Stochastic Modelling in Volatility and Its Applications in Derivatives

A Thesis

Presented in Candidacy for the Degree of

Doctor of Philosophy

Adam Smith Business School

University of Glasgow

by

Yihan Zou

September 2020

Abstract

STOCHASTIC MODELLING IN VOLATILITY AND ITS APPLICATIONS IN DERIVATIVES

by

Yihan Zou

This thesis consists of three articles concentrating on modelling stochastic volatility in commodity as well as equity and applying stochastic volatility models to evaluate financial derivatives and real options. Firstly, we introduce the general background and the incentive of considering stochastic volatility models.

In Chapter 2 we derive tractable analytic solutions for futures and options prices for a linear-quadratic jump-diffusion model with seasonal adjustments in stochastic volatility and convenience yield. We then calibrate our model to data from the fish pool futures market, using the extended Kalman filter and a quasi-maximum likelihood estimator and alternatively using an implied-state quasi-maximum likelihood estimator. We find no statistical evidence of jumps. However, we do find evidence for the positive correlation between salmon spot prices and volatility, seasonality in volatility and convenience yield. In addition we observe a positive relationship between seasonal risk premium and uncertainty within the EU salmon demand. We further show that our model produces option prices that are conform with the observation of implied volatility smiles and skews.

In Chapter 3, we introduce a linear quadratic volatility model with co-jumps and show how to calibrate this model to a rich dataset. We apply general method of moments (GMM) and more specifically match the moments of realized power and multi-power variations, which are obtained from high-frequency stock market data. Our model incorporates two salient features: the setting of simultaneous jumps in both return process and volatility process and the superposition structure of a continuous linear quadratic volatility process and a Lévy-driven Ornstein-Uhlenbeck process. We compare the quality of fit for several models, and show that our model outperforms the conventional jump diffusion or Bates model.

Besides that, we find evidence that the jump sizes are not normal distributed and that our model performs best when the distribution of jump-sizes is only specified through certain (co-) moment conditions. A Monte Carlo experiments is employed to confirm this.

Finally, in Chapter 4, we study the optimal stopping problems in the context of American options with stochastic volatility models and the optimal fish harvesting decision with stochastic convenience yield models, in the presence of drift ambiguity. From the perspective of an ambiguity averse agent, we transfer the problem to the solution of a reflected backward stochastic differential equation (RBSDE) and prove the uniform Lipschitz continuity of the generator. We then propose a numerical algorithm with the theory of RBSDEs and a general stratification technique, and an alternative algorithm without using the theory of RBSDEs. We test the accuracy and convergence of the numerical schemes. By comparing to the one dimensional case, we highlight the importance of the dynamic structure of the agent's worst case belief. Results also show that the numerical RBSDE algorithm with stratification is more efficient when the optimal generator is attainable.

To my love

Acknowledgements

During my doctoral studies, I feel honored to have the privilege to work with some brilliant minds that helped me to reach this point.

First and foremost, I would like to express my profound gratitude to my supervisor, Professor Christian-Oliver Ewald, for giving me the opportunity to commence my study. In countless inspiring discussions, his words have been a constant source of encouragement and the guiding light on my way pursuing the doctoral degree. The example of Professor Christian as a researcher and mentor has been and will be an invaluable lesson treasured in my career.

I am extremely grateful to my co-supervisor Dr. Ankush Agarwal, for selflessly teaching and leading my research. His patience and mathematically rigorous attitude have motivated me towards a meticulous researcher. Without his guidance this thesis would not have been completed.

I am greatly indebted to Professor Christian and Professor Marek Rutkowski for granting me to pay a two-month visit to the Department of Mathematics and Statistics at the University of Sydney, which expanded my research horizon and gave my career endless wonder and possibility. Talks with Professor Marek were enlightening experiences and his wisdom has deepened my knowledge.

Next, I would like to thank Yongjie Wang, who has the same supervisors as I do, for the great mutual encouragement along this tough way and numerous revealing discussions and reading groups.

I enjoyed being friends and flatmates with Dr. Hao Peng, who was a PhD candidate in Accounting and Finance Department during my time in the University of Glasgow. Hao's positive and humorous attitude really enriched my doctoral life. I still remember those days

we stayed up in the office till midnight.

I also had pleasure of being part of the “league of office’s secret”, in which other members are: Hualin Li, Jinpeng Liu, Dr. Zerun Liu, Dr. Hao Peng, Dr. Zhixiao Wang, Jiangteng Zhou, Yue Zhou, Dr. Zhekai Zhang. These fellows and friends have been my source of joy and given me lots of insightful advice so far.

I wish to thank all the following friends: Xiaofei Cao, Lingling Hua, Yichu Huang, Chaofan Niu, Shuqi Rao, Yayan Tang, Tianfang Xie, Hao Zhang, Xiangyu Zong and all flatmates in Flat 184, West Village, who experienced unnumbered happy memories together with me.

In addition, I would like to thank Dr. Vasilios Sogiakas, who was my previous co-supervisor but changed to another university, for those pleasant discussions and cooperation in teaching Advanced Portfolio Analysis.

Other people I would like to express my appreciation include Dr. Yuxiang Jiang, Professor Roderick McCrorie, Dr. Ruolan Ouyang, Dr. Pengcheng Song, Professor Serafeim Tsoukas, Professor Sjur Westgaard, Dr. Senyu Wang, Dr. Yuexiang Wu and Dr. Zhou Zhou who had wonderful and inspiring conversations with me.

Naturally, I would like to express my deepest gratitude to my parents and my family. Without their support and dedication, I would not be who I am today.

Declaration

I declare that, except where explicit reference is made to the contribution of others, that this thesis is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Printed Name: Yihan Zou

Signature: Yihan Zou

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Overview of Modelling Financial Markets

Contemporary theories on modelling financial markets with stochastic processes originated from the ground-breaking work of Bachelier (1900). In his article, "Théorie de la Spéculation", Louis Bachelier used a Brownian motion with drift to model price fluctuations of Paris stock market. However, it was unrealistic in allowing for negative stock prices. A more appropriate model with a geometric Brownian motion (hereafter GBM) was proposed by Samuelson (1965b), implying that the stock prices follow logarithmic normal distributions. The GBM model avoids the undesirable negativity of the Brownian motion, it also enjoys the geometric growth property, which is commonly observed in equity. It was 8 years before Black & Scholes (1973) demonstrated rigorously how to price European options under the GBM model. The GBM model, now called the Black&Scholes model, is also known as the Black-Scholes-Merton model due to the significant contributions by Merton (1973). Robert Merton and Myron Scholes were awarded the *Nobel Prize in Economics* in 1997 thanks to their foundational work. In their monographs, they used the risk-neutral pricing principle, and stated the standard assumptions under that principle:

- Securities can be traded continuously.
- Short-selling is allowed.
- There is no transaction cost.
- There is no arbitrage opportunity in the market.

Given those assumptions, the principle guarantees the existence of a risk-neutral measure,

which is stated in the first *Fundamental Theorem of Asset Pricing* (hereafter FTAP). It was crystallised by Harrison & Pliska (1981) in discrete-time version, and one can refer to Delbaen & Schachermayer (2006) for an overview. Generally, the FTAP expresses the relation between the absence of arbitrage and the existence of a risk-neutral measure (also know as an *equivalent martingale measure*). An equivalent martingale measure is a probability measure, under which discounted ex-dividends prices become martingales.

Under the above assumptions, Black & Scholes (1973) characterize the dynamics of stock prices S_t at time t under the equivalent martingale measure (hereafter EMM) Q by the following stochastic differential equation (SDE)

$$\mathrm{d}S_t = rS_t\mathrm{d}t + \sigma S_t \mathrm{d}W_t^Q, \tag{1.1}$$

where r is the risk-free interest rate, σ is the volatility and W_t^Q is a standard Brownian motion under Q measure. Black & Scholes (1973) give the related European call option price $C_t(S_t)$ at time t with strike price K and maturity time T as

$$C_t(S_t) = S_t N\big(d_+(S_t, T-t)\big) - Ke^{-r(T-t)} N\big(d_-(S_t, T-t)\big), \tag{1.2}$$

where

$$d_\pm(S_t, T-t) = \frac{\ln(S_t/K) + (r \pm \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

and $N(\cdot)$ is the standard Gaussian cumulative distribution function. Provided the analytical European option pricing formula (1.2), it should be noted that all parameters except the volatility σ can be observed directly with market data. On the other hand, given the market option price at time t with interest rate r, strike price K and maturity time T, the volatility can be inverted from the Black&Scholes formula. That volatility is named the *market's implied volatility*. Should the Black&Scholes model be correct, the *market's implied volatility* would be constant across options with different time to maturities and strike prices. However, this is not the case. There is substantial evidence from market data against the constant volatility assumption, claiming that the implied volatilities appear to have a "smile" or "smirk" shape.(To name a few, see Rubinstein (1985), Rubinstein (1994) and Dumas et al. (1998).) Those studies further imply that the implied volatility could be a function of time to maturity

and strike price. To be succinct, the Black&Scholes model is not perfect enough to explain the stylized volatility smile.

Succedent research concentrates on modelling volatility by providing tighter fits to the implied volatility surface. Derman & Kani (1994), Dupire (1994) and Rubinstein (1994) assume that the volatility of asset returns is deterministic functions of time and strike price, which are the so-called *local volatility functions*. The idea behind local volatility models is enabling practitioners to price exotic options using known vanilla options prices by making simplifying assumptions on the volatility. Nevertheless, the class of local volatility models are flawed to be inconsistent with market's index data, and then require daily re-calibration, according to Dumas et al. (1998).

Finally, we come to *stochastic volatility models*, which hypothesize that the volatility term σ in the Black&Scholes model (1.1) is a function of stochastic processes. The stochastic volatility models have achieved huge success in academia and in practice, both in fitting well to the volatility surface, and in characterizing the negative correlation between equity historical returns and its volatility, which is initially recorded in Black (1976) and known as the *leverage effect*. Though a wide class of stochastic volatility models there is, they share a common feature that the volatility process reverts to a long-run mean. Incorporating Brownian motions as the driven noise in the volatility function, for example, Hull & White (1987), Stein & Stein (1991) and Heston (1993). To capture the leverage effect, the Brownian motion driven model just assumes the correlation coefficient between the Brownian motions in returns and volatility is negative. One can refer to Fouque et al. (2000) for a review of stochastic volatility models and their financial applications. Despite the desirable properties for the stochastic volatility models, their capability to capture the tail events of underlying securities is limited. Cont & Tankov (2003) argue that those models driven by Gaussian noise is not as satisfactory as the simplest Markovian models with jumps (Lévy processes) in fitting to the heavy tails in returns and generating sudden, discontinuous moves in prices, which are observable in reality. Gatheral (2011) claims that diffusion volatility processes, i.e. Brownian motions driven models, cannot account for the large moves over very short timescales for in-the-money and out-of-the-money options. This justifies the practical importance of the jump diffusion stochastic volatility models. We will go into more specified details for stochastic

volatility models and investigate its applications both in financial products and in real life investment decisions.

1.2 Financial Derivatives and Real Options

In this section, we introduce the definition and mechanism of common financial derivatives and real options, as they are used in latter chapters.

1.2.1 Financial Derivatives

A financial derivative refers to a contract, of which the value is usually written based on another financial security, for instance, a stock or a bond. Derivatives are extensively used for hedging and speculation purpose. Some of them are traded over-the-counter (OTC), while others are traded in exchanges. We start introducing several kinds of derivatives that are related to this thesis.

