Applications of Lévy Semistationary Processes to Storable Commodities

by

Simon Lacoste-Bouchet

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Quantitative Finance

Waterloo, Ontario, Canada, 2025

©Simon Lacoste-Bouchet 2025

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Abstract

Volatility Modulated Lévy-driven Volterra (VMLV) processes have been applied by Barndorff-Nielsen, Benth and Veraart (2013) to construct a new framework for modelling spot prices of non-storable commodities, namely energy. In this thesis, we extend this framework to storable commodities by showing that successful classical models belong to the framework albeit under some parameter restrictions (a result which to our knowledge is new). Additionally, we propose a new model for spot prices of storable commodities which is built on the VMLV processes and their important subclass of so-called Lévy semi-stationary (LSS) processes. The main feature of the framework exploited in the model proposed in this thesis is the memory of the VMLV processes which is used judiciously to account for cumulative changes in inventory over time and the corresponding expected changes in prices and volatility. To the best of our knowledge, this is the first study which uses the LSS processes to investigate pricing in storable (as opposed to non-storable) commodity markets to account for the impact of inventory on pricing. To complement the theoretical development of the new model, we also provide in this thesis a companion set of calibration and empirical analyses to shed light on the new model's performance compared to previously established models in the literature.

Acknowledgements

I am very grateful to Professor Tony Wirjanto for his guidance and feedback throughout this project. His advice was instrumental in preparing this thesis and inspiring me to explore a number of alternatives throughout my research.

I am also indebted to the MQF program and Professor Alexander Schied for their continued support throughout this process, from my first admission in 2020 to my eventual readmission and completion of this thesis.

Dedication

To my family and friends who have supported throughout this process.

Table of Contents

Author's Declaration ii
Abstract iii
Acknowledgements iv
Dedication v
List of Figures viii
List of Tables x
1 Introduction 1
2 Commodity Markets and Spot Price Models 3
2.1 Properties of Commodity Markets, Prices, Forwards and Futures 3
2.1.1 Forwards & Futures Contracts 3
2.1.2 Pricing of Forwards and Futures Contracts 4
2.1.3 Commodity Forward Curves and Convenience Yield 5
2.2 Commodity Spot Price Models 7
2.2.1 Geometric Brownian Motion 7
2.2.2 One-Factor Schwartz Model 9
2.2.3 Two-Factor Models 11
2.2.4 Volatility Modulated Lévy-driven Volterra Framework 17
3 Application of VLMV to Storable Commodities 25
3.1 Proposed Model 25
3.1.1 Inventory Process 25
3.1.2 Spot Price Process 25
3.2 Semi-martingale Representation 26
3.3 Risk Neutral Measures & Market Price of Risk 28
3.4 Futures and Derivatives Valuation 29
3.4.1 Pricing PDE Representation 30
3.4.2 Monte Carlo Methods 34

4 Methodology and Empirical Analysis 36
4.1 Empirical Data 36
4.1.1 Spot and Forward Prices 36
4.1.2 Exchange Inventory 38
4.2 Calibration 40
4.2.1 The Filtering Problem 40
4.2.2 State Space Formulation 41
4.2.3 Probabilistic Setting 42
4.2.4 Filtering Techniques 43
4.3 Empirical Performance 48
4.3.1 Aluminum Market Results 48
4.3.2 Copper Markets Results 51
4.3.3 Nickel Market Results 54
4.3.4 Model Comparison 56
4.3.5 Discussion 66
5 Conclusion and Future Studies 68
References 70
Appendices 73
Numerical Methods for the Pricing PDE 74
Approaches to Valuation under the LSS Framework 85
Semi-martingale Approach and Density Process Criterions 88

List of Figures

Figure 1: Forward Curves at Various Interest Rates 5
Figure 2: Historical Aluminium Forward Curves 2020-2023 6
Figure 3: Historical Copper Forward Curves 2020-2023 7
Figure 4: Forward Curves of the Schwartz One-Factor Model with $\kappa = 1.5$, $\alpha * = 1.5$, $\sigma = 1$ 11
Figure 5: Forward Curves of the Schwartz and Smith Model in Backwardation 17
Figure 6: Forward Curves of the Schwartz and Smith in Contango 17
Figure 7: Aluminum Spot and Futures Prices 2020-2023 37
Figure 8: Copper Spot and Futures Prices 2020-2023 37
Figure 9: Nickel Spot and Futures Prices 2020-2023 38
Figure 10: Historical Aluminum Inventory Levels 39
Figure 11: Historical Copper Inventory Levels 39
Figure 12: Historical Nickel Inventory Levels 40
Figure 13: Aluminum Error Over Time under Filtered Factor Values 49
Figure 14: Aluminum Error Over Time under Actual Factor Values 49
Figure 15: Time Series of Actual and Filtered Aluminum Spot Prices 50
Figure 16: Time Series of Actual and Filtered Changes in Aluminum Inventories 51
Figure 17: Copper Error Over Time under Filtered Factor Values 52
Figure 18: Copper Error Over Time under Actual Factor Values 52
Figure 19: Time Series of Actual and Filtered Copper Spot Prices 53
Figure 20: Time Series of Actual and Filtered Changes in Copper Inventories 54
Figure 21: Nickel Error Over Time under Filtered Factor Values 55
Figure 22: Nickel Error Over Time under Actual Factor Values 55
Figure 23: Time Series of Actual and Filtered Nickel Spot Prices 56
Figure 24: Time Series of Actual and Filtered Changes in Nickel Inventories 56
Figure 25: Aluminum Market Error with Filtered Factor Values under the Gabillon Model 60
Figure 26: Aluminum Market Error with Actual Factor Values under the Gabillon Model 60
Figure 27: Actual and Filtered Aluminum Spot Prices under the Gabillon Model 61
Figure 28: Filtered Aluminum Spot and Long-Term Price under the Gabillon Model 61
Figure 29: Copper Market Error with Filtered Factor Values under the Gabillon Model 62
Figure 30: Copper Market Error with Actual Factor Values under the Gabillon Model 62
Figure 31: Actual and Filtered Copper Spot Prices under the Gabillon Model 63
Figure 32: Filtered Copper Spot and Long-Term Price under the Gabillon Model 64
Figure 33: Nickel Market Error with Filtered Factor Values under the Gabillon Model 64
Figure 34: Nickel Market Error with Actual Factor Values under the Gabillon Model 65
Figure 35: Actual and Filtered Nickel Spot Prices under the Gabillon Model 66
Figure 36: Filtered Nickel Spot and Long-Term Price under the Gabillon Model 66

Figure 37: Comparison of the Monte Carlo and PDE Solutions with Maturity of 2.25 Years 83
Figure 38: Comparison of the Monte Carlo and PDE Solutions with Maturity of 1.25 Years 83
Figure 39: Comparison of the Monte Carlo and PDE Solutions with Maturity of 0.25 Years 84

List of Tables

Table 1: Aluminum Basic Error Statistics....48
Table 2: Aluminum Futures Curve Contango and Backwardation Results....50
Table 3: Copper Basic Error Statistics....51
Table 4: Copper Futures Curve Contango and Backwardation Results....53
Table 5: Nickel Basic Error Statistics....54
Table 6: Nickel Futures Curve Contango and Backwardation Results....55
Table 7: Comparison of the Proposed and Gabillon Models under the Filtered Factors 59
Table 8: Comparison of the Proposed and Gabillon Models under the Actual Factors 59
Table 9: Basic Error Statistics of the Aluminum Market under the Gabillon Model....60
Table 10: Aluminum Futures Curve Shape Results under the Gabillon Model 61
Table 11: Basic Error Statistics of the Copper Market under the Gabillon Model....62
Table 12: Copper Futures Curve Shape Results under the Gabillon Model....63
Table 13: Basic Error Statistics of the Nickel Market under the Gabillon Model....64
Table 14: Nickel Futures Curve Shape Results under the Gabillon Model....65

Chapter 1

Introduction

Commodities can take a variety of forms but in principle can be described as “unprocessed or semi-processed goods” (Schofield, 2021). These goods are then consumed directly or used in some business process as an input towards a finished product before being sold to customers. The production, trade and use of commodities continues to be a fundamental part of the global economy and permeates every aspect of our lives including the food we eat, the infrastructure we use or the energy that powers our homes. In total, it is estimated that approximately US$10 trillion worth of commodities are produced on an annual basis (Trafigura Group, 2019). Due to their importance, commodity markets are often discussed in the media. For example, the war in Ukraine and the ensuing sanctions have distorted a number of commodity markets as Russia and Ukraine are major producers of energy, grain and metals. This has led to record increases in European natural gas prices as Russian supply dwindled, grain shortages in east Africa as Ukraine’s port were blockaded and has also led to Russian metals being banned from the London Metal Exchange (LME) and the Chicago Mercantile Exchange (CME).

Prior to the existence of modern markets, commodity markets had buyers and sellers physically meeting at a designated marketplace to complete transactions where the underlying product would be delivered immediately (Geman, 2005). As markets developed and grew in size, the complexity of transactions increased and a need for standardized products emerged. This has led to the creation of a number of exchanges such as the LME and the Chicago Board of Trade (CBOT) where every market participant would deal with the clearing house as counterparty. These new markets allowed any parties to enter into agreements to hedge their positions or speculate on future price movements. Nowadays, most participants in commodity markets are not interested in the physical commodities themselves and only a fraction of contracts actually leads to physical delivery of the underlying product.

Physical commodities can be classified into broad categories such as energy, industrial metals, precious metals and agricultural commodities. The behavior of commodity markets can vary substantially from one to another. For example, some commodities exhibit seasonal behavior due to seasonal supply or demand dynamics. On the demand side, natural gas demand tends to be greater during winter months and on the supply side agricultural commodity production is often concentrated to a harvest period with little supply for the rest of the year. Another, and perhaps most salient, distinction exists between commodities which are storable and those which are non-storable. Most physical commodities are storable with electricity being the notable exception. For

some commodities, major exchanges will provide a collection of storage facilities that can be used to settle contracts that reach maturity.

The behavior of commodity markets is distinct from that of most traditional financial assets which has led to a specialized and detailed literature on the topic. Amongst others, the literature includes a number of spot price models. The main topic of focus in this thesis is the application of a specific type of stochastic process known as Volatility Modulated Lévy-driven Volterra (VMLV) processes to model the spot price of storable commodities. To the best of our knowledge and it is also supported by leading scholars in the field (such as Fred Espen Beth from the Department of Mathematics at the University of Oslo, Norway), this is the first study in which the VMLV processes are applied to model the spot prices of storable commodities. This is in contrast with earlier works which focused instead on the widely studied spot prices of non-storable commodities due primarily to technical difficulties. The remainder of this thesis is organized as follows. We first start with a brief literature review covering relevant topics on commodity markets, spot price models and the VMLV framework. We show that a number of successful two factor spot price models belong to the VMLV framework, albeit with a limited parameter set. Next, we introduce a new spot price model for storable commodities and lastly, we outline a calibration procedure and an empirical analysis of the new model's performance compared to previously established models in the literature. The appendix includes a discussion of earlier unsuccessful attempts at utilizing numerical solutions to the pricing PDE of the new spot price model along with a discussion of other pricing approaches

Chapter 2

Commodity Markets and Spot Price Models

There is a large body of work on commodity markets in the literature ranging from theoretical models to statistical/econometric models as well as various models used to value commodity derivatives. For the purposes of this thesis, the literature review is selected to focus on key concepts relevant to and applied in the model proposed in this thesis. These concepts can be classified into two broad categories. The first category includes properties of commodity markets and prices while the second category concerns existing spot price models in the literature. Within the second category, particular attention is devoted to a class of processes known as Volatility Modulated Lévy-driven Volterra (VMLV) processes which are the main focus of study in this thesis.

2.1 Properties of Commodity Markets, Prices, Forwards and Futures

Commodities markets and prices exhibit behaviors distinct from conventional financial instruments. The spot market is seldom accessible in commodities and most of the trading is done via forward or futures contracts (Back & Prokopczuk, 2013). Forwards and futures contracts require the holder of the contract to purchase a commodity at a predetermined future date in accordance with the specifications set out in the contract (which include quantity, point of delivery product specifications, etc.).

2.1.1 Forwards & Futures Contracts

Futures and forward contracts are, economically, the same type of financial agreement whereby the holder agrees to purchase a commodity at a predetermined future date. However, in practice, they both arise in different situations. Forward contracts are negotiated directly between parties and their terms reflect the needs of the parties that enter the contracts. The parties are free to tailor the contract terms such as size, date of delivery, point of delivery, product specification etc. This contrasts with futures contracts which are standardized, exchange-traded contracts where each market participant transacts with a central clearing house. Exchanges will typically mark-to-market the futures contracts on a daily basis to limit the accumulation of large unrealized losses. To ensure that funds are available in the event of a daily loss, participants are required to maintain sufficient margins, or otherwise the clearinghouse will liquidate their positions. As futures contracts are standardized and exchange traded, they are the main instruments through which the market “expresses” its view on the price of a commodity. The majority of participants in futures markets is primarily interested in a financial exposure to commodities and not in a physical product. Hence most contracts are terminated before reaching maturity to avoid having to take delivery or deliver the corresponding commodity.

While futures are standardized and widely accessible, the specification of the contracts may not align completely with the situation faced by individual market participants. This difference can arise in a number of ways such as the location of delivery, the specification of the product, the time of delivery, etc. The difference between the business realities of market participants and the contracts available in the market can give rise to a so-called basis risk where the instruments used do not perfectly align with the needs of market participants. For example, a nickel producer may look to futures markets to hedge their exposure to nickel prices. However, common nickel futures contracts, such as those on the London Metal Exchange, are for 99.80% pure nickel metal (LME, 2024) while the producer may sell concentrates, nickel sulfides or other nickel products. The price received by the producer upon a sale will often be based on a widely used exchange benchmark such as the LME spot price. To hedge the price received the producer could take a short position in LME futures, in effect, selling 99.80% nickel on the exchange at a future date. Issues can arise if the position moves against the producer, as the nickel produced by the firm cannot be used to settle the futures contracts at maturity. As a result, the firm must provide cash to meet exchange margin requirements. A situation like this occurred in March 2022 when a producer (Tsingshan) with an enormous short position was forced to liquidate their holdings amidst increasing prices (Qi, Wu, Xu, & Li, 2024), (Jones, 2022). Due to the size of the Tsingshan's short position relative to the size of the market this has led to a short squeeze on the LME nickel market. Prices increased rapidly from US$25,000/mt to over US$100,000/mt. In the aftermath, the LME halted trading and subsequently cancelled some of the trades that took place during the squeeze. While this is considered to be an extreme event in the market, it does show the extent to which basis risk and short positions can impact market participants.

2.1.2 Pricing of Forwards and Futures Contracts

As noted above, futures and forward prices are economically the same contract, but differences can arise from mark-to-market and margin mechanisms. However, whenever the interest rate is assumed to be constant the value of equivalent forwards and futures are the same. As the modelling of interest rates is not the focus of this thesis, we assume that interest rates are constant meaning that for the purposes of this thesis forwards and futures are considered to be interchangeable.

Unlike most financial derivatives, forwards can be replicated without the necessity to specify an explicit spot price model, so that a simple no arbitrage argument can be used to determine their value. Consider a market with a risky asset S_t representing the price of the underlying at time t and a risk-free asset growing at the risk-free rate r. Let F be a forward with maturity T and delivery price K for one unit of S. Denote the value of F at time t as F_t. At maturity, the value of F becomes:

$$F_T = S_T - K$$

To price the forward, we construct a portfolio Π such that at maturity it has the same value as the forward: $\Pi_T = S_T - K$. To this end we take the portfolio Π to contain one unit of the underlying and a loan expiring at maturity with repayment equal to K. With this choice, we have $\Pi_T = F_T$ and the no arbitrage assumption then implies that the two must have the same value at $t = 0$. Hence:

$$F_0 = \Pi_0 = S_0 - Ke^{-rT}$$

From here on, we can define the notion of a forward or future price. The futures/forward price corresponds to the delivery price K such that the price of the forward at $t = 0$ is 0. Let X_0 denote the forward price for the forward F. Then X_0 satisfies the following equation:

$$S_0 - X_0 e^{-rT} = 0 \Rightarrow X_0 = S_0 e^{rT}$$

A useful tool in the futures markets is the futures curve. This curve plots the futures price as a function of time to maturity. In the simple market outlined above, this would be a plot of X_0 as a function of T. Figure 1 below shows the shape of the forward curve for various values of the interest rate r. Here we note that the shape of the forward curves depends on the value of r with an upwards slopping curve for $r > 0$, a flat curve for $r = 0$ and a downward sloping curve for $r < 0$. It is also important to note here that contracts traded on exchanges are entered at the current futures price and hence there is no cost to enter a position in these contracts.

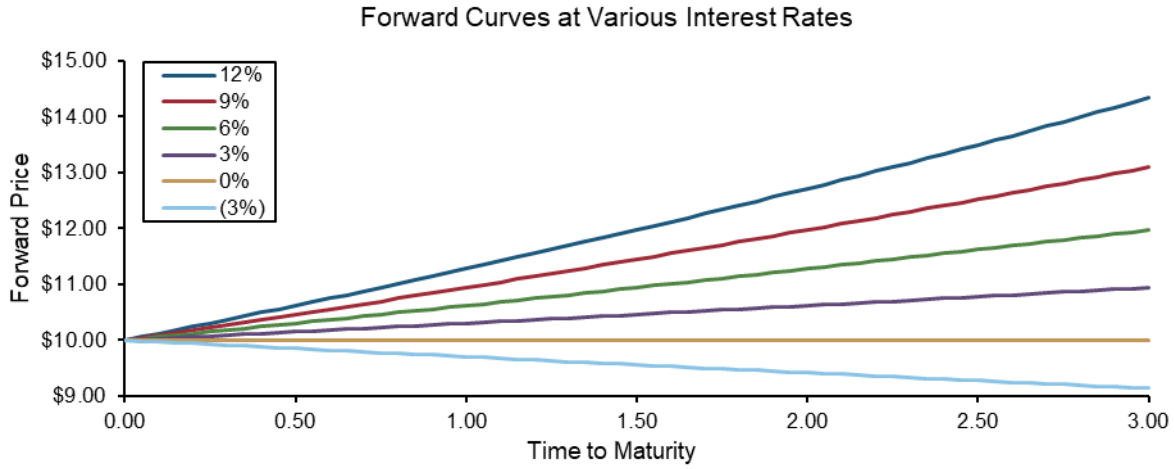


Figure 1: Forward Curves at Various Interest Rates

2.1.3 Commodity Forward Curves and Convenience Yield

Many theories have been developed in the literature to explain the dynamics of commodity prices and the shape of futures curves. The two main theories discussed in the literature are the theories of storage and hedging pressure. The later posits the existence of risk premiums in commodity futures markets and was first formulated in the seminal work of Keynes (1930) with his theory of normal backwardation. The theory postulates that hedgers in the market (typically producers) would hold a net short position to secure a selling price for their products and would be willing to accept lower than current spot prices to limit their exposure to price fluctuations. As a result, the futures curve would be expected to be in backwardation. This theory has subsequently been generalized in Cootner (1960) to posit the existence of a risk premium, which can be either positive or negative, based on the proportion of speculators and hedgers in the market. Since then, numerous authors have investigated the existence of these risk premiums across commodity markets. However, conflicting results have been reported with respect to this theory, and it remains a lively subject of an ongoing debate (Back & Prokopczuk, 2013).

There is however a broad consensus that is emerging in the literature concerning the theory of storage. This theory links inventory levels to spot prices and establishes an important notion of convenience yield. Convenience yield can be expressed as the advantages gained from holding physical inventory that are not available to an investor holding the corresponding futures

(Geman, 2005). Typically, convenience yield refers to a “net” convenience yield, meaning the yield stemming from holding inventory less the costs of holding said inventory. Hence convenience yield can in principle be both negative and positive. Convenience yield, as discussed in Brennan (1958) and Telser (1958), can be viewed as an “embedded timing option attached to the commodity” where the holder can sell when price is high and hold when prices are low, or inventory can be used to supply a business process with no further exposure to input costs. When inventories are low, the value of this embedded option is high implying that convenience yield is high and vice versa (Back & Prokopczuk, 2013). The role of inventory is also noted for its relationship to volatility of spot prices. Available empirical work to date has shown that for various commodities, spot price volatility increases as inventories decrease and vice versa (Fama & French, 1987)**,** (Geman, 2005). This can be understood in terms of supply and demand forces, where inventories act as a buffer for sudden changes in demand or supply. The capacity of this buffer decreases as the level of inventories falls. Prices are then directly affected by changes in supply or demand when the levels of inventories are low (Geman, 2005).

Commodity markets exhibit a wide variety of forward curves depending on market conditions and the commodities themselves. While it is difficult to characterize all possible future curves, three types of curves commonly arise in many commodity markets. The first and second curves are simply monotone; they include backwardation where the futures price is decreasing with maturity and its converse, known as contango, where future prices increase with maturity. The third type is of a humped shape type which is either convex up or down. Figure 2 and Figure 3 below show historical forward curves of the LME aluminium and copper markets.

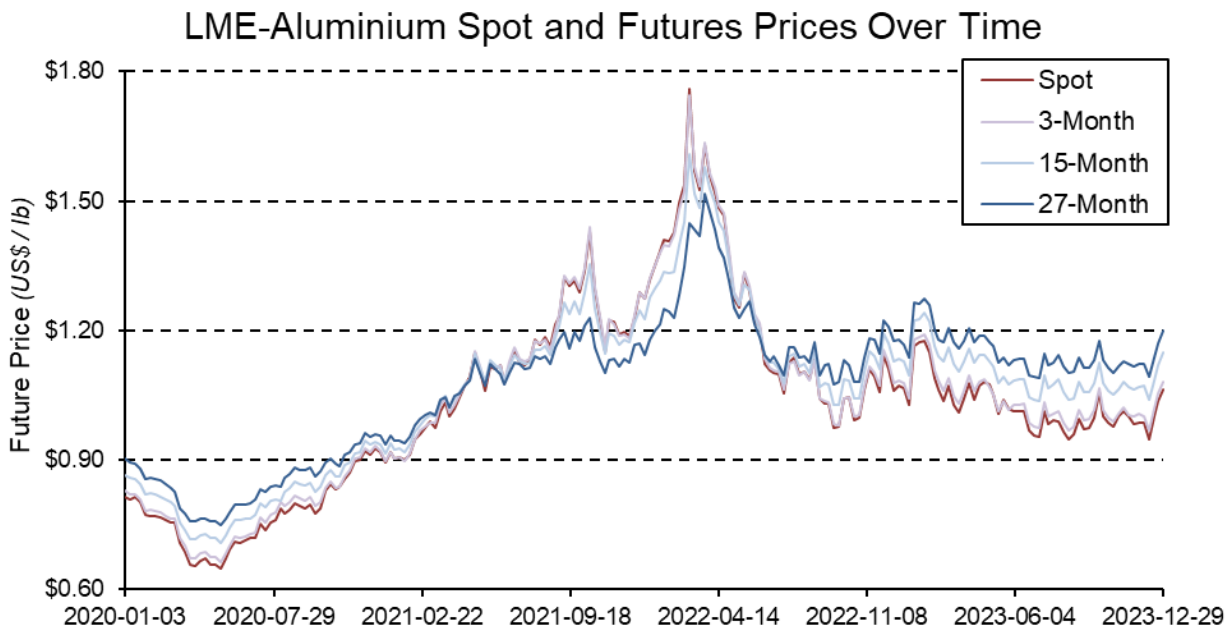


Figure 2: Historical Aluminium Forward Curves 2020-2023

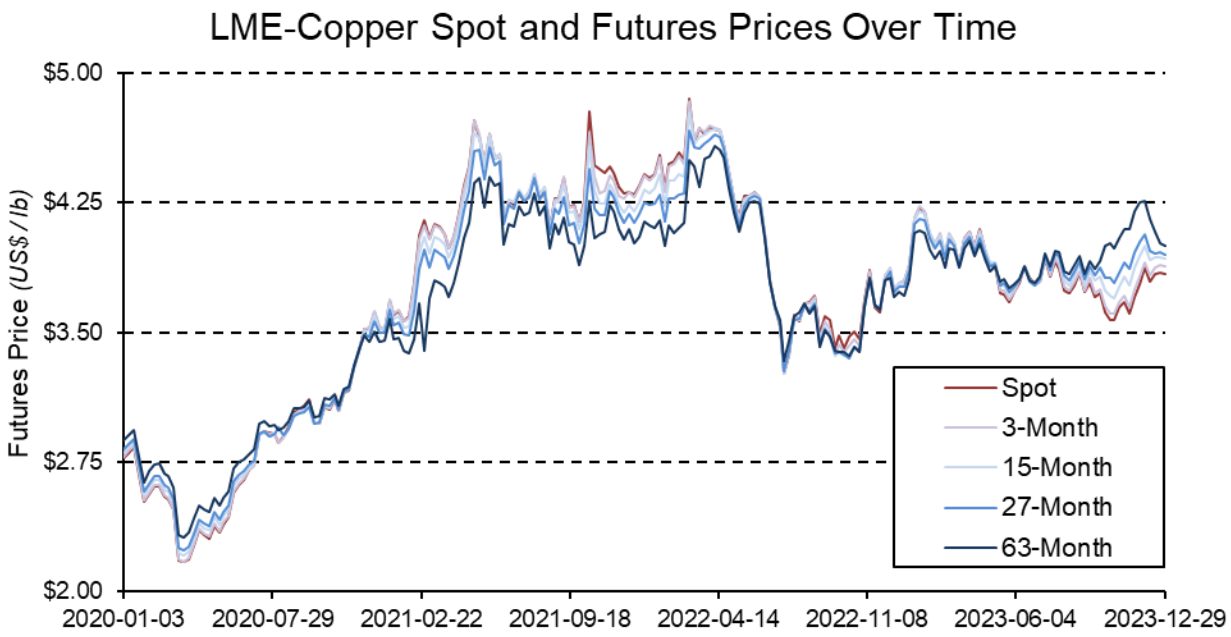


Figure 3: Historical Copper Forward Curves 2020-2023

Another key feature of commodity prices is their mean reversion property (Geman, 2005). This is expected from supply and demand driven arguments, where prices are expected to fall with excess supply leading to an eventual reduction in supply (with some lags) which is then expected to increase prices and vice versa. The mean reversion of commodity prices has been attested by various authors over time. However, there is still no clear-cut consensus emerging in the literature whether the level to which commodity prices revert to is constant or stochastic in nature (Back & Prokopczuk, 2013).

2.2 Commodity Spot Price Models

There is an extensive literature covering commodity spot price models with a great variation of model specifications. In this section we discuss a few prominent models and outline the key theoretical aspects underpinning their use in commodity markets.

2.2.1 Geometric Brownian Motion

Early models used in modelling commodity markets typically have utilized a geometric Brownian motion (GBM) approach similar to that of the Black-Scholes model in the options pricing literature. These models have been used to value natural resources investments such as mines and oilfields where commodity prices present a substantial risk to project viability (Brennan & Schwartz, 1979). The main difference between the original Black-Scholes model and this model is that the latter is used in the context of commodities with the addition of convenience yield. The inclusion of the convenience yield is based on the theory of storage whereby the marginal convenience yield is assumed to be inversely proportional to the level of inventories. Similarly, spot prices of the commodity are expected to be inversely proportional to the level of inventories and hence it may be reasonable to model a convenience yield as a function of spot prices. We define the convenience yield $C = C(S)$ as a function of the spot price S of a commodity. Assuming that the spot price follows a GBM we have a stochastic differential equation (SDE) of the following form:

$$dS_t = \mu S_t dt + \sigma S_t dB_t$$

for $\mu, \sigma \in \mathbb{R}$ and where B_t is a standard Brownian motion. To accommodate the presence of the convenience yield in the model we begin by setting $F = F(S,t)$ to be a forward price at time t with a maturity T. Applying Ito's lemma, we obtain the following result:

$$\begin{aligned} dF_t &= \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}(dS_t)^2 \\ &= \left(\frac{\partial F}{\partial t} + \mu S_t \frac{\partial F}{\partial S} + \frac{(\sigma S_t)^2}{2}\frac{\partial^2 F}{\partial S^2}\right)dt + \sigma S_t \frac{\partial F}{\partial S}dB_t \end{aligned}$$

Next, we consider an investor who holds a portfolio Π containing one unit of S and short x units of forward $F(S,t)$ such that the portfolio is risk free. As there is no cost to enter a forward contract, the value of the portfolio is simply given by $\Pi_t = S_t$. The instantaneous return of the portfolio at t can therefore be expressed as:

$$\frac{d\Pi_t}{\Pi_t} = \frac{1}{S_t}(dS_t + C(S_t)dt - xdF_t)$$

As the investor holds the commodity directly, their returns include the returns on the commodity price dS plus the convenience yield of holding the commodity which is $C(S)dt$. The last term in the equation is the short position of x in the future $F(S,t)$. As x is set such that the portfolio is risk free the diffusion term of Π must be zero and hence x must satisfy the following equation:

$$\sigma S_t - x\sigma S_t \frac{\partial F}{\partial S} = 0 \;\Rightarrow x = \left(\frac{\partial F}{\partial S}\right)^{-1}$$

where we assume that $\frac{\partial F}{\partial S} \neq 0$. With this choice of x the portfolio is risk free and by an arbitrage argument it must have an instantaneous return of r. This implies that:

$$\frac{1}{S_0}(dS + C(S)dt - xdF) = r$$

Substituting the expressions for dS and dF we get the following result:

$$\mu + \frac{C(S_t)}{S_t} - \frac{x}{S_t}\left(\frac{\partial F}{\partial t} + \mu S_t \frac{\partial F}{\partial S} + \frac{(\sigma S_t)^2}{2}\frac{\partial^2 F}{\partial S^2}\right) = r$$

which simplifies further to the following equation:

$$C(S_t)\frac{\partial F}{\partial S} - \frac{\partial F}{\partial t} - \frac{(\sigma S_t)^2}{2}\frac{\partial^2 F}{\partial S^2} = rS_t \frac{\partial F}{\partial S}$$

From here we obtain a PDE for the futures price of the form:

$$\big(rS_t - C(S_t)\big)\frac{\partial F}{\partial S} + \frac{(\sigma S_t)^2}{2}\frac{\partial^2 F}{\partial S^2} + \frac{\partial F}{\partial t} = 0$$

subject to a terminal condition given by $F(S,T) = S$ since the futures price converges to the spot price at maturity. We also note that this PDE does not include an undifferentiated term that is present in the classical Black-Scholes model since there is no cost to enter a futures contract.

The classical model further assumes for simplicity/convenience that $C(S_t) = cS_t$ for $c \in \mathbb{R}$ a constant. With this choice of convenience yield the solution to the above PDE becomes:

$$F(S,t) = Se^{-(c-r)(T-t)}$$

The main advantage of this model is its simplicity as it has few parameters and provides a closed form solution for future prices. However, this model has two main shortcomings which will be addressed in subsequent models. The first shortcoming is that this model posits that the price of the commodity will, on average, grow at a fixed rate. However empirical observations have shown that commodity prices tend to revert to a mean level (whether fixed or stochastic). The second shortcoming is the lack of flexibility of the model in producing sensible futures curves. As shown in Figure 1, this type of model only produces upwards or downwards sloping futures curves depending on the sign of $c - r$. This is at odds with the available empirical observation since commodity markets typically exhibit a variety of future curve shapes.

2.2.2 One-Factor Schwartz Model

To address the issue of mean reversion arising from the GBM model, Schwartz (1997) considers a model where the logarithm of the spot price follows an Ornstein-Uhlenbeck process. Letting the S_t denote the spot price at time t, the model is endowed with the following dynamics:

$$dS_t = \kappa(\theta - \ln(S_t))S_t dt + \sigma S_t dB_t$$

where $\kappa, \theta, \sigma \in \mathbb{R}$ are constants and $\kappa > 0$.

Taking $X_t = \ln(S_t)$ and applying Ito's lemma shows that X_t follows an Ornstein-Uhlenbeck process of the form (where $\mu = \theta - \sigma^2/2\kappa$):

$$dX_t = \kappa(\mu - X_t)dt + \sigma dB_t$$

The Ornstein-Uhlenbeck process has a well-known analytical solution. Starting with the substitution $P_t = X_t e^{\kappa t}$ and once again applying Ito's lemma we obtain the following result:

$$\begin{aligned} dP_t &= \frac{\partial P}{\partial t}dt + \frac{\partial P}{\partial X}dX_t + \frac{1}{2}\frac{\partial^2 P}{\partial X^2}(dX_t)^2 \\ &= \left(\kappa X_t e^{\kappa t} + e^{\kappa t}\kappa(\mu - X_t)\right)dt + e^{\kappa t}\sigma dB_t \\ &= \kappa\mu e^{\kappa t}dt + e^{\kappa t}\sigma dB_t \end{aligned}$$

We can now integrate the above equation to obtain an expression for P_t:

$$\begin{aligned} P_t - P_0 &= \int_0^t \kappa\mu e^{\kappa s}ds + \int_0^t e^{\kappa s}\sigma dB_s \\ &= \mu(e^{\kappa t} - 1) + N\left(0, \sigma^2 \int_0^t e^{2\kappa s}ds\right) \\ &= \mu(e^{\kappa t} - 1) + N\left(0, \frac{\sigma^2}{2\kappa}(e^{2\kappa t} - 1)\right) \end{aligned}$$

where $N(a, b)$ is used to denote a normal/Gaussian distribution with mean a and variance b. As the stochastic integral has a deterministic integrand, we conclude that it is normally distributed, and we can use Ito's isometry to compute its variance. From the definition of P_t we have that

$$X_t = e^{-\kappa t}\left(X_0 + \mu(e^{\kappa t} - 1) + N\left(0, \frac{\sigma^2}{2\kappa}(e^{2\kappa t} - 1)\right)\right)$$

$$= X_0 e^{-\kappa t} + \mu(1 - e^{-\kappa t}) + N\left(0, \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t})\right)$$

and we conclude that that X_t is normally distributed (assuming that X_0 is normally distributed or constant). Additionally, we observe the mean reverting nature of the process, as t increases the influence of X_0 decreases and the mean approaches the long-term mean level μ. Using the solution to X_t we obtain the result that S_t is log-normally distributed with the same parameters as X_t.

The model assumes that the spot price is not traded and as a result the model relies on a market price of risk for equilibrium pricing. The market price of risk is assumed to be constant and is denoted as λ. The risk neutral dynamics of X_t take the form:

$$dX_t = \kappa(\mu - \lambda - X_t)dt + \sigma dB_t$$

By setting $\alpha^* = \mu - \lambda$ we can use the solution outlined above. Future prices are computed by using the expected value of the spot price, and hence, the future price at time t with delivery at time $T > t$ is given by:

$$\mathbb{E}[S_T] = \exp\left(ln(S_t)\, e^{-\kappa(T-t)} + \alpha^*\left(1 - e^{-\kappa(T-t)}\right) + \frac{\sigma^2}{4\kappa}\left(1 - e^{-\kappa(T-t)}\right)\right)$$

by using the mean of the log normal distribution.

