolio is a dynamic hedge, and we must continually rebalance the portfolio by either adding or borrowing from the money market (easiest to assume this is a bank account earning the risk free rate). What’s important to note is that delta-hedging is not concerned with the amount of money invested in the money market because it is implicit in the self-financing strategy when purchasing and selling units of the stock in the amount as the delta is changing.
We present a representative subset of the MATLAB code utilized to generate the results of tables 4.11 and 4.12 (i.e., the code for the SDP formulation). We only present a representative subset because the other code is similar and will be repetitive.

The first section below (D.1) demonstrates the SDP formulation when we have one benchmark option and have calculated two moments of the underlying distribution. The second section (D.2) demonstrates the SDP formulation when we have two benchmark options and have calculated three moments of the underlying distribution. These are the simplest and most complex scenarios respectively, and the omitted formulations (one benchmark option with three moments, and two benchmark options with two moments) can be represented in a similar manner.

We have implemented a boolean variable named computeLowerBound into each program. When this is set to TRUE (i.e., 1), the resulting bound is the lower bound, whereas, if it is set to FALSE (i.e., 0), the resulting bound is the upper bound.

D.1 Sample Code 1

The SDP formulation with one benchmark options and two calculated moments of the underlying distribution

1 cvx_clear;
2
3 % Input data
4 k1=100; %benchmark strike
5 h1=1.18; %benchmark price
6 m1=92.76; %first moment
7 m2=8644.49; %second moment
8 k0=110; %illiquid call strike
%boolean – compute lower bound if true (1)
computeLowerBound=0;

%cvx indicator
cvx_begin

variable W(3,3) symmetric;
variable X(3,3) symmetric;
variable Z(3,3) symmetric;
W == semidefinite(3);
X == semidefinite(3);
Z == semidefinite(3);
variables lm(2) lh q1 q2 t s;
expressions y0 y1 y2;

if(convert.to_min)
  maximize( t+s );
else
  minimize( t+s );
end

subject to

%setting up W
y0 = s - lm(1)*m1 - lm(2)*m2;
y1 = lm(1); y2=lm(2);
if(convert.to_min)
  y0 = lm(1)*m1 + lm(2)*m2 - s;
y1 = -1*lm(1); y2=-1*lm(2);
end

%constraints for W
W(2,1) + W(1,2) == 0;
W(2,3) + W(3,2) == 0;
W(l,1) == y0;
W(l,3) + W(2,2) + W(3,1) == 2*y0 + y1*k1;
W(3,3) == y0 + y1*k1 +y2*k1*k1;

%setting up X
y0 = s - lh*k1 - lm(1)*m1 - lm(2)*m2;
y1 = lm(1) + lh; y2 = lm(2);
if(convert.to_min)
  y0 = lm(1)*m1 + lm(2)*m2 + lh*k1 - s;
y1 = -1*lm(1) - lh; y2 = -1*lm(2);
end
%constraints for X
X(2,1) + X(1,2) == 0;
X(2,3) + X(3,2) == 0;
X(1,1) == y0 + y1*k1 + y2*k1*k1;
X(1,3) + X(2,2) + X(3,1) == 2*y0 + y1*(k1+k0) + 2*y2*k1*k0;
X(3,3) == y0 + y1*k0 + y2*k0*k0;

%setting up Z
y0 = k0 - l*h*k1 - l*m(1)*m1 - l*m(2)*m2 + s;
y1 = l*m(1) + l*h - 1; y2 = l*m(2);
if(convert_to_min)
y0 = l*h*k1 + l*m(1)*m1 + l*m(2)*m2 - k0 - s;
y1 = 1 - l*m(1) - l*h; y2 = -1*l*m(2);
end
%constraints for Z
Z(2,1) + Z(1,2) == 0;
Z(2,3) + Z(3,2) == 0;
Z(1,1) == y0 + y1*k0 + y2*k0*k0;
Z(1,3) + Z(2,2) + Z(3,1) == y1 + 2*y2*k0;
Z(3,3) == y2;

%final inequalities
if(convert_to_min)
    h1*q1 - h1*q2 >= t
    q1 - q2 <= l*h;
    -l*q1 - q2 <= 1;
    q1<=0;
    q2<=0;
else
    h1*q1 - h1*q2 <= t
    q1 - q2 >= l*h;
    -l*q1 - q2 >= -1;
    q1>=0;
    q2>=0;
end

D.2 Sample Code 2

The SDP formulation with two benchmark options and three calculated moments of the underly-
cvx_clear;

% Input data
k1=100; %benchmark strike 1
k2=105; %benchmark strike 2
h1=1.18; %benchmark price 1
h2=.56; %benchmark price 2
m1=92.76; %first moment
m2=8644.49; %second moment
m3=810961.4885; %third moment
k0=110; %illiquid call strike

%boolean - compute lower bound if true (1)
computeLowerBound=0;