A *forward contract* is a contract between two parties, under which one party agrees to sell another a financial asset at a fixed future date for a specified delivery price. Usually a forward contract is used to lock a price or hedge of the underlying asset a priori. While it can be used for speculation purpose as well, as a forward contract does not require a advanced payment between two parties.

Futures contracts are similar to forward contracts, except that they are traded in exchanges. Hence, two trading parties have to comply with regulations of exchanges, hence are more protected from counter-party risk in some sense. The key rule of futures is to deposit a certain amount of money into the account in exchanges, i.e. margin account, and the account is monitored on daily basis, which is *market-to-market*. It costs nothing to enter futures and forward contracts.

A *call option* is a contract between two parties that gives the holder or buyer, the right, but not the obligation, to buy the underlying security at a fixed future date (maturity/exercise time) for a previous contracted price (strike price). Options can be written on stocks or other derivatives such as futures, options, etc. A *put option* is the same to a call option except it gives the holder to sell the underlying security. Vanilla options are differentiated into two

main categories according to their exercise styles. A *European option* entitles the holder to decide whether exercise or not only at the maturity date, while the holder of an *American option* can make the exercise decision at anytime prior to the maturity date. Generally, holding an option brings the holder some potential positive possibilities to receive a positive amount and pay nothing. Therefore, from the arbitrage-based perspectives the option's present value should be positive, i.e. entering an option contract comes with some cost.

1.2.2 Real Options

Real options refer to real life investment projects that resemble their financial counterparts and have similar payoff and exercise styles, hence they can be evaluated using option pricing techniques, for example, the evaluation of patents or ownership of lands, the optimal exploitation of natural resources, or optimal harvesting of forestry. Real options approach for projects evaluation dates back to Brennan & Schwartz (1985); McDonald & Siegel (1986); Dixit & Pindyck (1994). Note that there will not be considerations for arbitrage opportunity when evaluating real options, as there is no actively trading market for investment projects.

1.3 Preliminary Notations and Terminologies

We introduce some notations and terminologies used in latter context here in a less mathematically rigorous but more comprehensible way.

Generally we consider a *probability space* characterized by the triplet $(\Omega, \mathcal{F}, P)$, where Ω is the sample space containing all possible outcomes, $\mathcal{F}$ is a σ-algebra containing all subsets of zero probability elements of $\mathcal{F}$ and P is the reference probability measure function mapping from $\mathcal{F}$ to $\mathbb{R}_+$. Let $\{\mathcal{F}_t\}_{t\geq 0}$ be the natural non-decreasing right-continuous filtration on the probability space and $\mathcal{F} = \mathcal{F}_\infty = \vee_{t\geq 0}\mathcal{F}_t$.

We use, throughout the thesis, $\mathbb{R}_+ = (0, \infty)$, $\mathbb{N}_+ = \mathbb{R}_+ \cap \mathbb{N}$, where $\mathbb{N}$ is the set of all natural numbers. $\mathbb{E}[\cdot]$ is the expectation under the reference measure, abbreviated for $\mathbb{E}[\cdot|\mathcal{F}_0]$. If M and N are real numbers or real-valued functions, $M \vee N = \max(M, N)$, $M \wedge N = \min(M, N)$, $M^+ = \max(M, 0)$, $M^- = \max(-M, 0)$. For a stochastic process X that may have jumps, $\Delta X_t = X_t - X_{t-}$, where X_{t-} is the limit value of X at time t from

left hand.

Given a σ-finite measure ν defined on a measurable space $(E, \mathcal{E})$ for $E \subset \mathbb{R}^d$, a *Poisson random measure* μ with an *intensity measure* ν is mapping from $\Omega \times \mathcal{E}$ to $\mathbb{N}$, such that

- $\forall A \in \mathcal{E}$, $\mu(\cdot, A) = \mu(A)$ is a Poisson random variable with jump intensity $\nu(A)$ with $P(\mu(A) = k) = e^{-\nu(A)}(\nu(A))^k/k!, \forall k \in \mathbb{N}$.
- For disjoint sets $A_1, ..., A_n \in \mathcal{E}$, the variables $\mu(A_1), ..., \mu(A_n)$ are independent.

The corresponding *compensated Poisson random measure* $\tilde{\mu}$ is simply constructed by subtracting from μ it intensity measure ν: $\tilde{\mu} = \mu - \nu$. Specifically, if μ is an adapted and defined on $E = [0, T] \times \mathbb{R}^d \setminus \{0\}$, and $f : E \to \mathbb{R}^d$ satisfying $\int_{[0,T]} \int_{\mathbb{R}^d \setminus \{0\}} |f(s, x)| \nu(\mathrm{d}s \times \mathrm{d}x) < \infty$, then the stochastic process X_t that is an integral with respect to the compensated Poisson random measure

$$\begin{aligned} X_t &= \int_{[0,T]} \int_{\mathbb{R}^d \setminus \{0\}} f(s, x) \tilde{\mu}(\mathrm{d}s \times \mathrm{d}x) \\ &= \int_{[0,T]} \int_{\mathbb{R}^d \setminus \{0\}} f(s, x) \mu(\mathrm{d}s \times \mathrm{d}x) - \int_{[0,T]} \int_{\mathbb{R}^d \setminus \{0\}} f(s, x) \nu(\mathrm{d}s \times \mathrm{d}x) \end{aligned}$$

is a martingale.

A *compound Poisson process* X_t on with a jump intensity λ and the distribution g for the jump size has a intensity measure of such form: $\nu(\mathrm{d}t \times \mathrm{d}x) = \lambda g(\mathrm{d}x)\mathrm{d}t$.

Let $X_t \in \mathbb{R}^d$ be a jump diffusion process with initial value $X_0 = x_0 \in \mathbb{E}^d$ in such form

$$\mathrm{d}X_t = \alpha(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t + \int_{\mathbb{R}^d} f(s, X_{t-}, x) \tilde{\mu}(\mathrm{d}t, \mathrm{d}x), \tag{1.3}$$

where $\alpha : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times k}$, $f : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{n \times l}$, and W_t is a k-dimensional standard Brownian motion. Provided the conditions for the existence and uniqueness of the solution of the above stochastic differential equation (SDE) are satisfied, the *infinitesimal generator* $\mathcal{L}^X F$ of X_t is defined on $F : \mathbb{R}^d \to \mathbb{R}$ by

$$\mathcal{L}^X F(x) = \lim_{t \to 0^+} \frac{1}{t} \{\mathbb{E}^x[F(X_t)] - F(x)\}, \tag{1.4}$$

where $\mathbb{E}^x[F(X_t)] = \mathbb{E}[F(X_t^{(x)})], X_0^{(x)} = x$. Additionally, if F is twice differentiable, then $\mathcal{L}^X F(x)$ is given by

$$\begin{aligned} \mathcal{L}^X F(x) = \alpha^\intercal \frac{\partial F}{\partial x}(x) + \frac{1}{2} \sigma^\intercal \frac{\partial^2 F}{\partial x^2}(x) \sigma \\ + \int_{\mathbb{R}} \sum_{i=1}^{l} \nu_i(\mathrm{d}y) \Big(F(x + f_i(t, y)) - F(x) - f_i(t, y) \frac{\partial F}{\partial x} \Big), \end{aligned} \tag{1.5}$$

where $\alpha^{\intercal}$ is the transpose of $\alpha(\cdot,\cdot)$.

Other less frequently used definitions are in line with books such as Sato (1999), Cont & Tankov (2003) and Øksendal & Sulem (2007) for jump processes, Pham (2009) for stochastic control, Glasserman (2013) and Gobet (2016) for Monte Carlo methods, Hamilton (1994) for time series econometrics and Karatzas & Shreve (1998) for general stochastic processes.

Chapter 2

Analytic Formulas for Futures and Options for a Linear Quadratic Jump Diffusion Model with Stochastic Convenience Yield and Seasonality

2.1 Introduction

Fish Pool is a derivatives market for farmed Atlantic salmon in Norway. Established in 2006 it has witnessed substantial growth. Trading volume of financial salmon contracts at Fish Pool salmon exchange had been at $90,449$ tonnes in 2016 and almost doubled since 2012 (approximately 15% of annual production of farmed Atlantic salmon in Norway). Not only is the market vital to the Norwegian economy and the world Atlantic salmon trade, it has also been at the center of a large number of academic articles published in the last decade, see for example Ewald & Ouyang (2017), Ewald et al. (2016), Solibakke (2012), Bloznelis (2016) and Asche et al. (2016). We chose this particular market as it provides a good benchmark test for our model in terms of quality of fit, but of course our results are valid and useful in a much wider context. Ewald & Ouyang (2017) have indicated the presence of seasonality, Solibakke (2012) has investigated the presence of volatility in the spot price, but the presence of jumps has not been investigated so far. Our model presents the most comprehensive approach so far to understand Fish Pool prices.

This chapter focuses on studying the stochastic nature of financial salmon contracts and the provision of a tractable model with the best possible fit. We proceed by the following

steps. We start by designing a general linear-quadratic jump diffusion (LQJD) model with stochastic convenience yield and seasonality adjustment, where the volatility follows a seasonal mean-reverting process, i.e. the mean-reversion level of volatility is seasonal. We then derive general pricing formulas for futures and European options written on futures in analytical form. Finally, we fit our model to the fish pool market, estimate the parameters and test the fitting performance of our model. This agenda is in line with recent outputs in the Operations Research literature, in particular Recchioni & Sun (2016), Mrázek et al. (2016), Date & Islyaev (2016) and Rambeerich et al. (2013). However the models presented there do not capture seasonal effects in stochastic volatility and convenience yield which are fundamentally important in commodity markets, not only from an operational point of view. In fact, we examine the correlation structure between spot prices and volatility and the seasonal patterns in volatility and convenience yield and show that our model produces option prices that are conform with the traditional pattern of implied volatility smiles and skews, a feature that the celebrated Schwartz (1997) model for example fails to produce. Further, we investigate the dynamics of the state dependent seasonal risk premium of the spot price, and link it to the uncertainty in EU salmon demand.

A variety of models have been proposed to capture commodity spot and derivatives prices. Schwartz (1997) uses a three-factor model to emphasize the importance of incorporating stochastic convenience yield, the implied benefit accrued by holding commodities, and interest rate, in order to characterize the dynamics of futures prices. Allowing the market price to be an arbitrary affine function of state variables, Casassus & Collin-Dufresne (2005) propose a framework that covers the Schwartz (1997) three-factor model by applying the more flexible "essentially affine" specification, which is introduced by Duffee (2002). Chiu et al. (2015) propose a tractable model capturing the conitegration, stochastic volatility and stochastic correlations.

Beyond the stochastic convenience yield and market price of risk specification, it is crucial to understand the dynamics of volatility in commodities markets, in order to price and hedge derivatives. While the notion of stochastic volatility is widely accepted by academia, there are fewer works dealing with stochastic volatility in markets of commodities as compared to those in equity markets. Geman & Nguyen (2005) consider the possibility

of stochastic volatility in the soybean market. Richter & Sørensen (2002) and Trolle & Schwartz (2009) adopt the general affine class of stochastic processes to model the state variables, mainly because its analytical tractability (Duffie et al., 2000). Results by Casassus & Collin-Dufresne (2005) provide robust evidence for the inclusion of jumps in the spot price dynamics. Wong & Lo (2009) derive tractable solutions for prices of European options written on mean-reverting log-normal underlying processes with stochastic volatility. L. Andersen (2010) develops a general Markov model framework that can be used to build models which cover jumps, stochastic volatility and regime-switches. Islyaev & Date (2015) introduce a one-factor with stochastic volatility model and its closed-form pricing formula for futures in the electricity market.