While the model provides a mean reversion property, it does not have a sufficient degree of flexibility to appropriately model backwardation and contango. If we let maturity T go to infinity we find that the futures price converges to:

$$\lim_{T\to\infty} \exp\left(ln(S_t)\, e^{-\kappa(T-t)} + \alpha^*\left(1 - e^{-\kappa(T-t)}\right) + \frac{\sigma^2}{4\kappa}\left(1 - e^{-\kappa(T-t)}\right)\right) = \exp\left(\alpha^* + \frac{\sigma^2}{4\kappa}\right)$$

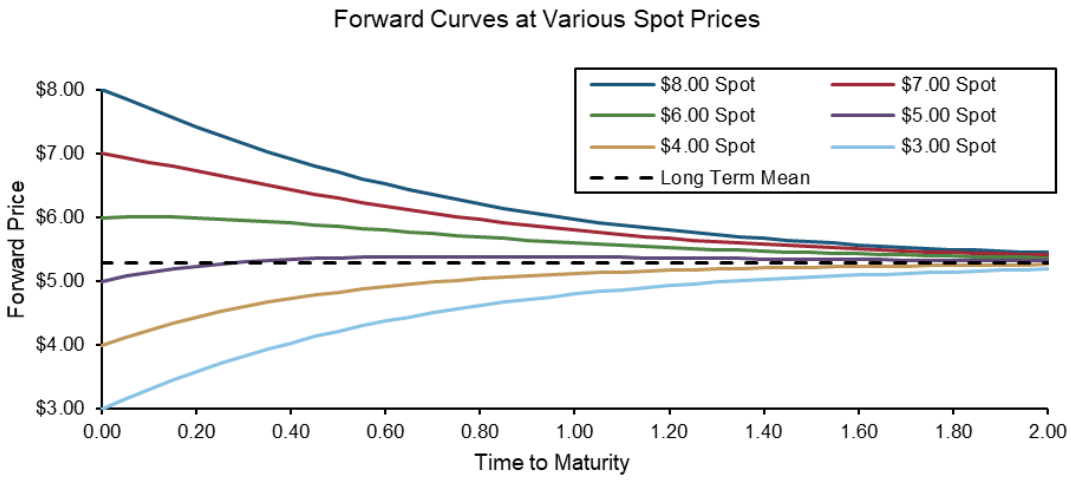


Figure 4: Forward Curves of the Schwartz One-Factor Model with $\kappa = 1.5$, $\alpha^* = 1.5$, $\sigma = 1$

The limit shows that the model exhibits backwardation whenever the logarithm of spot price exceeds α^* and it exhibits contango whenever the converse is true, as shown in Figure 4. This is an improvement over the GBM model; however it is clear from an inspection of empirical data that spot price, on its own, is not sufficient to determine whether a market is in backwardation or contango. This observation leads subsequently to a development of two factor models where the additional source of uncertainty is introduced to allow the model to more closely replicate futures curves observed in the market.

2.2.3 Two-Factor Models

In order to obtain models that exhibit a greater variety of futures curves, a number of authors have proposed two factor models and noted their ability to replicate observed futures curves. In these models, the number of factors refers to the number of sources of uncertainty included in the models. Whereas the GBM and Schwartz one-factor models only had one source of uncertainty (which is the spot price) two factor models will include a second source of uncertainty such as convenience yield or long-term price. In the next section, we describe three such models each with two factors that are expected to provide relatively better fits to commodity markets. It is worth pointing out that the framework for these models does not generally rely on their factors being tradeable and the derivative prices are based on equilibrium arguments where a market price of risk is used for non-tradeable factors under a risk neutral measure.

2.2.3.1 Stochastic Convenience Yield Model (Gibson and Schwartz, 1990)

This model was proposed in Gibson & Schwartz (1990) and contains two factors, which are spot price and convenience yield. As we have seen with the GBM model, the future curves were constrained to be either backwardation or contango depending on the sign of the difference between the convenience yield and the risk-free rate. In this model, the spot price continues to follow a GBM specification, but importantly, under the risk neutral measure, $\mathbb{Q}$, its instantaneous expected return is partly determined by a stochastic convenience yield. The factors of the model under the "real world" measure, $\mathbb{P}$, satisfy the following SDEs:

$$dS_t = \mu S_t dt + \sigma_S S_t dB_t^S$$

$$d\delta_t = \kappa(\alpha - \delta_t)dt + \sigma_\delta dB_t^\delta$$

$$dB_t^S dB_t^\delta = \rho dt$$

where S_t refers to spot price, δ_t is the convenience yield and $\mu, \sigma_S, \kappa, \alpha, \sigma_\delta, \rho$ are assumed to be constants.

Considering the value of an arbitrary European derivative $C = C(S, \delta, \tau)$ with time to maturity τ, by applying Ito's lemma to C we can express the dynamics of C as:

$$dC_t = \alpha(S, \delta, \tau)dt + \frac{\partial C}{\partial S}\sigma_S S_t dB_t^S + \frac{\partial C}{\partial \delta}\sigma_\delta dB_t^\delta$$

Further constructing a portfolio long one unit of C and short two other derivatives such that the portfolio is risk free, the following equilibrium condition must be satisfied by the price of C

$$\frac{\alpha(S, \delta, \tau)}{C} - r = \lambda_S(S, \delta)\left(\frac{1}{C}\frac{\partial C}{\partial S}\sigma_S S_t\right) + \lambda_\delta(S, \delta)\left(\frac{1}{C}\frac{\partial C}{\partial \delta}\sigma_\delta\right)$$

where λ_S and λ_δ are the market price of risk of S and δ respectively.

To characterize the market price of risk of S, we note that a physical unit should also satisfy the equilibrium condition. Clearly the value of such a derivative, which is a physical unit of S, at any given time is simply the spot price $C(S, \delta, \tau) = S$. The expected instantaneous return on physical units of the commodity at time t includes both the instantaneous increase in the price μ and the convenience yield δ_t. Hence:

$$\frac{\alpha(S, \delta, \tau)}{C} = \mu + \delta_t$$

Using the equilibrium condition above, substituting the expected instantaneous return and the derivatives of the price of the physical unit $C(S, \delta, \tau) = S$, we have:

$$\frac{\alpha(S, \delta, \tau)}{C} - r = \lambda_S(S, \delta)\left(\frac{1}{C}\frac{\partial C}{\partial S}\sigma_S S_t\right) + \lambda_\delta(S, \delta)\left(\frac{1}{C}\frac{\partial C}{\partial \delta}\sigma_\delta\right)$$

$$\mu + \delta_t - \mathrm{r} = \lambda_S(S, \delta)\left(\frac{1}{S_t}(1)\sigma_S S_t\right) + \lambda_\delta(S, \delta)\left(\frac{1}{S_t}(0)\sigma_\delta\right)$$

$$\mu + \delta_t - r = \lambda_S(S, \delta)\sigma_S$$

which implies that:

$$\lambda_S(S, \delta) = \frac{\mu + \delta_t - r}{\sigma_S}$$

For the convenience yield, the model assumes a constant market price of risk.

Using the market price of risks of the factors we further can express the risk neutral dynamics of the model as follows:

$$\mathrm{dS_t} = (\mathrm{r} - \delta_\mathrm{t})\mathrm{S_t dt} + \sigma_\mathrm{S}\mathrm{S_t dB_t^S}$$

$$d\delta_t = \kappa(\alpha - \delta_t) - \lambda_\delta dt + \sigma_\delta dB_t^\delta$$

$$dB_t^S dB_t^\delta = \rho dt$$

An analytical solution to futures prices under this model was developed by Bjerksund (1991) and can be shown to take the form below as used in Schwartz (1997):

$$F(S,\delta,\tau) = Sexp\left(-\delta\frac{(1-e^{-\kappa\tau})}{\kappa} + A(\tau)\right)$$

2.2.3.2 Stochastic Long-Term Price and Short-Term Deviation Model (Schwartz and Smith, 2000)

The second two-factor model discussed in this section was developed in Schwartz and Smith (2000). The model posits that the spot price of the commodity, S_t, follows a process given by:

$$\ln(S_t) = \chi_t + \xi_t$$

where ξ_t is referred to as the long-term price of the commodity and χ_t is referred to as the short-term variation from the long-term price. The long-term price follows a Brownian motion with drift while the short-term variation is mean reverting and follows an Ornstein-Uhlenbeck process. As neither factors (χ_t and ξ_t) are traded, a market price of risk is applied to the factors. The market price of risk of both factors is assumed to be constant. Under the risk neutral measure, χ_t and ξ_t satisfy the following SDEs:

$$d\chi_t = (-\kappa\chi_t - \lambda_\chi)dt + \sigma_\chi dB_t^\chi$$

$$d\xi_t = (\mu_\xi - \lambda_\xi)dt + \sigma_\xi dB_t^\xi$$

$$dB_t^\chi dB_t^\xi = \rho dt$$

where all of the parameters are assumed to be constant. The correlation between the driving Brownian motions implies that ξ_t and χ_t have a joint normal distribution and so it is sufficient to obtain their mean and covariance matrix to characterize their distribution. Using the solution to the Ornstein-Uhlenbeck process outlined in the previous section we have the result that χ_t has the mean and variance given respectively by:

$$\mathbb{E}[\chi_t] = \chi_0 e^{-\kappa t} + \frac{\lambda_\chi}{\kappa}(1-e^{-\kappa t})$$

$$Var(\chi_t) = \frac{\sigma_\chi^2}{2\kappa}(1-e^{-2\kappa t})$$

For the long-term price ξ_t we can simply integrate the SDE to obtain its mean and variance as follows

$$\xi_t = \xi_0 + \int_0^t \mu_\xi - \lambda_\xi ds + \int_0^t \sigma_\xi dB_s^\xi$$

$$= \xi_0 + t(\mu_\xi - \lambda_\xi) + N(0, t\sigma_\xi^2)$$

This gives us the following expressions:

$$\mathbb{E}[\xi_t] = \xi_0 + t(\mu_\xi - \lambda_\xi)$$

$$Var(\chi_t) = t\sigma_\xi^2$$

The only remaining item in the model is the covariance between the factors:

$$Cov(\chi_t, \xi_t) = \mathbb{E}[(\chi_t - \mathbb{E}[\chi_t])(\xi_t - \mathbb{E}[\xi_t])]$$
$$= \mathbb{E}\left[\left(\int_0^t e^{-\kappa(t-s)}\sigma_\chi dB_s^\chi\right)\cdot\left(\int_0^t \sigma_\xi dB_s^\xi\right)\right]$$

To compute the last expectation term, we express the Brownian motion B_t^ξ by using the Brownian motion B_t^χ and a Brownian motion independent to B_t^χ denoted as $B_t^\perp$:

$$B_t^\xi = \rho B_t^\chi + \sqrt{1-\rho^2}B_t^\perp$$

Using this result, we can split the Ito integral from ξ_t into two parts:

$$\int_0^t \sigma_\xi dB_s^\xi = \int_0^t \sigma_\xi \rho dB_s^\chi + \int_0^t \sigma_\xi\sqrt{1-\rho^2}B_s^\perp$$

Substituting this result into the previous covariance term we obtain the following expression:

$$Cov(\chi_t, \xi_t) = \mathbb{E}\left[\left(\int_0^t e^{-\kappa(t-s)}\sigma_\chi dB_s^\chi\right)\cdot\left(\int_0^t \sigma_\xi \rho dB_s^\chi\right)\right]$$
$$+ \mathbb{E}\left[\left(\int_0^t e^{-\kappa(t-s)}\sigma_\chi dB_s^\chi\right)\cdot\left(\int_0^t \sigma_\xi\sqrt{1-\rho^2}B_s^\perp\right)\right]$$

The second expectation in the above equation is equal to zero as the Brownian motions $B_s^\perp$ and B_s^χ are independent of each other. For the first expectation we can apply the technique of integration by parts to obtain a simpler expression for the covariance term:

$$Cov(\chi_t, \xi_t) = \int_0^t e^{-\kappa(t-s)}\sigma_\chi\sigma_\xi\rho ds$$
$$= \frac{\sigma_\chi\sigma_\xi\rho}{\kappa}(1-e^{-\kappa t})$$

Since $\xi_t + \chi_t$ is normally distributed we have the result that the spot price $S_t = \exp(\xi_t + \chi_t)$ has a log normal distribution and the futures price is therefore given by:

$$\mathbb{E}[S_t] = \exp\left(\mathbb{E}[\xi_t + \chi_t] + \frac{1}{2}Var(\xi_t + \chi_t)\right)$$

where the expectation and variance terms, respectively, are:

$$\mathbb{E}[\xi_t + \chi_t] = \xi_0 + t(\mu_\xi - \lambda_\xi) + \chi_0 e^{-\kappa t} + \frac{\lambda_\chi}{\kappa}(1-e^{-\kappa t})$$

$$Var(\xi_t + \chi_t) = Var(\xi_t) + Var(\chi_t) + 2Cov(\xi_t, \chi_t)$$

2.2.3.3 Long-Term Price and Spot Price Model (Gabillon, 1991)

Another two-factor model which incorporates a long-term price factor was outlined in Gabillon (1991). Unlike the previous model however, the model proposed by Gabillon specifies a spot price

directly instead of using a short-term deviation from the long-term price. The resulting model is described by the following SDEs:

$$dS_t = k(\ln(L_t) - \ln(S_t))S_t dt + \sigma_S S_t dB_t^S$$

$$dL_t = \mu L_t dt + \sigma_L L_t dB_t^L$$

$$dB_t^S dB_t^L = \rho dt$$

where S_t is the spot price, L_t is the long-term price and other parameters in the model are assumed to be constant. The model assumes that neither factor is tradeable and there is a market price of risk for each of them. A key assumption used for future prices in this model is that as maturity tends to infinity, future prices behave as the long-term price and the influence of the spot price diminishes.

Consider the futures price $F(S, L, \tau)$ as a function of the factors and time to maturity τ. By applying Ito's lemma, we can express the dynamics of the futures price F as:

$$dF_t = \alpha(S, L, \tau)dt + \gamma_S(S, L, \tau)dB_t^S + \gamma_L(S, L, \tau)dB_t^L$$

We can now construct a portfolio long one unit of F and short two other futures with distinct maturities such that the portfolio is risk free. This portfolio then implies an equilibrium condition that must be satisfied by the futures prices:

$$\alpha(S, L, \tau) = \lambda_S(S, L)\gamma_S(S, L, \tau) + \lambda_L(S, L)\gamma_L(S, L, \tau)$$

where λ_S and λ_L are the market prices of risk of spot and long-term price respectively.

The assumption on the behavior of the future prices as time to maturity goes to infinity then implies:

$$\lim_{\tau\to\infty} \alpha(S, L, \tau) = \lim_{\tau\to\infty} \lambda_S(S, L)\gamma_S(S, L, \tau) + \lambda_L(S, L)\gamma_L(S, L, \tau)$$

$$\mu L_t = \lambda_S(S, L) \times 0 + \lambda_L(S, L)\sigma_L L_t$$

$$\mu L_t = \lambda_L(S, L)\sigma_L L_t$$

where the limits follow from the assumption that the behavior of future prices becomes that of the long-term price as time to maturity tends to infinity. The last equation yields an expression for the market price of risk of long-term price in the model:

$$\lambda_L(S, L) = \frac{\mu}{\sigma}$$

The equilibrium condition in the model provides a PDE for the futures price:

$$\left(k\left(\ln\left(\frac{L_t}{S_t}\right)\right) - \lambda_S\sigma_S\right)S_t\frac{\partial F}{\partial S} + \frac{(\sigma_S S_t)^2}{2}\frac{\partial^2 F}{\partial S^2} + (\mu - \lambda_L\sigma_L)L_t\frac{\partial F}{\partial L} + \frac{(\sigma_L L_t)^2}{2}\frac{\partial^2 F}{\partial L^2} + \rho\sigma_S\sigma_L\frac{\partial^2 F}{\partial L\partial S} = \frac{\partial F}{\partial \tau}$$

Applying the assumption concerning the behavior of F as time to maturity tends to infinity, we can substitute the value for λ_L obtained from the limits given above and the PDE simplifies to the following expression:

$$\left(k\left(\ln\left(\frac{L_t}{S_t}\right)\right)-\lambda_S\sigma_S\right)S_t\frac{\partial F}{\partial S}+\frac{(\sigma_S S_t)^2}{2}\frac{\partial^2 F}{\partial S^2}+\frac{(\sigma_L L_t)^2}{2}\frac{\partial^2 F}{\partial L^2}+\rho\sigma_S\sigma_L\frac{\partial^2 F}{\partial L\partial S}=\frac{\partial F}{\partial \tau}$$

subject to the initial condition that $F(S,L,0)=S$.

This PDE has a known analytical solution of the form $F(S,L,\tau)=A(\tau)S^{B(\tau)}L^{1-B(\tau)}$ where A and B are the solutions to a system of ODEs in τ.

2.2.3.4 The Equivalence of the Models

The three two-factor models presented above can be shown to be equivalent to each other in the sense that their factors can be expressed as linear combinations of one another. The initial work by Schwartz and Smith (2000) showed that their model was formally equivalent to the model of Gibson and Schwartz. To this end they used a change of variable to obtain the following linear combinations:

$$\chi_t=\frac{1}{\kappa}(\delta_t-\alpha)$$

$$\xi_t=\ln(S_t)-\chi_t$$

Further work published by Wang (2012) extended this equivalency to the Gabillon model by showing it is formally equivalent to the model of Schwartz and Smith (2000). The equivalency was shown through the following linear combinations:

$$\chi_t=\ln(S_t)-\ln(L_t)$$

$$\xi_t=\ln(L_t)$$

Using the above linear combinations and applying the dynamics of the respective models, the parameters of each model can be expressed using the parameters of the others. These results show that while each model takes a different approach to modelling commodity markets, they nonetheless share the same mathematical structure. An example of the diversity in forward curves are shown below in Figure 5 and Figure 6 by using the Schwartz and Smith Model.

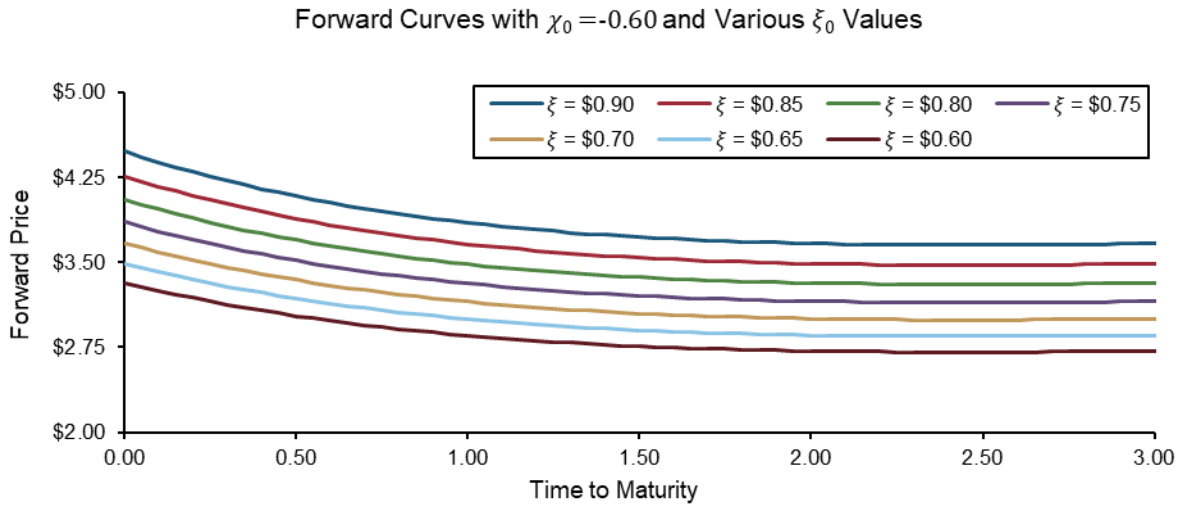


Figure 5: Forward Curves of the Schwartz and Smith Model in Backwardation

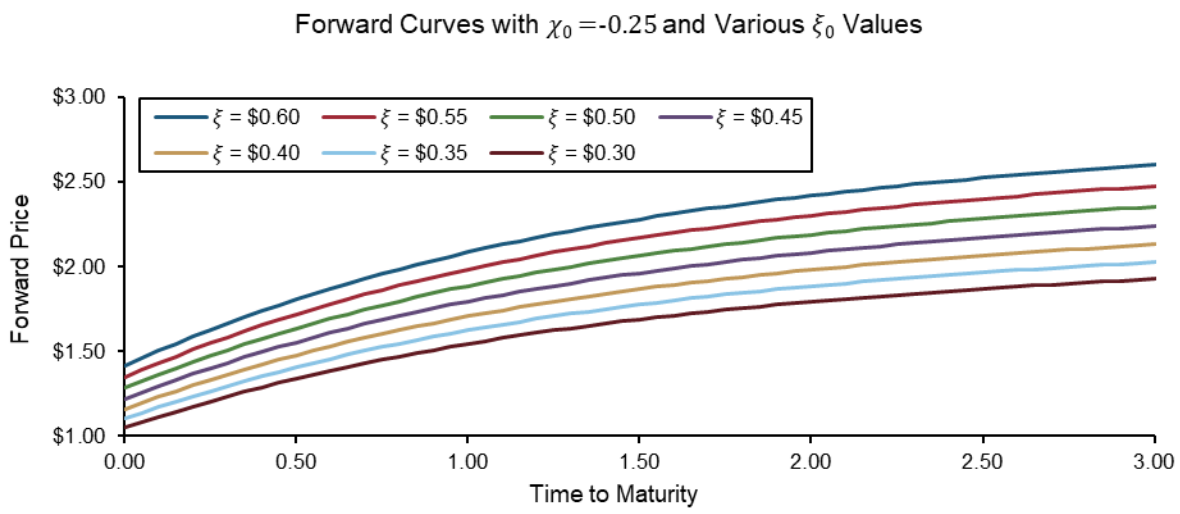


Figure 6: Forward Curves of the Schwartz and Smith in Contango

2.2.4 Volatility Modulated Lévy-driven Volterra Framework

The new model put forth in this thesis is an application of a framework developed in Barndorff-Nielsen, Benth, and Veraart (2013) which focuses on an application of VMLV processes to energy spot prices. VMLV processes were first introduced by Barndorff-Nielsen and Schmiegel (2008) in the context of modelling turbulence in physics and stochastic volatility in finance. The class of VMLV processes includes many classical spot price models of commodities such as the one-factor model in Schwartz (1997) as well as continuous time autoregressive moving average models (CARMA) models. An important subclass of VMLV models known as Lévy Semistationary (LSS) processes were introduced in Barndorff–Nielsen, Benth, and Veraart (2010). They utilized LSS processes in the context of non-storable commodities with applications focused primarily on electricity prices.

While the LSS processes in general are not semi-martingales, it can be shown that with additional conditions on their kernel functions and stochastic volatility terms, the LSS process does indeed have a semi-martingale representation (Barndorff-Nielsen, Benth, & Veraart, 2013). Before relaying the result in question, we first introduce the LSS processes.

A process Y_t is said to follow an LSS process if:

$$Y_t = \mu + \int_{-\infty}^{t} g(t-s)\omega_s dL_s + \int_{-\infty}^{t} q(t-s)\, a_s ds$$

where:

- μ is a constant
- L_s is a Lévy process
- $g, q: \mathbb{R} \to \mathbb{R}$, are non-negative functions such that $g(x) = q(x) = 0$ for $x \leq 0$
- ω_s and a_s are càdlàg processes

Proposition 3 from Barndorff-Nielsen, Benth, & Veraart (2013) provide a set of sufficient conditions to ensure that the LSS process is a semi-martingale. Let Y_t be an LSS process as defined above. Then if Y_t satisfies the following conditions it is a semi-martingale:

(i) $\mathbb{E}[|L_1|] < \infty$
(ii) The limits $g(0+)$ and $q(0+)$ exist and are finite
(iii) g is absolutely continuous and g' is square integrable
(iv) The process $(g'(t-s)\omega_{s-})_{s\in\mathbb{R}}$ is square integrable for any $t \in \mathbb{R}$
(v) The process $(q'(t-s)a_{s-})_{s\in\mathbb{R}}$ is integrable for any $t \in \mathbb{R}$
(vi) The processes ω_{s-} and a_{s-} are independent of L_s

We remark that, while the last property is not explicitly stated for this proposition, (vi) is assumed throughout (Barndorff-Nielsen, Benth, & Veraart, 2013). The supporting paper (Barndorff-Nielsen & Schmiegel, 2009) however utilizes a different condition instead, namely the stationarity of ω_{s-} and a_{s-}. We will proceed with our discussion by using property (vi) as opposed to using the concept of stationarity.

Under these conditions, proposition 3 implies that Y_t has the semi-martingale representation given below:

$$Y_t = Y_0 + g(0+)\int_0^t \omega_{s-} dL_s^* + \int_0^t A_s ds$$

$$A_s = g(0+)\omega_{s-}\mathbb{E}[L_1] + \int_{-\infty}^{s} g'(s-u)\omega_u dL_u + q(0+)a_s + \int_{-\infty}^{s} q'(s-u)\, a_u du$$

where $L_s^* = L_s - \mathbb{E}[L_s]$.

2.2.4.1 The Gabillon Model as an LSS Process

Using the sufficient conditions stated in the previous section, we will provide a new result showing that the Gabillon Model belongs to the LSS framework albeit with a reduced parameter set.

We start our discussion by specifying the kernel functions g and q where we set $g(t-s) = q(t-s) = e^{-\alpha(t-s)}$ for $t > s$ and 0 otherwise. The volatility process ω_s can be set to be a constant, $\omega_s = \sigma_s$. For the drift process, we let α_s be a Brownian motion with drift which can be thought of as the logarithm of the long-term price factor. Lastly, we let the driving Lévy process to be a Brownian motion independent of the process driving a_s and $\mu = 0$. The resulting process then becomes:

$$Y_t = \int_{-\infty}^{t} e^{-\alpha(t-s)}\sigma_s dB_s^Y + \int_{-\infty}^{t} e^{-\alpha(t-s)}\, a_s ds$$

$$da_s = \mu_a dt + \sigma_a dB_t^a$$

$$dB_t^a dB_s^Y = 0$$

Using the semi-martingale representation of proposition 3 from Barndorff-Nielsen, Benth, & Veraart (2013), we find that:

$$Y_t = Y_0 + g(0+)\int_0^t \sigma_s dB_s^Y + \int_0^t A_s\, ds$$

with A_s as

$$A_s = g(0+)\sigma_s \mathbb{E}[B_1^Y] - \alpha \int_{-\infty}^{s} e^{-\alpha(t-s)} \sigma_s dB_s^Y + q(0+)a_s - \alpha \int_{-\infty}^{s} e^{-\alpha(t-s)}\, a_s du$$

$$= 0 - \alpha \left(\int_{-\infty}^{s} e^{-\alpha(t-s)} \sigma_s dB_s^Y + \int_{-\infty}^{s} e^{-\alpha(t-s)}\, a_s du \right) + a_s$$

$$= \alpha \left(\frac{a_s}{\alpha} - Y_s \right)$$

As an SDE this can be written as:

$$dY_t = \alpha \left(\frac{a_t}{\alpha} - Y_t \right) dt + \sigma_s dB_s^Y$$

Further, the Gabillon model can be obtained by defining the spot price S_t and the long-term price L_t as follows and rearranging parameters:

$$S_t = \exp(Y_t)$$

$$L_t = \exp\left(\frac{a_s}{\alpha} \right)$$

It remains to show that conditions (i)-(vi) hold true:

(i) As B_1^Y is normally distributed we have that $\mathbb{E}\left[\left|B_1^Y\right|\right] < \infty$.

(ii) We have that $g(x) = q(x) = e^{-\alpha x} I(x > 0)$ hence $g(0+) = q(0+) = 1$ and both limits exist.

(iii) As g is an exponential function it is absolutely continuous on $\mathbb{R}$ and g' is square integrable since $g(x) = 0$ for $x < 0$.

(iv) As the process ω_s is constant, the square integrability of $(g'(t-s)\omega_{s-})_{s\in\mathbb{R}}$ follows from condition (iii).

(v) The process a_s is a Brownian motion with drift and we should show that $\left(-\alpha e^{-\alpha(t-s)} a_{s-}\right)_{s\in\mathbb{R}}$ is integrable for any $t \in \mathbb{R}$. The proof of this condition is shown below.

(vi) By construction we have that a_s is independent of B^Y and hence the condition is satisfied.

Before showing that condition (v) holds we first introduce additional definitions and notations. As we are working with the integrability of a stochastic process over $\mathbb{R}$ and not a finite interval we provide conditions equivalent to their integrability. Such conditions are outlined in Barndorff-Nielsen, Benth, & Veraart (2013).

Let L be a Lévy process with triplet (c, b, Π). The parameters of the triplet can be viewed through the lens of the Lévy-Ito decomposition where a Lévy process is shown to be a sum of three independent processes: a constant drift with slope $c \in \mathbb{R}$, a Brownian motion scaled by the volatility $b \in \mathbb{R}_{\geq 0}$ and a sum of an at most countable collection of compound Poisson processes characterized by the measure Π on $\mathbb{R} \setminus \{0\}$. To integrate over $\mathbb{R}$, we also need to extend the Lévy process to the entire real line. While a Lévy process L is typically only defined on $(0, \infty)$, it can be extended to $\mathbb{R}$ by considering an independent copy of L, denoted L^*. Using L^* we can then extend L to a two-sided Lévy process:

$$L_t = \begin{cases} L_t, & t \in (0, \infty) \\ -L^*_{-t}, & t \in (-\infty, 0\ ) \end{cases}$$

For $t \in \mathbb{R}$, a stochastic process $(\phi_s)_{s \le t}$ is integrable with respect to the two-sided Lévy process L if and only if the following conditions hold almost surely:

1. $\int_{-\infty}^{t} b\phi_s^2 \, ds < \infty$
2. $\int_{-\infty}^{t} \int_{\mathbb{R}} (1 \wedge |z\phi_s|^2) \Pi(dz) \, ds < \infty$
3. $\int_{-\infty}^{t} |c\phi_s + \int_{\mathbb{R}} (I(|z\phi_s| \le 1) - \phi_s I(|z| \le 1)) \Pi(dz)| \, ds < \infty$

In looking at condition (v) and the integrability of the process $\left(-\alpha e^{-\alpha(t-s)} a_{s-}\right)_{s \in \mathbb{R}}$, we can consider it as an integral with respect to a Lévy process defined by the triplet $(1, 0, 0)$. Such a Lévy process is simply a drift with a constant slope of 1. The integrability is then entirely determined by the third condition as $b = 0$ and $\Pi = 0$. It follows that it is sufficient to show that:

$$\int_{-\infty}^{t} \left| -\alpha e^{-\alpha(t-s)} a_{s-} \right| ds = \alpha \int_{-\infty}^{t} \left| e^{-\alpha(t-s)} (\mu_a s + \sigma_a B_s^a) \right| ds < \infty$$

holds almost surely.

Without loss of generality, we will show that the integrability condition holds for $t > 0$. Using the triangle inequality and the linearity of integration we can decompose the integral into stochastic and non-stochastic parts:

$$\alpha \int_{-\infty}^{t} \left| e^{-\alpha(t-s)} (\mu_a s + \sigma_a B_s^a) \right| ds \le \alpha \int_{-\infty}^{t} \left| e^{-\alpha(t-s)} \mu_a s \right| ds + \alpha \int_{-\infty}^{t} \left| e^{-\alpha(t-s)} \sigma_a B_s^a \right| ds$$

We begin the proof by showing that the non-stochastic integral must be finite. To this end, we further split the integral along its domain:

$$\int_{-\infty}^{t} \left| e^{-\alpha(t-s)} \mu_a s \right| ds = \int_{0}^{t} \left| e^{-\alpha(t-s)} \mu_a s \right| ds + \int_{-\infty}^{0} \left| e^{-\alpha(t-s)} \mu_a s \right| ds$$

The first term is the integral of a continuous function over an interval of finite measure and must be finite. For the second integral, an application of the method of integration by parts yields the following result:

$$\begin{aligned} \int_{-\infty}^{0} \left| e^{-\alpha(t-s)} \mu_a s \right| ds &= -\mu_a e^{-\alpha t} \int_{-\infty}^{0} e^{\alpha s} s \, ds \\ &= -\mu_a e^{-\alpha t} \left[\frac{e^{\alpha s}(\alpha s - 1)}{\alpha^2} \right]_{-\infty}^{0} \\ &= -\mu_a e^{-\alpha t} \left[-\frac{1}{\alpha^2} - \lim_{s \to -\infty} \frac{e^{\alpha s}(\alpha s - 1)}{\alpha^2} \right]_{-\infty}^{0} \\ &= \frac{\mu_a e^{-\alpha t}}{\alpha^2} < \infty \end{aligned}$$

Combining the two results shows that:

$$\int_{-\infty}^{t} \left| e^{-\alpha(t-s)} \mu_a s \right| ds < \infty$$

almost surely.

We now show that the stochastic part is finite almost surely. To show this result we rely on the law of iterated logarithm which states that if B_t is a standard Brownian motion then almost surely we have:

$$\limsup_{t \to \infty} \frac{B_t}{\sqrt{2t \log(\log(t))}} = 1$$

This implies that for any $\epsilon > 0$ there exists almost surely t_ϵ such that the Brownian motion is constrained to the envelope:

$$|B_t| \leq (1+\epsilon)\sqrt{2t \log(\log(t))}$$

for any $t > t_\epsilon$. The choice of t_ϵ depends on the path of the Brownian motion under consideration. This bound can readily be extended to a two-sided Brownian motion by simply mirroring the bound about $t = 0$. For a two-sided Brownian motion, the envelope can be expressed as:

$$|B_t| \leq (1+\epsilon)\sqrt{2|t| \log(\log(|t|))}$$

for any t such that $|t| > |t_\epsilon|$.

Taking $\epsilon > 0$ and using the law of iterated logarithm we can, almost surely, split the stochastic part as:

$$\begin{aligned} \int_{-\infty}^{t} \left| e^{-\alpha(t-s)} \sigma_a B_s^a \right| ds &= \sigma_a \int_{-\infty}^{t_\epsilon} e^{-\alpha(t-s)} |B_s^a| \, ds + \sigma_a \int_{t_\epsilon}^{t} e^{-\alpha(t-s)} |B_s^a| \, ds \\ &\leq \sigma_a (1+\epsilon) \int_{-\infty}^{t_\epsilon} e^{-\alpha(t-s)} \sqrt{2|s| \log(\log(|s|))} \, ds + \sigma_a \int_{t_\epsilon}^{t} |B_s^a| \, ds \end{aligned}$$

where in the first integral we have used the law of iterated logarithm to bound the absolute value of the Brownian motion. Using the two-sided nature of the Brownian motion, we assume without loss of generality that $t_\epsilon < 0$. Again, we proceed with each integral individually. Starting with the integral on $(-\infty, t_\epsilon)$ we have:

$$\begin{aligned} \sigma_a (1+\epsilon) \int_{-\infty}^{t_\epsilon} e^{-\alpha(t-s)} \sqrt{2|s| \log(\log(|s|))} \, ds &\leq \sigma_a (1+\epsilon) \int_{-\infty}^{t_\epsilon} e^{-\alpha(t-s)} \sqrt{2|s|^2} \, ds \\ &= -\sigma_a (1+\epsilon) e^{-\alpha t} \sqrt{2} \int_{-\infty}^{t_\epsilon} e^{\alpha s} s \, ds \end{aligned}$$

where we have used the result that $\log(t) < t$ for $t > 0$ and that $\log$ is a monotonically increasing function. The integral can now be computed by using integration by parts:

$$\int_{-\infty}^{t_\epsilon} e^{\alpha s} s \, ds = \left[\frac{e^{\alpha s}(\alpha t_\epsilon - 1)}{\alpha^2} - \lim_{s \to -\infty} \frac{e^{\alpha s}(\alpha s - 1)}{\alpha^2} \right]_{-\infty}^{t_\epsilon}$$

$$= \frac{e^{\alpha s}(\alpha t_\epsilon - 1)}{\alpha^2}$$

which then gives us:

$$\sigma_a(1+\epsilon)\int_{-\infty}^{t_\epsilon} e^{-\alpha(t-s)}\sqrt{2|s|\log(\log(|s|))}\ ds \leq -\sigma_a(1+\epsilon)e^{-\alpha t}\sqrt{2}\frac{e^{\alpha s}(\alpha t_\epsilon - 1)}{\alpha^2} < \infty$$

For the second integral:

$$\int_{t_\epsilon}^{t} |B_s^a|\ ds$$

we note that the symmetry of the two-sided Brownian motion about zero implies that it is sufficient to show the result for the interval $(0, t)$ with an arbitrary $t > 0$ and a standard Brownian motion B_t:

$$\int_0^t |B_s|\ ds < \infty$$

To establish this result, we will show that the expectation of this integral is finite. The expectation can be bounded as follows:

$$\mathbb{E}\left[\int_0^t |B_s|\ ds\right] \leq t\mathbb{E}\left[\sup_{s\in[0,t]} |B_s|\right]$$

$$= t\int_0^\infty \mathbb{P}\left(\sup_{s\in[0,t]} |B_s| > x\right) dx$$

where the equality holds as the supremum is positive almost surely. The integrand can be bounded since:

$$\left\{\sup_{s\in[0,t]} |B_s| > x\right\} \subset \left\{\sup_{s\in[0,t]} B_s > x\right\} \cup \left\{\inf_{s\in[0,t]} B_s < x\right\}$$

and hence, after a further application of the symmetry of the standard Brownian motion about zero, we obtain the following result:

$$\mathbb{P}\left(\sup_{s\in[0,t]} |B_s| > x\right) \leq \mathbb{P}\left(\sup_{s\in[0,t]} B_s > x\right) + \mathbb{P}\left(\inf_{s\in[0,t]} B_s < x\right)$$

$$= 2\mathbb{P}\left(\sup_{s\in[0,t]} B_s > x\right)$$

The reflection principle gives that:

$$\sup_{s\in[0,t]} B_s \overset{d}{=} |B_t|$$

and we can apply Chebyshev's inequality to complete the bound on the integrand as follows:

$$\mathbb{P}\left(\sup_{s\in[0,t]} B_s > x\right) = \mathbb{P}(|B_t| > x)$$

$$= \mathbb{P}\left(\left|\frac{B_t}{\sqrt{t}}\right| > \frac{x}{\sqrt{t}}\right)$$

$$\leq \left(\frac{\sqrt{t}}{x}\right)^2$$

Combining the observations made above shows that the expectation must be finite:

$$\mathbb{E}\left[\int_0^t |B_s|\, ds\right] \leq t\int_0^\infty \mathbb{P}\left(\sup_{s\in[0,t]} |B_s| > x\right) dx$$

$$\leq 2t\int_0^\infty \mathbb{P}\left(\sup_{s\in[0,t]} B_s > x\right) dx$$

$$\leq 2t\int_0^\delta \mathbb{P}\left(\sup_{s\in[0,t]} B_s > x\right) dx + 2t\int_\delta^\infty \mathbb{P}\left(\sup_{s\in[0,t]} B_s > x\right) dx$$

$$\leq 2t\int_0^\delta 1\, dx + 2t\int_\delta^\infty \frac{t}{x^2}\, dx$$

$$\leq 2t\delta + 2t^2\left[\frac{-1}{x}\right]_\delta^\infty$$

$$= 2t\delta + \frac{2t^2}{\delta}$$

$$< \infty$$

At last, we have shown that the stochastic part of our initial integral is almost surely finite:

$$\int_{-\infty}^t \left|e^{-\alpha(t-s)}\sigma_a B_s^a\right| ds < \infty$$

and combining this result with the non-stochastic part being finite almost surely completes the proof of the integrability of the process:

$$\left(-\alpha e^{-\alpha(t-s)} a_{s-}\right)_{s\in\mathbb{R}}$$

This concludes the proof of the validity of the semi-martingale representation of the LSS process used to show that the Gabillon model belongs to the LSS framework.