%cvx indicator
cvx_begin
variable W(4,4) symmetric;
variable X(4,4) symmetric;
variable Z(4,4) symmetric;
variable U(4,4) symmetric;
W == semidefinite(4);
X == semidefinite(4);
Z == semidefinite(4);
U == semidefinite(4);
variables lm(3) lh(2) q1 q2 q3 q4 t s;
expressions y0 y1 y2 y3;

if computeLowerBound
    maximize( t+s );
else
    minimize( t+s );
end

subject to

%setting up W
y0 = -1*(lm(1)*m1 + lm(2)*m2 + lm(3)*m3) + s;
y1 = lm(1); y2=lm(2); y3=lm(3);
if computeLowerBound
    y0=y0*1; y1=y1*1; y2=y2*1; y3=y3*1;
end
%constraints for W
    W(2,1) + W(1,2) == 0;
\begin{verbatim}
49    W(3,4) + W(4,3) == 0;
50    W(4,1) + W(3,2) + W(2,3) + W(1,4) == 0;
51    W(1,1) == y0;
52    W(1,3) + W(2,2) + W(3,1) == 3*y0 + y1*k1;
53    W(2,4) + W(3,3) + W(4,2) == 3*y0 + 2*y1*k1 + y2*k1*k1;
54    W(4,4) == y0 + y1*k1 + y2*k1*k1 + y3*k1*k1*k1;
55
56
57    %setting up U
58    y0 = -1*(lh(1)*k1 + lm(1)*m1 + lm(2)*m2 + lm(3)*m3) + s;
59    y1 = lm(1) + lh(1); y2 = lm(2); y3=lm(3);
60    if computeLowerBound
61        y0=y0*-1; y1=y1*-1; y2=y2*-1; y3=y3*-1;
62    end
63    %constraints for U
64    U(2,1) + U(1,2) == 0;
65    U(3,4) + U(4,3) == 0;
66    U(4,1) + U(3,2) + U(2,3) + U(1,4) == 0;
67    U(1,1) == y0 + y1*k1 + y2*k1*k1 + y3*k1*k1*k1;
68    U(1,3) + U(2,2) + U(3,1) == 3*y0 + y1*(2*k1+k2) + y2*(k1*k1+2*k2*k1) + 3*y3*(k1*k1*k2);
69    U(2,4) + U(3,3) + U(4,2) == 3*y0 + y1*(2*k2+k1) + y2*(k2*k2+2*k1*k2) + 3*y3*(k2*k2*k1);
70    U(4,4) == y0 + y1*k2 + y2*k2*k2 + y3*k2*k2*k2;
71
72
73    %setting up X
74    y0 = -1*(lh(1)*k1 + lh(2)*k2 + lm(1)*m1 + lm(2)*m2 + lm(3)*m3) + s;
75    y1 = lm(1) + lh(1) + lh(2); y2 = lm(2); y3=lm(3);
76    if computeLowerBound
77        y0=y0*-1; y1=y1*-1; y2=y2*-1; y3=y3*-1;
78    end
79    %constraints for X
80    X(2,1) + X(1,2) == 0;
81    X(3,4) + X(4,3) == 0;
82    X(4,1) + X(3,2) + X(2,3) + X(1,4) == 0;
83    X(1,1) == y0 + y1*k2 + y2*k2*k2 + y3*k2*k2*k2;
84    X(1,3) + X(2,2) + X(3,1) == 3*y0 + y1*(2*k2+k0) + y2*(k2*k2+2*k0*k2) + 3*y3*(k2*k2*k0);
85    X(2,4) + X(3,3) + X(4,2) == 3*y0 + y1*(2*k0+k2) + y2*(k0*k0+2*k2*k0) + 3*y3*(k0*k0*k2);
86    X(4,4) == y0 + y1*k0 + y2*k0*k0 + y3*k0*k0*k0;
87
88    %setting up Z
89    y0 = k0 - lh(1)*k1 - lh(2)*k2 - lm(1)*m1 - lm(2)*m2 - lm(3)*m3 + s;
90    y1 = lm(1) + lh(1) + lh(2) -1; y2 = lm(2); y3=lm(3);
91    if computeLowerBound
\end{verbatim}
```matlab
y0=y0*−1; y1=y1*−1; y2=y2*−1; y3=y3*−1;
end

% constraints for Z
Z(2,1) + Z(1,2) == 0;
Z(3,4) + Z(4,3) == 0;
Z(4,1) + Z(3,2) + Z(2,3) + Z(1,4) == 0;
Z(1,1) == y0 + y1*k0 + y2*k0*k0 + y3*k0*k0*k0;
Z(1,3) + Z(2,2) + Z(3,1) == y1 + 2*y2*k0 + 3*y3*k0*k0;
Z(2,4) + Z(3,3) + Z(4,2) == y2 + 3*y3*k0;
Z(4,4) == y3;

% final inequalities
if(computeLowerBound)
  h1*(q1−q2) + h2*(q3−q4) >= t
  q1  − q2  <= h(1);
  q3  − q4  <= h(2);
  −q1  − q2  <= 1;
  −q3  − q4  <= 1;
  q1<=0;
  q2<=0;
  q3<=0;
  q4<=0;
else
  h1*(q1−q2) + h2*(q3−q4) <= t
  q1  − q2  >= h(1);
  q3  − q4  >= h(2);
  −q1  − q2  >= −1;
  −q3  − q4  >= −1;
  q1>=0;
  q2>=0;
  q3>=0;
  q4>=0;
end
```

```preterp |x_end```