Some commodities differ from conventional equities in presenting seasonal patterns, especially agricultural and energy commodities. Ewald & Ouyang (2017) adopt a multi-factor model that incorporates seasonal adjustments in convenience yield for the fish pool market and find crucial evidence for seasonality in financial salmon contracts. Richter & Sørensen (2002) utilize a sum of trigonometric functions to capture the seasonality. Moreno et al. (2019) represent the mean reversion factor in spot prices of natural gas and seasonal factors in the stochastic convenience yield by Fourier series and highlight the importance of including seasonal patterns against the conventional Schwartz & Smith (2000) two-factor model.

It is notable that those features are employed to reveal the nature of economic phenomenon that are usually observed in financial or trading activities. We use jumps in price processes to capture possible large and abrupt changes in spot and futures prices. Jumps in prices are often attributed to tail events, such as natural disasters and financial crises that may trigger large downturn risks to investors and traders. Financial market data typically exhibit various forms of seasonality, such as the opening and closure effect of the markets, weekends effect or other day-of-the-week effect, and seasons effect that is commonly reported in agricultural commodity markets, typically subject to production and harvest cycles. These effects may lead to seasonal changes in the prices of financial products, which are reflected in the seasonality of convenience yield. While those effects may also impact on the trading volume in the market, then impact on the volatility of returns, resulting in the seasonality of

volatility.(see an overview in Arismendi et al. (2016))

Our model is able to capture seasonal stochastic volatility, so it is applicable to any commodity markets with this feature. In this chapter, we restrict the focus to fish pool market. Historical monthly volatility statistics are demonstrated in Table 2.1. It can be observed that the volatility fluctuates and is far from being constant. The futures contracts tend to exhibit a higher volatility during summer and fall. It peaks at 0.0952 in June and reaches the lowest point at 0.0588 in February.

Table 2.1: Historical Monthly Volatility of Salmon Futures

	JAN	FEB	MAR	APR	MAY	JUN
mean	0.0681	0.0588	0.0594	0.0715	0.0862	0.0952
std	0.0201	0.0164	0.0104	0.0231	0.0178	0.0235
	JUL	AUG	SEP	OCT	NOV	DEC
mean	0.0880	0.0885	0.0808	0.0786	0.0841	0.0796
std	0.0111	0.0275	0.0249	0.0264	0.0460	0.0392

Note: This table reports the mean and standard deviation for the annualized realized volatility levels for salmon futures. The realized volatility is calculated by taking standard deviation of weekly observations of futures returns for each month from 2007 to 2017.

In the empirical part, we first calibrate our model forcing jumps in spot prices to be zero. For this, we use the extended Kalman Filter (EKF) along with quasi-maximum likelihood (QML) estimation. This setup appeals from an operational point of view, since we have latent variables that are unobservable, and futures prices are non-linear with respect to the state variables. This estimation methodology has also been used by Richter & Sørensen (2002), Trolle & Schwartz (2009) and Ewald et al. (2019), where the estimator works well. Our results reveal that the model fits the financial salmon data well, especially for the contracts with short-term maturities. We find that volatility has higher persistence while convenience yield has higher volatility. We also find that the seasonal patterns are significant in convenience yield for both samples, and significant in volatility for the short-term sample. Small but positive correlation between spot prices and volatility, i.e. the "inverse leverage effect", and positive high correlation between the spot prices and the convenience yield are observed. In addition, the seasonal risk premium is found to be high at the beginning and end of each year, and low at every year's summertime. There is positive correlation between seasonal risk premium and variance of EU salmon consumption. We then calibrate our model allowing

for jumps in the spot prices, using the implied-state quasi-maximum likelihood (IS-QML) estimator introduced by Santa-Clara & Yan (2010). We do this due to the infeasibility of the EKF in the presence of jumps. We conclude from our results that the assumption of having jumps in the data generating process does not hold for the financial salmon prices.

This chapter differs from other research in this field in several aspects. First, we construct a new model that adopts a seasonal mean-reverting stochastic volatility with jumps and stochastic convenience yield for commodity markets, based on the two factor constant volatility model (Ewald & Ouyang, 2017). Precisely, we adopt seasonal stochastic volatility and show that our model outperform the two factor constant volatility model. Within this framework, the state variables of volatility and convenience yield follow linear-quadratic processes. There are many articles focusing on commodity markets, using models based on multi-factor approaches (Schwartz (1997)) to merely capture the term structure of futures or using stochastic volatility models (such as Trolle & Schwartz (2009)) to capture the term structure of implied volatility. While there are few incorporating some of the elements of stochastic volatility, jump-diffusions, stochastic convenience yield and seasonality, there is none so far covering them all. Ours is the first. Second, we develop tractable solutions for pricing commodity derivatives within our framework. Third, we apply our model to the fish pool market and investigate the dynamics of stochastic volatility and seasonal changes in the convenience yield. Moreover, we adopt a state dependent seasonal risk premium that allows a more flexible correlation structure of state variables, and investigate the relationship between the dynamics of seasonal risk premium and uncertainty of EU salmon demand. The latter is the first attempt in the context of salmon futures market, and is of fundamental importance in the application of our results to forecast salmon prices, which we will follow up in future work.

The chapter is organized as follows. In section 2 we set up our model under the pricing measure and look at the market price of risk specification and the corresponding physical measure. Section 3 contains the quasi-analytical pricing formulas for futures and European options written on futures. In Section 4 we present the empirical part, which consists of data processing, model calibration methodology and an analysis of calibration results and the seasonal risk premia. Section 5 summarizes the main findings of this chapter.

2.2 Model Specification

2.2.1 The model under the pricing measure

We start by setting up our linear-quadratic volatility with jump-diffusion model for commodity markets. P. Cheng & Scaillet (2007) introduce the general linear-quadratic jump-diffusion (LQJD) class and show that the conventional affine jump diffusion (AJD) class, defined in Duffie et al. (2000), is nested within the LQJD class. In line with the settings in P. Cheng & Scaillet (2007), we propose an m-dimensional state vector X_t driven by an n-dimensional Brownian motion W_t ($n \leq m$) and a pure jump process J_t given by

$$\mathrm{d}X_t = \zeta(X_t, t)\mathrm{d}t + \eta(X_t, t)\mathrm{d}W_t + \mathrm{d}J_t, \tag{2.1}$$

where $\zeta : \mathbb{R}^m \times \mathbb{R}^+ \to \mathbb{R}^m$ and $\eta : \mathbb{R}^m \times \mathbb{R}^+ \to \mathbb{R}^m$. For X_t to be an LQJD model, it is assumed that the drift coefficient matrix $\zeta(X_t, t)$, the covariance coefficient matrix $\Omega(X_t, t) = \eta(X_t, t)\eta(X_t, t)^\intercal$, and the jump intensity $\lambda(X_t, t)$ are linear quadratic with respect to X_t, i.e. of type $\frac{1}{2}X^\intercal F(t)X + G(t)^\intercal X + H(t)$ with $F(t)$, $G(t)$ and $H(t)$ appropriate deterministic functions.

In our specification of the LQJD class, the spot price is affected by two sources of risk: the continuous diffusive risk, denoted by a Brownian Motion, and discontinuous risk, denoted by a compound Poisson jump process. The convenience yield also follows a linear-quadratic stochastic process. This quadratic variance process is similar to Santa-Clara & Yan (2010), where they assume that diffusive volatility and the square root jump intensity follow two Gaussian OU processes. We place our model within a probability space $(\Omega,\ \mathcal{F},\ \{\mathcal{F}_t\}_{t\geqslant 0},\ P)$ equipped with the natural filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$. $S(t)$, $\delta(t)$ and $V(t)$ denote the spot price, convenience yield and variance process, respectively. Under the pricing measure Q, we specify our model as,

$$\frac{\mathrm{d}S(t)}{S(t-)} = \Big(r - \delta(t)\Big)\mathrm{d}t + e^{\alpha^*(t-)}\sigma(t-)\mathrm{d}W_s^Q(t) + \int_{\mathbb{R}} (e^y - 1)\tilde{\mu}(\mathrm{d}t, \mathrm{d}y), \tag{2.2}$$

$$V(t) = \sigma^2(t), \text{ where } \mathrm{d}\sigma(t) = (\mu - \kappa\sigma(t))\mathrm{d}t + \nu_\sigma \mathrm{d}W_\sigma^Q(t), \tag{2.3}$$

$$\mathrm{d}\delta(t) = (\alpha(t) - \beta\delta(t))\mathrm{d}t + \nu_\delta \mathrm{d}W_\delta^Q(t), \tag{2.4}$$

where $W_s^Q(t)$, $W_\sigma^Q(t)$ and $W_\delta^Q(t)$ are correlated Brownian motions with constant correlation matrix Σ containing the correlations $\rho_{s\sigma}, \rho_{s\delta}$ and $\rho_{\sigma\delta}$.

Here y is a variable defined on $\mathbb{R}$ and $\mu(\mathrm{d}t, \mathrm{d}y)$ is a time-homogeneous Poisson random measure. $\tilde{\mu}(\mathrm{d}t, \mathrm{d}y)$ is its compensated version, with a constant compensator $\nu(\mathrm{d}t, \mathrm{d}y) = \lambda f(y)\mathrm{d}t$ and $f(\cdot)$ the probability density function of y. The jump size y follows a normal distribution $y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ with mean μ_y and variance σ_y^2. This makes the jump process in the data generating process of returns to be a compound Poisson process. The functions $\alpha(t)$ and $\alpha^*(t)$ are deterministic and used to describe seasonal patterns in convenience yield and volatilities. We define $\alpha(t)$ and $\alpha^*(t)$ as,

$$\alpha^*(t) = \sum_{k=1}^{K_1} \big(\phi_k cos(2\pi kt) + \phi_k^* sin(2\pi kt)\big), \tag{2.5}$$

$$\alpha(t) = \alpha_0 + \sum_{k=1}^{K_2} \big(\psi_k cos(2\pi kt) + \psi_k^* sin(2\pi kt)\big). \tag{2.6}$$

Note that in order to be parsimonious on parameters, we follow the approach taken by Richter & Sørensen (2002), restricting K_1 and K_2 to be 2.

2.2.2 Market prices of risk specification

We need the dynamics of state variables under the real world measure P in order to calibrate the model. We denote with π_t the Radon-Nikodym derivative process transforming the real world measure into the markets pricing measure:

$$\pi_t = \frac{\mathrm{d}Q|_{\mathcal{F}_t}}{\mathrm{d}P|_{\mathcal{F}_t}} = e^{U_t}, \tag{2.7}$$

where

$$\begin{aligned} U_t = &\int_0^t \Xi(\tau)\mathrm{d}\bar{W}^P(\tau) - \frac{1}{2}\Xi(\tau)^2\mathrm{d}\tau + \int_0^t \int_{\mathbb{R}} \Phi(\Delta S_\tau)\mu'(\mathrm{d}\tau, \mathrm{d}y) \\ &- \int_0^t \int_{\mathbb{R}} (e^{\Phi(\Delta S_\tau)} - 1)\nu'(\mathrm{d}\tau, \mathrm{d}y), \end{aligned}$$

where $\bar{W}^P(\tau) = (\bar{W}_1^P(\tau), \bar{W}_2^P(\tau), \bar{W}_3^P(\tau))^\intercal$ is a vector of standard independent Brownian motions under the measure P. Here $\Xi(\tau) = (\Xi_1(\tau), \Xi_2(\tau), \Xi_3(\tau))$ is a vector of market prices for the Brownian shocks in the dynamics of the state variables defined by, $\Xi_1(\tau) = \lambda_1\sigma(\tau) + \lambda_1^c$, $\Xi_2(\tau) = \lambda_2\sigma(\tau) + \lambda_2^c$, $\Xi_3(\tau) = \lambda_3\delta(\tau) + \lambda_3^c$, where $\lambda_1, \lambda_1^c, \lambda_2, \lambda_2^c, \lambda_3, \lambda_3^c$ are

all constant coefficients. This kind of affine combination of state variables is referred to as a special case of the *essentially affine* class introduced by Duffee (2002). By specifying the market price of risk, we obtain the link between Brownian motions under the measures Q and P in form of $\mathrm{d}\bar{W}_i^Q(\tau) = \mathrm{d}\bar{W}_i^P(\tau) + \Xi_i(\tau)\mathrm{d}\tau,\ i = 1, 2, 3$.