The main limitation of the argument above is that this only shows that a subset of Gabillon models belong to the class of LSS process as we have assumed that the correlation between the two factors is zero (which corresponds to the restriction that $\rho = 0$). Nonetheless, this argument shows that this type of model can belong to the LSS framework albeit with a more restrictive set of parameters.

2.2.4.2 Other Features of the LSS Framework

In addition to including existing models, the LSS framework has a number of useful features when it comes to modelling commodity prices, namely a flexible autocorrelation structure, stochastic volatility and an arbitrary driving Lévy process. Through the choice of the kernel functions, the

framework provides significant control over the autocorrelation structure of the process (for example CARMA processes). The choice of the kernel also extends to the shape of the resulting Samuelson effect, where the volatility of futures is observed to decrease as time to maturity increases (Barndorff-Nielsen, Benth, & Veraart, 2013). By construction, the VMLV framework includes stochastic volatility, which as discussed above has been observed empirically in commodity markets (Back & Prokopczuk, 2013). Lastly, the framework applies with an arbitrary Lévy process and as a result can model jumps and discontinuities in the spot prices which have been used by a number of authors in the context of commodities (Back & Prokopczuk, 2013).

Chapter 3

Application of VMLV to Storable Commodities

In the previous chapter we have shown that the LSS framework includes a subset of common two-factor models such as the Gabillon model with some restriction on their parameters. In this section we propose a new model for commodity prices based on the LSS framework and the theory of storage. We provide an economic rationale, a semi-martingale representation along with two approaches to valuing derivatives under the proposed model.

3.1 Proposed Model

The proposed model relies on the LSS framework to link changes in inventory directly to changes in prices and volatility. Critically, the memory of the process is used in the proposed model to tally the changes of inventory over time and can be interpreted as the current inventory levels. The effect of inventory changes decays over time with the most recent changes carrying the most weight in the drift and volatility terms. The changes in inventory themselves are modelled as a mean reverting process to reflect the nature of inventory dynamics subject to demand and supply conditions over time.

3.1.1 Inventory Process

The changes in inventory process, X_t, is modelled using a mean reverting process. As we are modelling the changes in inventory, these changes can in principle be negative or positive and so a suitable candidate is given by an Ornstein-Uhlenbeck process:

$$dX_t = \kappa(\theta - X_t)dt + \sigma_X dB_t^X$$

3.1.2 Spot Price Process

The spot price process, S_t, is modelled as:

$$S_t = \exp(Y_t)$$

where Y_t is a LSS process:

$$Y_t = \int_{-\infty}^{t} g(t-s)\omega_s dL_s + \int_{-\infty}^{t} q(t-s)\, a_s ds$$

Using the càdlàg processes ω and a, we model the relationship between inventory changes and spot prices via the functions $H_1, H_2 : \mathbb{R} \to \mathbb{R}$ such that:

$$\omega_s := H_1(X_s)$$

$$a_s \coloneqq H_2(X_s)$$

These functions should satisfy the following properties in relation to inventory changes:

- The function H_1 should, while being positive on $\mathbb{R}$, take sufficiently small values when X_t is "large" and vice versa
    - In other words:
        - When inventories increase, volatility is expected to decrease
        - When inventories decrease, volatility is expected to increase
- The function H_2 should take on relatively large positive values when X_t is small and take relatively large negative values when X_t is small
    - This can be interpreted as follows:
        - When inventories increase, prices are expected to decrease
        - When inventories decrease, prices are expected to increase

For simplicity we will let the functions H_1 and H_2 to take the following form for the remainder of this thesis:

$$H_i(X) = \frac{a_i}{1 + e^{c_i(X - d_i)}} + b_i \quad i = 1,2$$

Further, we will assume that H_1 is bounded strictly above zero. Hence for some $m > 0$ we have:

$$H_1(X) > m$$

for any $X \in \mathbb{R}$. As will be shown later, this technical assumption will be helpful to characterize the risk neutral measure.

In terms of the kernel functions g and q, we are looking to model the decay over time of the effect of a change in inventory on spot prices. To this end, we take both g and q to be the same function:

$$g(t-s) = q(t-s) = \begin{cases} e^{-\alpha(t-s)} & t - s > 0 \\ 0 & otherwise \end{cases}$$

where α is a positive constant. With this choice, the contribution of ω_s and a_s to their corresponding integral for a fixed s will be at its maximum when $t = s$ and decay to zero as $t - s$ increases. Moreover, as will be shown in the next section, this choice of the kernel function has convenient technical implications for the semi-martingale representation of Y_t.

Lastly, to fully characterize the process Y_t, we should also specify a driving Lévy process. For simplicity we will let the Lévy process be a Brownian motion ($L \coloneqq B^Y$). As the formulation of the model requires that the càdlàg processes ω and a be independent of L, we will let the Brownian motion B^Y be independent of the Brownian motion in the inventory process (B^X).

Putting all of the above results together, we obtain the following LSS model for Y_t:

$$Y_t = \int_{-\infty}^{t} e^{-\alpha(t-s)} H_1(X_s) dB_s^Y + \int_{-\infty}^{t} e^{-\alpha(t-s)} H_2(X_s) ds$$

3.2 Semi-martingale Representation

Using proposition 3 in Barndorff-Nielsen, Benth, and Veraart (2013), the proposed model can be expressed as a semi-martingale provided that conditions (i)-(vi) in the previous section are

satisfied. To obtain changes of measure where the density process can be shown to be a martingale for pricing purposes, we will use the assumption that the function H_1 is bounded as follows:

$$0 < m < H_1(x) < M$$

for any $x \in \mathbb{R}$ and where M and m are constants. These assumptions, combined with the choice of the kernel functions and the Lévy process are sufficient to ensure the existence of the semi-martingale representation as shown below:

(i) $\mathbb{E}[|L_1|] < \infty$
   a. As the Lévy process is a standard Brownian motion, $L_1 = B_1$ and $\mathbb{E}[|B_1|] = \frac{\sqrt{2}}{\sqrt{\pi}} < \infty$

(ii) The limits $g(0+)$ and $q(0+)$ exist and are finite
   a. This is clear as $g(x) = q(x) = \exp(-\alpha x)$ which is continuous at 0

(iii) g is absolutely continuous and g' is square integrable
   a. g is absolutely continuous as it is either constant or exponential and since $\alpha > 0$ we have that g' is square integrable

(iv) The process $(g'(t-s)\omega_{s-})_{s\in\mathbb{R}}$ is square integrable for any $t \in \mathbb{R}$
   a. The square integrability follows from condition (iii) as $\omega_s = H_1(X_s)$ is bounded

(v) The process $(q'(t-s)a_{s-})_{s\in\mathbb{R}}$ is integrable for any $t \in \mathbb{R}$
   a. The integrability holds as a is bounded and q' is integrable

(vi) The processes ω_{s-} and a_{s-} are independent of L_s
   a. Satisfied by construction as we take B_s^X independent of $L_s = B_s^Y$

The proposition implies that the semi-martingale representation takes the form below:

$$Y_t = Y_0 + g(0+)\int_0^t H_1(X_s)dL_s + \int_0^t A_s\,ds$$

with A_s given by:

$$A_s = g(0+)H_1(X_s)\mathbb{E}[L_1] - \alpha\int_{-\infty}^{s} e^{-\alpha(s-u)}H_1(X_u)dL_u + q(0+)H_2(X_s) - \alpha\int_{-\infty}^{s} e^{-\alpha(s-u)}H_2(X_u)du$$

Since $\mathbb{E}[L_1] = \mathbb{E}[B_1^Y] = 0$ and given that $g(0+) = q(0+) = 1$, we obtain the following simplification of A_s:

$$A_s = -\alpha Y_s + H_2(X_s)$$

with the resulting semi-martingale representation being written as:

$$Y_t = Y_0 + \int_0^t H_1(X_s)dL_s + \int_0^t -\alpha Y_s + H_2(X_s)\,ds$$

which can further be expressed as the following Stochastic Differential Equation (SDE):

$$dY_t = \alpha\left(\frac{H_2(X_t)}{\alpha} - Y_t\right)dt + H_1(X_t)dB_t^Y$$

$$= \alpha\left(H_2^*(X_t) - Y_t\right)dt + H_1(X_t)dB_t^Y$$

where $H_2^*(X_t) = \frac{H_2(X_t)}{\alpha}$ is simply a re-parameterization of $H_2(X)$.

The resulting model can be viewed as a generalization of the one-factor Schwartz model.

3.3 Risk Neutral Measures & Market Price of Risk

To use the model in the context of pricing derivatives, we outline a "risk neutral" version of the model where each factor (S and X) is adjusted by a market price of risk. To apply the market price of risk, we conduct a change of measure from the original model under the "real world" measure to the intended "risk neutral" measure. Following the example set out in Schwartz and Smith (2000), we assume that each factor is subject to a constant market price of risk. The resulting change of measure for the proposed model is implemented via Girsanov's theorem.

Under the "real world" measure $\mathbb{P}$, the proposed model satisfies the following system of SDEs:

$$dY_t = \alpha(H_2(X_t) - Y_t)dt + H_1(X_t)dB_t^{Y,\mathbb{P}}$$

$$d\mathrm{X}_t = \kappa(\theta - X_t)dt + \sigma dB_t^{X,\mathbb{P}}$$

$$dB_t^Y dB_t^X = 0$$

which can be expressed as a single two-dimensional SDE

$$d\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} \alpha(H_2(X_t) - Y_t) \\ \kappa(\theta - X_t) \end{bmatrix} dt + \begin{bmatrix} H_1(X_t) & 0 \\ 0 & \sigma \end{bmatrix} d\begin{bmatrix} B_t^{Y,\mathbb{P}} \\ B_t^{X,\mathbb{P}} \end{bmatrix}$$

We now make the ansatz that under the risk neutral measure $\mathbb{Q}$ we have a process l_t such that

$$\begin{bmatrix} B_t^{Y,\mathbb{P}} \\ B_t^{X,\mathbb{P}} \end{bmatrix} = \begin{bmatrix} B_t^{Y,\mathbb{Q}} \\ B_t^{X,\mathbb{Q}} \end{bmatrix} - \int_0^t l_s \, ds$$

where the process l_t is a two-dimensional process and $B_t^{Y,\mathbb{Q}}$, $B_t^{X,\mathbb{Q}}$ are a two-dimensional Brownian motion:

$$l_t = \begin{bmatrix} l_t^Y \\ l_t^X \end{bmatrix}$$

After substituting the change of measure into the original SDE we obtain the dynamics of the resulting process under the measure $\mathbb{Q}$:

$$d\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} \alpha(H_2(X_t) - Y_t) - H_1(X_t)l_t^Y \\ \kappa(\theta - X_t) - \sigma l_t^X \end{bmatrix} dt + \begin{bmatrix} H_1(X_t) & 0 \\ 0 & \sigma \end{bmatrix} d\begin{bmatrix} B_t^{Y,\mathbb{Q}} \\ B_t^{X,\mathbb{Q}} \end{bmatrix}$$

Using our assumption that the market price of risk is a constant, the process l_t^Y and l_t^X must satisfy the following equations

$$H_1(X_t)l_t^Y = \lambda_Y \Rightarrow l_t^Y = \frac{\lambda_Y}{H_1(X_t)}$$

$$\sigma l_t^X = \lambda_X \Rightarrow l_t^X = \frac{\lambda_X}{\sigma}$$

where $\lambda_Y, \lambda_X \in \mathbb{R}$ are the constant market price of risks of Y and X respectively. The resulting density process for this choice of market price of risk is then:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}\left(-\int_0^t l_s \, d\begin{bmatrix} B_s^{Y,\mathbb{P}} \\ B_s^{X,\mathbb{P}} \end{bmatrix}\right) = \exp\left(-\int_0^t l_s^Y \, dB_s^{Y,\mathbb{P}} - \int_0^t l_s^X \, dB_s^{X,\mathbb{P}} - \frac{1}{2}\int_0^t (l_s^X)^2 + (l_s^Y)^2 \, ds\right)$$

The density process, as a Doléan-Dade exponential, is in general only a local martingale. To show that the density process above is indeed a martingale, we rely on the Novikov criterion and show that:

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t (l_s^X)^2 + (l_s^Y)^2 \, ds\right)\right] < \infty$$

From the definition we have that l_t^x is constant and hence trivially bounded. For l_t^Y we have that $H_1 > m > 0$ which implies that:

$$|l_t^Y| < \frac{\lambda^Y}{m}$$

Using the boundedness of l^X and l^Y we can show that the Novikov criterion is satisfied:

$$\begin{aligned}\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t (l_s^X)^2 + (l_s^Y)^2 \, ds\right)\right] &< \mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t \left(\frac{\lambda_X}{\sigma}\right)^2 + \left(\frac{\lambda^Y}{m}\right)^2 ds\right)\right] \\ &= \exp\left(\frac{t}{2}\left(\left(\frac{\lambda_X}{\sigma}\right)^2 + \left(\frac{\lambda^Y}{m}\right)^2\right)\right) \\ &< \infty\end{aligned}$$

and we conclude that the density process $\frac{d\mathbb{Q}}{d\mathbb{P}}$ must be a martingale.

The Girsanov theorem then gives that under the measure $\mathbb{Q}$ the following process:

$$\begin{bmatrix} B_s^{Y,\mathbb{P}} \\ B_s^{X,\mathbb{P}} \end{bmatrix} + \int_0^t l_s \, ds$$

is a Brownian motion which validates the dynamics of the proposed model under the risk neutral dynamics.

3.4 Futures and Derivatives Valuation

The risk neutral version of the model can be used to price derivatives. Let $B = B(S_t, X_t, t)$ be the value of a European derivative with a payoff function $P(S_T, X_T)$ at maturity $T > t$. Let r be the risk-free rate. As we are assuming that the risk-free rate is constant, the value of the derivative B is given by the conditional expectation below:

$$B(S_t, X_t, t) = e^{-r(T-t)}\mathbb{E}[P(S_T, X_T)|S_t, X_t]$$

There are two common methods to evaluate the above conditional expectation. The first is to work with the expectation directly and applying a probabilistic argument to compute the expectation. The second method is to derive and solve a PDE that must be satisfied by the conditional

expectation. In instances where there are no known analytical solutions to the probabilistic or PDE approaches, numerical methods can be used to estimate the solution. For the probabilistic method this entails Monte Carlo simulations and for the PDE approach the use of numerical schemes to approximate the solution of the PDE. The following section derives a PDE representation for the proposed model and outlines a simple Monte Carlo method to approximate the price of European derivatives.

3.4.1 Pricing PDE Representation

In order to obtain European derivative prices, we can derive an equilibrium pricing PDE with its solution being the price of the derivative. We proceed by following the work of Gibson and Schwartz (1990) where they applied the methodology from fixed income markets outlined in Brennan and Schwartz (1979). The work by these authors focused on pricing methodologies with two sources of uncertainty (factors) where some of the factors may not be tradable. This is directly applicable to the proposed model as the factors are not assumed to be traded.

The approach considers an arbitrary European derivative B with a set time to maturity τ and looks to construct a risk-free portfolio to hedge the derivative B. The portfolio is constructed with two other derivatives F_1 and F_2 with respective maturities τ_1 and τ_2 such that $\tau_2 > \tau_1 > \tau$. Two derivatives are needed to hedge B as its value will, in general, depend on both factors.

3.4.1.1 Definitions

Recall the dynamics of Y_t and X_t. The spot price S_t is given by $S_t = e^{Y_t}$ and

$$dY_t = \alpha(H_2(X_t) - Y_t)dt + H_1(X_t)dB_t^Y$$

$$d\mathrm{X}_t = \kappa(\theta - X_t)dt + \sigma dB_t^X$$

$$dB_t^Y dB_t^X = 0$$

For brevity, we can express the dynamics as follows (the dynamics of S_t can be obtained from Ito's lemma applied to $S_t = \ln(Y_t)$) where μ_S, μ_X, σ_S, σ_X are functions of S, X or t in general.

$$dS_t = \mu_S dt + \sigma_S dB_t^Y$$

$$d\mathrm{X}_t = \mu_X dt + \sigma_X dB_t^X$$

3.4.1.2 Dynamics of an Arbitrary Derivative

Considering an arbitrary European derivative with a value given by some function of the factors $C = C(S_t, X_t, t)$ with time to maturity τ and maturity T. Here we consider C a function of t but, as in the literature, we will use a derivative with respect to time to maturity $\tau = T - t$. Applying Ito's lemma, we obtain:

$$dC_t = -\frac{\partial C}{\partial \tau}dt + \frac{\partial C}{\partial S}dS_t + \frac{\partial C}{\partial X}dX_t + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}(dS_t)^2 + \frac{1}{2}\frac{\partial^2 C}{\partial X^2}(dX_t)^2 + \frac{\partial^2 C}{\partial S \partial X}dS_t dX_t$$

$$= \left(-\frac{\partial C}{\partial \tau} + \mu_S\frac{\partial C}{\partial S} + \mu_x\frac{\partial C}{\partial X} + \frac{\sigma_s^2}{2}\frac{\partial^2 C}{\partial S^2} + \frac{\sigma_X^2}{2}\frac{\partial^2 C}{\partial X^2}\right)dt + \sigma_S\frac{\partial C}{\partial S}dB_t^Y + \sigma_X\frac{\partial C}{\partial X}dB_t^X$$

In the second step we remove the covariation term since the model assumes that $dB_t^Y dB_t^X = 0$.

For brevity, and noting the dependence on time to maturity τ, we define the following quantities:

$$\mu(\tau) = \frac{1}{C_t}\left(-\frac{\partial C}{\partial \tau} + \mu_S \frac{\partial C}{\partial S} + \mu_x \frac{\partial C}{\partial X} + \frac{\sigma_s^2}{2}\frac{\partial^2 C}{\partial S^2} + \frac{\sigma_X^2}{2}\frac{\partial^2 C}{\partial X^2}\right)$$

$$\sigma_1(\tau) = \frac{1}{C_t}\sigma_S \frac{\partial C}{\partial S}$$

$$\sigma_2(\tau) = \frac{1}{C_t}\sigma_X \frac{\partial C}{\partial X}$$

Using these expressions, the dynamics of an arbitrary derivative can be expressed in an exponential form as:

$$dC = \mu(\tau)Cdt + \sigma_1(\tau)CdB_t^Y + \sigma_2(\tau)CdB_t^X$$

3.4.1.3 Hedging Portfolio

We now consider an arbitrary European derivative B with time to maturity τ. Using the above result, its dynamics can be expressed as:

$$dB = \mu(\tau)Bdt + \sigma_1(\tau)BdB_t^Y + \sigma_2(\tau)BdB_t^X$$

Lastly, we consider two other derivatives F_1 and F_2 with respective times to maturity given by τ_1 and τ_2. Without loss of generality, we can assume the times to maturity satisfy $\tau_2 > \tau_1 > \tau$. This ensures that the maturities are distinct and that the hedge can be applied at any time before the maturity of B. As F_1 and F_2 are themselves derivatives, their dynamics are given by:

$$dF_1 = \mu(\tau_1)F_1dt + \sigma_1(\tau_1)F_1dB_t^Y + \sigma_2(\tau_1)F_1dB_t^X$$

$$dF_2 = \mu(\tau_2)F_2dt + \sigma_1(\tau_2)F_2dB_t^Y + \sigma_2(\tau_2)F_2dB_t^X$$

We now construct a portfolio Π that is short one unit of B and long x_1, x_2 units of F_1 and F_2 respectively. The value of the portfolio is given by:

$$\Pi = x_1F_1 + x_2F_2 - B$$

From this result, we obtain the dynamics of Π as:

$$d\Pi = x_1dF_1 + x_2dF_2 - dB$$

$$d\Pi = \big(x_1F_1\mu(\tau_1) + x_2F_2\mu(\tau_2) - B\mu(\tau)\big)dt + \big(x_1F_1\sigma_1(\tau_1) + x_2F_2\sigma_1(\tau_2) - B\sigma_1(\tau)\big)dB_t^Y$$

$$+\big(x_1F_1\sigma_2(\tau_1) + x_2F_2\sigma_2(\tau_2) - B\sigma_2(\tau)\big)dB_t^x$$

As we are looking to use F_1 and F_2 to hedge the portfolio, we are looking for x_1 and x_2 such that the two diffusion terms are zero. With such a choice of x_1 and x_2, the resulting portfolio is now risk free and a no arbitrage argument implies that Π should grow at the risk-free rate r. This reasoning yields three equations to be stated below. Since there are three equations and two coefficients, these equations are over-identified and, hence, do not have a unique solution in general.

$$x_1F_1\mu(\tau_1) + x_2F_2\mu(\tau_2) - B\mu(\tau) = r\Pi$$

$$x_1F_1\sigma_1(\tau_1) + x_2F_2\sigma_1(\tau_2) - B\sigma_1(\tau) = 0$$

$$x_1F_1\sigma_2(\tau_1) + x_2F_2\sigma_2(\tau_2) - B\sigma_2(\tau) = 0$$

The two equations from the diffusion terms can be expressed as a linear system of the form $Ax = b$ as:

$$\begin{bmatrix} F_1\sigma_1(\tau_1) & F_2\sigma_1(\tau_2) \\ F_1\sigma_2(\tau_1) & F_2\sigma_2(\tau_2) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} B\sigma_1(\tau) \\ B\sigma_2(\tau) \end{bmatrix}$$

Cramer's rule can be applied to solve this system of equations with the solution shown below:

$$x_1 = \left(\frac{B}{F_1}\right)\left(\frac{\sigma_1(\tau)\sigma_2(\tau_2) - \sigma_1(\tau_2)\sigma_2(\tau)}{\sigma_1(\tau_1)\sigma_2(\tau_2) - \sigma_1(\tau_2)\sigma_2(\tau_1)}\right)$$

$$x_2 = \left(\frac{B}{F_2}\right)\left(\frac{\sigma_1(\tau_1)\sigma_2(\tau) - \sigma_1(\tau)\sigma_2(\tau_1)}{\sigma_1(\tau_1)\sigma_2(\tau_2) - \sigma_1(\tau_2)\sigma_2(\tau_1)}\right)$$

3.4.1.4 Equilibrium Condition

The equation from the drift term of Π can be further rewritten as follows, after substituting the expression for Π and simplifying terms:

$$x_1F_1(\mu(\tau_1) - r) + x_2F_2(\mu(\tau_2) - r) - B(\mu(\tau) - r) = 0$$

From here the solution for x_1 and x_2 in the previous section can be substituted in the drift term equation. We note that the factors of F_1, F_2 and B can be simplified which gives:

$$\left(\frac{\sigma_1(\tau)\sigma_2(\tau_2) - \sigma_1(\tau_2)\sigma_2(\tau)}{\sigma_1(\tau_1)\sigma_2(\tau_2) - \sigma_1(\tau_2)\sigma_2(\tau_1)}\right)(\mu(\tau_1) - r) + \left(\frac{\sigma_1(\tau_1)\sigma_2(\tau) - \sigma_1(\tau)\sigma_2(\tau_1)}{\sigma_1(\tau_1)\sigma_2(\tau_2) - \sigma_1(\tau_2)\sigma_2(\tau_1)}\right)(\mu(\tau_2) - r)$$

$$= \mu(\tau) - r$$

Expressions in both brackets can be reorganized to single out the terms which depend on τ. For brevity, define $D(\tau_1, \tau_2) = \sigma_1(\tau_1)\sigma_2(\tau_2) - \sigma_1(\tau_2)\sigma_2(\tau_1)$ for the denominators:

$$\left(\sigma_1(\tau)\frac{\sigma_2(\tau_2)}{D(\tau_1, \tau_2)} - \sigma_2(\tau)\frac{\sigma_1(\tau_2)}{D(\tau_1, \tau_2)}\right)(\mu(\tau_1) - r) + \left(\sigma_2(\tau)\frac{\sigma_1(\tau_1)}{D(\tau_1, \tau_2)} - \sigma_1(\tau)\frac{\sigma_2(\tau_1)}{D(\tau_1, \tau_2)}\right)(\mu(\tau_2) - r)$$

$$= \mu(\tau) - r$$

Grouping the left-hand side terms of $\sigma_1(\tau)$ and $\sigma_2(\tau)$ we get:

$$\sigma_1(\tau)\left(\frac{\sigma_2(\tau_2)(\mu(\tau_1) - r)}{D(\tau_1, \tau_2)} - \frac{\sigma_2(\tau_1)(\mu(\tau_2) - r)}{D(\tau_1, \tau_2)}\right) + \sigma_2(\tau)\left(\frac{\sigma_1(\tau_1)(\mu(\tau_2) - r)}{D(\tau_1, \tau_2)} - \frac{\sigma_1(\tau_2)(\mu(\tau_1) - r)}{D(\tau_1, \tau_2)}\right)$$

$$= \mu(\tau) - r$$

And at last, we define λ_S, λ_X as a "market price of risk" of each factor as:

$$\lambda_S = \left(\frac{\sigma_2(\tau_2)(\mu(\tau_1) - r)}{D(\tau_1, \tau_2)} - \frac{\sigma_2(\tau_1)(\mu(\tau_2) - r)}{D(\tau_1, \tau_2)}\right)$$

$$\lambda_X = \left(\frac{\sigma_1(\tau_1)(\mu(\tau_2) - r)}{D(\tau_1, \tau_2)} - \frac{\sigma_1(\tau_2)(\mu(\tau_1) - r)}{D(\tau_1, \tau_2)}\right)$$

Using these expressions, we have the following condition which is equivalent to the existence of a solution to the system of 3 linear equations from the previous sections:

$$\sigma_1(\tau)\lambda_S + \sigma_2(\tau)\lambda_X = \mu(\tau) - r$$

We observe that since the choice of τ_1 and τ_2 is arbitrary and as the RHS does not depend on τ_1 and τ_2 we must have that λ_S and λ_X are free from τ_1 and τ_2. Hence this equilibrium results holds for any time to maturity τ. Moreover as B was an arbitrary European derivative, this holds for any European derivative.

3.4.1.5 Pricing Partial Differential Equation

Using the equilibrium condition, we obtain the PDE that must be satisfied by the value of the derivative B:

$$\frac{1}{B}\sigma_S\lambda_S\frac{\partial B}{\partial S} + \frac{1}{B}\sigma_X\lambda_X\frac{\partial B}{\partial X} = \frac{1}{\mathrm{B}}\left(-\frac{\partial B}{\partial \tau} + \mu_S\frac{\partial B}{\partial S} + \mu_X\frac{\partial B}{\partial X} + \frac{\sigma_s^2}{2}\frac{\partial^2 B}{\partial S^2} + \frac{\sigma_X^2}{2}\frac{\partial^2 B}{\partial X^2}\right) - r$$

This can be simplified as:

$$\frac{\sigma_s^2}{2}\frac{\partial^2 B}{\partial S^2} + \frac{\sigma_X^2}{2}\frac{\partial^2 B}{\partial X^2} + (\mu_S - \sigma_S\lambda_S)\frac{\partial B}{\partial S} + (\mu_X - \sigma_X\lambda_X)\frac{\partial B}{\partial X} - \frac{\partial B}{\partial \tau} - Br = 0$$

Of note, this equation relies on the market prices of risk (λ_S and λ_X) of both factors.

3.4.1.6 Futures Contracts

The PDE representation can be tailored to pricing future prices. As previously discussed, the cost of entering a futures contract is zero. Hence when considering a portfolio long a futures contract and hedged by two other futures contracts with distinct maturities, the value of the portfolio is also zero as it can be entered at no cost. The growth rate of the value of the hedged portfolio is then zero. Using the notation outlined in the previous section, the resulting system of equations for the drift and diffusion of the hedging portfolio can be expressed respectively as:

$$x_1F_1\mu(\tau_1) + x_2F_2\mu(\tau_2) - B\mu(\tau) = 0$$

$$x_1F_1\sigma_1(\tau_1) + x_2F_2\sigma_1(\tau_2) - B\sigma_1(\tau) = 0$$

$$x_1F_1\sigma_2(\tau_1) + x_2F_2\sigma_2(\tau_2) - B\sigma_2(\tau) = 0$$

Proceeding as we did with an arbitrary derivative in the previous section the equilibrium condition becomes:

$$\sigma_1(\tau)\lambda_S + \sigma_2(\tau)\lambda_X = \mu(\tau)$$

Substituting the value of each term we obtain a PDE for the futures prices:

$$\frac{1}{B}\sigma_S\frac{\partial B}{\partial S}\lambda_S + \frac{1}{B}\sigma_X\frac{\partial B}{\partial X}\lambda_X = \frac{1}{\mathrm{B}}\left(-\frac{\partial B}{\partial \tau} + \mu_S\frac{\partial B}{\partial S} + \mu_X\frac{\partial B}{\partial X} + \frac{\sigma_s^2}{2}\frac{\partial^2 B}{\partial S^2} + \frac{\sigma_X^2}{2}\frac{\partial^2 B}{\partial X^2}\right)$$

which can be simplified to the following expression:

$$\frac{\sigma_s^2}{2}\frac{\partial^2 B}{\partial S^2} + \frac{\sigma_X^2}{2}\frac{\partial^2 B}{\partial X^2} + (\mu_S - \sigma_S\lambda_S)\frac{\partial B}{\partial S} + (\mu_X - \sigma_X\lambda_X)\frac{\partial B}{\partial X} - \frac{\partial B}{\partial \tau} = 0$$

subject to the initial condition $B(S, X, 0) = S$.

We note that the difference with the PDE of the price of an arbitrary derivative is the presence of the undifferentiated term rB.

3.4.2 Monte Carlo Methods

As an alternative to the PDE representation, the conditional expectation under the risk neutral measure can also be approximated with Monte Carlo methods. This is done by simulating the paths of the system of SDEs several times to approximate the conditional expected value.

3.4.2.1 Euler-Maruyama Method

The Euler-Maruyama method is a recursive numerical scheme that approximates the solution to an SDE with the lowest order of stochastic convergence. Consider a process Z_t with an initial value $Z_0 = z$ satisfying the SDE below:

$$dZ_t = \mu(Z_t, t)dt + \sigma(Z_t, t)dW_t$$

where W_t is a standard Brownian motion. To approximate the solution on the interval $[0, T]$, the interval is partitioned into a collection of subintervals $0 = t_0 < t_1 < \cdots < t_{P-1} < t_P = T$ with interval sizes $\Delta t_{i+1} = t_{i+1} - t_i$ for $i = 0 \ldots P - 1$. On each subinterval the solution is estimated according to the following equations:

$$\int_{t_i}^{t_i+1} dZ_t = \int_{t_i}^{t_i+1} \mu(Z_t, t)dt + \int_{t_i}^{t_i+1} \sigma(Z_t, t)dW_t$$

$$Z_{t_{i+1}} - Z_{t_i} \approx \mu\big(Z_{t_i}, t_i\big)\Delta t_{i+1} + \sigma\big(Z_{t_i}, t_i\big)N(0, \Delta t_{i+1})$$

where $N(a, b)$ denotes a value sampled from a normal distribution with mean a and variance b. Solving for $Z_{t_{i+1}}$ we have:

$$Z_{t_{i+1}} \approx Z_{t_i} + \mu\big(Z_{t_i}, t_i\big)\Delta t_{i+1} + \sigma\big(Z_{t_i}, t_i\big)N(0, \Delta t_{i+1})$$

Using this approximation, we can proceed recursively by starting with $Z_0 = z$ and sample the appropriate normal distribution at each step to simulate a path of the solution on the interval $[0, T]$.

This method can be generalized to systems of $n \geq 2$ SDEs. Let $Z_t = [Z_t^1 \ldots Z_t^n]$ be an n-dimensional process satisfying the system of equations below:

$$dZ_t = \begin{bmatrix} \mu_1(Z_t, t) \\ \vdots \\ \mu_n(Z_t, t) \end{bmatrix} dt + \begin{bmatrix} \sigma_{1,1}(Z_t, t) & \cdots & \sigma_{1,n}(Z_t, t) \\ \vdots & \ddots & \vdots \\ \sigma_{n,1}(Z_t, t) & \cdots & \sigma_{n,n}(Z_t, t) \end{bmatrix} dB_t$$

where $B_t = [B_t^1 \ldots B_t^n]$ is an n-dimensional Brownian motion. Using the Euler-Maruyama method the value of the Z^k component on a subinterval can then be approximated by

$$Z_{t_{i+1}}^k \approx Z_{t_i}^k + \mu_k\big(Z_{t_i}, t_i\big)\Delta t_{i+1} + \sum_{j=1}^{n} \sigma_{k,j}\big(Z_{t_i}, t_i\big)\sqrt{\Delta t_{i+1}}N^j(0, \mathrm{I})$$

In this case $N^j(0, I)$ refers to the j-th component of an n-dimensional vector sampled from a normal distribution with a mean zero and an identity matrix serving as a covariance matrix.

3.4.2.2 Estimating Derivative Valuations with Monte Carlo Methods

Using the numerical scheme of the previous section we can simulate paths of the proposed model. Given initial values S_0 and X_0 at $t = 0$, paths can be simulated on the interval $[0, T]$ where on each subinterval $[t_i, t_{i+1}]$ the change in value is estimated by the following recursive equations.

$$X_{t_{i+1}} - X_{t_i} = \left(\kappa\left(\theta - X_{t_i}\right) - \lambda_X\right)\Delta t_{i+1} + \sigma\sqrt{\Delta t_{i+1}}N^X(0,1)$$

$$S_{t_{i+1}} - S_{t_i} = S_{t_i}\left(\alpha\left(H_2\left(X_{t_i}\right) - \ln\left(S_{t_i}\right)\right) - \lambda_Y\right)\Delta t_{i+1} + H_1\left(X_{t_i}\right)S_{t_i}\sqrt{\Delta t_{i+1}}N^S(0,1)$$

where the normal variable samples $N^X(0,1)$ and $N^S(0,1)$ are assumed to be independent of each another.