The difference $\Delta S_\tau = S(\tau) - S(\tau-)$ in equation (2.7) stands for the jump magnitude at time τ, $\Phi(\cdot) = \log\left(\frac{\mathrm{d}\nu}{\mathrm{d}\nu'}\right)$, and $\mu'(\mathrm{d}\tau, \mathrm{d}y)$ and $\nu'(\mathrm{d}\tau, \mathrm{d}y)$ the Poisson random measure under measure P and its compensator, respectively. By assuming the jump process to be a compound Poisson process, we gain great flexibility in transforming measures, since the jump intensity and distribution of jump magnitudes in $\nu'(\mathrm{d}\tau, \mathrm{d}y) = \lambda' f'(y)\mathrm{d}t$ can be arbitrary. To be parsimonious in parameters, we choose only to change the mean of jump magnitudes, and the others to remain the same, i.e. $\lambda' = \lambda$, and y follows a normal distribution $y \sim \mathcal{N}(\mu'_y, \sigma_y^2)$ under the measure P. This means the jump intensity does not change under different measures, only the jump magnitude does. A similar approach has been used by Pan (2002).

We find that the dynamics of state variables under the measure P is given as follows:

$$\begin{aligned}\frac{\mathrm{d}S(t)}{S(t-)} &= \Big(r - \delta(t) + \Lambda_1(t)e^{\alpha^*(t)}\sigma(t) - \int_{\mathbb{R}}(e^y - 1)(\nu - \nu')(\mathrm{d}y)\Big)\mathrm{d}t \\ &\quad + e^{\alpha^*(t)}\sigma(t)\mathrm{d}W_s^P(t) + \int_{\mathbb{R}}(e^y - 1)\tilde{\mu}'(\mathrm{d}t, \mathrm{d}y),\end{aligned} \tag{2.8}$$

$$\mathrm{d}\sigma(t) = (\mu - \kappa\sigma(t) + \nu_\sigma\Lambda_2(t))\mathrm{d}t + \nu_\sigma \mathrm{d}W_\sigma^P(t), \tag{2.9}$$

$$\mathrm{d}\delta(t) = (\alpha(t) - \beta\delta(t) + \nu_\delta\Lambda_3(t))\mathrm{d}t + \nu_\delta \mathrm{d}W_\delta^P(t), \tag{2.10}$$

where the correlation structure is obtained through the Cholesky decomposition.[1] Here we

[1]Formally

$$\begin{bmatrix}\mathrm{d}W_\delta^Q(t)\\ \mathrm{d}W_\sigma^Q(t)\\ \mathrm{d}W_s^Q(t)\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ \rho_{\sigma\delta} & \sqrt{1-\rho_{\sigma\delta}^2} & 0\\ \rho_{s\delta} & \frac{\rho_{s\sigma}-\rho_{s\delta}\rho_{\sigma\delta}}{\sqrt{1-\rho_{\sigma\delta}^2}} & \sqrt{1-\rho_{s\sigma}^2-\left(\frac{\rho_{s\sigma}-\rho_{s\delta}\rho_{\sigma\delta}}{\sqrt{1-\rho_{\sigma\delta}^2}}\right)^2}\end{bmatrix} \cdot \begin{bmatrix}\mathrm{d}\bar{W}_3^P(t) - \big(\lambda_3\delta(t) + \lambda_3^c\big)\mathrm{d}t\\ \mathrm{d}\bar{W}_2^P(t) - \big(\lambda_2\sigma(t) + \lambda_2^c\big)\mathrm{d}t\\ \mathrm{d}\bar{W}_1^P(t) - \big(\lambda_1\sigma(t) + \lambda_1^c\big)\mathrm{d}t\end{bmatrix},$$

$$\begin{bmatrix}\mathrm{d}W_\delta^P(t)\\ \mathrm{d}W_\sigma^P(t)\\ \mathrm{d}W_s^P(t)\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ \rho_{\sigma\delta} & \sqrt{1-\rho_{\sigma\delta}^2} & 0\\ \rho_{s\delta} & \frac{\rho_{s\sigma}-\rho_{s\delta}\rho_{\sigma\delta}}{\sqrt{1-\rho_{\sigma\delta}^2}} & \sqrt{1-\rho_{s\sigma}^2-\left(\frac{\rho_{s\sigma}-\rho_{s\delta}\rho_{\sigma\delta}}{\sqrt{1-\rho_{\sigma\delta}^2}}\right)^2}\end{bmatrix} \cdot \begin{bmatrix}\mathrm{d}\bar{W}_3^P(t)\\ \mathrm{d}\bar{W}_2^P(t)\\ \mathrm{d}\bar{W}_1^P(t)\end{bmatrix}.$$

use the following notation,

$$\Lambda_1(t) = \sqrt{1-\rho_{s\sigma}^2 - \big(\frac{\rho_{s\sigma}-\rho_{s\delta}\rho_{\sigma\delta}}{\sqrt{1-\rho_{\sigma\delta}^2}}\big)^2} \cdot \big(\lambda_1\sigma(t)+\lambda_1^c\big) + \frac{\rho_{s\sigma}-\rho_{s\delta}\rho_{\sigma\delta}}{\sqrt{1-\rho_{\sigma\delta}^2}} \cdot \big(\lambda_2\sigma(t)+\lambda_2^c\big)$$
$$+ \rho_{s\delta}\big(\lambda_3\delta(t)+\lambda_3^c\big),$$
$$\Lambda_2(t) = \sqrt{1-\rho_{\sigma\delta}^2} \cdot \big(\lambda_2\sigma(t)+\lambda_2^c\big) + \rho_{\sigma\delta}\big(\lambda_3\delta(t)+\lambda_3^c\big) \text{ and } \Lambda_3(t) = \lambda_3\delta(t)+\lambda_3^c.$$

It should be noted that continuous risk premia for the return risk are parametrized by $\Lambda_1(t)e^{\alpha^*(t)}\sigma(t)$, which contains a quadratic polynomial of state variables $\sigma(t)$ and $\delta(t)$. For the jump risk premia, we could either incorporate premia for the jump size uncertainty by changing the mean and variance or the law for jump sizes, or incorporate premia for the jump times if we change the intensity. Concerning the potential difficulties of identifying risk premia for jump intensity and jump size separately, the jump risk premia are only characterized by the jump size uncertainty, i.e. we set $\lambda = \lambda'$. Therefore, the jump risk premia in our model is simply $\lambda(\mu_y' - \mu_y)$.

Volatility risk premia and convenience yield risk premia are parametrized by $\nu_\sigma\Lambda_2(t)$ and $\nu_\delta\Lambda_3(t)$, correspondingly. These risk premia are not as transparent as risk premia in the return process, since they are not directly observable. However, they are still reflected in the futures prices and options prices, as we will see in the quadratic structure of state variables in the pricing formulas for financial claims in the next chapter.

2.3 Pricing of of Contingent Claims

2.3.1 Price of Futures

Let $F(S,\sigma,\delta,t,T_0)$ (abbreviated as $F(t,T_0)$) denote the time t price of a futures contract that matures at time T_0, that is $F(t,T_0) = \mathbb{E}^Q\big[S(T_0)|\mathcal{F}_t\big]$. In the absence of arbitrage, the futures price is a martingale under the pricing measure, according to Duffie (2001). We also know that the time discounted price of an option written on a futures contract is a martingale under the pricing measure.

Proposition 2.3.1. *The futures price at time t for a contract expiring at time T_0 $F(t,T_0)$ follows a partial integro-differential equation (hereafter PIDE) with a boundary condition*

$F(T_0, T_0) = S(T_0)$ *as such*

$$\begin{aligned}
&\frac{\partial F}{\partial t} + \Big(r - \delta(t)\Big)S(t-)\frac{\partial F}{\partial S} + \Big(\mu - \kappa\sigma(t)\Big)\frac{\partial F}{\partial \sigma} + \Big(\alpha(t) - \beta\delta(t)\Big)\frac{\partial F}{\partial \delta} \\
&+ \frac{1}{2}e^{2\alpha^*(t-)}\sigma^2(t-)S^2(t-)\frac{\partial^2 F}{\partial S^2} + \frac{1}{2}\nu_\sigma^2\frac{\partial^2 F}{\partial \sigma^2} + \frac{1}{2}\nu_\delta^2\frac{\partial^2 F}{\partial \delta^2} \\
&+ \rho_{s\sigma}\nu_\sigma e^{\alpha^*(t-)}\sigma(t-)S(t-)\frac{\partial^2 F}{\partial S\partial \sigma} + \rho_{s\delta}\nu_\delta e^{\alpha^*(t-)}\sigma(t-)S(t-)\frac{\partial^2 F}{\partial S\partial \delta} \\
&+ \rho_{\sigma\delta}\nu_\sigma\nu_\delta\frac{\partial^2 F}{\partial \sigma\partial \delta} + \int_{\mathbb{R}}\nu(\mathrm{d}y)\Big(F(t, Se^y) - F(t, S) - (e^y - 1)S(t-)\frac{\partial F}{\partial S}\Big)) = 0.
\end{aligned} \tag{2.11}$$