Let $B = B(S, X, t)$ be the value of a European derivative with maturity $T > t$ and $P(S, X)$ be the payoff function at maturity. Simulating N paths over the interval $[t, T]$ yields a collection $\left\{\left(S_T^i, X_T^i\right)\right\}_{i=1}^N$ of equiprobable values of the factors at maturity which can be used to estimate the value of the derivative:

$$B(S_0, X_0, t) \approx \frac{e^{-r(T-t)}\sum_{i=1}^N P\left(S_T^i, X_T^i\right)}{N}$$

where r is the risk-free rate.

While Monte Carlo methods can be very versatile and simple to implement, they suffer from two main issues. The first is the high computational cost of the simulations required to obtain reliable estimates of the solution. The second is that the use of simulations introduce noise to our observations and the computed values can therefore vary every time the value is approximated.

Chapter 4

Methodology and Empirical Analysis

In this chapter, we calibrate and test the model presented in the previous section using empirical data. The commodity markets used for the empirical analysis are based on widely traded industrial/base metals. We start our analysis by first providing an overview of the market data used which includes both historical forward curves and inventory data. We then outline a calibration procedure for the model based on an Unscented Kalman Filter and present the model's fit to empirical data. Lastly, we implement the Gabillon model, a two-factor model from the literature, to provide a comparison of the proposed model's performance.

4.1 Empirical Data

The empirical analysis will be conducted using base/industrial metals. These commodities possess several advantageous properties relative to other commodities. They do not exhibit seasonality which simplifies the modelling and the exchanges on which these commodities are traded provide regular updates on levels of inventories in their warehouses. Moreover, the products deliverable on exchanges are standardized across the industry and do not decay over time.

The data used in this study covers the period from January 3rd, 2020 to December 29th, 2023 in weekly increments which represents 209 observations for each data set. The time period includes two large scale events with a pronounced effect on the commodities: the COVID-19 crisis and the start of the war in Ukraine. For nickel in particular this time period includes the short squeeze that occurred in early March of 2022.

4.1.1 Spot and Forward Prices

The source of the forward curve data is the London Metal Exchange (LME) which is the physical metal exchange with the largest volume globally. The LME reports daily futures prices for their product with constant maturities. The daily reported futures include spot contracts which refer to contracts on their last day of trading and for which settlement must be made within the next 2 days. In our calibration exercise, we have used the "spot" contracts as being the spot price of the underlying commodity. The other fixed maturities data includes maturities of 3, 15 and 27 months for aluminum and nickel. Copper has an additional contract with a maturity of 63 months. The constant time to maturity of the contract data facilitates the calibration exercise. Below we show the historical forward and spot price of aluminum, copper and nickel over the period used in this study. In each figure the darker the red the shorter the time to maturity and the darker the blue the longer the time to maturity.

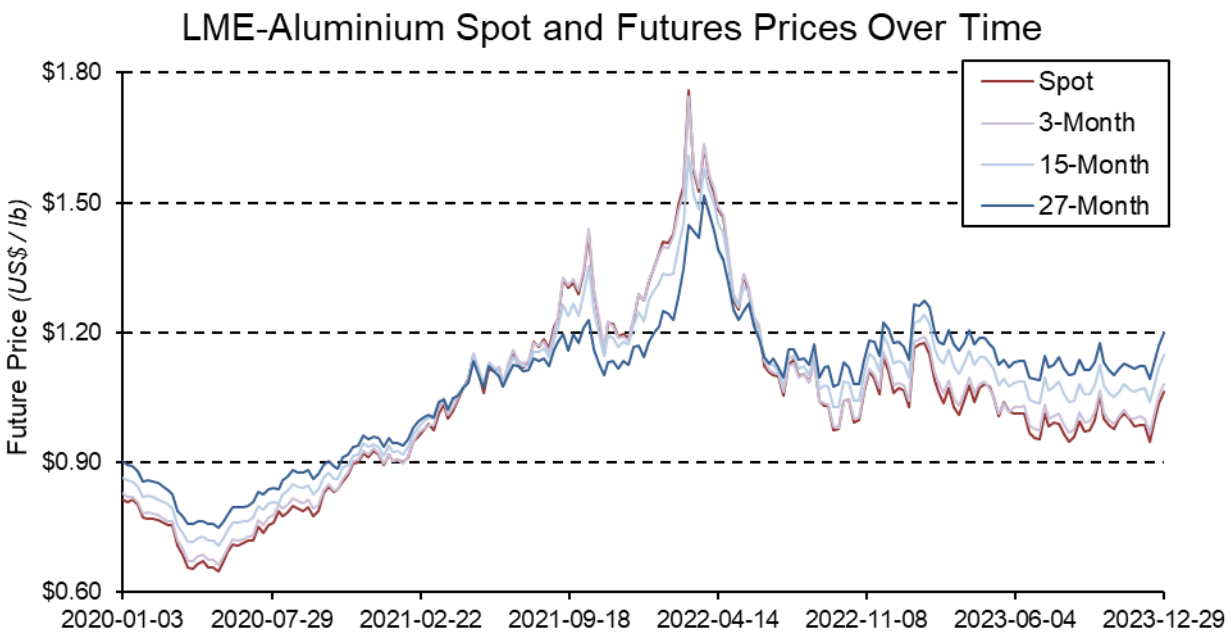


Figure 7: Aluminum Spot and Futures Prices 2020-2023

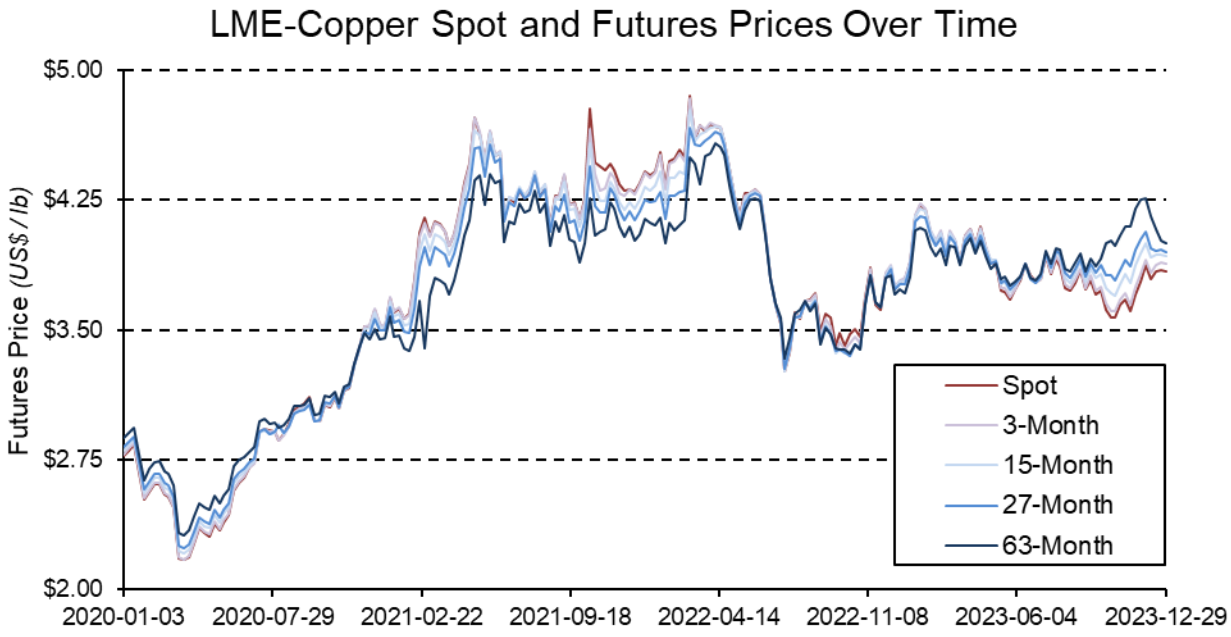


Figure 8: Copper Spot and Futures Prices 2020-2023

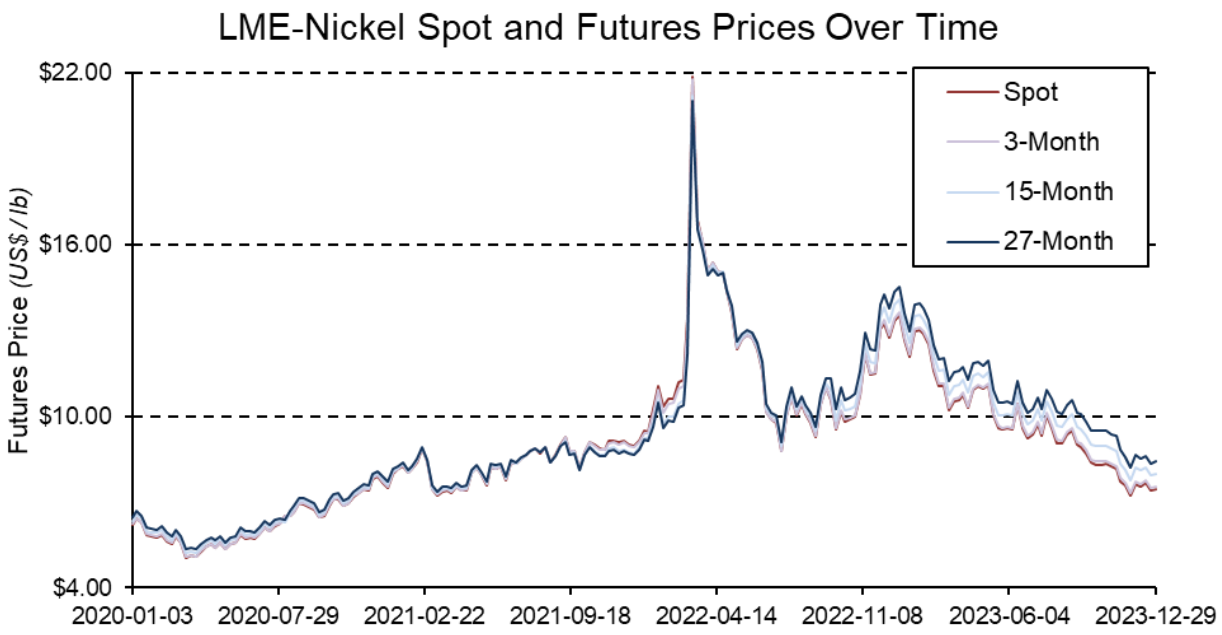


Figure 9: Nickel Spot and Futures Prices 2020-2023

The historical prices and forward curves shown in Figure 7, Figure 8 and Figure 9 exhibit various properties common to commodity markets. First and foremost, we note the mean reversion of the spot prices in all markets. For example, the spot price after the short squeeze in nickel returns to lower levels, the elevated copper prices between August 2021 and June 2022 reverts back to earlier levels and the high aluminum prices following the invasion of Ukraine in early 2022 returns to lower prices. Moreover this mean reversion also presents itself in the forward curve, as the spot price increases or decreases, the longer term forwards tend to remain closer to the earlier prices. Another property we note from the figures is that the spot price is not sufficient to determine the shape of the futures curve as different future curves arise at the same spot price in a number of instances. These properties motivate the use of at least two factors in modelling commodity prices. Lastly, these charts show that the most common shapes for forward curves in these markets are contango and backwardation with other types of curves being much less common.

4.1.2 Exchange Inventory

A key component of the proposed model is the relationship between inventory and commodity prices. While inventories in general are not observable as they are held by various business and warehouses, metal exchanges report the levels of inventory in their warehouses on a regular basis. The levels of these inventories will be used as a proxy for global inventory levels. To include as much information as possible in the inventory values used, we combine the inventory levels from three exchanges: London Metal Exchange, Chicago Mercantile Exchange (CME) and the Shanghai Futures Exchange (SHFE). The combined inventory levels for aluminum, copper and nickel are shown below.

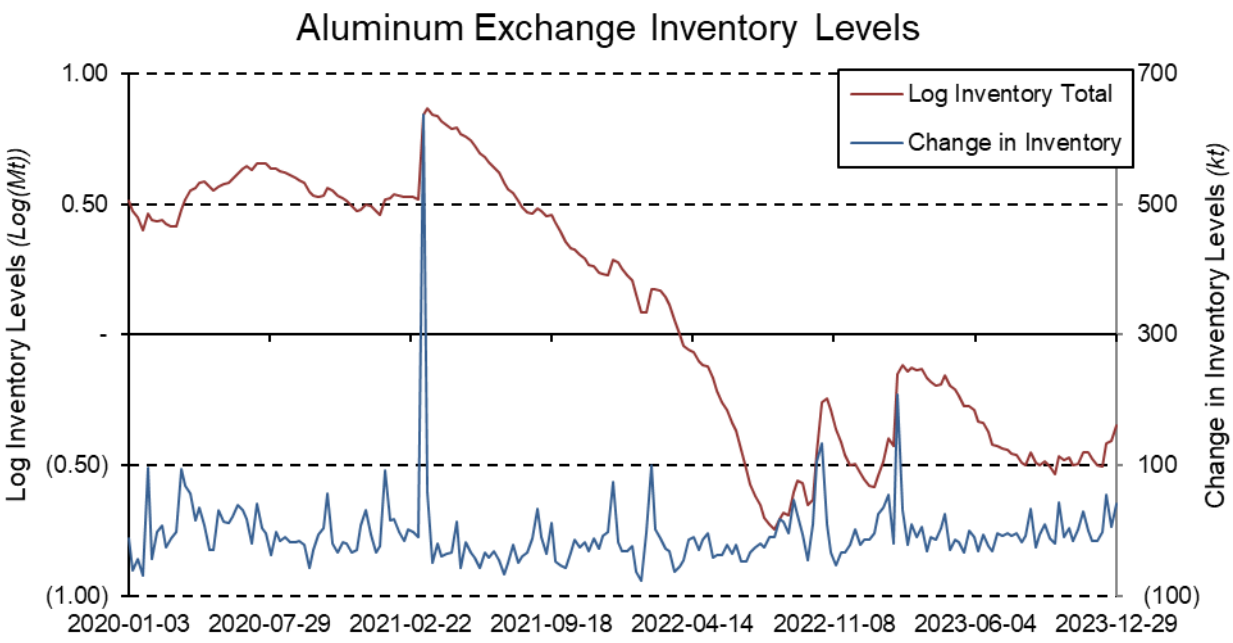


Figure 10: Historical Aluminum Inventory Levels

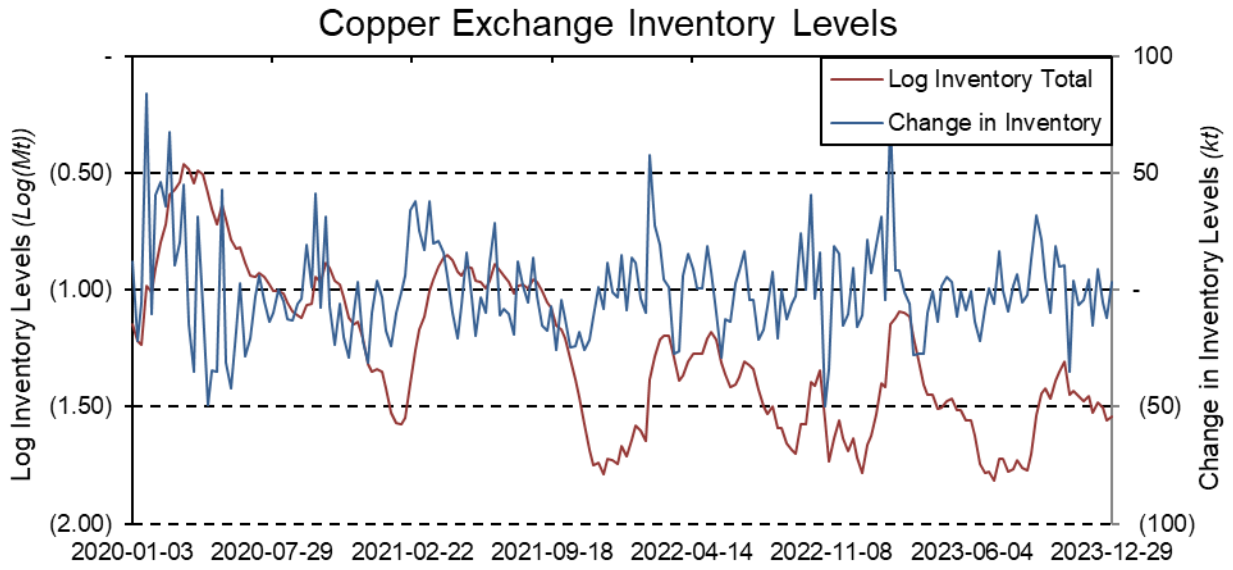


Figure 11: Historical Copper Inventory Levels

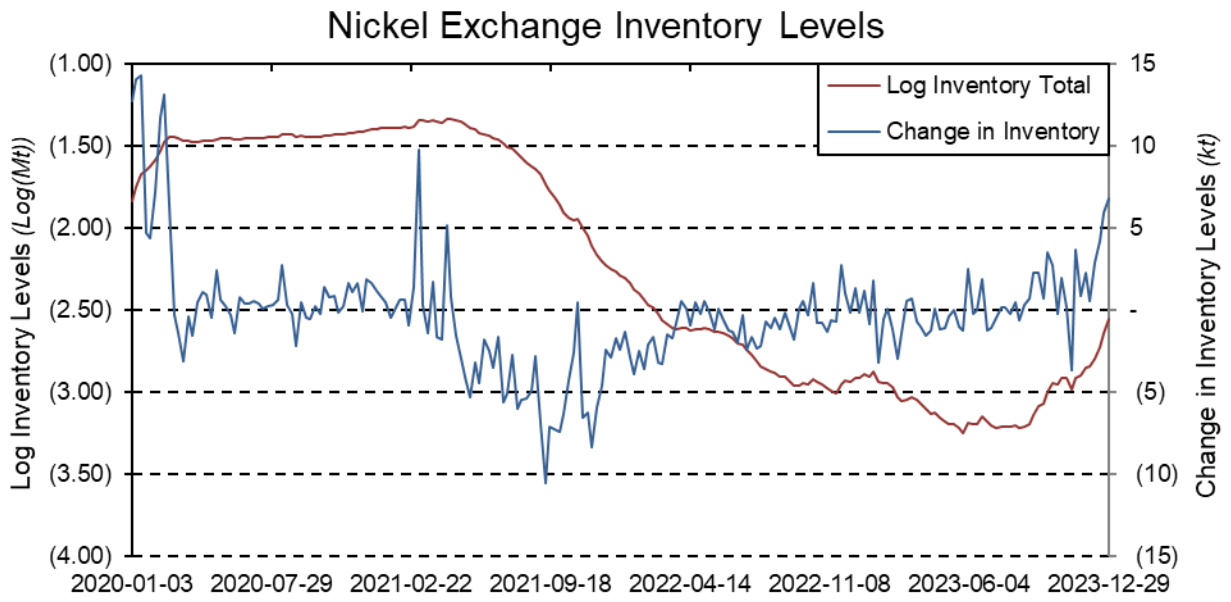


Figure 12: Historical Nickel Inventory Levels

The change in inventory appears to be mean reverting in all three markets. This behavior is expected based on supply and demand arguments and justifies the choice of the Ornstein Uhlenbeck process to model its dynamics.

4.2 Calibration

Before testing the proposed model with empirical data, it first needs to be calibrated. This section outlines the use of filtering methods to calibrate the model. We start with a brief overview of the filtering problem and the filtering techniques, then we discuss the probabilistic context of these calibration methods and lastly, we present the methodology used to calibrate the model by using an Unscented Kalman Filter.

4.2.1 The Filtering Problem

The main objective of filtering methods is to provide suitable estimates of the states of an unobserved (or state space) process given observations and dynamics of this process. The process and observations are expressed in discrete time and at each step the system takes a new state and an observation is made. The observations and the state dynamics of the process are subject to uncertainty or noise at each step. The general framework of this problem can be expressed recursively in a state space form as:

$$x_{t+1} = F(x_t, v_t)$$

$$y_t = G(x_t, n_t)$$

The first equation represents the dynamics of an unobserved process where x_t is the state of the process at time t and v_t is the noise of the process. Given the state of the process at time t, the state of the process at time $t+1$ is given by a function F of the previous state x_t and the noise of the process v_t. The second equation relates the states of the process at time t to the observations, y_t, through a function G. The observations, like the dynamics, are subject to noise which is denoted n_t. Various methodologies have been developed to estimate the state of the process

given the available information at every step. The process by which the estimates are obtained is known as filtering.

These methods are of particular interest in calibrating derivative pricing models for a number of reasons. The underlying factors in a derivative pricing model may not be observed, and while the models posit a dynamic for these factors, these dynamics are assumed to be perturbed by noise. The noisy dynamics of unobserved factors are precisely the type of problem for which filtering methods were developed. Moreover, while the factors in general may not be observed, the value of the derivatives themselves are observed in the market and hence can be treated as the observation in the filtering problem to inform on the values of the factors and model parameters.

4.2.2 State Space Formulation

To make use of the filtering framework in calibrating the model, we first provide a state space formulation for the proposed model. Recall that the proposed model takes the form:

$$S_t = \exp(Y_t)$$

$$dY_t = \alpha(H_2(X_t) - Y_t)dt + H_1(X_t)dB_t^Y$$

$$d\mathrm{X}_t = \kappa(\theta - X_t)dt + \sigma dB_t^X$$

$$dB_t^Y dB_t^X = 0$$

where S_t is the spot price, X_t is the change in inventory levels and B^X, B^Y are the standard Brownian motions.

4.2.2.1 State Dynamics

To characterize the dynamics of the factors, we use the same discretization of the Euler-Maruyama scheme. For a time step of size Δt from step t to $t+1$, this gives:

$$Y_{t+1} - Y_t = \alpha(H_2(X_t) - Y_t)\Delta t + H_1(X_t)N^Y(0, \Delta t)$$

$$\mathrm{X}_{t+1} - \mathrm{X}_t = \kappa(\theta - X_t)\Delta t + \sigma N^X(0, \Delta t)$$

where $N(a, b)$ represents a normal distribution with mean a and variance b. The states of the system are expressed as

$$p_t = \begin{bmatrix} Y_t \\ X_t \end{bmatrix}$$

Using the discretization scheme, we will treat the drift terms as the expected change and the diffusion terms as the innovation/error terms. The next state p_{t+1} can then be expressed as a function of the current state:

$$p_{t+1} = \begin{bmatrix} Y_{t+1} \\ X_{t+1} \end{bmatrix} = \begin{bmatrix} Y_t \\ X_t \end{bmatrix} + \begin{bmatrix} \alpha(H_2(X_t) - Y_t) \\ \kappa(\theta - X_t) \end{bmatrix} \Delta t + \begin{bmatrix} N^Y \\ N^X \end{bmatrix} = F\left(\begin{bmatrix} Y_t \\ X_t \end{bmatrix}\right) + v_t$$

$$\begin{bmatrix} N^Y \\ N^X \end{bmatrix} \sim N\left(0, \begin{bmatrix} \mathrm{H_1(X_t)}^2 \Delta t & 0 \\ 0 & \sigma^2 \Delta t \end{bmatrix}\right)$$

where F in the expression above is used to denote the nonlinear dynamics of the states and v_t is the random noise of the state dynamics. Note that the dynamics of the model are nonlinear due to the presence of the $H_2(X_t)$ function in the drift term.

4.2.2.2 Observation Function

The observations will be used to calibrate the model. At each point in time, the observation consists of the forward curve and the change in the inventory level. For the proposed model however, we do not have an analytical solution for the price of a forward. To obtain a function for future prices, we will use the Monte Carlo method outlined in the previous section. A discussion of numerical methods as an alternative approach for the pricing PDE is relegated to the appendix along with details of its implementation.

In general, for the observations, at any point in time t we have N forwards each with different times to maturity $\{T_i\}, i = 1 \dots N$ and the observed change in exchange inventories. The observation at time t is expressed as a vector y_t. The value at time t of the forward with time to maturity T_i is denoted as B_t^i. Using the current state of the factors we have the following observation equation:

$$y_t = \begin{bmatrix} B_t^1 \\ \vdots \\ B_t^N \\ X_t \end{bmatrix} = \begin{bmatrix} B(S_t, X_t, T_1) \\ \vdots \\ B(S_t, X_t, T_N) \\ X_t \end{bmatrix} + N(0, rI) = G(S, X) + n_t$$

where the function G captures a nonlinear relationship between the factors and the observations. The last entry in the observation vectors are changes in exchange inventory levels while the first entry are the "spot" contracts which have the shortest time to maturity. The "spot" contracts have a maturity of 2 days and are used as the spot price. As a result, the two factors of the model (changes in exchange inventory levels and spot price) are deemed to be observed.

The last component is the observation error, n_t which is meant to represent a bid-ask spread and other minor mispricing in the market. Following the literature (Schwartz, 1997), we let the observation errors be uncorrelated with a relatively small variance r. Allowing for small uncorrelated errors ensures that the changes in observations are mostly driven by the factors.

To ensure that the calibration prioritizes the parameters of the model rather than the filtered values of the factors (which are observed), we assign a smaller observation error to the factors which are observed relative to the other forward prices. This ensures that the filtered values of the factors remain close to their observed values.

4.2.3 Probabilistic Setting

Model calibration in the state space context can be seen through the lens of Maximum a Posteriori (MAP) estimation (Wan & Nelson, 2001). Letting X be a sequence of states of the system, Y be a sequence of observations and Θ be a set of model parameters to be estimated. The objective of the estimation process is to maximize the likelihood of the observations:

$$P_{X,\theta|Y}(X = x, \Theta = \theta | Y = y)$$

The solution to the estimation problem can be carried out by two main methods. The first method involves a joint estimation framework. With this method both the states and the parameters are optimized jointly with the estimates being a solution to the joint optimization program:

$$(\hat{X}, \hat{\Theta}) = argmax_{X,\Theta}\, P_{X,\theta|Y}$$

The second approach is known as a marginal estimation method where the conditional distribution $P_{X,\theta|Y}$ is split in the following components:

$$P_{X,\theta|Y} = P_{X|\Theta,Y} P_{\Theta|Y}$$

The state and parameter estimates are then obtained by optimizing each factor separately:

$$\hat{\Theta} = argmax_{\Theta}\, P_{\Theta|Y}$$

$$\hat{X} = argmax_{X}\, P_{X|\Theta,Y}$$

The first step of this approach consists of maximizing the likelihood of the observations over the parameter space. To the extent the state of the process is needed for the optimization of the parameters, the current set of parameters is used to estimate the states of the model. In other words, the estimated states of the process are found by maximizing $P_{X|\Theta,Y}$ whereas the set of parameters Θ is first found by maximizing $P_{\Theta|Y}$.

This marginal estimation approach can be found in a number of papers in the literature, for example: (1) On the calibration of the Schwartz two-factor model to WTI crude oil options and the extended Kalman Filter (Ewald, Zhang, & Zong, 2019), (2) The stochastic behavior of commodity prices: Implications for valuation and hedging (Schwartz, 1997), and (3) Affine Term-Structure Models: Theory and Implementation (Bolder, 2001).

The calibration method used in this thesis is based on the marginal estimation method where the likelihood of the observation is maximized as a function of the parameters and then the filtered states are computed based on the observations and the parameters. This is accomplished via filtering methods whereby the filter provides a method to compute the likelihood of the observations as a function of the parameters (meaning the value of $P_{\Theta|Y}$) as well as the resulting filtered estimates conditioned on the parameters and observations (which is exactly the quantity $\hat{X} = argmax_X P_{X|\Theta,Y}$).

4.2.4 Filtering Techniques

Filtering techniques are used to provide estimates of the state of the unobserved system. A number of methods have been developed in the literature. In the next section we first briefly discuss the simple case of the Kalman filter for a linear Gaussian model. We then provide a full description of the Unscented Kalman Filter and its application to the proposed model which is non-linear.

4.2.4.1 Kalman Filter

The Kalman filter can be applied for a special case of the filtering problem above when the state dynamics are linear, the observations are a linear function of the states and the noise in both the observations and state dynamics are additive and Gaussian. While this is more restrictive than the general framework of the filtering problem, the restrictions to linear functions and additive Gaussian noise provide a known closed form formula for the maximum likelihood estimate (MLE) of the states of the process.

The Kalman filter applies to models of the form:

$$x_{t+1} = F x_t + v_t$$

$$y_t = Hx_t + n_t$$

where F and H are linear maps and the noise v_t and n_t being Gaussian with a mean of zero and covariance matrices Q and R (respectively).

Under the assumption of an additive Gaussian noise, the MLE is equivalent to estimates obtained from minimizing a mean squared error. The resulting estimate, $\hat{x}_{t+1}$, can be computed recursively and takes the following form:

$$\hat{x}_{t+1} = \hat{x}'_{t+1} + K_t(y_t - H\hat{x}'_{t+1})$$

where K_t is a matrix known as the Kalman gain and $\hat{x}'_{t+1} = F\hat{x}_t$ is the prior estimate.

For calibration purposes, the likelihood refers to the difference between the observation expected based on the prior and the actual observation. The expected observations is denoted as $\hat{y}'_t$ and computed by applying the observation function to the prior hence: $\hat{y}'_t = H\hat{x}'_t$. The difference between the observation and the expected observation have a distribution given by:

$$\hat{y}'_t - y'_t \sim N(0, HP_{t|t-1}H^T + R)$$

where $P_{t|t-1}$ is the covariance matrix of the prior estimate. The MLE is then obtained by using the probability density function of the corresponding distribution at each step.

The Kalman filter has been used in the literature to calibrate several models in finance. This includes the models in Schwartz (1997), the model in Schwartz and Smith (2000) as well as a number of models for fixed income markets as discussed, for example, in Bolder (2001).

4.2.4.2 Unscented Kalman Filter

In working with filtering problems exhibiting non-linear dynamics and a non-Gaussian noise, there are typically no known analytical solutions for optimal estimates of the model states. Several approaches have been developed to approximate the solution to this problem. In this thesis we focus on the Unscented Kalman Filter which provides a “second order” estimate of the filtered distribution while the commonly used Extended Kalman Filter is only accurate up to a “first order”. This increase in accuracy is achieved through a deterministic procedure known as the Unscented Transform and does not require the computation/approximation of Jacobians which reduces the number of approximations needed.

The Unscented Transformation, in general, is used to estimate the resulting distribution of a non-linear function applied to a random variable. Taking X to be an arbitrary random variable and G an arbitrary non-linear function, the Unscented Transformation can be used to approximate the distribution of $G(X)$. The method proceeds by sampling the distribution of $G(X)$ using a fixed collection of points determined by the distribution of X. These sampled points are known as Sigma points. The sampled points are then used to compute an estimated mean and variance of $G(X)$.

Consider a state space formulation of the new spot price model for storable commodities proposed in this thesis. Let $\{X_k\}$ be the states of the process over time with their dynamics being expressed as:

$$X_{k+1} = F(X_k) + v_k$$

Let $\{Y_k\}$ be the observations at each time k which are linked to the state of the process by:

$$Y_k = G(X_k) + n_k$$

where v_k and n_k are the noises of the dynamics and the observation respectively. Both n_k and v_k have means of zero. Denote their covariance matrices as R^{n_k} and R^{v_k} respectively.

4.2.4.2.1 Initialization

At the beginning of each step $k \geq 1$ we have an estimate of the state at $k-1$ and its covariance matrix. The estimate is denoted $\hat{X}_{k-1|k-1}$ as it is the estimate for state X_{k-1} with all information available at time $k-1$. The covariance matrix of this estimate is denoted $P_{k-1|k-1}$ with a subscript following the same convention as $\hat{X}_{k-1|k-1}$

4.2.4.2.2 Prior Estimate

The first step of the process is to obtain a prior estimate. This entails dealing with the nonlinear function F of the state dynamics. This will be done by using the Unscented Transform.

This procedure starts with computing sigma points characterizing the distribution of $\hat{X}_{k-1|k-1}$. These sigma points are encoded in a matrix where each column is a sigma point. In total $2L+1$ sigma points are needed where L is the number of dimensions of X_k (this also implies that $P_{k-1|k-1}$ is an LxL matrix). The sigma points take the following form:

$$\Sigma_{k-1} = \left[\hat{X}_{k-1|k-1} \quad | \quad \hat{X}_{k-1|k-1} + \gamma\sqrt{P_{k-1|k-1}} \quad | \quad \hat{X}_{k-1|k-1} - \gamma\sqrt{P_{k-1|k-1}}\right]$$

The square root of the matrix $P_{k-1|k-1}$ is a lower triangular Cholesky decomposition of $P_{k-1|k-1}$ and $\sqrt{P_{k-1|k-1}}$ is itself a LxL matrix.

The choice of sigma points depends on the scale parameter $\gamma = \sqrt{L+\lambda}$ which itself depends on $\lambda = \alpha^2(L+\kappa) - L$. The choice of parameter α determines the spread of the sigma points around the mean and is typically taken to be small (on the order of 10^{-4}). The remaining parameter is κ which is a "secondary scaling parameter" usually set to $3-L$.

To use the Unscented Transform, we should also determine the weights used for the estimation of the transformed mean vector and covariance matrix. The computation of the weights depends on a parameter β which represents prior knowledge of the distribution of X. In the Gaussian case, $\beta = 2$ is optimal.

Using as the i-th sigma point the i-th column of Σ_{k-1}, the weights used to estimate the mean vector of the transformed distribution are denoted $W_i^{(m)}$ and are computed as:

$$W_1^{(m)} = \frac{\lambda}{L+\lambda}$$

$$W_i^{(m)} = \frac{1}{2(L+\lambda)} \; for \; i \in \{2 \dots 2L+1\}$$

The formula for the weights used to estimate the covariance matrix of the transformed distribution (denoted $W_i^{(c)}$) are:

$$W_1^{(c)} = \frac{\lambda}{L+\lambda} + 1 - \alpha^2 + \beta$$

$$W_i^{(c)} = \frac{1}{2(L+\lambda)} \; for \; i \in \{2 \dots 2L+1\}$$

Using the weights and the sigma points we can now obtain the mean vector and covariance matrix of the prior estimate. Denote $\Sigma_{\mathrm{k-1}}^{(\mathrm{i})}$ the i-th column of the matrix of sigma points (this is also the i-th sigma point). The prior estimate is then simply:

$$\hat{X}_{k|k-1} = \sum_{i=1}^{2L+1} w_i^{(m)} F\left(\Sigma_{\mathrm{k-1}}^{(\mathrm{i})}\right)$$

$$P_{k|k-1} = \sum_{i=1}^{2L+1} w_i^{(c)} \left(F\left(\Sigma_{\mathrm{k-1}}^{(\mathrm{i})}\right) - \hat{X}_{k|k-1}\right)\left(F\left(\Sigma_{\mathrm{k-1}}^{(\mathrm{i})}\right) - \hat{X}_{k|k-1}\right)^T + R^{v_k}$$

4.2.4.2.3 Expected Observation

Using the prior estimate, we then look for the expected observation and its covariance matrix. Once again, we will use the Unscented Transform. The simplest way to proceed is to compute sigma points characterizing the distribution of the prior $\hat{X}_{k|k-1}$ and then applying the methodology.