Proof. Let $\mathcal{L}^X$ be the infinitesimal generator of the state process $X_t = \{S, \sigma, \delta\}$, then

$$\frac{\partial F}{\partial t} + \mathcal{L}^X F = 0,$$

With $a(t, X)$ the coefficients vector of the drift terms and $b(t, X)$ the coefficients of the diffusion terms, the infinitesimal generator of X_t is given as

$$\begin{aligned}
\mathcal{L}^X F(X) = a^\intercal(t, X)\frac{\partial F}{\partial X} + \frac{1}{2}b^\intercal(t, X)\frac{\partial^2 F}{\partial X^2}b(t, X) \\
+ \int_{\mathbb{R}}\nu(\mathrm{d}y)\Big(F(t, Se^y) - F(t, S) - (e^y - 1)S(t-)\frac{\partial F}{\partial S}\Big).
\end{aligned}$$

which is equivalent to (2.11) □

Proposition 2.3.2. *The futures price $F(x, \sigma, \delta, t, T_0)$ in the model (2)-(4) is given as*

$$F(x, \sigma, \delta, t, T_0) = e^{x + A(\tau) + B(\tau)^\intercal U + U^\intercal G(\tau) U}, \tag{2.12}$$

where $x_t = \log(S_t)$ and $\tau = T_0 - t$. Here $U \equiv \left[\begin{smallmatrix}\sigma\\ \delta\end{smallmatrix}\right]$ is a vector of state variables, and $U^\intercal$ stands for the transpose of U. $A(\tau)$ is a scalar function. $B(\tau) = \left[\begin{smallmatrix}B_\sigma\\ B_\delta\end{smallmatrix}\right]$ and $G(\tau) = \left[\begin{smallmatrix}G_{\sigma\sigma} & G_{\sigma\delta}\\ G_{\sigma\delta} & G_{\delta\delta}\end{smallmatrix}\right]$ are 2×1 and 2×2 matrices of functions, respectively. $A(\tau)$, $B(\tau)$ and $G(\tau)$ solve a Riccati system of ordinary differential equations (ODEs) with initial condition $A(0) = 0$, $B(0) = \left[\begin{smallmatrix}0\\ 0\end{smallmatrix}\right]$, and $G(0) = \left[\begin{smallmatrix}0 & 0\\ 0 & 0\end{smallmatrix}\right]$:

$$\frac{\mathrm{d}A(\tau)}{\mathrm{d}\tau} = r + \Theta^\intercal B(\tau) + B(\tau)^\intercal \Pi B(\tau) + tr\big(\Pi G(\tau)\big), \tag{2.13}$$

$$\frac{\mathrm{d}B(\tau)}{\mathrm{d}\tau} = \begin{bmatrix}0\\ -1\end{bmatrix} + (-\Gamma + \Lambda)B(\tau) + 2G(\tau)\Pi B(\tau) + 2G(\tau)\Theta, \tag{2.14}$$

$$\frac{\mathrm{d}G(\tau)}{\mathrm{d}\tau} = (-\Gamma + \Lambda)G(\tau) + G(\tau)(-\Gamma^\intercal + \Lambda^\intercal) + 2G(\tau)\Pi G(\tau), \tag{2.15}$$

where Θ, Γ, Λ *and* Π *are defined as,*

$$\Theta = \begin{bmatrix} \mu \\ \alpha_t \end{bmatrix}, \ \Gamma = \begin{bmatrix} \kappa & 0 \\ 0 & \beta \end{bmatrix}, \ \Lambda = \begin{bmatrix} \rho_{s\sigma}\nu_\sigma e^{\alpha_t^*} & \rho_{s\delta}\nu_\delta e^{\alpha_t^*} \\ 0 & 0 \end{bmatrix}, \ \Pi = \begin{bmatrix} \nu_\sigma^2 & \rho_{\sigma\delta}\nu_\sigma\nu_\delta \\ \rho_{\sigma\delta}\nu_\sigma\nu_\delta & \nu_\delta^2 \end{bmatrix}.$$

Proof. Let $x_t = \log(S_t)$, then it is not difficult to get the PIDE with respect to the logarithmic spot price x_t (with the boundary condition $F(x, \sigma, \delta, T_0, T_0) = e^{x_{T_0}}$),

$$\begin{aligned} &\frac{\partial F}{\partial t} + \Big(r - \delta(t) - \frac{1}{2}e^{2\alpha^*(t)}\sigma(t)^2 - \int_{\mathbb{R}}(e^y - 1)\nu(\mathrm{d}y)\Big)\frac{\partial F}{\partial x} + \Big(\mu - \kappa\sigma(t)\Big)\frac{\partial F}{\partial \sigma} \\ &+ \Big(\alpha(t) - \beta\delta(t)\Big)\frac{\partial F}{\partial \delta} + \frac{1}{2}e^{2\alpha^*(t)}\sigma^2(t)\frac{\partial^2 F}{\partial x^2} + \frac{1}{2}\nu_\sigma^2\frac{\partial^2 F}{\partial \sigma^2} + \frac{1}{2}\nu_\delta^2\frac{\partial^2 F}{\partial \delta^2} \\ &+ \rho_{s\sigma}\nu_\sigma e^{\alpha^*(t)}\sigma(t)\frac{\partial^2 F}{\partial x \partial \sigma} + \rho_{s\delta}\nu_\delta e^{\alpha^*(t)}\sigma(t)\frac{\partial^2 F}{\partial x \partial \delta} \\ &+ \rho_{\sigma\delta}\nu_\sigma\nu_\delta\frac{\partial^2 F}{\partial \sigma \partial \delta} + \int_{\mathbb{R}}\nu(\mathrm{d}y)\Big(F(t, x+y) - F(t, x)\Big) = 0. \end{aligned} \tag{2.16}$$

Inspired from solutions to other pricing models, we may assume that a solution is of the form:

$$F(x, \sigma, \delta, t, T_0) = e^{x + A(\tau) + B(\tau)^\intercal U + U^\intercal G(\tau) U}.$$

Substituting this into the PIDE (2.16), we have the following equation[2],

$$\begin{aligned} &\frac{\mathrm{d}A(\tau)}{\mathrm{d}\tau} + \frac{\mathrm{d}B(\tau)^\intercal}{\mathrm{d}\tau} + U^\intercal\frac{\mathrm{d}G(\tau)}{\mathrm{d}\tau}U = r + \mu B_\sigma + \alpha_t B_\delta + \nu_\sigma^2 G_{\sigma\sigma} + \nu_\delta^2 G_{\delta\delta} + 2\rho_{\sigma\delta}\nu_\sigma\nu_\delta G_{\sigma\delta} \\ &+ \frac{1}{2}\nu_\sigma^2 B_\sigma^2 + \frac{1}{2}\nu_\delta^2 B_\delta^2 + \rho_{\sigma\delta}\nu_\sigma\nu_\delta B_\sigma B_\delta + (-\kappa B_\sigma + \rho_{s\sigma}\nu_\sigma e^{\alpha_t^*}B_\sigma + 2\mu G_{\sigma\sigma} + \rho_{s\delta}\nu_\delta e^{\alpha_t^*}B_\delta \\ &+ 2\alpha_t G_{\sigma\delta} + 2\nu_\sigma^2 B_\sigma^2 G_{\sigma\sigma} + 2\nu_\delta^2 B_\delta^2 G_{\sigma\delta} + 2\rho_{\sigma\delta}\nu_\sigma\nu_\delta B_\sigma G_{\sigma\delta} + 2\rho_{\sigma\delta}\nu_\sigma\nu_\delta B_\delta G_{\sigma\sigma})\sigma_t \\ &+ (-2\kappa G_{\sigma\sigma} + 2\rho_{s\sigma}\nu_\sigma e^{\alpha_t^*}G_{\sigma\sigma} + 2\rho_{s\delta}\nu_\delta e^{\alpha_t^*}G_{\sigma\delta} + 2\nu_\sigma^2 G_{\sigma\sigma}^2 + 2\nu_\delta^2 G_{\sigma\delta}^2 + 4\rho_{\sigma\delta}\nu_\sigma\nu_\delta G_{\sigma\sigma}G_{\sigma\delta})\sigma_t^2 \\ &+ (-1 + 2\mu G_{\sigma\delta} - \beta B_\delta + 2\alpha_t G_{\delta\delta} + 2\nu_\sigma^2 B_\sigma^2 G_{\sigma\delta} + 2\nu_\delta^2 B_\delta^2 G_{\delta\delta} + 2\rho_{\sigma\delta}\nu_\sigma\nu_\delta B_\sigma G_{\delta\delta} \\ &+ 2\rho_{\sigma\delta}\nu_\sigma\nu_\delta B_\delta G_{\sigma\delta})\delta_t + (-2\beta G_{\delta\delta} + 2\nu_\delta^2 G_{\delta\delta}^2 + 2\nu_\sigma^2 G_{\sigma\delta}^2 + 4\rho_{\sigma\delta}\nu_\sigma\nu_\delta G_{\delta\delta}G_{\sigma\delta})\delta_t^2 \\ &+ (-2\kappa C_{\sigma\delta} + 2\rho_{s\sigma}\nu_\sigma e^{\alpha_t^*}G_{\sigma\sigma} + 2\rho_{s\delta}\nu_\delta e^{\alpha_t^*}G_{\sigma\delta} - 2\beta G_{\sigma\delta} + 4\rho_{\sigma\delta}\nu_\sigma\nu_\delta G_{\sigma\delta}^2 \\ &+ 4\nu_\sigma^2 G_{\sigma\sigma}G_{\sigma\delta} + 4\nu_\delta^2 G_{\delta\delta}G_{\sigma\delta} + 4\rho_{\sigma\delta}\nu_\sigma\nu_\delta G_{\sigma\sigma}G_{\delta\delta})\sigma_t\delta_t, \end{aligned}$$

after matching terms with the same power, it is not difficult to obtain the above Riccati matrix ODEs (2.13), (2.14) and (2.15). We may also notice that the jump terms are eliminated from the futures price excluding the jumps in the spot price. □

[2]Note that we have $\frac{\partial F}{\partial t} = -\frac{\partial F}{\partial \tau}$

Note that by observing the futures pricing formula (2.12)-(2.15), we figure out that the jump parameters are not reflected in the futures pricing formula explicitly, while the stochastic volatility parameters are reflected. Additionally, by applying a change of variable to (2.12), we are able to get the stochastic differential equation (SDE) for the futures price $F(t, T_0)$ under the Q measure,

$$\begin{aligned}\frac{\mathrm{d}F(t,T_0)}{F(t-,T_0)} &= e^{\alpha^*(t-)}\sigma(t-)\mathrm{d}W_s^Q(t) + \left(B(\tau)^\intercal + 2U^\intercal G(\tau)\right)\begin{bmatrix}\nu_\sigma \mathrm{d}W_\sigma^Q(t)\\ \nu_\delta \mathrm{d}W_\delta^Q(t)\end{bmatrix}\\ &\quad + \int_{\mathbb{R}}(e^y - 1)\tilde{\mu}(\mathrm{d}y, \mathrm{d}t). \end{aligned} \tag{2.17}$$

To obtain this we set the drift coefficient to be zero, knowing that the futures price is a martingale under the Q measure as stated before.

2.3.2 Price of European Option Written on Futures

We use the Fourier inversion method to get the price of options written on a futures contract. The Fourier inversion method has been extensively used to price derivatives (see for example Carr & Madan (1999) and Lewis (2000)), and it is of substantial efficiency when the characteristic function of the underlying asset price process is known analytically but the distribution function is not.

2.3.3 Price of European Option Written on Futures

We use the Fourier inversion method to obtain the prices of options written on a futures contract. The Fourier inversion method has been extensively used to price derivatives (see for example Carr & Madan (1999) and Lewis (2000)).