Denote the new sigma points using the matrix $\Sigma_{k|k-1}$ where each column i refers to a sigma point:

$$\Sigma_{k|k-1} = \left[\hat{X}_{k|k-1} \quad | \quad \hat{X}_{k|k-1} + \gamma\sqrt{P_{k|k-1}} \quad | \quad \hat{X}_{k|k-1} - \gamma\sqrt{P_{k|k-1}}\right]$$

As we have the same number of dimensions as in the first use of the Unscented Transform, we can compute the weights by using the same formulas as that for the prior estimate. However, for the expectation function we have chosen to use a larger value of α as we are dealing with a solution estimated by Monte Carlo values. By increasing the spread of the sigma point we reduce the effect of the noise in the computation of the observation function. Similarly to the prior, the expected observation and its covariance matrix are computed as follows:

$$\hat{Y}_{k|k-1} = \sum_{i=1}^{2L+1} w_i^{(m)} G\left(\Sigma_{\mathrm{k|k-1}}^{(\mathrm{i})}\right)$$

$$P_{Y_k|k-1} = \sum_{i=1}^{2L+1} w_i^{(c)} \left(G\left(\Sigma_{\mathrm{k|k-1}}^{(\mathrm{i})}\right) - \hat{Y}_{k|k-1}\right)\left(G\left(\Sigma_{\mathrm{k|k-1}}^{(\mathrm{i})}\right) - \hat{Y}_{k|k-1}\right)^T + R^{n_k}$$

where $\hat{Y}_{k|k-1}$ is the expected observation and $P_{Y_k|k-1}$ is its covariance matrix.

4.2.4.2.4 Update and Kalman Gain

In this step we compute the Kalman gain and obtain the posterior estimate of X_k as well as its covariance. The last matrix needed is the covariance matrix of the prior estimate of the state $\hat{X}_{k|k-1}$ with the expected observation $\hat{Y}_{k|k-1}$:

$$P_{Y_k,X_k} = \sum_{i=1}^{2L+1} w_i^{(c)} \left(F\left(\Sigma_{\mathrm{k-1}}^{(\mathrm{i})}\right) - X_{k|k-1}\right)\left(G\left(\Sigma_{\mathrm{k|k-1}}^{(\mathrm{i})}\right) - \hat{Y}_{k|k-1}\right)^T$$

Using this matrix, the Kalman gain is computed as:

$$\mathcal{K}_k = P_{Y_k,X_k}\left(P_{Y_k|k-1}\right)^{-1}$$

At last, we now have the posterior mean and covariance matrix of X_k which completes the recursive step of the algorithm:

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + \mathcal{K}_k\left(Y_k - \hat{Y}_{k|k-1}\right)$$

$$P_{k|k} = P_{k|k-1} - \mathcal{K}_k P_{Y_k|k-1}\mathcal{K}_k^T$$

4.2.4.2.5 Likelihood Function

Using the algorithm outlined above, we can compute an estimated likelihood function ℓ for the observations as a function of a set of parameters Θ. Denote the residual at each time step $e_k = Y_k - \hat{Y}_{k|k-1}$. The likelihood function is then a product of the likelihood of the residuals which are themselves assumed to be Gaussian with mean zero (as the noise has a mean of zero) and a covariance matrix $\Sigma_{\mathrm{k}}^e = P_{Y_k|k-1}$ (as the observation is known at time k, the only variance remaining is that of the estimate). This yields a likelihood function of the following form:

$$\ell(\Theta) = \sum_{k=1}^{M} \ln\left(\left(\frac{1}{2\pi}\right)^{\frac{N}{2}} \frac{1}{\sqrt{det\left(\Sigma_{\mathrm{k}}^e\right)}} exp\left(-\frac{1}{2}\left(e_k^T(\Sigma_{\mathrm{k}}^e)^{-1}e_k\right)\right)\right)$$

Model calibration then seeks to optimize this likelihood function over the space of possible Θ.

4.2.4.2.6 Likelihood Maximization Methodology

Maximization of the likelihood faces two key constraints. The first constraint is that the objective function can only be "observed" through Monte Carlo methods and is thus overly noisy and not sufficiently smooth. This precludes the use of a gradient based method for optimization purposes. The second constraint is the high run time required by Monte Carlo methods to evaluate the objective function. To obtain results within a reasonable time frame, the number of iterations or observations that can be made is then necessarily limited. Moreover, in addition to dealing with the aforementioned issues, we also want to have a sufficient degree of flexibility for an exploration of the parameter space to ensure that the method remains robust relative to the starting point of the algorithm.

To manage these challenging difficulties and explore the parameter space, a two-stage optimization process is used. The first stage is based on a simulated annealing method where in the early stage of the methodology the parameter space is explored through a random perturbation of the proposed solution and the possibility of moving to a sub-optimal point. This first stage explores the parameter space and provides a suitable starting point for a more focused method. The second stage is based on a direct search method where the starting point is the solution of the simulated annealing method. The purpose of the second stage is to continue the optimization process with a more structured approach. While this second stage is more structured in its search for an optimal solution, it does not have the same capacity to explore the parameter space.

The optimization is implemented in MATLAB by using the *"simulannealbnd"* routine for the simulated annealing and *"fminsearch"* for the direct search methods. Both methods are used with

300 iterations as the noise in the objective function renders common termination criteria difficult to attain.

4.3 Empirical Performance

Before discussing the results of the calibration process and the model's ability we first describe the form of the results obtained and the metrics by which we will evaluate its performance. The result of the calibration process has two outputs: a set of parameters and a time series of filtered values of the factors. The set of parameters characterizes the model. In the case of the proposed model, both factors are presumed to be observed, and we can view the filtered factor estimates of the factors as a by-product of the calibration process. When assessing the performance of the model, we will consider its performance under both the actual values of the factors and the filtered values of the factors.

To evaluate the performance of the model, we look at three criteria. The first one is the error of the futures prices relative to the data. This is expressed as the proportional and dollar value difference between the futures prices computed by the model and the actual futures values in the data set. The second criterion is the ability of the model to accurately model the shape of the forward curve. For this criterion we look at whether the futures with a maturity greater than three months match the data in whether they are above or below the spot price. This provides an indication of the model's ability to appropriately model the shape of the forward curve. In the last criterion, we compare the resulting filtered estimate of the factors to their observed values. This last criterion indicates the extent to which the factors were effective in modelling the futures curves.

4.3.1 Aluminum Market Results

The first step of the optimization procedure (which is the simulated annealing method) resulted in a log-likelihood value of 495. The second step (which is a direct search method) further increased the log likelihood value to 542.

Looking at the error in the forward curve over time, we observe that when using the actual values, longer maturities typically have a larger error and that the error of all contracts tends to be cyclical, i.e., they tend to increase/decrease together. When we use the filtered values however, there is no discernible trend in the error between different maturities or there is no cyclicality observed in the error. It is interesting to note that the average error across all maturities was lower when using the actual of the factors rather than the filtered values. This is likely due to the fact that the actual forward curve in the data "starts" at the actual spot price as opposed to the filtered spot price.

Table 1: Aluminum Basic Error Statistics

	Inputs		Inputs	
Error Basic Statistics	Actual	Filtered	Actual	Filtered
	(%)	*(%)*	*($)*	*($)*
3-M Average Difference	0.87%	4.59%	$0.01	$0.05
15-M Average Difference	3.22%	3.28%	$0.03	$0.03
27-M Average Difference	5.37%	4.17%	$0.06	$0.04
Total Average Difference	3.16%	4.01%	$0.03	$0.04
Greatest Difference	11.01%	17.31%	$0.12	$0.27

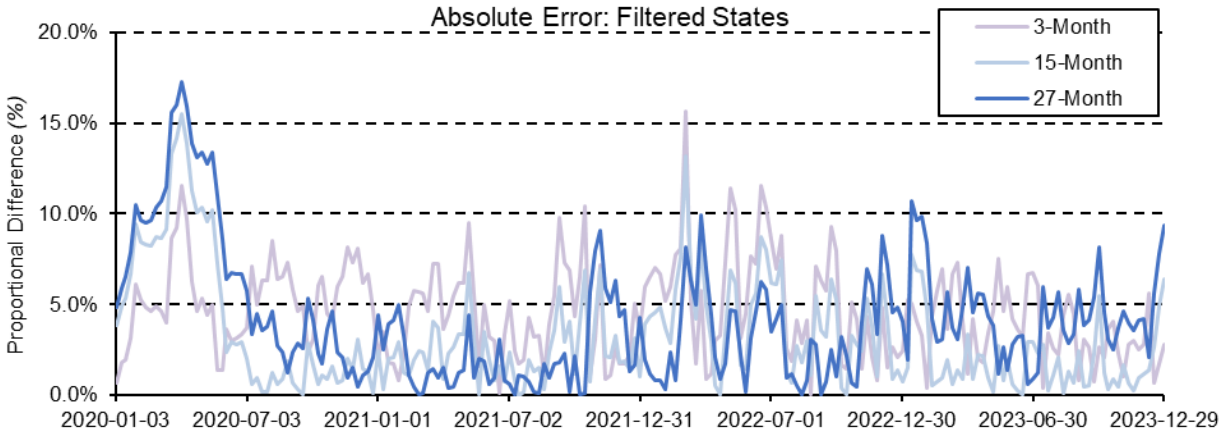


Figure 13: Aluminum Error Over Time under Filtered Factor Values

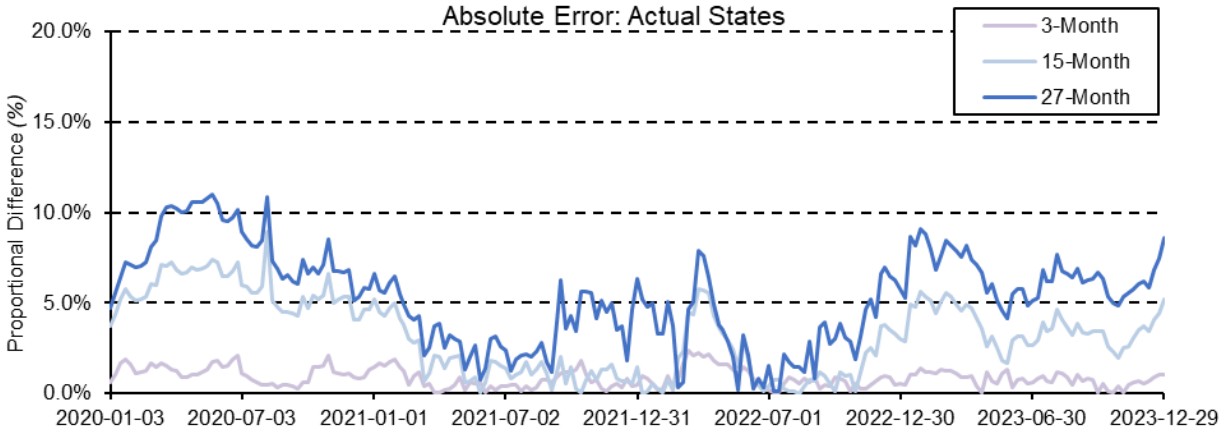


Figure 14: Aluminum Error Over Time under Actual Factor Values

The fitted model for aluminum performs quite well in identifying contango and backwardation in the sample. Nearly all instances of contango are properly obtained (>94%) and while backwardation is not as accurately modelled the success rate remains greater than 75% in all instances. Similarly, for the error, the accuracy is greater when we use actual factor values relative to the filtered factor values. Moreover, as shown in the subsequent sections, the accuracy of the aluminum model is much greater than that of the copper or nickel markets in terms of backwardation and contango.

Table 2: Aluminum Futures Curve Contango and Backwardation Results

Contango	Inputs	
	Actual	Filtered
15-Month Sample	161	161
15-Month Correct Shape	154	153
Proportion	**95.7%**	**95.0%**
27-Month Sample	153	153
27-Month Correct Shape	149	144
Proportion	**97.4%**	**94.1%**

Backwardation	Inputs	
	Actual	Filtered
15-Month Sample	48	48
15-Month Correct Shape	39	36
Proportion	**81.3%**	**75.0%**
27-Month Sample	56	56
27-Month Correct Shape	47	42
Proportion	**83.9%**	**75.0%**

For the fidelity of filtered values relative to the actual values, we see that the spot price time series are quite close to one another but that the filtered estimates do not retain the volatility of the actual values. For the changes in inventory however, we see that while the trend is the same between filtered values and actual values (both tend to remain near zero), the filtered values do not exhibit any variability over the sample period. In this case it appears that under the filtered estimates, the model reverted to the one-factor Schwartz model. This result can be taken to indicate that the changes in exchange inventory are poor indicators for forward curve shapes.

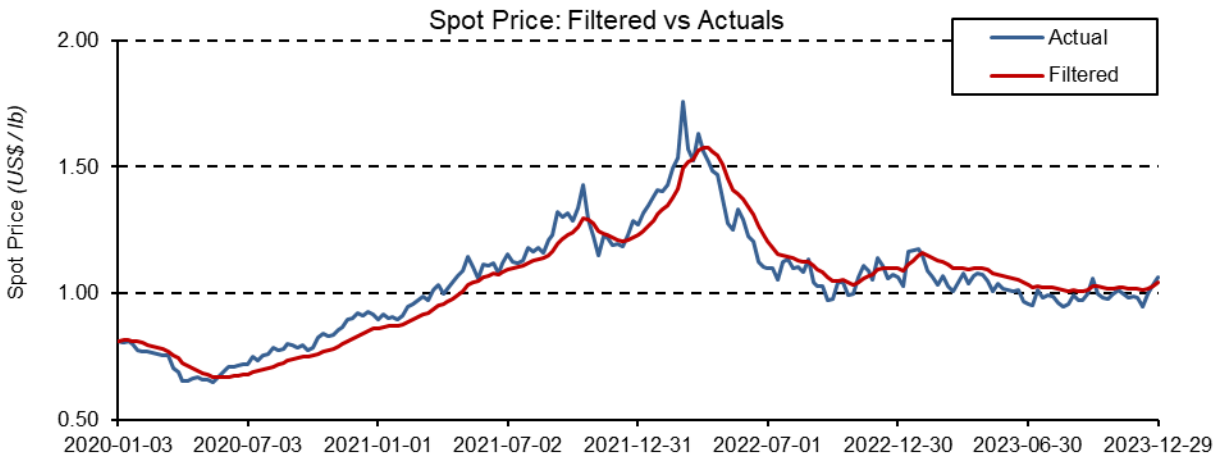


Figure 15: Time Series of Actual and Filtered Aluminum Spot Prices

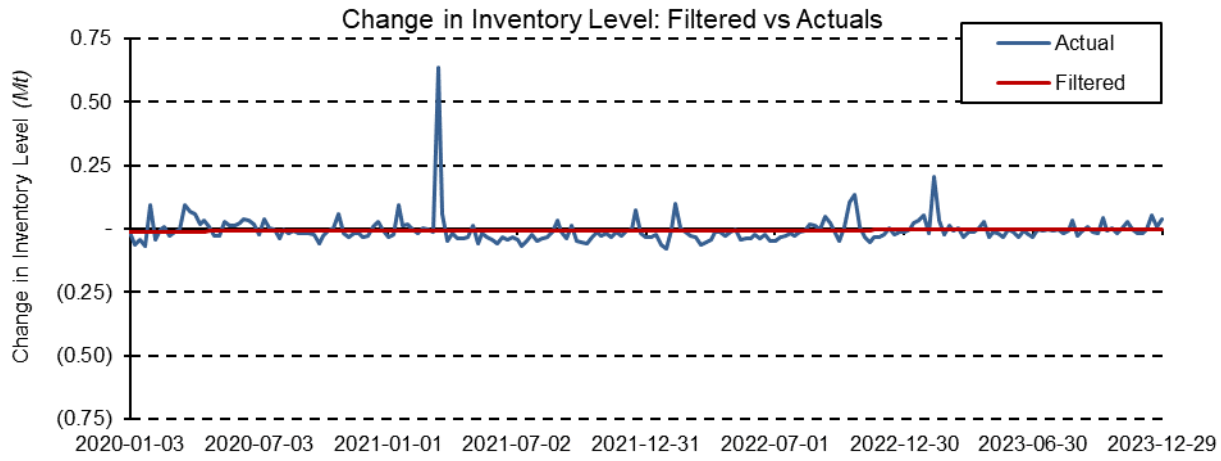


Figure 16: Time Series of Actual and Filtered Changes in Aluminum Inventories

4.3.2 Copper Markets Results

Starting with the optimization procedure, the simulated annealing algorithm ended with a log likelihood of 127 and the direct search method increased the log likelihood to a value of 415.

Looking at the error over time, both the actual and filtered inputs provide a fairly good fit to the observed data. The error for the actual values tends to occur in bursts with the errors being spread across the entire forward curves and generally being larger for longer maturities. This can be explained by the difference between the filtered and actual values of inventory changes as the filtered values do not reflect the volatility of the actual values. When using the filtered values as inputs we find that the errors are quite mixed across the forward curves although a few "bursts" do still occur. Similar to the results for the aluminium market, the model under the actual factor has a smaller average error across all maturities compared to the filtered values.

Table 3: Copper Basic Error Statistics

	Inputs		Inputs	
Error Basic Statistics	Actual	Filtered	Actual	Filtered
	(%)	*(%)*	*($)*	*($)*
3-M Average Difference	0.51%	2.27%	$0.02	$0.08
15-M Average Difference	1.26%	1.78%	$0.05	$0.06
27-M Average Difference	1.91%	1.70%	$0.07	$0.06
63-M Average Difference	3.35%	2.62%	$0.12	$0.09
Total Average Difference	1.76%	2.09%	$0.06	$0.07
Greatest Difference	17.21%	14.84%	$0.58	$0.43

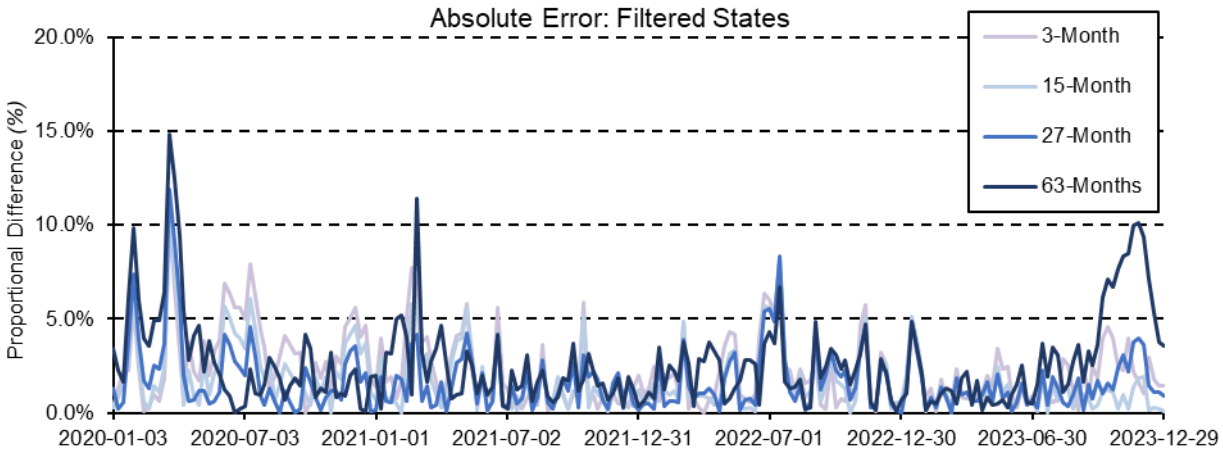


Figure 17: Copper Error Over Time under Filtered Factor Values

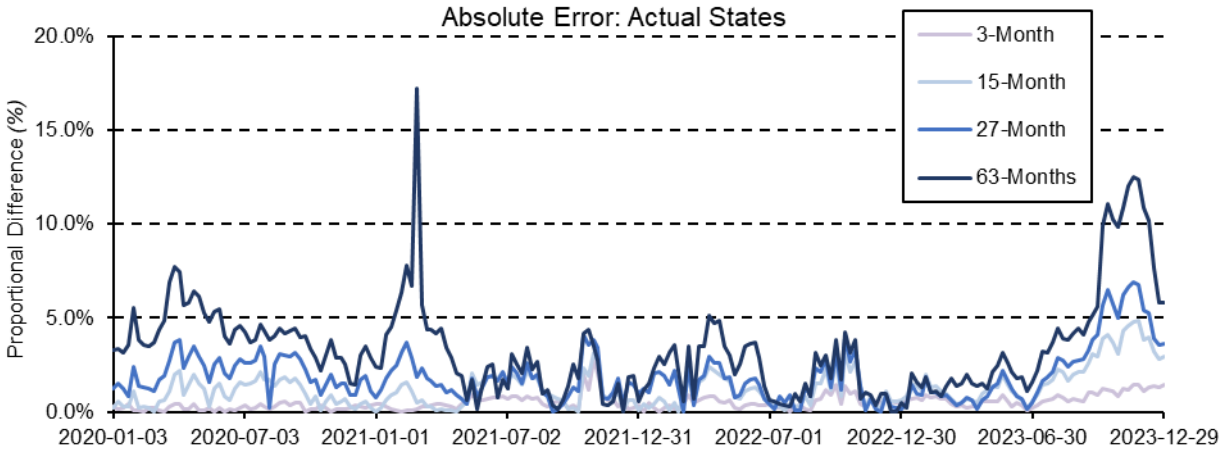


Figure 18: Copper Error Over Time under Actual Factor Values

While the model is able to attain a fairly good fit to the sample data, its ability to match the observed term structure is found to be rather limited. The performance of the actual inputs and the filtered inputs are similar with a slightly better performance with the actual inputs. The performance is particularly poor in identifying contango as only about half of contango instances, depending on maturity, are modelled correctly.

Table 4: Copper Futures Curve Contango and Backwardation Results

Contango	Inputs	
	Actual	Filtered
15-Month Sample	101	101
15-Month Correct Shape	41	41
Proportion	**40.6%**	**40.6%**
27-Month Sample	80	80
27-Month Correct Shape	45	42
Proportion	**56.3%**	**52.5%**
63-Month Sample	86	86
63-Month Correct Shape	53	49
Proportion	**61.6%**	**57.0%**

Backwardation	Inputs	
	Actual	Filtered
15-Month Sample	108	108
15-Month Correct Shape	90	77
Proportion	**83.3%**	**71.3%**
27-Month Sample	129	129
27-Month Correct Shape	100	93
Proportion	**77.5%**	**72.1%**
63-Month Sample	123	123
63-Month Correct Shape	103	97
Proportion	**83.7%**	**78.9%**

For the filtered states, we see that the filtered spot values remain quite close to the actual values. However, for the inventory changes we see that while the filtered estimates trend towards the actual values they do not have the same volatility and the amplitude as the observed values. As with the aluminum market, this result may indicate that the change in exchange inventory levels is a poor predictor of the shape of the futures curve.

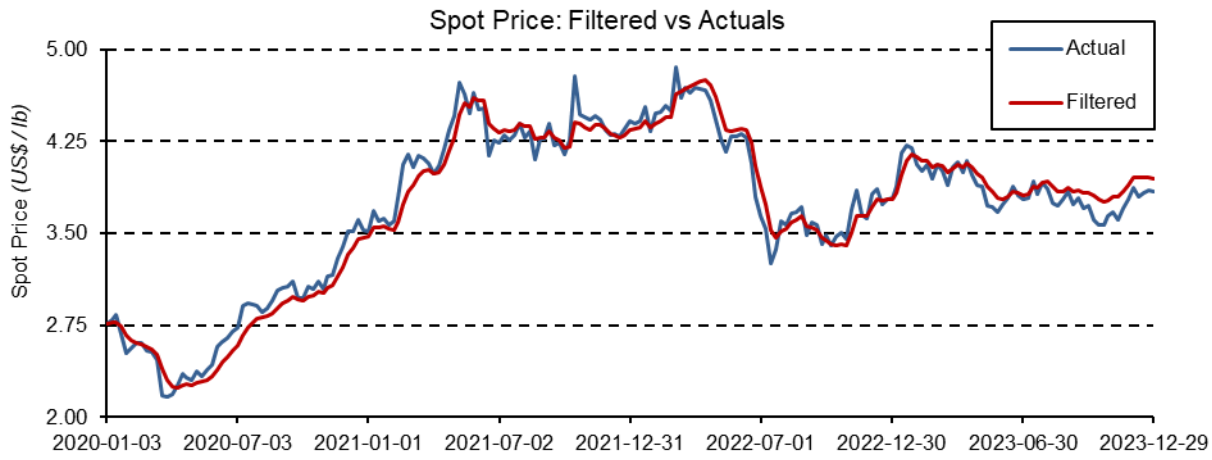


Figure 19: Time Series of Actual and Filtered Copper Spot Prices

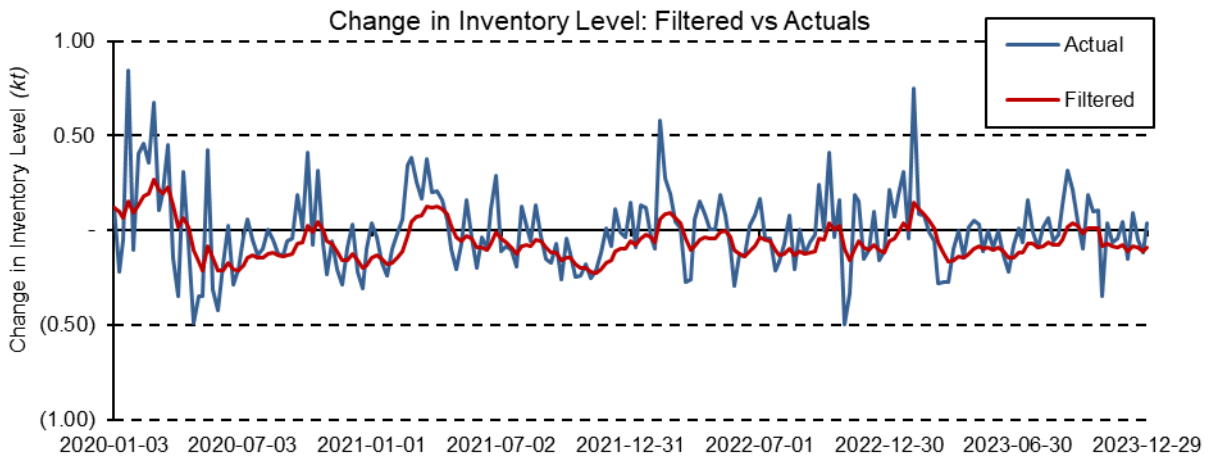


Figure 20: Time Series of Actual and Filtered Changes in Copper Inventories

4.3.3 Nickel Market Results

The initial simulated annealing algorithm reached a log likelihood value of 241. The direct search method then increased the log likelihood value to 260.

When we look at the proportional error over time, the behavior of the model under the actual values is markedly different from its behavior under the filtered states. Under the actual states we see that the error tends to increase with maturity and that it is strongly correlated across the futures curve. We can also observe that there is a period from July 2021 to March 2022 where there are consistently large errors particularly for the longer 27-month contract. The end of this period coincides with the short squeeze that occurred in March 2022 and the invasion of Ukraine by Russia (a major producer of nickel). This behavior contrasts with that of the filtered states where there is no clear relationship between the magnitude of the error and the maturity. Moreover, the filtered states do not exhibit any period of substantially larger error with only a few spikes above the 6% mark, the largest of which is due to the short squeeze in March 2022.

Table 5: Nickel Basic Error Statistics

	Inputs		Inputs	
Error Basic Statistics	Actual	Filtered	Actual	Filtered
	(%)	*(%)*	*($)*	*($)*
3-M Average Difference	0.64%	2.98%	$0.06	$0.28
15-M Average Difference	3.07%	1.98%	$0.28	$0.19
27-M Average Difference	5.09%	3.10%	$0.48	$0.30
Total Average Difference	2.93%	2.69%	$0.27	$0.26
Greatest Difference	14.41%	22.34%	$1.75	$4.79

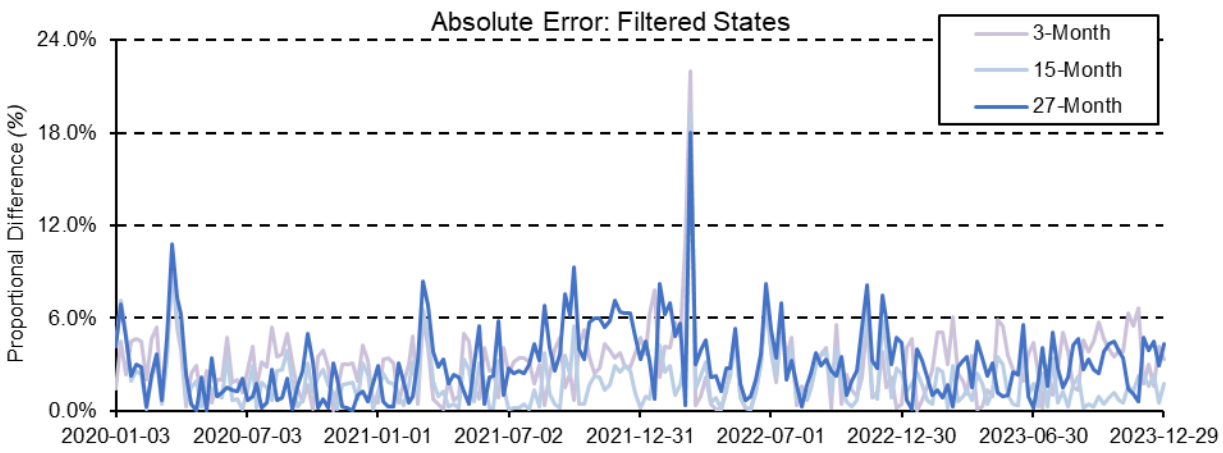


Figure 21: Nickel Error Over Time under Filtered Factor Values

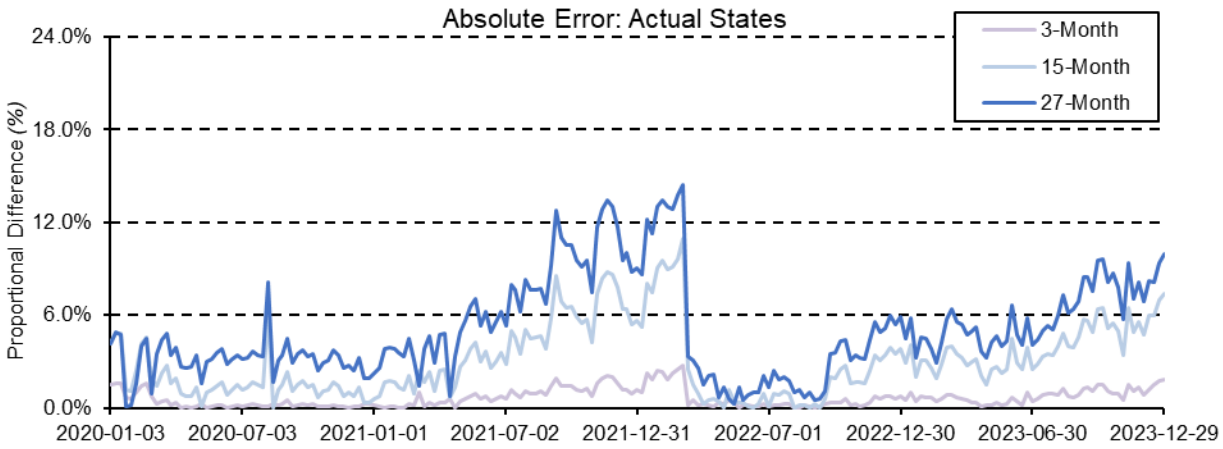


Figure 22: Nickel Error Over Time under Actual Factor Values

When it comes to identifying contango and backwardation, there is little difference between using the filtered or actual states. The model is nearly always in contango and fails to recognize backwardation. The sample is mostly in contango itself but evidently the model is not able to identify periods of backwardation with less than 10% of those instances being correctly modelled under either actual or filtered states.

Table 6: Nickel Futures Curve Contango and Backwardation Results

Contango	Inputs	
	Actual	Filtered
15-Month Sample	169	169
15-Month Correct Shape	159	160
Proportion	**94.1%**	**94.7%**
27-Month Sample	173	173
27-Month Correct Shape	171	172
Proportion	**98.8%**	**99.4%**

Backwardation	Inputs	
	Actual	Filtered
15-Month Sample	40	40
15-Month Correct Shape	3	0
Proportion	**7.5%**	**0.0%**
27-Month Sample	36	36
27-Month Correct Shape	1	0
Proportion	**2.8%**	**0.0%**

Similar to the copper market, the filtered estimates of the spot price are quite close to the actual spot price with the greatest error occurring during the short squeeze. For changes in the inventory level, we see that the filtered values have a similar trend to the actual value but do not exhibit the same level of volatility. This is particularly apparent for the larger movements in the inventory levels.

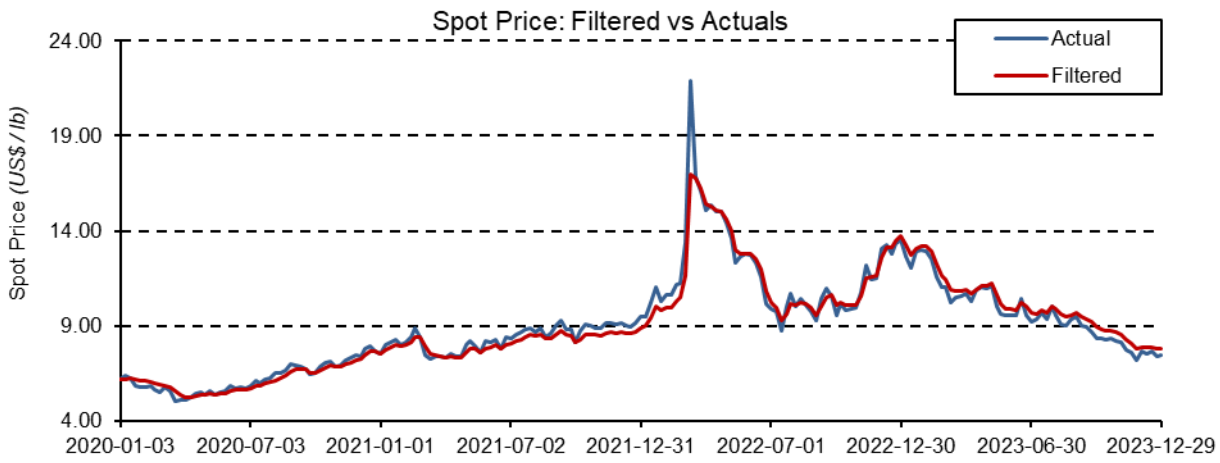


Figure 23: Time Series of Actual and Filtered Nickel Spot Prices

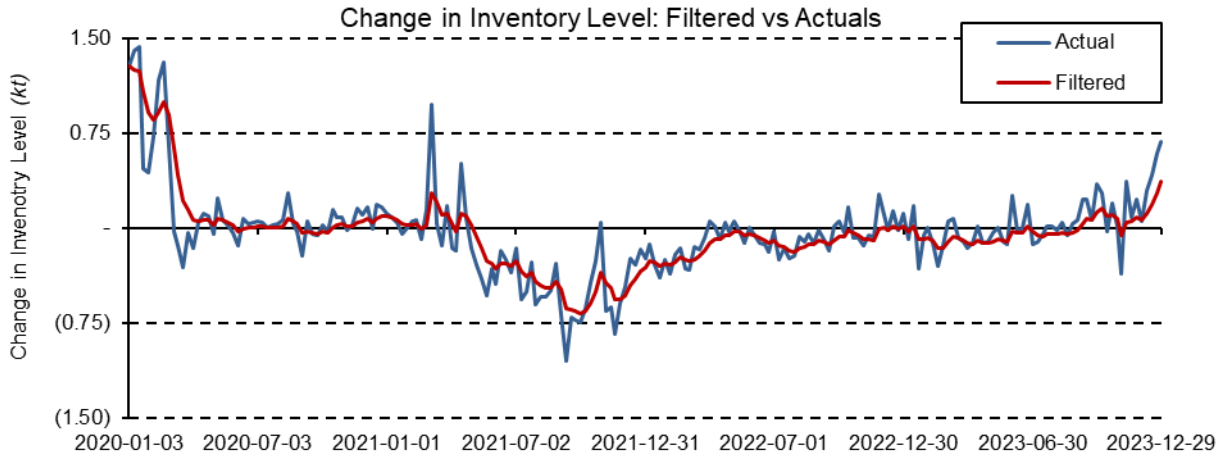


Figure 24: Time Series of Actual and Filtered Changes in Nickel Inventories

4.3.4 Model Comparison

To benchmark the proposed model and compare the empirical results of the previous section, a two-factor model from the literature is implemented. The three two-factor models discussed earlier in section 2.2.3 (Gibson & Schwartz, Schwartz & Smith and Gabillon) are known to be formally equivalent in the sense that their factors are linear combinations of one another. This implies that any of them will be well suited for comparison purposes. Due to the similarity in formulation to the proposed model, the Gabillon model is used for comparison purposes.

Recall that the two factors in the Gabillon model are the spot price and the long-term price denoted S_t and L_t respectively. The model posits that the factors satisfy the following system of SDEs under the “real world” measure:

$$dS_t = k(\ln(L_t) - \ln(S_t))S_t dt + \sigma_S S_t dB_t^S$$

$$dL_t = \mu L_t dt + \sigma_L L_t dB_t^L$$

$$dB_t^S dB_t^L = \rho dt$$

Under the assumptions that the long-term price is finite, and the behavior of the future prices approaches that of the long-term price as maturity goes to infinity, future prices $F(S, L, \tau)$ can be computed with the following closed form expressions (Gabillon, 1991):

$$F(S, L, \tau) = A(\tau)S^{B(\tau)}L^{1-B(\tau)}$$

$$A(\tau) = \exp\left(\frac{\lambda_S \sigma_S}{\kappa}(e^{-\kappa\tau} - 1) + \frac{v}{4\kappa}(2e^{-\kappa\tau} - e^{-2\kappa\tau} - 1)\right)$$

$$B = e^{-\kappa\tau}$$

where $v = \sigma_S^2 + \sigma_L^2 - 2\rho\sigma_S\sigma_L$ and λ_S is the market price of risk of the spot price.

4.3.4.1 Calibration and State Space Form

The Gabillon model can be expressed in a state space form with linear dynamics and an additive Gaussian noise. As such, the Kalman filter can be used to calibrate the model. To obtain a linear state space model with an additive Gaussian noise, we apply Ito's lemma to express the model by using the logarithm of both the spot price, $X_t \coloneqq \ln(S_t)$ and long-term price $Y_t = \ln(L_t)$. This gives the system of SDEs below under the "real world" measure:

$$dX_t = \left(k(\mathrm{Y_t} - \mathrm{X_t}) - \frac{\sigma_S^2}{2}\right)dt + \sigma_S dB_t^S$$

$$dY_t = \left(\mu - \frac{\sigma_L^2}{2}\right)dt + \sigma_L dB_t^L$$

$$dB_t^S dB_t^L = \rho dt$$

Under the "risk neutral" measure, assuming the market price of risk of spot price (λ_S) is constant and the assumptions regarding long-term price, the model satisfies the following SDEs:

$$dX_t = \left(k(\mathrm{Y_t} - \mathrm{X_t}) - \frac{\sigma_S^2}{2} - \lambda_S\right)dt + \sigma_S dB_t^S$$

$$dY_t = \left(\frac{-\sigma_L^2}{2}\right)dt + \sigma_L dB_t^L$$

$$dB_t^S dB_t^L = \rho dt$$

Proceeding similarly to the proposed model, we treat the drift terms of the SDEs as the dynamics of the model and the diffusion terms as the noise of the dynamics under the Euler-Maruyama discretization scheme. Under the "risk neutral" measure, the dynamics of the state space form over a time interval of Δt from t to $t + 1$ takes the form:

$$\begin{bmatrix} X_{t+1} \\ Y_{t+1} \end{bmatrix} = \begin{bmatrix} -\Delta t\left(\frac{\sigma_S^2}{2} + \lambda_S\right) \\ -\Delta t\left(\frac{\sigma_L^2}{2}\right) \end{bmatrix} + \begin{bmatrix} (1-\kappa\Delta t) & \kappa\Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_t \\ Y_t \end{bmatrix} + N\left(0, \begin{bmatrix} \sigma_S^2 \Delta t & \rho\sigma_S\sigma_L\Delta t \\ \rho\sigma_S\sigma_L\Delta t & \sigma_L^2\Delta t \end{bmatrix}\right)$$

where $N(a, b)$ denotes a Gaussian distribution with mean a and covariance matrix b.

The observations at each time step corresponds to the future price curve. Using the closed form solution of the future prices under the Gabillon model, we have the following relationship between the factors:

$$\begin{aligned} \ln\big(F(S, L, \tau)\big) &= \ln\big(A(\tau)\big) + B(\tau)\ln(S) + \big(1 - B(\tau)\big)\ln(L) \\ &= \ln\big(A(\tau)\big) + B(\tau)X + \big(1 - B(\tau)\big)\mathrm{Y} \end{aligned}$$

This can also be expressed in a matrix form as:

$$\begin{bmatrix} \ln\big(F_t(\tau_1)\big) \\ \vdots \\ \ln\big(F_t(\tau_n)\big) \end{bmatrix} = \begin{bmatrix} \ln\big(A(\tau_1)\big) \\ \vdots \\ \ln\big(A(\tau_n)\big) \end{bmatrix} + \begin{bmatrix} B(\tau_1) & \big(1 - B(\tau_1)\big) \\ \vdots & \vdots \\ B(\tau_n) & \big(1 - B(\tau_n)\big) \end{bmatrix} \begin{bmatrix} X_t \\ Y_t \end{bmatrix} + N\left(0, \begin{bmatrix} \Sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Sigma_n \end{bmatrix}\right)$$

where $F_t(\tau_i)$ is the futures price with maturity τ_i at time t. The noise in the observation is kept small to ensure that the likelihood is determined by the parameters. As the data includes "spot" contracts which is also one of the factors of the model, the noise associated with this contract is kept smaller than the other future prices. Moreover, to ensure that the comparison between the model is done consistently, the noise of the observations in the filtering process is taken to be the same as that of the proposed model.

Using the model dynamics and observation function described in this section, the likelihood can be determined by using the Kalman filter. To calibrate the model, we then seek to maximize the likelihood over the parameter space. To this end we rely on the FMINCON routine in MATLAB which provides a solution to constrained non-linear optimization problems. Unlike the proposed model, the closed form solution to the Gabillon model materially reduces the computational cost of the calibration/optimization process.

4.3.4.2 Results

The calibration of the Gabillon model is implemented by using the same dataset as the proposed model. Moreover, the performance of the model will be based on the same criteria as the proposed model, i.e., the proportional error over time, the replication of the shape of the futures curve and the fidelity of filtered estimate to the actual values. For the last criterion, the spot price is observed but the long-term price is not observed and is determined solely through the filtering procedure. The calibration procedure provides two outputs: a set of model parameters and the filtered state estimates. In the case of the Gabillon model the filtered states are the estimated values of the spot and long-term price. As the spot price is observed, we can test the model's fit with calibrated parameters using the filtered estimates of the spot price or the actual spot prices. In both cases the spot prices (filtered or actual) will be used with the filtered long-term price (as it is not observed). This is referred to as filtered or actual inputs in the results presented in the next sections.

The results for the aluminum, copper and nickel markets under the Gabillon model are presented in detail in the upcoming sections. Table 7 and Table 8 below show a summary of the comparison between the models in terms of various metrics. The last metric comparing the fit of the filtered spot price to the actual spot price is based on the L^1-distance between the two curves. Overall, we note that the Gabillon model provides a much better fit to the empirical data for all metrics outlined below. In particular it achieved substantially larger likelihood values. The error is reduced, on average, for every market under both the actual and filtered factors. Importantly, the Gabillon model exhibits a much greater capacity to model the shape of the forward curve as shown below in the proportion of backwardation and contango properly represented across the forward curves. The filtered spot price is much closer to the actual spot price under the Gabillon model in all three markets.

Table 7: Comparison of the Proposed and Gabillon Models under the Filtered Factors

	Aluminum		Copper		Nickel	
Comparison (Filtered Factors)	Proposed	Gabillon	Proposed	Gabillon	Proposed	Gabillon
Calibration						
Likelihood	542.06	720.41	415.32	852.45	262.73	629.10
Sample Error						
Average Error	4.01%	1.41%	2.09%	1.46%	2.69%	1.58%
Greatest Error	17.31%	7.30%	14.84%	14.77%	22.34%	9.08%
Forward Curves						
Backwardation	75.0%	100.0%	74.2%	82.5%	0.0%	19.7%
Contango	94.6%	96.2%	49.4%	82.4%	97.1%	99.7%
Observed Factor Fit						
Spot Price	0.2082	0.0432	0.3637	0.1390	1.2982	0.2968

Table 8: Comparison of the Proposed and Gabillon Models under the Actual Factors

	Aluminum		Copper		Nickel	
Comparison (Actual Factors)	Proposed	Gabillon	Proposed	Gabillon	Proposed	Gabillon
Calibration						
Likelihood	542.06	720.41	415.32	852.45	262.73	629.10
Sample Error						
Average Error	3.16%	1.17%	1.76%	1.20%	2.93%	1.48%
Greatest Error	11.01%	6.41%	17.21%	14.92%	14.41%	8.94%
Forward Curves						
Backwardation	82.7%	99.0%	81.4%	86.9%	5.3%	19.7%
Contango	96.5%	95.9%	52.1%	78.3%	96.5%	99.7%
Observed Factor Fit						
Spot Price	0.2082	0.0432	0.3637	0.1390	1.2982	0.2968

4.3.4.3 Gabillon Model: Aluminum Market Results

In terms of error over the sample, the Gabillon model yields a demonstrably better fit than the proposed model. This is illustrated most clearly by the “Total Average Difference” and “Greatest Difference” metric in Table 1 and Table 9. These figures show that the average error (Actual Factors/Filtered Factors) is 3.16%/4.01% in the proposed model compared to 1.17%/1.41% under the Gabillon model. Moreover, the largest error over the sample decreases from 11.01%/17.31% in the proposed model to 6.41%/7.30% under the Gabillon model. Not only does the Gabillon model produce a better fit overall this increased accuracy does not occur at the cost of poor localized performance.

Table 9: Basic Error Statistics of the Aluminum Market under the Gabillon Model

	Inputs		Inputs	
Error Basic Statistics	Actual	Filtered	Actual	Filtered
	(%)	*(%)*	*($)*	*($)*
3-M Average Difference	0.67%	1.14%	$0.01	$0.01
15-M Average Difference	1.03%	1.25%	$0.01	$0.01
27-M Average Difference	1.80%	1.84%	$0.02	$0.02
Total Average Difference	1.17%	1.41%	$0.01	$0.02
Greatest Difference	6.41%	7.30%	$0.09	$0.12

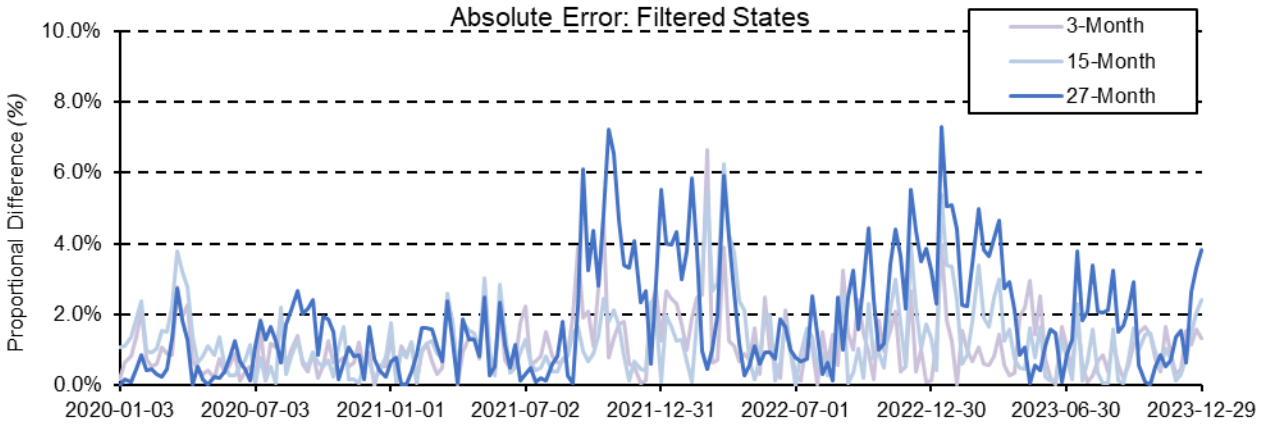


Figure 25: Aluminum Market Error with Filtered Factor Values under the Gabillon Model

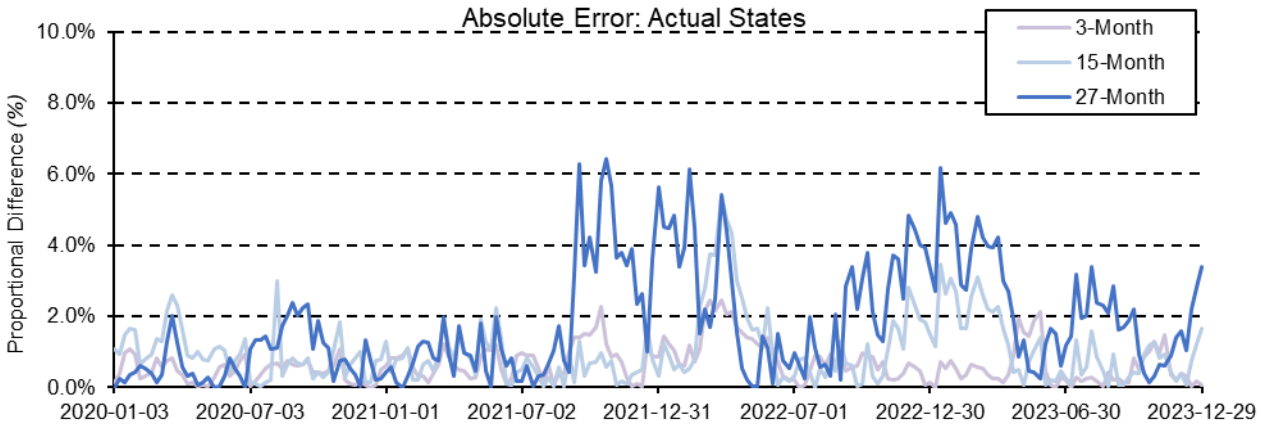


Figure 26: Aluminum Market Error with Actual Factor Values under the Gabillon Model

The Gabillon model also outperforms the proposed model in modelling the shape of the futures curve. While the performance in identifying contango is slightly greater for the proposed model, the performance in identifying backwardation is significantly better for the Gabillon model. This illustrates the greater flexibility of the Gabillon model in terms of the shape of the futures curve.

Table 10: Aluminum Futures Curve Shape Results under the Gabillon Model

Contango	Inputs		Backwardation	Inputs	
	Actual	Filtered		Actual	Filtered
15-Month Sample	161	161	15-Month Sample	48	48
15-Month Correct Shape	151	152	15-Month Correct Shape	48	48
Proportion	**93.8%**	**94.4%**	**Proportion**	**100.0%**	**100.0%**
27-Month Sample	153	153	27-Month Sample	56	56
27-Month Correct Shape	150	150	27-Month Correct Shape	55	56
Proportion	**98.0%**	**98.0%**	**Proportion**	**98.2%**	**100.0%**

Spot price is the only factor that is observable in both the Gabillon model and the empirical data. As such, when comparing the fidelity of the filtered factors to the actual values, only the spot price can be examined. Looking at Figure 27 and Figure 15 it is clear that the Gabillon model produces filtered estimates that are much closer to those of the proposed model. The last figure, Figure 28, shows the filtered spot and long-term price and a comparison to Figure 7 we see that the long-term price behaves similarly to the longer term maturity contracts, as is expected when the Gabillon model is calibrated based on futures prices.

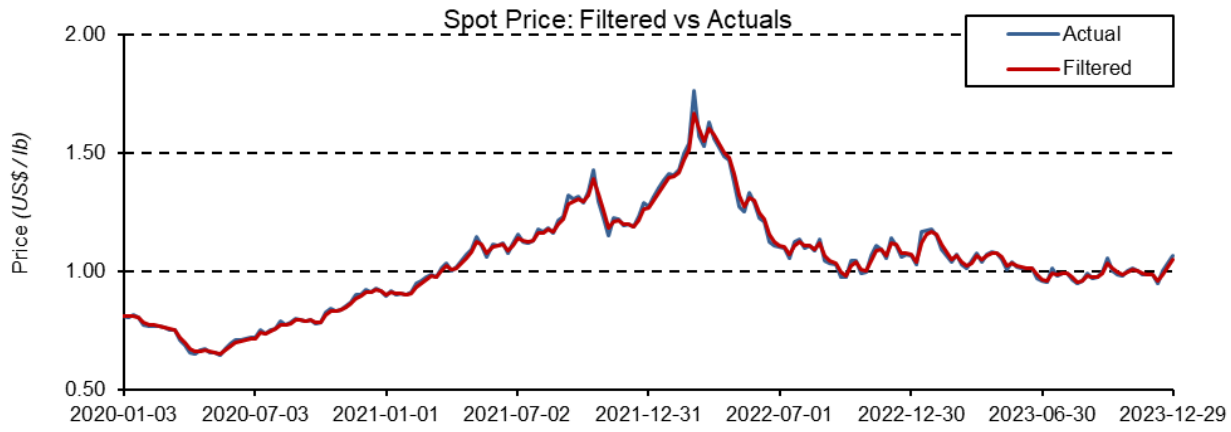


Figure 27: Actual and Filtered Aluminum Spot Prices under the Gabillon Model

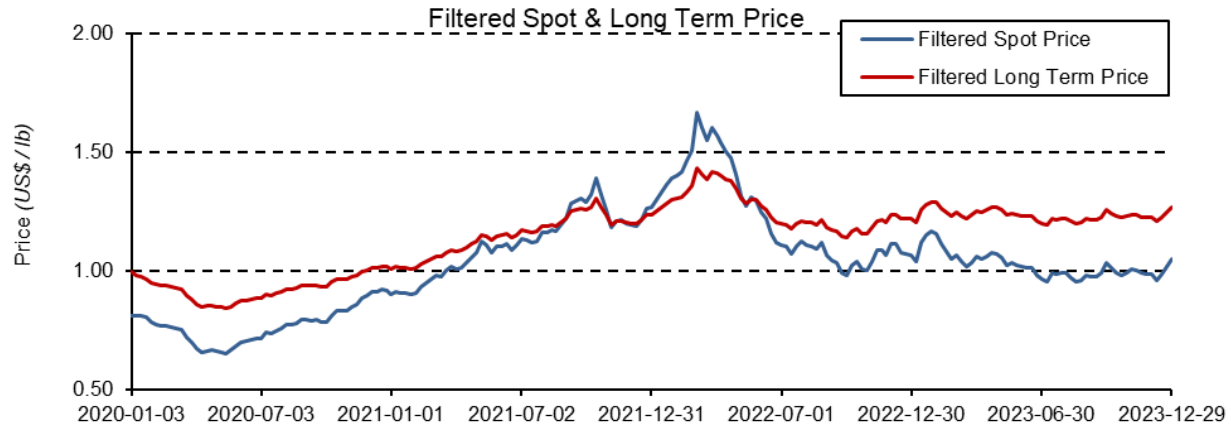


Figure 28: Filtered Aluminum Spot and Long-Term Price under the Gabillon Model

4.3.4.4 Gabillon Model: Copper Market Results

As is the case with the aluminum market, the Gabillon model provides a better fit to the copper market than the proposed model. This can be seen in both the average error across the futures curve and the greatest difference over the sample. The average error decreases from 1.76%/2.09% (Actual Factors/Filtered Factors) in the proposed model to 1.20%/1.46% in the Gabillon model. The greatest error decreases from 17.21%/14.84% in the proposed model to 14.92%/14.77% in the Gabillon model.

Table 11: Basic Error Statistics of the Copper Market under the Gabillon Model

	Inputs		Inputs	
Error Basic Statistics	Actual	Filtered	Actual	Filtered
	(%)	*(%)*	*($)*	*($)*
3-M Average Difference	0.45%	1.04%	$0.02	$0.04
15-M Average Difference	1.02%	1.29%	$0.04	$0.05
27-M Average Difference	1.26%	1.40%	$0.05	$0.05
63-M Average Difference	2.07%	2.09%	$0.08	$0.08
Total Average Difference	1.20%	1.46%	$0.05	$0.05
Greatest Difference	14.92%	14.77%	$0.51	$0.50

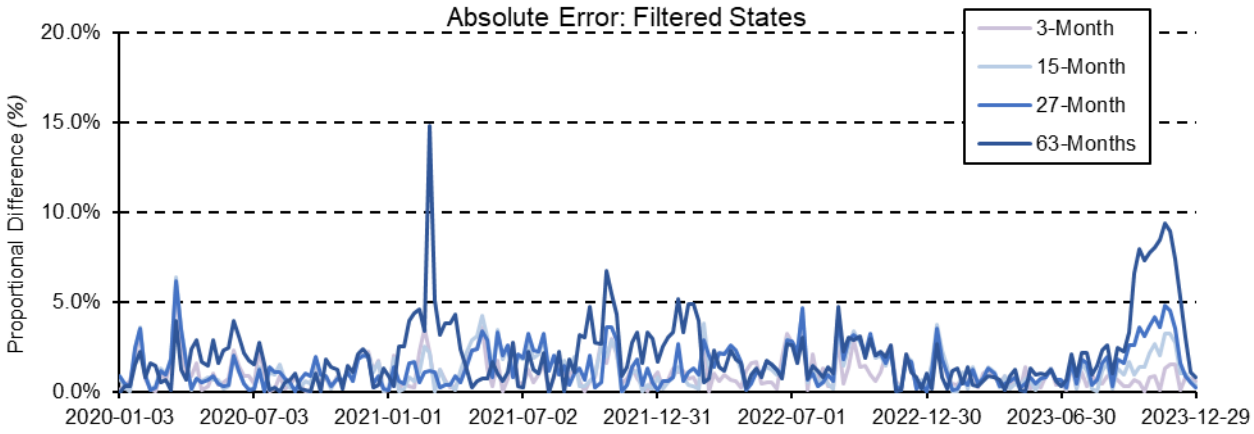


Figure 29: Copper Market Error with Filtered Factor Values under the Gabillon Model

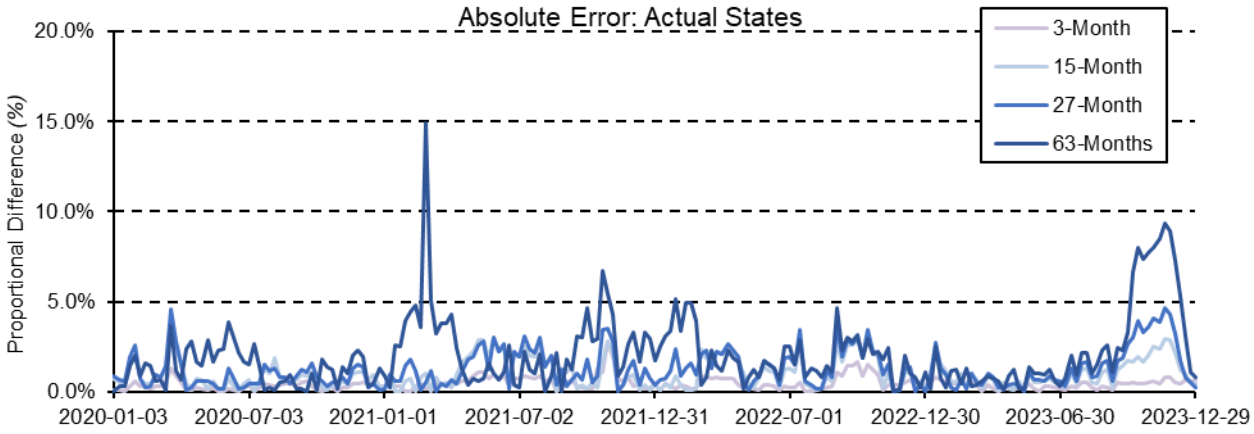


Figure 30: Copper Market Error with Actual Factor Values under the Gabillon Model

The Gabillon model performs much better in accurately modelling the shape of the futures curve over the duration of the sample. While the improvement can be seen in all metrics shown in Table 4 and Table 12, the greatest increase is observed for instances of the sample that are in contango.

Table 12: Copper Futures Curve Shape Results under the Gabillon Model

Contango	Inputs		Backwardation	Inputs	
	Actual	Filtered		Actual	Filtered
15-Month Sample	101	101	15-Month Sample	108	108
15-Month Correct Shape	70	72	15-Month Correct Shape	92	84
Proportion	**69.3%**	**71.3%**	**Proportion**	**85.2%**	**77.8%**
27-Month Sample	80	80	27-Month Sample	129	129
27-Month Correct Shape	66	71	27-Month Correct Shape	110	105
Proportion	**82.5%**	**88.8%**	**Proportion**	**85.3%**	**81.4%**
63-Month Sample	86	86	63-Month Sample	123	123
63-Month Correct Shape	73	77	63-Month Correct Shape	111	108
Proportion	**84.9%**	**89.5%**	**Proportion**	**90.2%**	**87.8%**

A comparison of Figure 19 and Figure 31 shows that the filtered spot price remained closer to the actual spot price under the Gabillon model. Lastly, as is expected in the Gabillon model, the long-term price behaves similarly to the longer-term maturity futures.

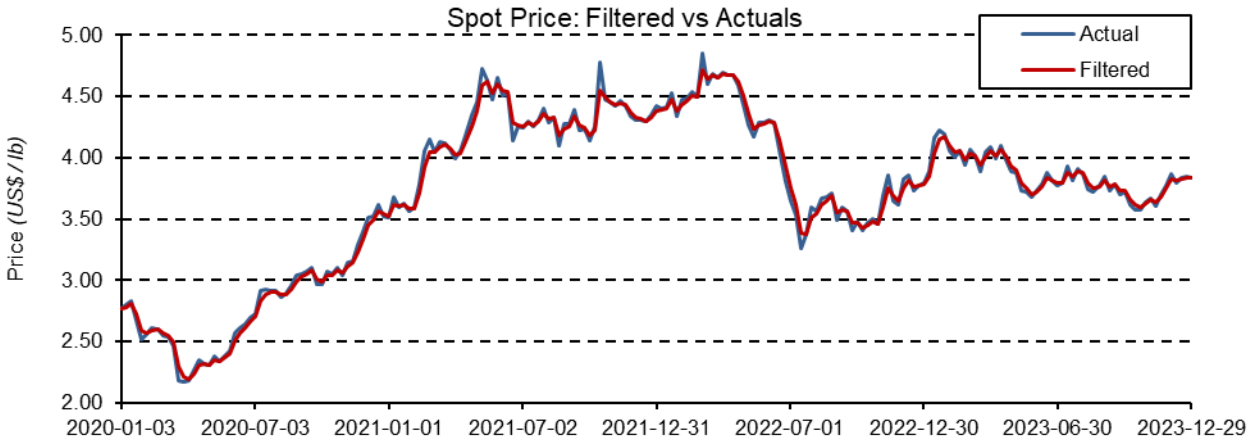


Figure 31: Actual and Filtered Copper Spot Prices under the Gabillon Model

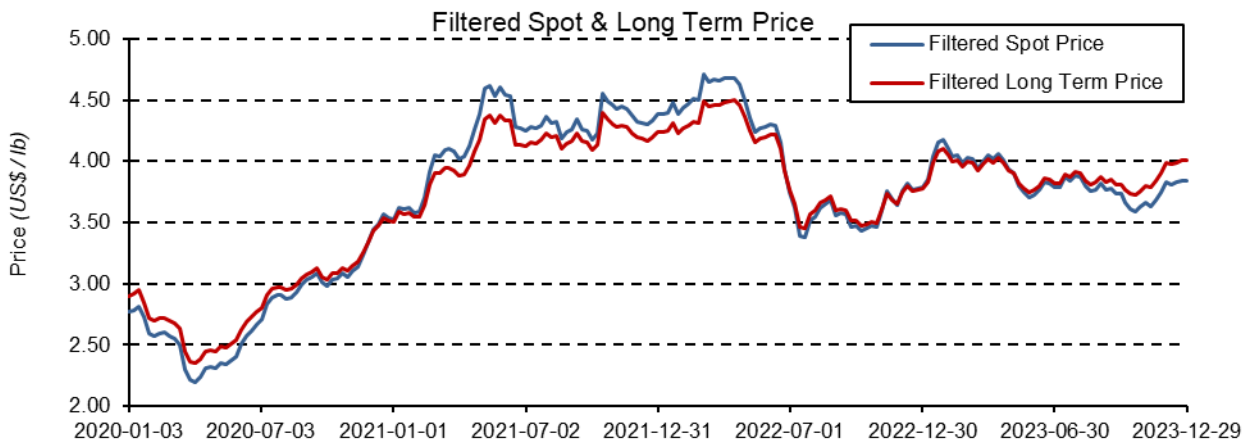


Figure 32: Filtered Copper Spot and Long-Term Price under the Gabillon Model

4.3.4.5 Gabillon Model: Nickel Market Results

The Gabillon model outperforms the proposed model for the nickel market as well. This includes a significant decrease in average and largest error under both actual and filtered factors. The large error followed by a rapid decrease in early 2022 under both the actual and filtered factors corresponds to unusual market conditions around a short squeeze.

Table 13: Basic Error Statistics of the Nickel Market under the Gabillon Model

	Inputs		Inputs	
Error Basic Statistics	Actual	Filtered	Actual	Filtered
	(%)	*(%)*	*($)*	*($)*
3-M Average Difference	0.42%	0.72%	$0.04	$0.07
15-M Average Difference	1.51%	1.51%	$0.14	$0.14
27-M Average Difference	2.53%	2.51%	$0.25	$0.25
Total Average Difference	1.48%	1.58%	$0.14	$0.16
Greatest Difference	8.94%	9.08%	$0.87	$1.90

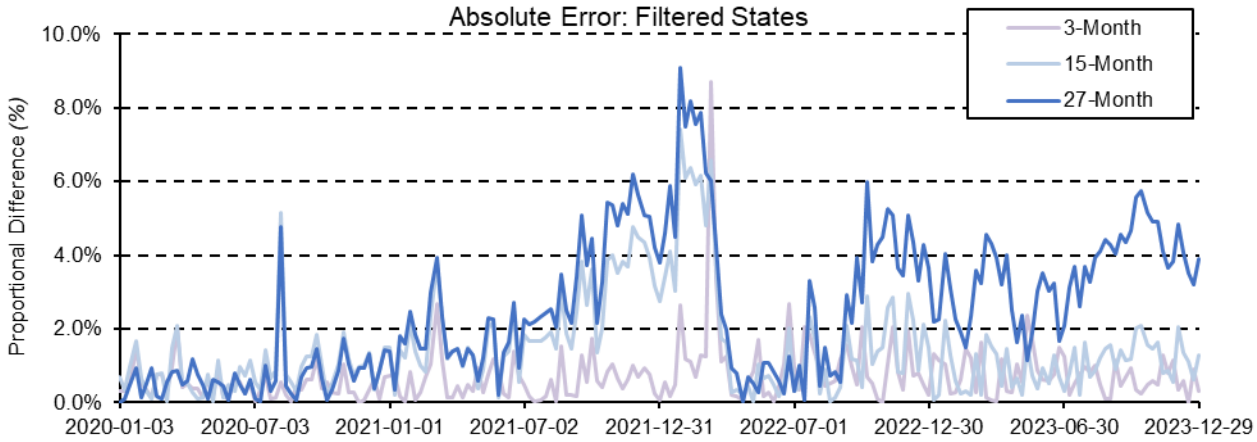


Figure 33: Nickel Market Error with Filtered Factor Values under the Gabillon Model

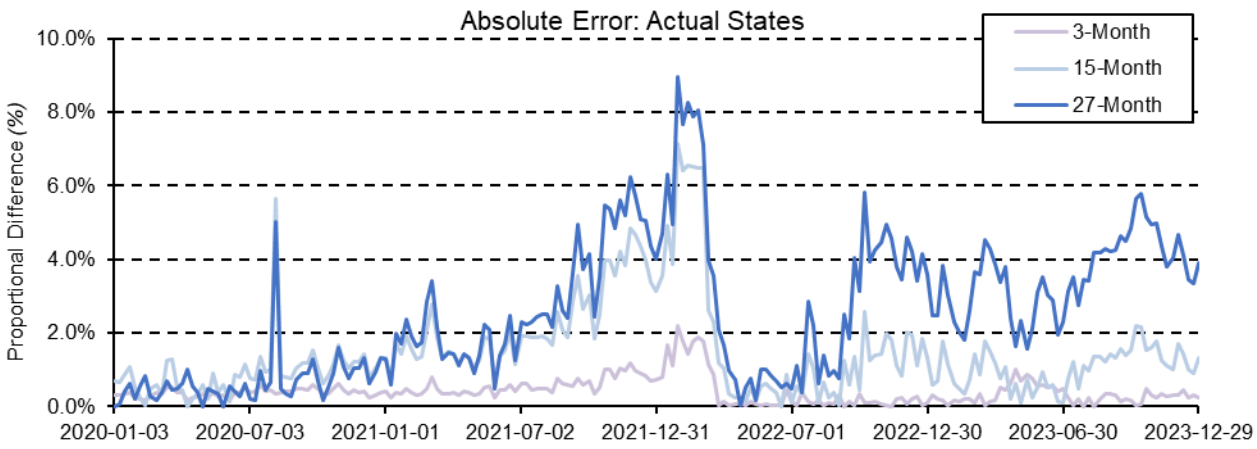


Figure 34: Nickel Market Error with Actual Factor Values under the Gabillon Model

In terms of modelling the shape of the forward curve, the Gabillon model also struggles for the nickel market as shown in Table 14. The performance for the contango is comparable to that of the proposed model (Table 6) while identification of backwardation is marginally better but still poor for the Gabillon model.

Table 14: Nickel Futures Curve Shape Results under the Gabillon Model

Contango	Inputs		Backwardation	Inputs	
	Actual	Filtered		Actual	Filtered
15-Month Sample	169	169	15-Month Sample	40	40
15-Month Correct Shape	168	168	15-Month Correct Shape	7	7
Proportion	**99.4%**	**99.4%**	**Proportion**	**17.5%**	**17.5%**
27-Month Sample	173	173	27-Month Sample	36	36
27-Month Correct Shape	173	173	27-Month Correct Shape	8	8
Proportion	**100.0%**	**100.0%**	**Proportion**	**22.2%**	**22.2%**

By comparing the fit of the filtered spot price to the actual spot price of the proposed model (Figure 23) and that of the Gabillon model (Figure 35), we see that a much closer fit for the Gabillon model. When looking at the long-term price and filtered spot price (Figure 36) to the historical forward curve (Figure 9), we see that some sections of the long-term price do not reflect the shape of the forward curve around the short squeeze of March 2022 and this resulted in the poor performance in identifying backwardation.