Proposition 2.3.3. *The time t price of a European call option expiring at time T_1 with strike price K written on the futures price for a contract expiring at time T_0 ($t \leq T_1 \leq T_0$),*

$C(F,\sigma,\delta,t,T_1,T_0,K)$, *satisfies the PIDE*

$$
\begin{aligned}
rC &= \frac{\partial C}{\partial t} + \big(\mu - \kappa\sigma(t)\big)\frac{\partial C}{\partial \sigma} + \big(\alpha(t) - \beta\delta(t)\big)\frac{\partial C}{\partial \delta} + \frac{1}{2}\nu_\sigma^2\frac{\partial^2 C}{\partial \sigma^2} + \frac{1}{2}\nu_\delta^2\frac{\partial^2 C}{\partial \delta^2} \\
&+ \frac{1}{2}\Big(e^{\alpha^*_{t-}}\sigma(t-) + \sigma_F^2\nu_\sigma^2 + \delta_F^2\nu_\delta^2 + 2\rho_{s\sigma}\nu_\sigma e^{\alpha^*_{t-}}\sigma(t-)\sigma_F + 2\rho_{s\delta}\nu_\delta e^{\alpha^*_{t-}}\sigma(t-)\delta_F \\
&+ 2\rho_{\sigma\delta}\nu_\sigma\nu_\delta\sigma_F\delta_F\Big)F(t-,T_0)^2\frac{\partial^2 C}{\partial F^2} + \rho_{\sigma\delta}\nu_\sigma\nu_\delta\frac{\partial^2 C}{\partial\sigma\partial\delta} + \nu_\sigma\big(\rho_{s\sigma}e^{\alpha^*_{t-}}\sigma(t-) + \nu_\sigma\sigma_F \\
&+ \rho_{\sigma\delta}\nu_\sigma\delta_F\big)F(t-,T_0)\frac{\partial^2 C}{\partial F\partial\sigma} + \nu_\delta\big(\rho_{s\delta}e^{\alpha^*_{t-}}\sigma(t-) + \rho_{\sigma\delta}\nu_\sigma\sigma_F + \nu_\delta\delta_F\big)F(t-,T_0)\frac{\partial^2 C}{\partial F\partial\delta} \\
&+ \int_{\mathbb{R}}\nu(\mathrm{d}y)\Big(C(Fe^y,t) - C(F,t) - (e^y-1)F(t-,T_0)\frac{\partial C}{\partial F}\Big),
\end{aligned}
\tag{2.18}
$$

with $\sigma_F = B_\sigma + 2G_{\sigma\sigma}\sigma(t) + 2G_{\sigma\delta}\delta(t)$ *and* $\delta_F = B_\delta + 2G_{\delta\delta}\delta(t) + 2G_{\sigma\delta}\sigma(t)$ *and boundary condition* $C(F,\sigma,\delta,T_1,T_1,T_0,K) = \big(F(T_1,T_0) - K\big)^+$.

Proof. From the SDE for the futures price the PIDE of the European call option price (2.18) can be obtained in a similar way as for the futures price. Further, the PIDE of the option price with respect to the logarithmic futures price $f(t,T_0) = \log\big(F(t,T_0)\big)$ can be obtained as

$$
\begin{aligned}
rC &= \frac{\partial C}{\partial t} + \big(\mu - \kappa\sigma(t)\big)\frac{\partial C}{\partial \sigma} + \big(\alpha(t) - \beta\delta(t)\big)\frac{\partial C}{\partial \delta} + \frac{1}{2}\nu_\sigma^2\frac{\partial^2 C}{\partial \sigma^2} + \frac{1}{2}\nu_\delta^2\frac{\partial^2 C}{\partial \delta^2} \\
&- \frac{1}{2}D_{FF}\frac{\partial C}{\partial f} + \frac{1}{2}D_{FF}\frac{\partial^2 C}{\partial f^2} + \nu_\sigma\big(\rho_{s\sigma}e^{\alpha^*_{t-}}\sigma(t-) + \nu_\sigma\sigma_F + \rho_{\sigma\delta}\nu_\sigma\delta_F\big)\frac{\partial^2 C}{\partial f\partial\sigma} \\
&\nu_\delta\big(\rho_{s\delta}e^{\alpha^*_{t-}}\sigma(t-) + \rho_{\sigma\delta}\nu_\sigma\sigma_F + \nu_\delta\delta_F\big)\frac{\partial^2 C}{\partial f\partial\delta} + \rho_{\sigma\delta}\nu_\sigma\nu_\delta\frac{\partial^2 C}{\partial\sigma\partial\delta} \\
&+ \int_{\mathbb{R}}\nu(\mathrm{d}y)\Big(C(f+y,t) - C(f,t) - (e^y-1)\frac{\partial C}{\partial f}\Big),
\end{aligned}
\tag{2.19}
$$

where we use the notation $D_{FF} = e^{\alpha^*_{t-}}\sigma(t-) + \sigma_F^2\nu_\sigma^2 + \delta_F^2\nu_\delta^2 + 2\rho_{s\sigma}\nu_\sigma e^{\alpha^*_{t-}}\sigma(t-)\sigma_F + 2\rho_{s\delta}\nu_\delta e^{\alpha^*_{t-}}\sigma(t-)\delta_F + 2\rho_{\sigma\delta}\nu_\sigma\nu_\delta\sigma_F\delta_F$. □

Proposition 2.3.4. *The time t price of a European call option expiring at time T_1 with strike price K written on the futures price for a contract expiring at time T_0 is given as*

$$
C(f,\sigma,\delta,\tau_1,T_0,K) = -\frac{Ke^{-r\tau_1}}{2\pi}\int_{ik_i-\infty}^{ik_i+\infty} e^{-ikX}\frac{H(k,\sigma,\delta,\tau_1,T_0,K)}{k^2 - ik}\mathrm{d}k, \tag{2.20}
$$

where $\tau_1 = T_1 - t$, $X = f - \log(Ke^{-r\tau_1}) - \tau_1\int_{\mathbb{R}}(e^y - 1)\nu(\mathrm{d}y)$, $f(t,T_0) = \log\big(F(t,T_0)\big)$ *and* $k = k_r + ik_i$ *is a complex number, in which* k_r *and* k_i *are the real and imaginary part respectively.*[3] *Note that the expiration price may not exist un-*

[3] To compute the complex integral, we fix the imaginary part k_i of k first, and then integrate the left part with respect to k_r within the real domain. The result is a real number, although the integration includes complex numbers.

less $k_i > 1$. $H(k,\sigma,\delta,\tau_1,T_0,K)$ *is called a fundamental transform and is given by* $H(k,\sigma,\delta,\tau_1,T_0,K) = e^{a(\tau_1)+b(\tau_1)^\intercal U+U^\intercal g(\tau_1)U}$, *with* $a(\tau_1)$ *a scalar function. The functions* $b(\tau_1) = \left[\begin{smallmatrix} b_\sigma \\ b_\delta \end{smallmatrix}\right]$ *and* $g(\tau_1) = \left[\begin{smallmatrix} g_{\sigma\sigma} & g_{\sigma\delta} \\ g_{\sigma\delta} & g_{\delta\delta} \end{smallmatrix}\right]$ *represent* 2×1 *and* 2×2 *matrices of functions, respectively. The functions* $a(\tau_1)$, $b(\tau_1)$ *and* $g(\tau_1)$ *solve a Riccati system of ordinary differential equations (ODEs) with initial condition* $a(0)=0$, $b(0)=\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$, *and* $g(0)=\left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$:

$$\begin{aligned}
\frac{\mathrm{d}a(\tau_1)}{\mathrm{d}\tau_1} &= \frac{1}{2}(1+i)kB(\tau)^\intercal \Pi B(\tau) + b(\tau_1)^\intercal\big(\Theta - ik\Pi B(\tau)\big) \\
&\quad + tr\big(\Pi g(\tau_1)\big) + \int_{\mathbb{R}} (e^y - 1)\nu(\mathrm{d}y), \qquad (2.21)\\
\frac{\mathrm{d}b(\tau_1)}{\mathrm{d}\tau_1} &= 2(1+i)kG(\tau)\Pi B(\tau) + (1+i)k\Lambda B(\tau) + \big(\Gamma + ik\Lambda \\
&\quad + 2iG(\tau)\Pi\big)b(\tau_1) + 2g(\tau_1)\big(\Theta - ik\Pi B(\tau)\big), \qquad (2.22)\\
\frac{\mathrm{d}g(\tau_1)}{\mathrm{d}\tau_1} &= (1+i)k\Big(\left[\begin{smallmatrix} e^{2\alpha_t^*} & 0 \\ 0 & 0 \end{smallmatrix}\right] + 2G(\tau)\Pi G(\tau) + \Lambda G(\tau) + G(\tau)\Lambda^\intercal\Big) \qquad (2.23)\\
&\quad - \big(\Gamma + ik\Lambda + 2ikG(\tau)\Pi\big)g(\tau_1) - g(\tau_1)\big(\Gamma + ik\Lambda + 2ikG(\tau)\Pi\big)^\intercal \\
&\quad + 2g(\tau_1)\circ\Pi\circ g(\tau_1),
\end{aligned}$$

where $A(\tau)$, $B(\tau)$, $G(\tau)$, Θ, Π, Γ *and* Λ *have been defined previously.*

Proof. This pricing formula is an inverse Fourier transform of an exponential of a quadratic function of state variables. We will show the steps in obtaining the final prices and dynamic system of ODEs. The process resembles that of Lewis (2000), who introduces the fundamental transform to price a variety of contingent claims.

Since we already have the PIDE of the option price regarding the logarithmic futures price (2.19), we consider the generalized Fourier transform of the option price

$$\hat{C}(k,\sigma,\delta,t,T_1,T_0,K) = \int_{-\infty}^{\infty} e^{ikf} C(f,\sigma,\delta,t,T_1,T_0,K)\mathrm{d}f.$$

One salient feature of this process is that we can separate the jump term in the option price and get rid of the partial derivative of f by transforming the PIDE for the option price in the time domain to the frequency domain. The option price will be the inverse Fourier transform of $\hat{C}(k,\sigma,\delta,t,T_1,T_0,K)$,

$$C(f,\sigma,\delta,t,T_1,T_0,K) = \frac{1}{2\pi}\int_{ik_i-\infty}^{ik_i+\infty} e^{-ikf}\hat{C}(k,\sigma,\delta,t,T_1,T_0,K)\mathrm{d}k. \qquad (2.24)$$

By taking the time derivative on both sides of the above equation (2.24) and changing the order of integrals respecting the jump term, the previous PIDE (2.19) is transformed into a PIDE for $\hat{C}(k,\sigma,\delta,t,T_1,T_0,K)$,

$$\begin{aligned} r\hat{C} &= \frac{\partial \hat{C}}{\partial t} + \frac{1}{2}(1+i)kD_{FF}\hat{C} + ik\int_{\mathbb{R}}(e^y-1)\nu(\mathrm{d}y)\hat{C} + (\mu-\kappa\nu_t)\frac{\partial \hat{C}}{\partial \sigma} \\ &+ (\alpha_t-\beta\delta_t)\frac{\partial \hat{C}}{\partial \delta} + (-ik)\nu_\sigma(\rho_{s\sigma}e^{\alpha^*_{t-}}\sigma_{t-} + \nu_\sigma\sigma_F + \rho_{\sigma\delta}\nu_\delta\delta_F)\frac{\partial \hat{C}}{\partial \sigma} \\ &+ (-ik)\nu_\delta(\rho_{s\delta}e^{\alpha^*_{t-}}\sigma_{t-} + \rho_{\sigma\delta}\nu_\sigma\sigma_F + \nu_\delta\delta_F)\frac{\partial \hat{C}}{\partial \delta} + \frac{1}{2}\nu_\sigma^2\frac{\partial^2 \hat{C}}{\partial \sigma^2} + \frac{1}{2}\nu_\delta^2\frac{\partial^2 \hat{C}}{\partial \delta^2} \\ &+ \rho_{\sigma\delta}\nu_\sigma\nu_\delta\frac{\partial^2 \hat{C}}{\partial\sigma\partial\delta} + \int_{\mathbb{R}}(e^{-iky}-1)\nu(\mathrm{d}y)\hat{C}. \end{aligned} \tag{2.25}$$

The boundary condition changes to

$$\hat{C}(k,\sigma,\delta,T_1,T_1,T_0,K) = \left(\frac{e^{(ik+1)f}}{ik+1} - K\frac{e^{ikf}}{ik}\right)\Bigg|_{f=\log(K)}^{f=\infty}. \tag{2.26}$$