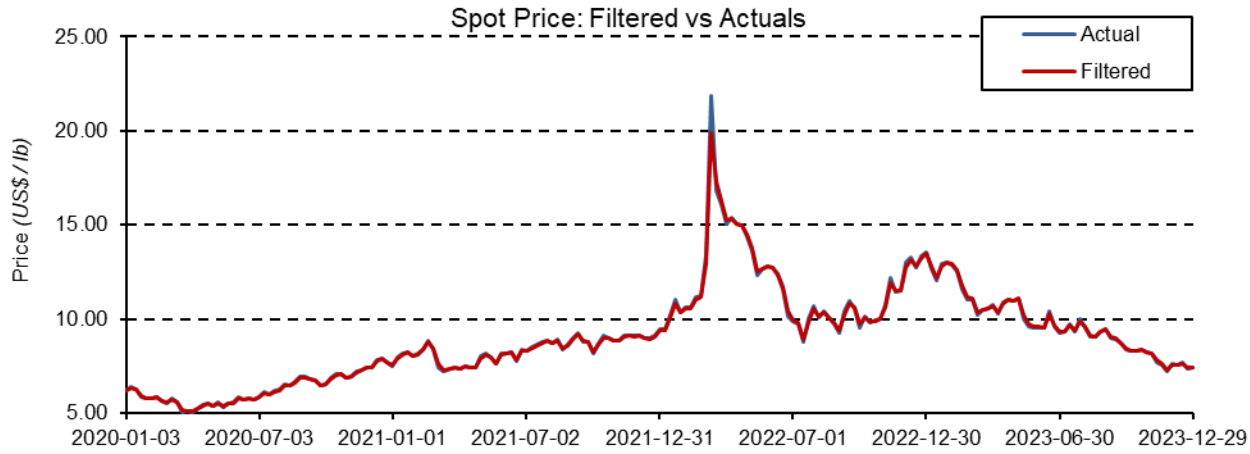


Figure 35: Actual and Filtered Nickel Spot Prices under the Gabillon Model

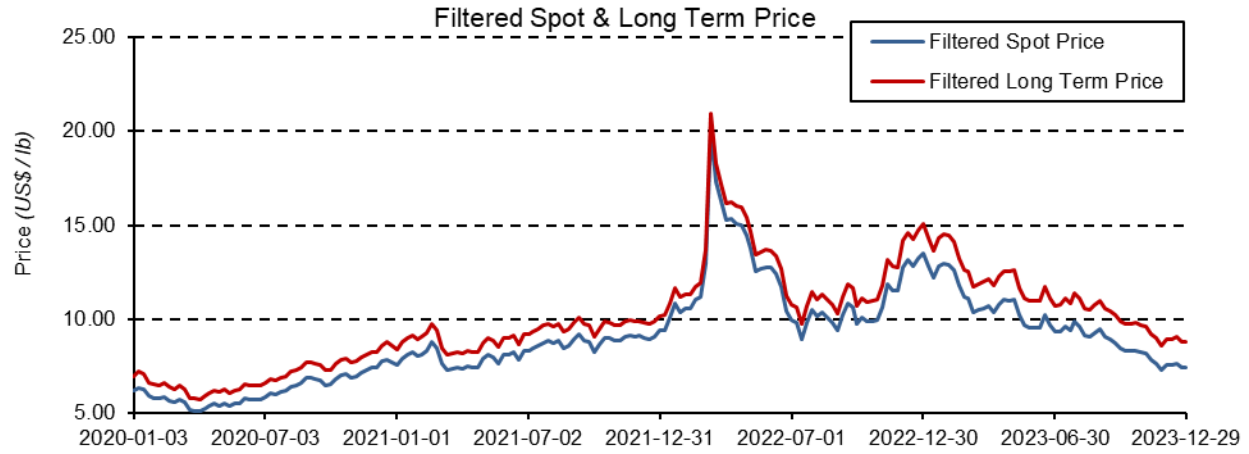


Figure 36: Filtered Nickel Spot and Long-Term Price under the Gabillon Model

4.3.5 Discussion

The gain in performance of the Gabillon model highlights one of the key features of the commodity models, namely the reliance on latent, unobserved factors. The models in the literature posit the existence of these factors and provides supposed dynamics of these factors based on ex-post empirical evidence or drawn ex-ante from economics arguments. For the Gabillon and Schwartz & Smith models this occurs through the inclusion of a long-term price which is assumed to follow a GBM. Similarly, the model of Gibson & Schwartz includes a convenience yield which follows a mean reverting Ornstein-Uhlenbeck process. This latent or “free” parameter can then be used to provide a close fit to empirical data as the filtering process by which it is determined is based on the observed forward curves. The second factor is free in as much as its value can be determined based on the observation. However, the choice of parameters for the dynamics of this second factor plays an important role as the likelihood is determined based on the expected observation based on the prior value of the factor. Leaving the second factor to be “free” contrasts with the proposed model which relied on a second factor, changes in inventory levels, which is deemed to be observable. As the second factor is observable, it is not “free” to take a value to be determined by the observation made at each step and hence the quality of the fit was adversely affected.

As previously discussed, exchange inventory levels are used as a proxy for actual global inventory levels which cannot be observed. In the empirical results for the proposed model, a trend in all three markets tested empirically is that there is poor fidelity between the filtered changes in inventory and the actual changes in inventory. The low fidelity of the filtered values implies that the changes in inventory are poor indicators of the shape of the forward curves. The original underlying rationale for the model is to cumulate the changes in inventory, scaled by the functions H_1 and H_2, with the effect of changes in inventory decaying over time. However, the results in (Barndorff-Nielsen, Benth, & Veraart, 2013) show that this is equivalent to the system of SDEs provided in earlier sections where only the current level of the change in inventory is sufficient to determine the dynamics of the system. With this formulation, it becomes easier to see that the changes in inventory are not well suited to model the behavior of the spot price and by extension the forward curve.

The relationships between inventory levels and spot prices, however, are generally based on global inventory levels which cannot be observed as opposed to exchange inventory levels. As such, the use of exchange inventory levels as proxies introduces additional noise in the empirical evaluation and there is no way to confirm the quality of the proxy being used. As such, while the empirical results based on exchange inventory levels are poor, it remains unclear whether actual global inventory levels will be well suited to model spot prices and forward curves.

In light of the empirical results obtained and their comparison to the Gabillon model we have seen that the proposed model largely fails to appropriately replicate the observed forward curves. These difficulties are mainly driven by the changes in exchange inventory levels which, when used as a weighted moving average are poor indicators for the shape of the forward curves. To remedy this shortcoming, we can look for other sources of information on inventory that may be more representative of global inventory. Alternatively, we could look towards the commonly used “free” second factor and presume that inventories cannot be observed. In doing so we can have the second factor take on a value which is representative of the forward curve observed. The primary difference with the other two-factor model in this case would then be the interpretation of the second factor as being changes in inventory.

Chapter 5

Conclusion and Future Studies

In this thesis we have explored spot price models used in commodity markets. In particular we have explored the application of the LSS modelling framework to storable commodities. While the application of LSS processes has been applied to electricity and power markets, those are not storable commodities. However, little attention has been paid to their use for storable commodities as attested by leading researchers in the field due to technical difficulties. We have two main contributions to the application of LSS processes to storable commodities. The first is to show that existing, successful, two-factor spot price models from the literature can be embedded within the LSS framework. The second contribution is a newly proposed model for the spot price of storable commodities that relies on the "memory" property of the LSS framework and the theory of storage.

Literature on the LSS framework, as utilized for commodity markets, had noted that a number of existing spot price models are embedded in the LSS framework. However, when it comes to storable commodities these observations are limited in scope and do not include traditionally successful two-factor spot price models. In this thesis we have shown that the LSS framework embeds a number of these successful spot price models. We proceeded by first showing that the model from Gabillon belongs to the LSS framework, albeit with a reduced parameter set as the correlation between the factors was assumed to be zero. Results from the literature have shown that the Gabillon model is equivalent to those of Gibson & Schwartz and Schwartz & Smith in the sense that their factors are linear combinations of one another. As such, by showing that the Gabillon model with a reduced parameter set belongs to the LSS framework, we can infer that the Gibson & Schwartz and Schwartz & Smith models also belong to the LSS framework albeit subject to certain restrictions on their parameter sets. To the best of our knowledge this result is new and enables significantly more flexibility in using the LSS framework to model storable commodity markets.

In addition to extending the LSS framework to existing models in the literature, this thesis has also explored applications of the LSS framework to spot price models of storable commodities by proposing a new model. The proposed model has two factors: spot price and changes in inventory levels. The model applies implications of the theory of storage whereby spot prices and their volatility are expected to increase as inventory levels decrease and vice versa. The "memory" of LSS processes is then used to tally the changes in inventory to impart the resulting effect to spot price of the corresponding changes in inventory levels. In general, global inventory levels cannot be observed, however futures exchanges may publish the amount of inventory held in

their warehouses. The exchange inventory data is used in our empirical study as a proxy for global inventory levels. We then proceed to develop a calibration approach by using the Unscented Kalman Filter and apply it to test the proposed model using empirical data from the aluminum, copper and nickel markets. The two main criteria by which the fit of the model was evaluated were the proportional error of the forward curves and the position of futures prices relative to spot prices. The proposed model however has mixed results in terms of fit to the empirical data and its ability to model the shape of the forward curve.

To benchmark the performance of the proposed model against other two factor models from the literature, the Gabillon model was implemented and applied to the same data sets. When comparing the results of the proposed model to that of the Gabillon model, it is clear that the Gabillon model substantially outperformed the proposed model across all of the metrics we used. The difficulties in the performance for the proposed model stem from two main sources. The first goes back to the theory of storage and its reliance on global inventories whereas only exchange inventories were observable. As there is no way to ascertain the relationship between exchange inventories and global inventories, it is unclear to what extent the use of exchange inventories is an appropriate proxy. The empirical results nonetheless show that exchange inventory levels are fairly poor indicators of forward curves. The second source of difficulties is that the second factor, changes in inventory levels, is tied to data as opposed to being left unobserved. This differs from other two factor models in the literature which include a latent factor to be determined solely by the calibration procedure as opposed to being treated as observable. The estimated value of this latent factor could then be used to ensure that the model's forward curve has a shape similar to the observed data. By keeping the second factor observable, this limits the flexibility of the model to accurately fit the empirical data unlike the models from the literature.

In short, in this thesis we have provided two avenues through which the LSS framework can be applied to storable commodities. The first is through the existing models as we showed that they belong to the LSS framework under a reduced parameters set and the second is that we propose a new model for spot prices of storable commodities. There remain however several directions for future research in this topic. Perhaps, as the results used provided only sufficient conditions for the semi-martingale condition, it could be possible to relax the restriction on the parameters of the two-factor models that are shown to belong to the LSS framework. For the proposed model, it may be possible to obtain broader data on inventories that provide better indicators on the forward curves or alternatively to experiment with similar models and leaving the second factor to be unobservable as is the case with a number of existing models in the literature.

References

Barndorff-Nielsen, O. E., & Schmiegel, J. (2008). Time Change, Volatility, and Turbulence. *Mathematical Control Theory and Finance* (pp. 29-53). Springer.

Back, J., & Prokopczuk, M. (2013). Commodity Price Dynamics and Derivative Valuation: A Review. *International Journal of Theoretical and Applied Finance*, Vol. 16, No. 6 1350032 (30 pages).

Barndorff-Nielsen, O. E., & Schmiegel, J. (2009). Brownian semistationary processes and volatility / intermittency. *Advanced Financial Modelling; Radon Series on Computational and Applied Mathematics*, Volume 8, 1-25.

Barndorff-Nielsen, O. E., Benth, F. E., & Veraart, A. D. (2010). *Modelling energy spot prices by Levy semistationary processes.* Aarhus: CREATES, School of Economics and Management.

Barndorff-Nielsen, O. E., Benth, F. E., & Veraart, A. D. (2013). Modelling energy spot prices by volatility modulated Lévy-driven Volterra processes. *Bernoulli Society for Mathematical Statistics and Probability*, Vol 19, Issue 3, 803-845.

Bjerksund, P. (1991). *Contingent Claims Evaluation when the Convenience Yield is Stochastic: Analytical Results.* Norwegian School of Economics and Business Administration.

Bolder, D. J. (2001). *Affine Term-Structure Models: Theory and Implementation.* Ottawa: Bank of Canada - Banque du Canada.

Brennan, M. J. (1958). The supply of storage. *American Economic Review*, 48, 50–72.

Brennan, M., & Schwartz, E. (1979). A Continuous Time Approach to the Pricing of Bonds. *Journal of Banking and Finance*, Vol 3; Issue 2; Pages 133-155.

Brennan, M., & Schwartz, E. (1985). Evaluating Natural Resource Investment. *The Journal of Business, Volume 58, Issue 2*, 135-157.

Cootner, P. H. (1960). Returns to Speculators: Telser versus Keynes. *Journal of Political Economy*, vol. 68, issue 4, 396.

Ewald, C.-O., Zhang, A., & Zong, Z. (2019). On the calibration of the Schwartz two-factor model to WTI crude oil options and the extended Kalman Filter. *Annals of Operations Research*, Issue 1-2; Pages 119-130.

Fama, E., & French, K. (1987). Commodity futures prices: Some evidence on forecast power, premiums and the theory of storage. *Journal of Business*, 60, 55–73.

Feller, W. (1965). *An Introduction to Probability Theory and Its Applications.* Princeton: John Wiley & Sons.

Gabillon, J. (1991). *The Term Structures of Oil Futures Prices.* Oxford Institute for Energy Studies.

Geman, H. (2005). *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy.* John Wiley & Sons.

Gibson, R., & Schwartz, E. (1990). Stochastic Convenience Yield and the Pricing of Oil Contingent Claims. *The Journal of Finance*, Vol. 45, No. 3, 959-976.

Guasoni, P., Miklós, R., & Walter, S. (2008). Consistent Price Systems and Face-Lifting Pricing under Transaction Costs. *The Annals of Applied Probability*, Vol. 18, No. 2, Pages 491-520.

Heston, S. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Option. *The Review of Financial Studies*, Vol. 6, No. 2, 327-343.

Jones, A. (2022). The Nickel Short Squeeze: What Happened? *International Banker*.

Keynes, J. M. (1930). *A Treatise on Money Vol. II: The Applied Theory of Money.*

LME. (2024, November). *LME Nickel - Contract specifications*. Retrieved from LME: https://www.lme.com/en/metals/non-ferrous/lme-nickel/contract-specifications

Nielsen, M., & Schwartz, E. (2004). Theory of Storage and the Pricing of Commodity Claims. *Review of Derivatives Research*, Vol. 7, 5–24.

Pakkanen, M. S. (2011). Brownian Semistationary Processes and Conditional Full Support. *International Journal of Theoretical and Applied Finance*, Vol. 14, No. 04, pp. 579-586.

Qi, J., Wu, T., Xu, Z., & Li, Z. (2024). An Even Analysis: The Nickel Short Squeeze by Tsingshan and Lessons for Enterprises. *2nd International Conference on Management Research and Economic Development.*

Schofield, N. C. (2021). *Commodity derivatives : markets and applications.* John Wiley & Sons.

Schwartz, E. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, Volume 52, Issue 3, 923-973.

Schwartz, E., & Smith, J. (2000). Short-Term Variations and Long-Term Dynamics in Commodity Prices. *Management Science*, Vol 46, No 7, 893-911.

Telser, L. (1958). Futures trading and the storage of cotton and wheat. *Journal of Political Economy*, 66, 233–255.

Trafigura Group. (2019). *Commodities Demystified: A Guide To Trading And The Global Supply Chain.*

Wan, E., & Nelson, A. T. (2001). Chapter 5: Dual Extended Kalman Filter Methods. In *Kalman Filtering and Neural Networks* (pp. 123-173). Hamilton: John Wiley & Sons, Inc.

Wan, E., & van der Merwe, R. (2001). Chapter 7: The Unscented Kalman Filter. In *Kalman Filtering and Neural Networks* (pp. 221-280). Hamilton: John Wiley & Sons, Inc.

Wang, Z. (2012, June 26). *Commodity Derivatives: Modeling and Pricing.* Paris: L'École nationale supérieure des mines de Paris.

Appendices

The appendices contain a few sections concerning the model and provides further details for some sections of the thesis. A number of the issues in the appendix do not have a resolution and are included to indicate the avenues explored in the preparation of this thesis.

Appendix A

Numerical Methods for the Pricing PDE

In this appendix, we present in detail a numerical method to approximate the solution to the pricing PDE in this thesis. We provide a full description of the domain discretization, boundary conditions and the resulting system of equations. However, as we currently do not have an appropriate functional form of the boundary conditions for the discretized domain, we have not been able to replicate the results of the simpler, albeit slower, Monte Carlo method presented in the main text.

A.1 Pricing PDE and Initial Conditions

Under the proposed model, the price of a general European derivative is a function of three variables: the spot prices S, the change in inventory denoted X and time to maturity τ. The relationship between the price and the variables is expressed as a PDE:

$$\frac{\sigma_S^2}{2}\frac{\partial^2 U}{\partial S^2} + \frac{\sigma_X^2}{2}\frac{\partial^2 U}{\partial X^2} + (\mu_S - \sigma_S\lambda_1)\frac{\partial U}{\partial S} + (\mu_X - \sigma_X\lambda_2)\frac{\partial U}{\partial X} - \frac{\partial U}{\partial \tau} - rU = 0$$

where $U(S, X, \tau)$ is the price of the derivative as a function of the three variables. The coefficients in the equation are taken from the SDEs governing the behavior of S and X, namely:

$$dS_t = \alpha(H_2(X_t) - ln(S_t))S_t dt + H_1(X_t)S_t dB_t^Y$$

$$d\mathrm{X}_t = \kappa(\theta - X_t)dt + \sigma dB_t^X$$

and for convenience, we use the following notations:

- $\mu_S = \alpha(H_2(X_t) - ln(S_t))S_t$
- $\sigma_S = H_1(X_t)S_t$
- $\mu_X = \kappa(\theta - X_t)$
- $\sigma_X = \sigma$

The solution is defined for any values that can be taken by the three variables: the spot price must remain positive hence $S \in [0, \infty)$, the change in inventory level takes values in $\mathbb{R}$ meaning $X \in (-\infty, \infty)$ and lasty time to maturity must be positive implying $\tau \in [0, \infty]$. The price of the derivative is known at maturity ($\tau = 0$) via a pay-off function $U(S, X, 0) = P(S, X)$ and this serves as an initial condition. Boundary conditions will be discussed in a later section.

For futures prices, the PDE does not include the undifferentiated term and becomes:

$$\frac{\sigma_S^2}{2}\frac{\partial^2 U}{\partial S^2} + \frac{\sigma_X^2}{2}\frac{\partial^2 U}{\partial X^2} + (\mu_S - \sigma_S\lambda_1)\frac{\partial U}{\partial S} + (\mu_X - \sigma_X\lambda_2)\frac{\partial U}{\partial X} - \frac{\partial U}{\partial \tau} = 0$$

The initial conditions for futures prices are set to $U(S, X, 0) = S$ as the futures price converges to spot as time to maturity goes to zero. The remainder of this thesis is focused on estimating the solution for futures prices but the method can be adjusted for the additional term when pricing European derivatives in general.

A.2 Domain Discretization

The domain discretization discussed in this section is described for the purposes of the calibration procedure outlined in the thesis. The domain discretization can be tailored as needed to particular problems or other derivatives. When calibrating the model, we are working with a historical time series of spot prices and changes in inventory levels. This time series has spot prices in a fixed range $[S_{min}^h, S_{max}^h]$ and inventory level changes ranging from $[X_{min}^h, X_{max}^h]$. To calibrate the model, forwards must be evaluated at all spot prices and inventory level changes of the historical time series. As such we should ensure that the grid used encompasses all historical values and provides sufficient distance between historical values and the boundaries.

For the spot price we consider a range of prices exceeding by 50% the maximum of the historical prices on the upper end and zero on the lower end, meaning that the grid will stretch from 0 to $S_{max} = 1.5 \times S_{max}^h$. As none of the spot price values have any particular meaning for futures prices (unlike, for example, a strike price for an option) we will use a uniform grid with N_S partitions meaning $N_S + 1$ points.

We proceed similarly with inventory level changes as we did for the spot price, exceeding the historical levels by 50% in absolute value meaning that the grid will stretch from $X_{min} = -1.5 \times \left|X_{min}^h\right|$ to $X_{max} = 1.5 \times X_{max}^h$. We again use a uniform grid as there is no particularly meaningful inventory change level. The discretization has N_x partitions meaning that the grid will contain $N_x + 1$ points.

For time to maturity, we are working with forward curves that stretches to 63-month. We use weekly increments from $\tau = 0$ to the greatest maturity of the forward curve with each time increment being equivalent to $1/52$ years.

A.2.1 Spatial Boundary Conditions

We have discussed above the initial conditions described by the pay-off function. This initial condition serves as the value of the solution at $\tau = 0$. As we are estimating the solution numerically, we are limited to a finite domain, and we must specify the behavior of the solution at the boundaries of the discretized domain. We refer to the variables S and X as spatial variables and to τ as a time variable. This distinction will be helpful when describing the methodology used to estimate the solution to the PDE.

There are two common or standard boundary conditions: Dirichlet and Neuman boundary conditions. For Dirichlet boundary conditions the value of the solution is specified at the boundary while for Neuman boundary conditions the value of the derivative of the solution normal to the boundary is specified. In the case where the discretized domain is aligned with the variables of the problem, the derivatives specified by Neuman boundary conditions are simply partial derivatives with respect to the variables in question.

For the spot price, we apply Neuman conditions at S_{max}, where the underlying assumption is that the futures curve will move in accordance with changes in the spot price. This is meant to provide a boundary which would not have a significant effect on the regions that will be used for calibration. This condition can be written:

$$\frac{\partial U}{\partial S}(S_{max}, X, \tau) = 1$$

On the lower end of the grid, we have that $S = 0$. For this boundary we will use a Dirichlet boundary condition and set the value of the function at zero. The rationale here is that if the spot price of the commodity were to reach zero then there would no longer be any value in this commodity and the resulting forward prices would also be zero. This last condition takes the form:

$$U(0, X, \tau) = 0$$

We should also provide boundary conditions for the inventory levels. For this factor, however it is not clear what the behavior of the function should be at the boundary. As we do not know the value of the function nor its behavior, we will set these to be Neuman boundary condition based on some functions $F_{min}(\tau, S)$ at X_{min} and $F_{max}(\tau, S)$ at X_{max}.

A.3 Finite Difference Scheme

The solution we are estimating solves a parabolic PDE initial value problem with two spatial variables (S and X). The solution proceeds from the known value at $\tau = 0$ given by the initial value. The solution then proceeds to larger values of τ in a sequence of steps, each being computed based on the previous step. To utilize the relationship outlined by the PDE, we first need to use estimates of the corresponding derivatives. Before describing the finite difference schemes, we first provide some additional notation.

For the grid, as we are using a uniform grid for all variables, we have the following increments:

- $\Delta S = \frac{S_{max} - S_{min}}{N_S}$ as the spot price increment
- $\Delta X = \frac{X_{max} - X_{min}}{N_X}$ as the inventory change increment
- Δt as the time increment

The estimated solution is expressed as $u_{i,j}^k \approx U(i\Delta S, X_{min} + j\Delta X, k\Delta t)$ where the i, j and k are indexes referring to the grid point at which the solution is evaluated/estimated.

A.3.1 Spatial Derivatives

For the spatial derivative, we use a central difference scheme for both the first and second derivatives. These schemes take the form:

$$\frac{\partial U}{\partial S}(i\Delta S, X_{min} + j\Delta X, k\Delta t) \approx \frac{u_{i+1,j}^k - u_{i-1,j}^k}{2\Delta S}$$

$$\frac{\partial^2 U}{\partial S^2}(i\Delta S, X_{min} + j\Delta X, k\Delta t) \approx \frac{u_{i+1,j}^k - 2u_{i,j}^k + u_{i-1,j}^k}{\Delta S^2}$$

For brevity we denote the finite differences by using the δ operators which refer to:

$$\delta_S u_{i,j}^k = u_{i+1,j}^k - u_{i-1,j}^k$$

$$\delta_S^2 u_{i,j}^k = u_{i+1,j}^k - 2u_{i,j}^k + u_{i-1,j}^k$$

The same arguments apply analogously for the finite difference approximations of the derivatives with respect to X. Using Taylor series arguments, these approximation schemes can be shown to be second order accurate with the error terms being on the order of ΔS^2 or ΔX^2.

The schemes above are well suited for all interior points of the grid but need to be revised for the boundaries subject to Neuman boundary conditions as points on both sides of the edges of the discretized domain are not available. This is not necessary for boundaries subject to Dirichlet boundary conditions as the solution does not need to be estimated at these boundaries. For Neuman boundary conditions, the first derivative is known as it is imposed by the boundary condition itself, however the second derivative must still be estimated. To this end, a central difference scheme can still be used at the boundary but a "ghost" point outside of the grid is used. While the value of the solution is not known/not well defined for this "ghost point", the value of the solution there can be estimated by using the Neuman boundary condition. Assuming a central difference scheme at a boundary, for example with $u_{-1,j}^k$ as the "ghost" point and b the value of the derivative of the Neuman boundary condition we have:

$$\frac{\partial U}{\partial S}(0, X_{min} + j\Delta X, k\Delta t) = b \approx \frac{u_{1,j}^k - u_{-1,j}^k}{2\Delta S}$$

This implies:

$$u_{-1,j}^k \approx u_{1,j}^k - 2\Delta S b$$

Using this estimate for $u_{-1,j}^k$ we can now approximate the second derivative on the boundary:

$$\frac{\partial^2 U}{\partial S^2}(0, X_{min} + j\Delta X, k\Delta t) \approx \frac{u_{1,j}^k - 2u_{0,j}^k + u_{-1,j}^k}{\Delta S^2} \approx \frac{u_{1,j}^k - 2u_{0,j}^k + (u_{1,j}^k - 2\Delta S b)}{\Delta S^2}$$

A similar argument applies to any Neuman boundary conditions.

A.3.2 Time Derivative

As the solution methodology progresses forward in time from the initial value, we will not have access to sufficiently many points for a central difference scheme. Instead, we will use a first order scheme to estimate the time derivative:

$$\frac{\partial U}{\partial \tau} \approx \frac{u_{i,j}^{k+1} - u_{i,j}^k}{\Delta t}$$

A.4 Numerical Solution

Using the above finite difference schemes, we can obtain a system of linear equations at each time step. For convenience we define the following coefficients in the pricing PDE:

- $a_S = \frac{\sigma_S^2}{2}$
- $a_X = \frac{\sigma_X^2}{2}$
- $b_S = (\mu_S - \sigma_S \lambda_1)$
- $b_X = (\mu_X - \sigma_X \lambda_2)$

The resulting PDE for futures prices then has the form given below where in general each coefficient can be a function of S, X, and τ.

$$\frac{\partial U}{\partial \tau} = a_S \frac{\partial^2 U}{\partial S^2} + a_X \frac{\partial^2 U}{\partial X^2} + b_S \frac{\partial U}{\partial S} + b_X \frac{\partial U}{\partial X}$$

At every value of τ, we have to estimate the solution at all points on a two-dimensional grid by solving a system of linear equations obtained from the PDE. To express this system of equations into matrix form, the grid points must be converted to a vector. The resulting system of equations, for each time step will then have $N_S \times N_X$ equations. The matrix representation of the system of the equation will depend on the ordering of the grid as a vector. In the matrices developed below, the vector of the estimated solution at time k will be constructed by the vertical concatenation of vectors with constant values of X and S ranging from 0 to S_{max}. We refer to each vector with a fixed value of X and the full range of values of S as a "stack".

$$u^k = \begin{bmatrix} u^k_{0,0} \\ \vdots \\ u^k_{N_S,0} \\ \vdots \\ \\ \vdots \\ u^k_{0,N_X} \\ \vdots \\ u^k_{N_S,N_X} \end{bmatrix}$$

and we denote the solution vector at step k as u^k.

In obtaining the solution to the PDE, we will use the implicit method. This choice is motivated by its relative simplicity and stability, but it is somewhat demanding computationally. The computational cost of the implicit method can be mitigated by using Alternating Direction Implicit (ADI) methods. The implicit method, like other numerical methods, proceeds recursively from the initial conditions where the solution is known and takes "steps" forward in time by solving a system of linear equations. As such, to utilize the implicit method it is sufficient to describe the procedure by which these "steps" are taken. In the following description we assume that $u^k_{i,j}$ is known for all i and j and we want to outline the implicit method to obtain the values of $u^{k+1}_{i,j}$ for all i and j.

The systems of equations are constructed by using the finite difference approximations outlined in the previous section. Starting with the interior points (those not on the boundary of the grid), we have the following equation, simply by replacing the partial derivatives with the finite difference approximation. For the implicit method, the spatial derivatives are estimated based on the values of $u^{k+1}_{i,j}$ resulting in the following expression:

$$\frac{u^{k+1}_{i,j} - u^k_{i,j}}{\Delta t} = \frac{a_S \delta_S^2 u^{k+1}_{i,j}}{\Delta S^2} + \frac{a_X \delta_X^2 u^{k+1}_{i,j}}{\Delta X^2} + \frac{b_S \delta_S u^{k+1}_{i,j}}{2\Delta S} + \frac{b_X \delta_X u^{k+1}_{i,j}}{2\Delta X}$$

Expanding the δ operators would then provide a linear equation relating the values of $u^{k+1}_{i,j}$, $u^{k+1}_{i+1,j}$, $u^{k+1}_{i-1,j}$, $u^{k+1}_{i,j+1}$ and $u^{k+1}_{i,j-1}$. This expression can be extended to $u^{k+1}_{i,j}$ for all $i = 0 \ldots N_S$ and $j = 0 \ldots N_X$ (after accounting for boundary conditions) which yields a system of $(N_S + 1) \times (N_x + 1)$ linear equations to be solved at each time step.

The right-hand side of the finite difference equation can be expressed as a matrix. Denote the matrix representing the finite difference expression of the spatial derivatives as A. Using this we can simplify the expressions of the implicit method:

$$\frac{u^{k+1} - u^k}{\Delta t} = Au^{k+1} + B$$

where B is a vector of constants dependent on the boundary conditions. At least we now have a system of linear equations to solve for the estimated solution at the next time step (u^{k+1}):

$$(I - \Delta t A)u^{k+1} = u^k + B\Delta t$$

A.5 Linear System of Equations

Using the PDE and the boundary conditions, we can now provide a matrix representation for the linear system of equations at each step. To simplify the construction of the system of equation and the vector of constants, the matrix used for the system of equation is split into the matrices $A = A_S + A_X$ where A_S will include the derivatives with respect to S and A_X is defined analogously. Similarly, the vectors of constants used for the boundary conditions will be split into $B = B_S + B_X$. The matrices will be defined in accordance with the vector defined above where the grid is expressed in "stacks" of constant X with S ranging from 0 and S_{max}. The matrices will be expressed as block matrices where each block will be related to a "stack" of constant X.

Starting with the matrix A_S we have the following terms to include in the system for the point i, j where $i \neq 0, N_S$:

$$\begin{aligned} a_S \frac{\partial^2 U}{\partial S^2} + + b_S \frac{\partial U}{\partial S} &= a_S \frac{u^k_{i+1,j} - 2u^k_{i,j} + u^k_{i-1,j}}{\Delta S^2} + b_S \frac{u^k_{i+1,j} - u^k_{i-1,j}}{2\Delta S} + \frac{r u^k_{i,j}}{2} \\ &= u^k_{i+1,j}\left(\frac{a_S}{\Delta S^2} + \frac{b_S}{2\Delta S}\right) + u^k_{i,j}\left(-\frac{2a_S}{\Delta S^2} + \frac{r}{2}\right) + u^k_{i-1,j}\left(\frac{a_S}{\Delta S^2} - \frac{b_S}{2\Delta S}\right) \\ &= u^k_{i+1,j} c^{+1}_{i,j} + u^k_{i,j} c^0_{i,j} + u^k_{i-1,j} c^{-1}_{i,j} \end{aligned}$$

where in the last line we have defined the coefficients $c^{+1}_{i,j}$, $c^0_{i,j}$ and $c^{-1}_{i,j}$ for brevity.

Each stack has two points on the boundary, the first point at $S = 0$ and the last point at $S = S_{max}$. For the point $S = 0$, we have Dirichlet boundary conditions which are constant in time implying that:

$$u^k_{0,j} = u^{k-1}_{0,j} = \cdots u^0_{0,j} = 0$$

for any $j \in \{0 \ldots N_X\}$.

For $S = S_{max}$ we have Neuman boundary conditions such that:

$$\frac{\partial U}{\partial S}(S_{max}, X, \tau) = 1$$

To obtain the equation at the boundary, we can proceed with the same method as the interior point but the constant vector will be needed due to the Neuman boundary condition. To obtain an equation for the second derivative at the boundary we use a "ghost" point ($u^k_{N_S+1,j}$) by assuming a central difference scheme at the boundary:

$$\frac{\partial U}{\partial S}(S_{max}, X, \tau) = \frac{u^k_{N_S+1,j} - u^k_{N_S-1,j}}{2\Delta S} = 1$$

which further implies that:

$$u^k_{N_S+1,j} = 2\Delta S \ + u^k_{N_S-1,j}$$

Using this result we can apply a central difference scheme at the boundary for the second derivative term:

$$\begin{aligned} a_S \frac{\partial^2 U}{\partial S^2}(S_{max}, X, \tau) &= a_S \left(\frac{u^k_{N_S+1,j} - 2u^k_{N_S,j} + u^k_{N_S-1,j}}{\Delta S^2} \right) \\ &= a_S \left(\frac{\left(2\Delta S \ + u^k_{N_S-1,j}\right) - 2u^k_{N_S,j} + u^k_{N_S-1,j}}{\Delta S^2} \right) \\ &= \frac{2a_S}{\Delta S} + u^k_{N_S,j}\left(-\frac{2a_S}{\Delta S^2}\right) + u^k_{N_S-1,j}\left(\frac{2a_S}{\Delta S^2}\right) \end{aligned}$$

This last equation, combined with the term from the first derivative provides the equation for the last element of each interior stack:

$$\begin{aligned} a_S \frac{\partial^2 U}{\partial S^2} + b_S \frac{\partial U}{\partial S} &= \frac{2a_S}{\Delta S} + u^k_{N_S,j}\left(-\frac{2a_S}{\Delta S^2}\right) + u^k_{N_S-1,j}\left(\frac{2a_S}{\Delta S^2}\right) + b_S(1) \\ &= \left(\frac{2a_S}{\Delta S} + b_S\right) + u^k_{N_S,j}\left(-\frac{2a_S}{\Delta S^2}\right) + u^k_{N_S-1,j}\left(\frac{2a_S}{\Delta S^2}\right) \end{aligned}$$

We can now construct the matrix A_S with the only nonzero entries in the $\boldsymbol{C}$ block matrices

$$A_S = \begin{bmatrix} \boldsymbol{C_0} & \mathbf{0} & & & \\ \mathbf{0} & \boldsymbol{C_1} & \cdots & \mathbf{0} & \\ \vdots & & \ddots & \vdots & \\ & & \boldsymbol{C_{N_X-1}} & \mathbf{0} & \\ \mathbf{0} & & \cdots & \mathbf{0} & \boldsymbol{C_{N_X}} \end{bmatrix}$$

With each $\boldsymbol{C_j}$ defined as:

$$\boldsymbol{C_j} = \begin{bmatrix} 1 & 0 & 0 & & & & \\ c^{-1}_{1,j} & c^{0}_{1,j} & c^{+1}_{1,j} & \cdots & & \mathbf{0} & \\ 0 & c^{-1}_{2,j} & c^{0}_{2,j} & & & & \\ & \vdots & & \ddots & & \vdots & \\ & & & & c^{0}_{N_S-2,j} & c^{+1}_{N_S-2,j} & 0 \\ & \mathbf{0} & & \cdots & c^{-1}_{N_S-1,j} & c^{0}_{N_S-1,j} & c^{+1}_{N_S-1,j} \\ & & & & 0 & 2a_S/\Delta S^2 & -2a_S/\Delta S^2 \end{bmatrix}$$

In the matrix $\boldsymbol{C_j}$, the PDE coefficients are evaluated at the corresponding index i, j for all terms (for example $c^0_{2,j}$ has the PDE coefficients computed using the grid point values $S = 0 + 2\Delta S$ and $X = X_{min} + j\Delta X$).

Finally, we can define the vector B_S associated with A_S

$$B_S = \begin{bmatrix} \boldsymbol{\beta}_0^S \\ \vdots \\ \boldsymbol{\beta}_{N_X} \end{bmatrix}$$

where each $\boldsymbol{\beta_S}$ is a vector with N_S entries with values:

$$\boldsymbol{\beta}_i^S = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \frac{2a_S}{\Delta S} + b_S \end{bmatrix}$$

The remaining matrix to construct contains the conditions imposed by the partial derivatives with respect to X. For the point i,j where $j \neq 0, N_X$ we proceed with the same method as the A_S matrix:

$$\begin{aligned} a_X \frac{\partial^2 U}{\partial X} + + b_X \frac{\partial U}{\partial X} &= a_X \frac{u_{i,j+1}^k - 2u_{i,j}^k + u_{i,j-1}^k}{\Delta X^2} + b_X \frac{u_{i,j+1}^k - u_{i,j-1}^k}{2\Delta X} \\ &= u_{i,j+1}^k \left(\frac{a_X}{\Delta X^2} + \frac{b_X}{2\Delta X} \right) + u_{i,j}^k \left(-\frac{2a_X}{\Delta X^2} \right) + u_{i,j-1}^k \left(\frac{a_X}{\Delta X^2} - \frac{b_X}{2\Delta X} \right) \\ &= u_{i,j+1}^k d_{i,j}^{+1} + u_{i,j}^k d_{i,j}^0 + u_{i,j-1}^k d_{i,j}^{-1} \end{aligned}$$

where we define the coefficients $d_{i,j}^{+1}$, $d_{i,j}^0$ and $d_{i,j}^{-1}$ for brevity.

The boundary conditions associated with the variable S are included in the matrix A_S and the vector B_S and do not need to be reflected in either A_X or B_X. The only remaining boundary conditions to be included are those at X_{min} and X_{max}. Here we rely on arbitrary functions $F_{min}(\tau, S)$ at X_{min} $F_{max}(\tau, S)$ at X_{max} for the Neuman boundary conditions as it is not clear how the solution behaves at these boundaries. Using "ghost" points we obtain the following equation for X_{min} at the level $S = i\Delta S$

$$\begin{aligned} a_X \frac{\partial^2 U}{\partial X} + + b_X \frac{\partial U}{\partial X} &= -\frac{2a_X}{\Delta X} F_{min}(\tau, i\Delta S) + u_{i,0}^k \left(-\frac{2a_X}{\Delta X^2} \right) + u_{i,1}^k \left(\frac{2a_X}{\Delta X^2} \right) + b_X F_{min}(\tau, i\Delta S) \\ &= \left(b_X F_{min}(\tau, i\Delta S) - \frac{2a_X}{\Delta X} F_{min}(\tau, i\Delta S) \right) + u_{i,0}^k \left(-\frac{2a_X}{\Delta X^2} \right) + u_{i,1}^k \left(\frac{2a_X}{\Delta X^2} \right) \\ &= d_{i,0}^{min} + u_{i,0}^k d_{i,0}^0 + u_{i,1}^k d_{i,0}^{+1} \end{aligned}$$

where $d_{i,0}^{min}$ is added to the constant vector B_X.

Similarly, for X_{max} we have:

$$\begin{aligned} a_X \frac{\partial^2 U}{\partial X} + + b_X \frac{\partial U}{\partial X} &= \frac{2a_X}{\Delta X} F_{max}(\tau, i\Delta S) + u_{i,N_x}^k \left(-\frac{2a_X}{\Delta X^2} \right) + u_{i,N_X-1}^k \left(\frac{2a_X}{\Delta X^2} \right) + b_X F_{max}(\tau, i\Delta S) \\ &= \left(b_X F_{max}(\tau, i\Delta S) + \frac{2a_X}{\Delta X} F_{max}(\tau, i\Delta S) \right) + u_{i,N_X}^k \left(-\frac{2a_X}{\Delta X^2} \right) + u_{i,N_X-1}^k \left(\frac{2a_X}{\Delta X^2} \right) \end{aligned}$$

$$= d_{i,N_x}^{max} + u_{i,N_X}^{k} d_{i,N_x}^{0} + u_{i,N_X}^{k} d_{i,N_x}^{-1}$$

where d_{i,N_x}^{max} is a constant in the vector B_X.

Putting all of the above together we can express the block matrix A_X as follows:

$$A_X = \begin{bmatrix} \boldsymbol{D_0^0} & \boldsymbol{D_0^{+1}} & \mathbf{0} & & & & \\ \boldsymbol{D_1^{-1}} & \boldsymbol{D_1^0} & \boldsymbol{D_1^{+1}} & \dots & & \mathbf{0} & \\ \mathbf{0} & \boldsymbol{D_2^{-1}} & \boldsymbol{D_2^0} & & & & \\ & \vdots & & \ddots & & \vdots & \\ & & & & \boldsymbol{D_{N_X-2}^0} & \boldsymbol{D_{N_X-2}^{+1}} & \mathbf{0} \\ & \mathbf{0} & & \dots & \boldsymbol{D_{N_X-1}^{-1}} & \boldsymbol{D_{N_X-1}^0} & \boldsymbol{D_{N_X-1}^{+1}} \\ & & & & \mathbf{0} & \boldsymbol{D_{N_X}^{-1}} & \boldsymbol{D_{N_X}^0} \end{bmatrix}$$

where each type of block has the form:

$$\boldsymbol{D_j^{-1}} = \begin{bmatrix} d_{0,j}^{-1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & d_{N_S,j}^{-1} \end{bmatrix}$$

$$\boldsymbol{D_j^{0}} = \begin{bmatrix} d_{0,j}^{0} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & d_{N_S,j}^{0} \end{bmatrix}$$

$$\boldsymbol{D_j^{+1}} = \begin{bmatrix} d_{0,j}^{+1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & d_{N_S,j}^{+1} \end{bmatrix}$$

The constant vector B_X only has non-zero entries on first and last "stack" and can be written as:

$$B_X = \begin{bmatrix} \boldsymbol{\beta_X^{min}} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \boldsymbol{\beta_X^{max}} \end{bmatrix}$$

where we have:

$$\boldsymbol{\beta_X^{min}} = \begin{bmatrix} d_{0,0}^{min} \\ \vdots \\ d_{N_S,0}^{min} \end{bmatrix}$$

$$\boldsymbol{\beta_X^{max}} = \begin{bmatrix} d_{0,N_x}^{max} \\ \vdots \\ d_{N_S,N_X}^{max} \end{bmatrix}$$

A.6 Example

Using the methodology outlined in the previous section we in principle can estimate the solution to the PDE for futures prices. However, as we currently do not have an appropriate boundary

condition for the X_{min} and X_{max} boundaries, the results will not be consistent with the Monte Carlo method used in the main text. Below we present an example of the results of the methodology and a comparison to the value obtained via the Monte Carlo method.

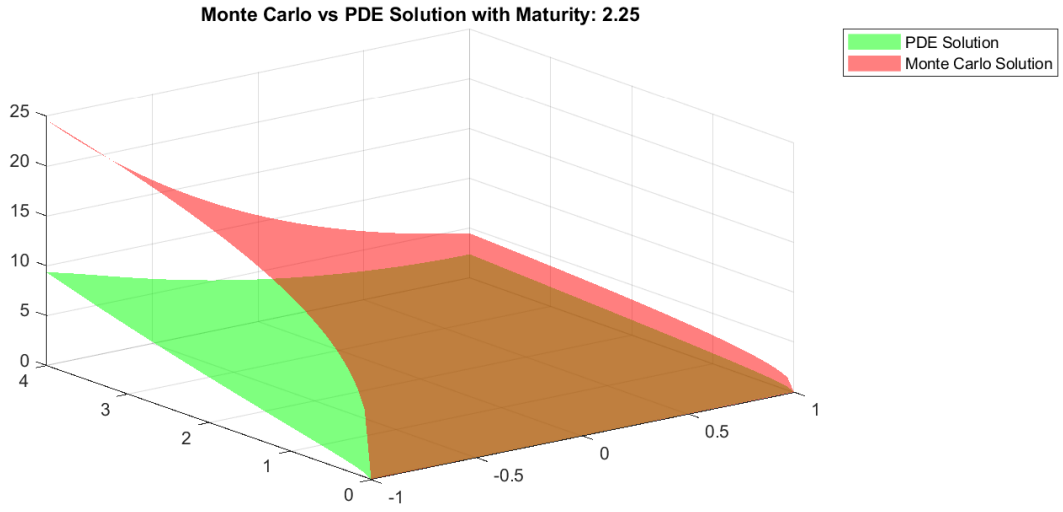


Figure 37: Comparison of the Monte Carlo and PDE Solutions with Maturity of 2.25 Years

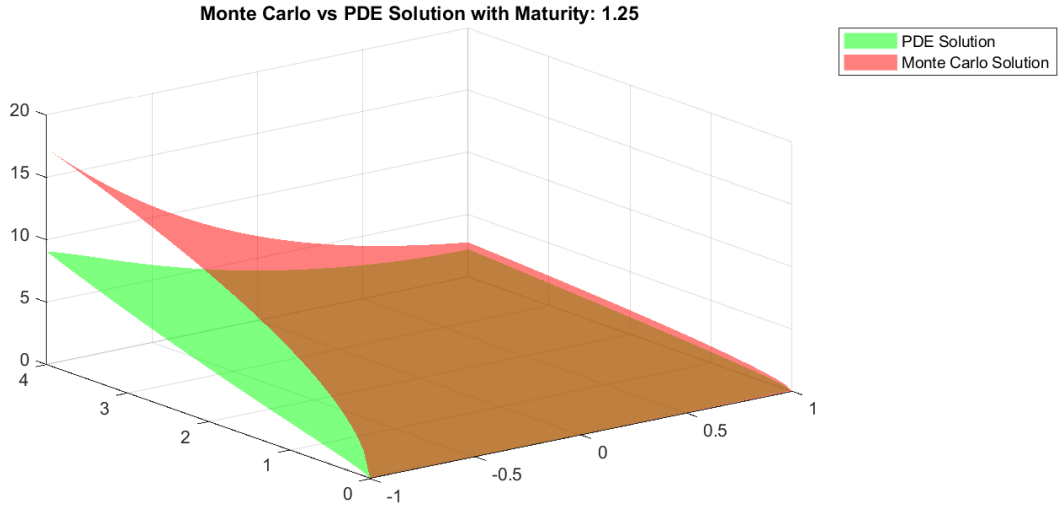


Figure 38: Comparison of the Monte Carlo and PDE Solutions with Maturity of 1.25 Years

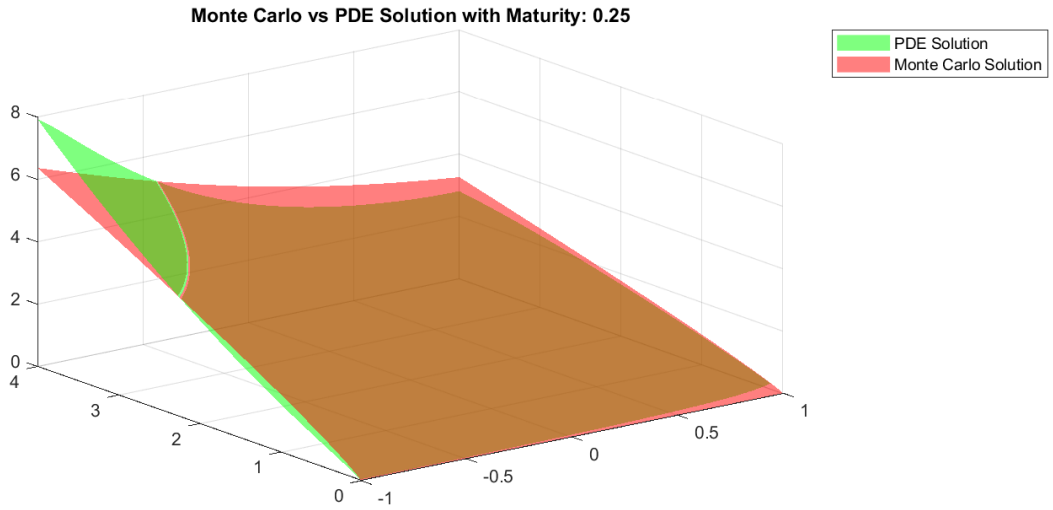


Figure 39: Comparison of the Monte Carlo and PDE Solutions with Maturity of 0.25 Years

Appendix B

Approaches to Valuation under the LSS Framework

In this appendix we outline two approaches for valuing contingent claims in the LSS Framework. The first is based on a semi-martingale property of the model and looks at the classical risk neutral measure where the discounted asset price is a martingale. The second approach to valuation is based on the existence of consistent price systems where a friction is assumed in the market and prices are obtained by superreplication. The friction assumed in the market can be reduced to zero to obtain asymptotic superreplication prices which is similar to assuming no market friction.

B.1 Semi-martingale Representation

Recall the semi-martingale representation shown in the main text of this thesis and reproduced for convenience:

$$Y_t = Y_0 + \int_0^t H_1(X_s) dL_s + \int_0^t -\alpha Y_s + H_2(X_s)\, ds$$

This model can also be expressed by using an SDE as:

$$dY_t = \alpha(H_2(X_t) - Y_t)\, dt + H_1(X_t) dB_t^Y$$

up to a rescaling of the H_2 function.

B.2 Risk Neutral Measure

The price dynamics are expressed as $S_t = \exp(Y_t)$.

For the change of measure, we want to find a risk neutral measure $\mathbb{Q}$ such that the risk neutral dynamics of Y_t would be $dY_t = H_1(X_t) dL_t + \left(r - \left(\frac{1}{2}\right) H_1(X_t)^2\right) dt$. This choice of the measure $\mathbb{Q}$ would ensure that the drift term of the price process, S_t, is given by $rS_t dt$. Fundamentally, this change of measure follows from dynamic hedging arguments used in the classical Black-Scholes model.

To find such a change of measure, we make the ansatz, for $\overline{B_t^Y}$ a $\mathbb{Q}$-Brownian motion and b_t an adapted process such that:

$$\overline{B_t^Y} = B_t^Y - \int_0^t b_s ds$$

Using the ansatz, under the measure $\mathbb{Q}$, Y_t has the following dynamics:

$$dY_t = H_1(X_t)d\overline{B_t^Y} + (-\alpha Y_t + H_2(X_t) + H_1(X_t)b_t)dt$$

From here on, we are looking for a process b_t such that the drift term becomes $\left(r - \left(\frac{1}{2}\right)H_1(X_t)^2\right)$. This implies:

$$b_t = \frac{r}{H_1(X_t)} + \frac{\alpha Y_t}{H_1(X_t)} - \frac{H_2(X_t)}{H_1(X_t)} - \frac{H_1(X_t)}{2}$$

It then remains to verify that the resulting process b_t is square integrable and that the resulting density process is a true martingale:

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left(\int_0^t b_s dL_s - \frac{1}{2}\int_0^t b_s^2 ds\right)$$

where $\mathcal{F}_t$ is the underlying filtration and $\mathbb{P}$ is the "real world measure.

We then look to apply a criterion, such as a Novikov or Kazamaki criterion, to ensure that b_t is a true martingale.

B.3 Consistent Price Systems

In section 3.3 of Barndorff-Nielsen, Benth, & Veraart (2013), the authors note that under market frictions, the no arbitrage condition is not required and that instead consistent price systems would be sufficient. Moreover, they cite the condition of "conditional full support" as sufficient to ensure the existence of consistent price systems.

The rationale to look for consistent price systems as opposed to a semi-martingale approach would be that transactions costs are pervasive for storable commodities which include, amongst others, storage costs, transportation, insurance as well as traditional transaction costs.

In the new model proposed in this thesis, if the driving Lévy process is a Brownian motion, we can apply the result from Pakkanen (2011) to show that the proposed model would have a conditional full support which in turn implies the existence of consistent price systems.

The conditions required for a conditional full support in Pakkanen (2011) appear to be fairly mild, with condition (iv) likely being the most restrictive. For condition (iv), the following choice would likely be sufficient (in the notation of Pakkanen (2011)): taking $\beta = 0$, taking $\overline{W}^{\perp}$ an independent Brownian motion from $H_1(X_t)$ and $H_2(X_t)$ and lastly letting $\overline{W}$ be any Brownian motion (perhaps appearing as a semi-martingale) on a filtration with respect to which $H_1(X_t)$ and $H_2(X_t)$ are adapted would satisfy the condition. Of note, the main result (which is given in Theorem 1.3) showing the conditional full support of the process assumes independence between $(H_1(X_t), H_2(X_t))$ and the driving Brownian motion but not independence of $H_1(X_t)$ and $H_2(X_t)$.

B.4 Discussion & Future Research

For the semi-martingales and the risk neutral measure, future research should be directed at establishing that the resulting density process is a martingale. With respect to consistent price systems, future research could explore the argument regarding the conditional full support of the model as well. Additionally, the existing literature is quite sparse on the application of consistent price systems in practice or the related empirical results. The use of consistent price systems

would require an evaluation of transactions unless an asymptotic approach is used as described in Guasoni, Miklós, & Walter (2008).

Appendix C

Semi-martingale Approach and Density Process Criterions

This section considers the change of drift process linked to the classical dynamic hedging argument and attempts to show that the ensuing density process is a martingale. We look at the problem through the lens of both the Novikov and Kazamaki criteria and explore steps towards obtaining the final result. Additionally, this section provides a simple upper bound on the spot price process.

C.1 Characterization of the Density Process

Recall the change in drift process b_t which ensures that the spot price S_t rises at the risk-free rate under the risk neutral measure.

$$b_t = \frac{r}{H_1(X_t)} + \frac{\alpha Y_t}{H_1(X_t)} - \frac{H_2(X_t)}{H_1(X_t)} - \frac{H_1(X_t)}{2}$$

where H_1 and H_2 are bounded for any $X \in \mathbb{R}$:

$$0 < m_1 < H_1(X) < M_1$$

$$m_2 < H_2(X) < M_2$$

This implies a density process of the form:

$$\mathcal{E}\left(\int_0^t b_s dB_s^Y\right) = \exp\left(\int_0^t b_s dB_s^Y - \int_0^t b_s^2 ds\right)$$

In the subsections below, we look at the implications of this correction on the Novikov and Kazamaki criteria.

C.2 Upper Bound of the Y_t Process

C.2.1 Solution to Y_t

We begin by applying the usual transformation used in the solution of the Ornstein-Uhlenbeck process, namely $F_t = e^{\alpha t} Y_t$. Applying Ito's lemma we get:

$$dF_t = \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial Y} dY_t + \frac{1}{2}\frac{\partial^2 F}{\partial Y^2}(dY_t)^2$$

$$dF_t = \alpha e^{\alpha t} Y_t dt + e^{\alpha t}(\alpha(H_2(X_t) - Y_t)dt + H_1(X_t)dB_t^Y) + 0$$

$$dF_t = \alpha e^{\alpha t} H_2(X_t)dt + e^{\alpha t} H_1(X_t) dB_t^Y$$

Expressing this in an integral form provides a solution for Y_t:

$$Y_t = e^{-\alpha t} Y_0 + \int_0^t \alpha e^{-\alpha(t-s)} H_2(X_s) ds + \int_0^t e^{-\alpha(t-s)} H_1(X_s) dB_s^Y$$

C.2.2 Bound on Integral with Respect to Time

As the integrand of the first integral is bounded, we can easily bound it in absolute value. Let $M_2^* = \max(|M_2|, |m_2|)$ where M_2 and m_2 are upper and lower bounds of H_2 respectively:

$$\left| \int_0^t \alpha e^{-\alpha(t-s)} H_2(X_s) ds \right| \leq \int_0^t \left| \alpha e^{-\alpha(t-s)} H_2(X_s) \right| ds \leq t\alpha M_2^*$$

C.2.3 Bound on Stochastic Integral

To bound the stochastic integral, we can apply an argument using the Dambis-Dubins–Schwarz (DDS) theorem. In particular, we rely on a version of this theorem that applies to processes that have a finite quadratic variation with a positive probability.

Define a local martingale M_t and the corresponding Brownian motion W_t^1 from the DDS theorem:

$$M_t = \int_0^t e^{\alpha s} H_1(X_s) dB_s^Y$$

$$M_t = W^1{}_{\langle M \rangle_t}$$

To bound the integral, we can then proceed as follows:

$$\begin{aligned}
\left| e^{-\alpha t} \int_0^t e^{\alpha s} H_1(X_s) dB_s^Y \right| &\leq e^{-\alpha t} \sup_{u \in [0,t]} \left| \int_0^u e^{\alpha s} H_1(X_s) dB_s^Y \right| \\
&= e^{-\alpha t} \sup_{u \in [0,t]} \left| W^1_{\langle M \rangle_u} \right| \\
&\leq e^{-\alpha t} \sup_{u \in \left[0, \frac{M_1^2}{2\alpha}(e^{2\alpha t} - 1)\right]} |W_u^1| \\
&= \sup_{u \in \left[0, \frac{M_1^2}{2\alpha}(e^{2\alpha t} - 1)\right]} |e^{-\alpha t} W_u^1|
\end{aligned}$$

where the upper bound of u in the last line is based on a bound for the quadratic variation of M_t:

$$\langle M \rangle_t = \int_0^t \left(e^{\alpha s} H_1(X_s) \right)^2 ds \leq \frac{M_1^2}{2\alpha} (e^{2\alpha t} - 1)$$

We can apply a similar argument with the local martingale $(e^{-\alpha t} W_u^1)_{u \geq 0}$, where t is a constant, which has the quadratic variation:

$$\langle e^{-\alpha t} W^1 \rangle_s = e^{-2\alpha t} s$$

Let W_t^2 be a Brownian motion from the DDS theorem such that:

$$e^{-\alpha t}W_s^1 = W^2{}_{\langle e^{-\alpha t}W^1\rangle_s}$$

To complete the upper bound, we then have:

$$\begin{aligned}\left|e^{-\alpha t}\int_0^t e^{\alpha s}H_1(X_s)dB_s^Y\right| &\leq \sup_{u\in\left[0,\frac{M_1^2}{2\alpha}(e^{2\alpha t}-1)\right]}|e^{-\alpha t}W_u^1|\\ &= \sup_{u\in\left[0,\frac{M_1^2}{2\alpha}(e^{2\alpha t}-1)\right]}\left|W^2_{\langle e^{-\alpha t}W^1\rangle_u}\right|\\ &= \sup_{u\in\left[0,\frac{M_1^2}{2\alpha}(1-e^{-2\alpha t})\right]}|W_u^2|\\ &\leq \sup_{u\in\left[0,\frac{M_1^2}{2\alpha}\right]}|W_u|\end{aligned}$$

We observe that this upper bound is free from t and applies for any t.

C.2.4 Bound on the process Y_t

Using the above result and the triangle inequality we have the following bound on the process Y_t

$$|Y_t| \leq |e^{-\alpha t}Y_0\,| + \left|\int_0^t \alpha e^{-\alpha(t-s)}H_2(X_s)ds\right| + \left|\int_0^t e^{-\alpha(t-s)}H_1(X_s)dB_s^Y\right|$$

$$|Y_t| \leq |Y_0\,| + t\alpha M_2^* + \Sigma$$

$$|Y_t| \leq K_t + \Sigma$$

where, for convenience we define:

$$\Sigma = \sup_{u\in\left[0,\frac{M_1^2}{2\alpha}\right]}|W_u|$$

$$K_t = |Y_0\,| + t\alpha M_2^*$$

Note that K_t is increasing and that Σ does not depend on t.

C.3 Novikov Criterion

Using the characterization of the density process, to satisfy the Novikov criterion we should show that:

$$\mathbb{E}\left[\exp\left(\left(\frac{1}{2}\right)\langle D\rangle_t\right)\right] < \infty$$

where $D_t = \int_0^t b_s dB_s^Y$ is the density process.

C.3.1 Quadratic Variation of Density Process

To find the quadratic variation of the density process, we first rely on the linearity of a stochastic integration to express the density process as the sum of four processes:

$$D_t = \int_0^t b_s dB_s^Y = \int_0^t \frac{r}{H_1(X_s)} dB_s^Y + \int_0^t \frac{\alpha Y_s}{H_1(X_s)} dB_s^Y + \int_0^t -\frac{H_2(X_s)}{H_1(X_s)} dB_s^Y + \int_0^t -\frac{H_1(X_s)}{2} dB_s^Y$$

$$P_1 = \int_0^t \frac{r}{H_1(X_s)} dB_s^Y$$

$$P_2 = \int_0^t \frac{\alpha Y_s}{H_1(X_s)} dB_s^Y$$

$$P_3 = \int_0^t -\frac{H_2(X_s)}{H_1(X_s)} dB_s^Y$$

$$P_4 = \int_0^t -\frac{H_1(X_s)}{2} dB_s^Y$$

We will apply the polarization identity with the two sums below:

$$D_t = (P_2 + P_4) + (P_1 + P_3)$$

and for each sum we have:

$$\langle P_i + P_j \rangle_t = \langle P_i \rangle_t + \langle P_j \rangle_t + 2\langle P_i, P_j \rangle_t$$

C.3.2 Quadratic variation of $(\boldsymbol{P_2 + P_4})$

$$\langle P_2 \rangle_t = \int_0^t \left(\frac{\alpha Y_s}{H_1(X_s)}\right)^2 ds = \int_0^t Y_s^2 \left(\frac{\alpha}{H_1(X_s)}\right)^2 ds$$

$$\langle P_4 \rangle_t = \int_0^t \left(-\frac{H_1(X_s)}{2}\right)^2 ds$$

$$\langle P_2, P_4 \rangle_t = \int_0^t \left(-\frac{H_1(X_s)}{2}\right)\left(\frac{\alpha Y_s}{H_1(X_s)}\right) ds = \int_0^t -\frac{\alpha Y_s}{2} ds$$

The polarization identity gives:

$$\langle P_2 + P_4 \rangle_t = \int_0^t Y_s^2 \left(\frac{\alpha}{H_1(X_s)}\right)^2 - \alpha Y_s + \left(-\frac{H_1(X_s)}{2}\right)^2 ds$$

C.3.3 Quadratic variation of $(\boldsymbol{P_1 + P_3})$

$$\langle P_1 \rangle_t = \int_0^t \left(\frac{r}{H_1(X_s)}\right)^2 ds$$

$$\langle P_3 \rangle_t = \int_0^t \left(-\frac{H_2(X_s)}{H_1(X_s)}\right)^2 ds$$

$$\langle P_1, P_3 \rangle_t = \int_0^t \left(\frac{r}{H_1(X_s)}\right)\left(-\frac{H_2(X_s)}{H_1(X_s)}\right) ds = \int_0^t -\frac{rH_2(X_s)}{H_1(X_s)^2} ds$$

The polarization identity then gives:

$$\langle P_1 + P_3 \rangle_t = \int_0^t \left(\frac{r}{H_1(X_s)}\right)^2 + \left(-\frac{H_2(X_s)}{H_1(X_s)}\right)^2 - 2\frac{rH_2(X_s)}{H_1(X_s)^2}ds$$

C.3.4 Covariation of $(P_2 + P_4)$ and $(P_1 + P_3)$

$$\langle P_2 + P_4, P_1 + P_3 \rangle_t = \int_0^t \left(\frac{\alpha Y_s}{H_1(X_s)} - \frac{H_1(X_s)}{2}\right)\left(\frac{r}{H_1(X_s)} - \frac{H_2(X_s)}{H_1(X_s)}\right)ds$$

$$= \int_0^t Y_s\left(\frac{\alpha r}{H_1(X_s)^2} - \frac{\alpha H_2(X_s)}{H_1(X_s)^2}\right) - \frac{1}{2}\big(r - H_2(X_s)\big)ds$$

C.3.5 Quadratic Variation of D_t

Using the result on the individual components, we can apply the polarization identity to obtain the quadratic variation of D_t.

$$\langle D \rangle_t = \int_0^t Y_s^2\left(\frac{\alpha}{H_1(X_s)}\right)^2 - \alpha Y_s + \left(-\frac{H_1(X_s)}{2}\right)^2 + \left(\frac{r}{H_1(X_s)}\right)^2 + \left(-\frac{H_2(X_s)}{H_1(X_s)}\right)^2 - 2\frac{rH_2(X_s)}{H_1(X_s)^2}ds$$

$$+ \int_0^t Y_s\left(\frac{\alpha r}{H_1(X_s)^2} - \frac{\alpha H_2(X_s)}{H_1(X_s)^2}\right) - \frac{1}{2}\big(r - H_2(X_s)\big)ds$$

After simplifying terms, we obtain a second-degree polynomial in Y_t:

$$\langle D \rangle_t = \int_0^t Y_s^2\left(\frac{\alpha}{H_1(X_s)}\right)^2 + Y_s\left(\frac{\alpha r}{H_1(X_s)^2} - \frac{\alpha H_2(X_s)}{H_1(X_s)^2} - \alpha\right)ds$$

$$+ \int_0^t \left(-\frac{H_1(X_s)}{2}\right)^2 + \left(\frac{r}{H_1(X_s)}\right)^2 + \left(-\frac{H_2(X_s)}{H_1(X_s)}\right)^2 - 2\frac{rH_2(X_s)}{H_1(X_s)^2}$$

$$- \frac{1}{2}\big(r - H_2(X_s)\big)ds$$

this polynomial can be expressed as:

$$\langle D \rangle_t = \int_0^t C_2(X_s)Y_s^2 + C_1(X_s)Y_s + C_0(X_s)ds$$

with coefficients $C_2(X_s)$, $C_1(X_s)$ and $C_0(X_s)$ defined as:

$$C_2(X_s) = \left(\frac{\alpha}{H_1(X_s)}\right)^2$$

$$C_1(X_s) = \left(\frac{\alpha r}{H_1(X_s)^2} - \frac{\alpha H_2(X_s)}{H_1(X_s)^2} - \alpha\right)$$

$$C_0(X_s) = \left(-\frac{H_1(X_s)}{2}\right)^2 + \left(\frac{r}{H_1(X_s)}\right)^2 + \left(-\frac{H_2(X_s)}{H_1(X_s)}\right)^2 - 2\frac{rH_2(X_s)}{H_1(X_s)^2} - \frac{1}{2}\big(r - H_2(X_s)\big)$$

C.3.6 Upper Bound on the Quadratic Variation of D_t

Combining the expression of the quadratic variation with the upper bound on Y_t we obtain the bound below:

$$\begin{aligned}\langle D\rangle_t &= \int_0^t C_2(X_s)Y_s^2 + C_1(X_s)Y_s + C_0(X_s)ds \\ &\leq \int_0^t C_2(X_s)(K_s+\Sigma)^2 + C_1(X_s)(K_s+\Sigma) + C_0(X_s)ds \\ &\leq (K_t+\Sigma)^2\int_0^t C_2(X_s)ds + (K_t+\Sigma)\int_0^t C_1(X_s)ds + \int_0^t C_0(X_s)ds\end{aligned}$$

where we have relied on the fact that Σ does not depend on t and that K_t is increasing. As the integrands of the remaining integrals are all bounded they must all be finite.

C.3.7 Novikov Criterion

With the above result, we can finally look at the Novikov Criterion more closely:

$$\begin{aligned}\mathbb{E}\left[\exp\left(\left(\frac{1}{2}\right)\langle D\rangle_t\right)\right] &= \mathbb{E}\left[\exp\left(\left(\frac{1}{2}\right)\int_0^t C_2(X_s)Y_s^2 + C_1(X_s)Y_s + C_0(X_s)ds\right)\right] \\ &\leq \mathbb{E}\left[\exp\left(\left(\frac{1}{2}\right)\left((K_t+\Sigma)^2\int_0^t C_2(X_s)ds + (K_t+\Sigma)\int_0^t C_1(X_s)ds + \int_0^t C_0(X_s)ds\right)\right)\right]\end{aligned}$$

We can now split the expectation between the regions of the underlying measure space where $(K_s+\Sigma)\leq 1$ and its complement. Where $(K_s+\Sigma)\leq 1$, the resulting expectation is clearly bounded, and we denote this by B_1:

$$\begin{aligned}B_1 &= \mathbb{E}\left[\exp\left(\left(\frac{1}{2}\right)\left((K_t+\Sigma)^2\int_0^t C_2(X_s)ds + (K_t+\Sigma)\int_0^t C_1(X_s)ds + \int_0^t C_0(X_s)ds\right)\right)I(K_t+\Sigma)\leq 1)\right] \\ &\leq \mathbb{E}\left[\exp\left(\left(\frac{1}{2}\right)\left((1)^2\int_0^t C_2(X_s)ds + (1)\int_0^t C_1(X_s)ds + \int_0^t C_0(X_s)ds\right)\right)\right]\end{aligned}$$

Define B_2 to be an upper bound of the term that is free from Y_t:

$$B_2 = \int_0^t C_0(X_s)ds$$

Using B_1 and B_2, we have that:

$$\begin{aligned}\mathbb{E}\left[\exp\left(\left(\frac{1}{2}\right)\langle D\rangle_t\right)\right] &\leq \mathbb{E}\left[\exp\left(\left(\frac{1}{2}\right)(K_t+\Sigma)^2\int_0^t C_2(X_s)ds + (K_t+\Sigma)\int_0^t C_1(X_s)ds + \int_0^t C_0(X_s)ds\right)\right] \\ &\leq B_1 + B_2\mathbb{E}\left[\exp\left(\left(\frac{1}{2}\right)(K_t+\Sigma)^2\int_0^t C_2(X_s) + C_1(X_s)ds\right)\right]\end{aligned}$$

Here we recognize that the remaining expectation is a moment generating function (MGF) of the random variable $(K_t + \Sigma)^2$ evaluated at $\left(\frac{1}{2}\right)\int_0^t C_2(X_s) + C_1(X_s)ds$ (this last integral can be bounded). The main difficulty that we are facing here is that Σ is the supremum of the absolute value of a Brownian motion which does not have a simple/well known distribution. Moreover, we are interested in the MGF of Σ^2 which is larger than Σ.

The event $\{\Sigma < a\}$ can be seen to be equivalent to the event $\{\tau_a^{2a} > \frac{M_1^2}{2a}\}$ for the stopping time:

$$\tau_a^{2a} = \inf\{t \geq 0 | W_t^a \notin [0, 2a]\}$$

where W_t^a is a standard Brownian motion starting at a. Letting $t^* = \frac{M_1^2}{2a}$, the probability of $\{\Sigma < a\}$ is given by (from equation 5.9 on page 342 of Feller (1965)):

$$P(\Sigma < a) = P(\tau_a^{2a} > t^*) = \frac{4}{\pi}\sum_{n\geq 0}\frac{1}{2n+1}\exp\left(-\frac{(2n+1)\pi^2}{2(2a)^2}t^*\right)\sin\left(\frac{(2n+1)\pi a}{2a}\right)$$

$$= \frac{4}{\pi}\sum_{n\geq 0}\frac{(-1)^n}{2n+1}\exp\left(-\frac{(2n+1)\pi^2}{8a^2}t^*\right)$$

Given that the Cumulative Distribution Function (CDF) of Σ includes the level a of Σ in the denominator of the exponential terms, the MGF of Σ^2 is only defined for negative parameters meaning the expectation of the upper bound above is not finite.

C.4 Kazamaki Criterion

Using the formula of the density process, we can also look at a Kazamaki Criterion. This criterion is satisfied if:

$$\sup_{s\in[0,t]} \mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^s b_u dB_u^Y\right)\right] < \infty$$

We can express the density process as:

$$\int_0^s b_u dB_u^Y = \int_0^s \frac{r}{H_1(X_u)} - \frac{H_2(X_u)}{H_1(X_u)} - \frac{H_1(X_u)}{2}dB_u^Y + \int_0^s \frac{\alpha Y_u}{H_1(X_u)}dB_u^Y$$

The first integral has a bounded integrand and as such has a bounded quadratic variation and we can apply the same argument as above with the DDS theorem to obtain a bound.

Define $C \in \mathbb{R}$ as:

$$\left|\frac{r}{H_1(X_u)} - \frac{H_2(X_u)}{H_1(X_u)} - \frac{H_1(X_u)}{2}\right| \leq C$$

The quadratic variation of the first term is then bounded by:

$$\int_0^s \left(\frac{r}{H_1(X_u)} - \frac{H_2(X_u)}{H_1(X_u)} - \frac{H_1(X_u)}{2}\right)^2 du \leq C^2 s$$

and finally by the DDS theorem we have the following for W_u a Brownian motion:

$$\int_0^s \frac{r}{H_1(X_u)} - \frac{H_2(X_u)}{H_1(X_u)} - \frac{H_1(X_u)}{2} dB_u^Y \leq \sup_{u\in[0,C^2 s]} W_u$$

Let $\Sigma_s = \sup_{u\in[0,C^2 s]} W_u$

To show the criterion, it is then sufficient to show that:

$$\sup_{s\in[0,t]} \mathbb{E}\left[\exp\left(\frac{\Sigma_s}{2} + \frac{1}{2}\int_0^s \frac{\alpha Y_u}{H_1(X_u)} dB_u^Y\right)\right]$$

However, at this time we have not found a viable way to proceed further.

C.5 Discussion & Future Research

On the Novikov criterion, future research can look for smaller bounds as a supremum of the process is itself quite a large upper bound. The main difficulty for the MGF stems from having to bound the absolute value of Y_t and perhaps via other methods this could be avoided. Second, for the Kazamaki Criterion, future research could look profitably towards methods which allow us to manage the MGF of stochastic integrals or perhaps methods that enable us to characterize the distribution of these integrals as they pertain to the proposed model.