Note that we need to impose a restriction on the imaginary part of k to guarantee the existence of the upper limit $f=\infty$ of (2.26). According to Lewis (2000), we set $k_i > 1$, then

$$\hat{C}(k,\sigma,T_1,T_1,T_0,K) = -\frac{K^{1+ik}}{k^2-ik}. \tag{2.27}$$

To remove the dependence on the risk-free interest rate and jump part in (2.25), let $\tau_1 = T_1 - t$ and

$$\hat{C}(k,\sigma,\delta,\tau_1,T_0,K) = -\frac{K^{1+ik}}{k^2-ik}\exp\left\{-r\tau_1 + ik\tau_1\int_{\mathbb{R}}\nu(\mathrm{d}y)(e^y-1)\right\}H(k,\sigma,\tau_1,T_0,K).$$

PIDE (2.25) transforms into

$$\begin{aligned} \frac{\partial H}{\partial \tau_1} &= \frac{1}{2}(1+i)kD_{FF}H + \left(\mu-\kappa\nu_t - ik\nu_\sigma(\rho_{s\sigma}e^{\alpha^*_{t-}}\sigma_{t-} + \nu_\sigma\sigma_F + \rho_{\sigma\delta}\nu_\delta\delta_F)\right)\frac{\partial H}{\partial \sigma} \\ &+ \left(\alpha_t-\beta\delta_t - ik\nu_\delta(\rho_{s\delta}e^{\alpha^*_{t-}}\sigma_{t-} + \rho_{\sigma\delta}\nu_\sigma\sigma_F + \nu_\delta\delta_F)\right)\frac{\partial H}{\partial \delta} + \frac{1}{2}\nu_\sigma^2\frac{\partial^2 H}{\partial \sigma^2} + \frac{1}{2}\nu_\delta^2\frac{\partial^2 H}{\partial \delta^2} \\ &+ \rho_{\sigma\delta}\nu_\sigma\nu_\delta\frac{\partial^2 H}{\partial\sigma\partial\delta} + \int_{\mathbb{R}}(e^{-iky}-1)\nu(\mathrm{d}y)H. \end{aligned} \tag{2.28}$$

The boundary condition changes to the initial condition $H(k,\sigma,\delta,0,T_0,K) = 1$. We propose a solution of exponential quadratic form $H(k,\sigma,\delta,\tau_1,T_0,K) = e^{a(\tau_1)+b(\tau_1)^\intercal U + U^\intercal g(\tau_1)U}$. Substituting this into (2.28) and matching coefficients of terms with the same power, we obtain the Riccati matrix system of ODEs (2.21), (2.22) and (2.23). It can be verified that this solution satisfies the above PIDE (2.28), and $a(\tau_1)$, $b(\tau_1)$ and $g(\tau_1)$ are solutions with the

initial conditions $a(0) = 0$, $b(0) = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$, and $g(0) = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$, respectively. Thus the option price formula is given by,

$$C(f, \sigma, \delta, t, T_1, T_0, K) = \frac{e^{-r\tau_1}}{2\pi} \int_{ik_i - \infty}^{ik_i + \infty} \frac{-K^{ik+1}}{k^2 - ik} e^{-ik\left(f - \tau_1 \int_{\mathbb{R}} (e^y - 1)\nu(\mathrm{d}y)\right)} H(k, \sigma, \delta, \tau_1, T_0, K) \mathrm{d}k.$$

By simple substitutions of variables we get (2.20). □

2.4 Empirical Implementation

2.4.1 Data

To demonstrate that our model is fully functional, we fit the model using a data set of weekly salmon forward/futures prices traded on the Fish Pool ASA commodity exchange. Later we use these, jointly with the calibrated models to evaluate a hypothetical European option contract[4]. Forward contracts differ from futures contracts in so far as a futures contract is a forward contract which is cleared through an exchange on a daily basis. However, at least in theory, the futures price will coincide with the corresponding forward price if the interest rate is uncorrelated with commodity prices. Contracts traded at the Fish pool before 19/07/2007 had been of forward type, but close to 100% of contracts traded afterward had been cleared daily via Fish Pool's link with NASDAQ, and hence are of futures type, according to Ewald & Ouyang (2017). In this chapter, we assume that the interest rate is constant, and in conclusion we do not need to differentiate between futures and forward prices. With regards to the spot price, the traditional approach is to use the closest-to-maturity futures contract as a proxy, since the spot price is mostly unobservable. Here we take the Fish Pool Index as a proxy of salmon spot prices. The Fish Pool Index (hereafter FPI) is a weighted synthesis of the NASDAQ Salmon Index, Fish Pool European Buyers Index and Statistics Norway customs statistics. Note that since only weekly FPI data are available, we choose to estimate our model at weekly frequency. We also sample futures data at weekly frequency, even though futures contract are settled daily. We adopt the approach taken by Santa-Clara & Yan (2010), i.e. choose the futures prices on Wednesday each week, and obtain prices from, in order

[4] "Hypothetical" option in the sense that in the whole history of Fish Pool ASA only a dozen or so options have been traded OTC

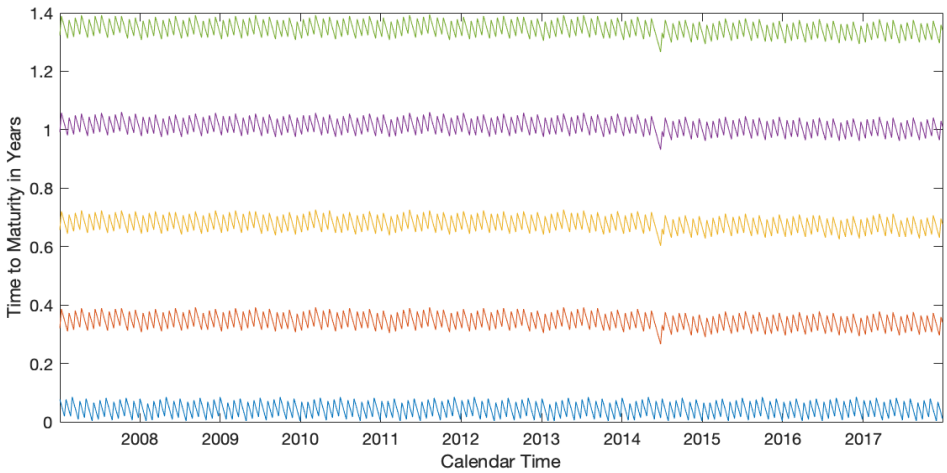
of preference, Tuesday, Thursday, Monday, Friday. Moreover, taking weekly observations addresses liquidity considerations and can also eliminate market micro-structure effects to some extent.

Table 2.2: Statistics of Contracts From January 2007 to December 2017

Contract	Mean Price (Standard Deviation)	Mean Maturity (Standard Deviation)
FPI	$38.63 (13.47)	0 (0) years
F1	38.52 (13.28)	0.059 (0.024)
F3	38.23 (13.04)	0.244 (0.025)
F5	37.99 (12.79)	0.475 (0.025)
F7	37.76 (12.72)	0.704 (0.025)
F9	37.41 (12.38)	0.934 (0.025)
F11	37.08 (11.99)	1.164 (0.025)
F13	36.69 (11.70)	1.394 (0.025)
F17	35.92 (11.18)	1.853 (0.025)
F25	34.80 (10.38)	2.772 (0.025)

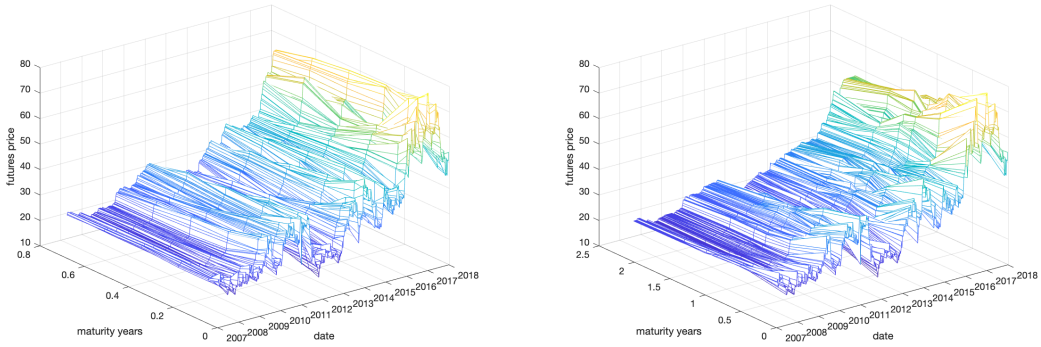
Note: We use the Fish Pool Index (FPI) as the spot price and F1 as the futures closest to maturity and F25 furthest to maturity.

Figure 2.1: Maturities For Contract F1, F3, F5, F7, F9



Ranging from January 2007 to December 2017, our data consist of 574 weekly observations, which contains FPI as spot prices and 25 futures contracts maturing at different dates. Since each futures contract has a fixed maturity date, the time to maturity varies as time goes by. We use notations F1-F25 to stand for different futures, of which maturities range from 0.043 years to 2.0123 years. We choose F1, F3, F5, F7 and F9 as the short-term sub-sample and F7, F11, F13, F17 and F25 as the long-term sub-sample. The average maturity for the short-term futures sample is 0.483 years, and that for the long-term futures sample is 1.577

Figure 2.2: Term Structure of Salmon Futures - Actual Forward Curves



(a) Short-term Maturity Sample (b) Long-term Maturity Sample

Note: These three dimensional figures show the salmon futures actual forward curves from 2007 to 2017.

years. From the definition of the spot price, we know that it is the value of an expiring futures contract. Here we add the FPI into each sample and set its maturity to be 0. Table 2.2 gives a summary statistics of prices and maturities regarding different futures. It can be observed from Figure 2.1 that maturities of short-term futures vary in a narrow range. Those of long-term futures present similar patterns, we omit these. We depict the actual term structure of futures contract, i.e. the actual forward curves across different times to maturities[5], in Figure 2.2.[6] A common ascending trend in prices can be observed across all types of futures.

2.4.2 Calibration without Jump Component in Spot Prices

Using Maximum likelihood estimation, we start this part by calibrating our model when forcing the parameters for jump intensity and jump size to be zero. Conventionally, the maximum likelihood estimator is combined with Kalman filtering in order to estimate parameters for fully specified models. The Kalman filter is efficient in extracting information from observations to infer on latent variables when the relationship of variables is known. An additional advantage of the Kalman Filter is that it is the best linear estimator if the state variables are Gaussian[7]. The Kalman filter in combination with maximum likelihood estimation belongs to a class of estimation approaches that works for state variables that are

[5]Odd looking shapes for the longer maturities futures are likely caused by low liquidity. The same is observed in Ewald & Ouyang (2017).

[6]These forward curves are analogous to interest rate term structures.

[7]The Kalman Filter is indeed the best filter among the set of all filters when the noise processes are Gaussian, and the best linear filter among the set of all linear filters when the noise processes are not Gaussian, as argued in the last section of Chapter 3 in Anderson & Moore (1979).

only observed up to unknown parameters in the model. This issue occurs in modeling equity derivatives, multi-factor interest rates and so on. For example, Pan (2002) applies an implied-state generalized method of moments to data of spot prices and options. Cortazar & Schwartz (2003) propose a two step least-square method to imply latent state variables for oil futures markets. However, it is well known that the results are sensitive to the choice of moments when implementing the generalized method of moments, and the significance of parameters cannot be inferred by the two step least-square method.

We employ a quasi-maximum likelihood (QML) estimator jointly with the extended Kalman filter (EKF) to calibrate our model. Here we use the EKF and QML estimator, since formulas for claims and data generating process are nonlinear. This approach has also been followed by other authors, see for example Chang & Kim (2001), Richter & Sørensen (2002) and Trolle & Schwartz (2009). The Kalman filter consists of a transition equation and a measurement equation. The measurement equation depicts the relationship between prices of claims and state variables. In our case, it is the relationship between futures prices and state variables $X_t = [x_t, \sigma_t, \delta_t]^\intercal$, where $x_t = \log(S_t)$. The transition equation in state space form can be considered as a discrete-time version of the data generating process. Following Santa-Clara & Yan (2010), we approximate our model (2.8), (2.9) and (2.10) by the following discrete-time system:

$$\begin{aligned}\Delta \log(S_t) = \Big(r - \delta(t-1) - \frac{1}{2}e^{2\alpha^*(t-1)}\sigma^2(t-1) + \Lambda_1(t-1)e^{\alpha^*(t-1)}\sigma(t-1) \\ - \lambda\mu_y'\Big)\Delta t + e^{\alpha^*(t-1)}\sigma(t-1)\epsilon_{s,t}\sqrt{\Delta t} + Y_t B_t,\end{aligned} \tag{2.29}$$

$$\Delta\sigma(t) = (\mu - \kappa\sigma(t-1) + \nu_\sigma\Lambda_2(t-1))\Delta t + \nu_\sigma\epsilon_{\sigma,t}\sqrt{\Delta t}, \tag{2.30}$$

$$\Delta\delta(t) = \alpha(t-1) - \beta\delta(t-1) + \nu_\delta\Lambda_3(t-1))\Delta t + \nu_\delta\epsilon_{\delta,t}\sqrt{\Delta t}, \tag{2.31}$$

where $\Delta\log(S_t) = \log(S_t) - \log(S_{t-1})$, and similar for $\Delta\sigma_t$ and $\Delta\delta_t$.[8] The variable Δt denotes the length of the time step in real time, which is 7/365 in our case. The variables $\epsilon_{s,t}$, $\epsilon_{\sigma,t}$ and $\epsilon_{\delta,t}$ are three time-independent Gaussian random variables with mean 0 and covariance structure Σ. We denote $B_t \sim i.i.d.\mathscr{P}(\lambda\Delta t)$, where $\mathscr{P}(\cdot)$ is a truncated Poisson

[8] In the discrete time approximation, the step from t to $t+1$ reflects a step of one unit in time, where the unit can be chosen arbitrarily small. Here t and $t+1$ denote indizes rather than actual times. The difference in time between t and $t+1$ is in fact the time unit for going forward Δt. We prefer this notation to the more cumbersome t to $t+\Delta$.

distribution. We further have Y_t i.i.d. Gaussian distribution with mean μ_y and variance σ_y^2, and independent of other variables. In this section, we force jump parameters to be 0, so the spot price process reduces to

$$\Delta \log(S_t) = \Big(r - \delta(t-1) - \frac{1}{2}e^{2\alpha^*(t-1)}\sigma^2(t-1) + \Lambda_1(t-1)e^{\alpha^*(t-1)}\sigma(t-1)\Big)\Delta t + e^{\alpha^*(t-1)}\sigma(t-1)\epsilon_{s,t}\sqrt{\Delta t}. \tag{2.32}$$

Given the discrete-time system (2.30), (2.31) and (2.32), we can now state the transition equation for the EKF as

$$X_{t+1} = T_t(X_t) + w_t, \tag{2.33}$$

where the $(m \times 1)$ state variable vector is given by $X_t = [x_t, \sigma_t, \delta_t]^\intercal$, m is the number of state variables (here $m = 3$), and $T(\cdot)$ stands for the nonlinear mapping from X_t to X_{t+1}. We have w_t an $(m \times 1)$ vector of serially uncorrelated disturbances at time t. By equation (2.12), the measurement equation can be denoted as

$$F_t = Z_t(X_t) + v_t, \tag{2.34}$$

where F_t stands for the $(n \times 1)$ vector of selected logarithmic futures prices, n is the number of futures with different maturities in each sample, and v_t is an $(n \times 1)$ vector of measurement errors (noise) added into the measurement equation in order to allow imperfections of observations. In the equation above $Z_t(\cdot)$ reflects the nonlinear mapping from X_t to F_t. To make the EKF functional, we linearize (2.33) and (2.34) by taking the Taylor expansions up to first order:

$$\hat{T}_t = \frac{\partial T_t(x)}{\partial x}\bigg|_{x=X_{t|t}}, \quad \hat{Z}_t = \frac{\partial Z_t(x)}{\partial x}\bigg|_{x=X_{t|t-1}},$$

where $(\cdot)_{t|t}$ and $(\cdot)_{t|t-1}$ denote the prediction conditional on information up to time t and $t-1$, respectively (the available information are the futures prices). The expressions $\hat{T}_t$ and $\hat{Z}_t$ are of type $(m \times m)$ matrix and $(n \times m)$ matrix respectively. Closed forms for $\hat{T}_t$ and $\hat{Z}_t$ can be obtained by computing the Jacobian determinants. Following this we can then

approximate (2.33) and (2.34) by,

$$X_{t+1} = \hat{T}_t X_t + c_t + w_t,$$

$$F_t = \hat{Z}_t X_t + d_t + v_t,$$

$$\mathbb{E}(w_t) = 0, \ \ \text{Var}(w_t) = Q_t,$$

$$\mathbb{E}(v_t) = 0, \ \ \text{Var}(v_t) = R_t,$$

where

$$Q_t = \begin{bmatrix} e^{2\alpha^*(t)}\sigma^2(t)\Delta t & \rho_{s\sigma}\nu_\sigma e^{\alpha^*(t)}\sigma^(t)\Delta t & \rho_{s\delta}\nu_\delta e^{\alpha^*(t)}\sigma^(t)\Delta t \\ \rho_{s\sigma}\nu_\sigma e^{\alpha^*(t)}\sigma^(t)\Delta t & \nu_\sigma^2\Delta t & \rho_{\sigma\delta}\nu_\sigma\nu_\delta\Delta t \\ \rho_{s\delta}\nu_\delta e^{\alpha^*(t)}\sigma^(t)\Delta t & \rho_{\sigma\delta}\nu_\sigma\nu_\delta\Delta t & \nu_\delta^2\Delta t \end{bmatrix},$$

$$R_t = diag\{\xi_1^2, \xi_2^2, \xi_3^2, \xi_4^2, \xi_5^2, \xi_6^2\},$$

where Q_t is the covariance-variance matrix of the transition errors at time t. Here R_t is the covariance-variance matrix of the measurement errors. We assume that the measurement errors of 6 futures observations in each sample at time t follow i.i.d. normal distributions with mean 0 and variance ξ_i^2, $i = 1, 2, 3, 4, 5, 6$, so R_t is a diagonal matrix. By adding serially and cross-sectionally uncorrelated disturbances with zero mean into the measurement equation, we can take into account market micro-structure factors such as bid-ask spreads and price limits. The variables c_t and d_t can be computed by

$$c_t = T_t(X_t) - \hat{T}_t X_t,$$

$$d_t = Z_t(X_t) - \hat{Z}_t X_t.$$

For formal results on the EKF, see for example Anderson & Moore (1979). Since we approximate the system by a linearization, the measurement errors v_t and noises w_t in the transition equation (2.32) follow a Gaussian distribution. This makes our estimator a QML estimator.

In the traditional way, we present the EKF as follows,

$$
\begin{aligned}
X_{t|t-1} &= T_t(X_{t-1|t-1}), \\
P_{t|t-1} &= \hat{T}_t P_{t-1|t-1} \hat{T}_t^{\intercal} + Q_t, \\
\varepsilon_{t|t-1} &= F_t - \hat{Z}_t X_{t|t-1}, \\
H_{t|t-1} &= \hat{Z}_t P_{t|t-1} \hat{Z}_t^{\intercal} + R_t, \\
K_t &= P_{t|t-1} \hat{Z}_t^{\intercal} (H_t)^{-1}, \\
X_{t|t} &= X_{t|t-1} + K_t \varepsilon_{t|t-1}, \\
P_{t|t} &= P_{t|t-1} - K_t \hat{Z}_t P_{t|t-1},
\end{aligned}
$$

where the iteration begins with a Gaussian random variable $X_0 = X_{0|0}$ with mean $\tilde{X}_0$ and covariance matrix P_0. Through the above recursion algorithm, the EKF can generate the updated state variable $X_{t|t}$ and covariance matrix $P_{t|t}$. Hence, we estimate parameters by maximizing the log-likelihood of the measurement errors, which is given by,

$$
\begin{aligned}
\mathscr{L}(\theta) &= \sum_{t=1}^{T} \log \big(Pr(F_t | \Psi_{t-1}) \big), \\
&= \sum_{t=1}^{T} \log \left((2\pi |H_{t|t-1}|)^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} \varepsilon_{t|t-1}{}^{\intercal} (H_{t|t-1})^{-1} \varepsilon_{t|t-1} \right\} \right).
\end{aligned}
$$

The QML parameter set θ is obtained from the following optimization target,

$$
\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \; -\mathscr{L}(\theta), \quad \text{where} \quad \theta := \left\{ \begin{array}{l} \mu, \kappa, \nu_\sigma, \beta, \nu_\delta, \phi_1, \phi_1^*, \phi_2, \phi_2^*, \alpha_0, \\ \psi_1, \psi_1^*, \psi_2, \psi_2^*, \lambda, \mu_y, \sigma_y, \rho_{s\sigma}, \rho_{s\delta}, \rho_{\sigma\delta}, \\ \lambda_1, \lambda_2, \lambda_3, \mu_y', \xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6 \end{array} \right\}.
$$

Our optimization procedure works as follows: Step 1, we choose an initial parameter set. Step 2, we run the EKF and get a log-likelihood value for our model. Step 3, we use the built in algorithm *fmincon* in MATLAB. This is based on the interior-point algorithm, in order to find a new parameter set that obtains the highest log-likelihood value. Eventually convergence (up to a predefined level of fluctuation) is achieved. We call this the optimal parameter set. Through the last run of the EKF using the optimal set, we also get the filtered latent state variables as a byproduct. Using the filtered latent state variables, we compute the filtered futures prices.

Numerical Considerations

Given a large enough sample size, we know that the variance-covariance matrix of the QML estimator can be approximated by the inverse of the Hessian matrix, see Hamilton (1994) for example. Hence we use the Hessian matrix to approximate the standard error of parameters in our model. We compute the Hessian matrix by the Broyden–Fletcher–Goldfarb–Shanno algorithm. Note that Trolle & Schwartz (2009) use the outer product of the first derivatives of the likelihood function due to the numerical instability of the Hessian matrix. Moreover, we employ the Cholesky factorization method to take the inverse of $H_{t|t-1}$ and the Joseph form of the covariance update formula $P_{t|t}$ to avoid potential numerical stability issues with the EKF. The Joseph form of the covariance update formula is denoted as, $P_{t|t} = (I - K_t Z_t) P_{t|t-1} (I - K_t Z_t)^\intercal + K_t R_t K_t$, where I is an identity matrix.

One iteration in the calibration of our model takes about five minutes; it takes about two days to calibrate our model. It should be pointed out that we use the Runge-Kutta method of fourth order accuracy to solve the Riccati system in order to compute the futures prices. This is sufficiently fast without losing too much accuracy. The computational burden mainly lies in solving ODE systems, running the Kalman filter iterations and computing Fourier integrals when incorporating options. As for previous papers, Richter & Sørensen (2002) calibrate a model with thirty parameters with about three weeks. Actually we believe the computational time is still moderate, and it might be of practical use such as empirical analysis. As the spot data are daily observations, a two-day calibration time might be enough for empirical analysis and trading purpose. [2]

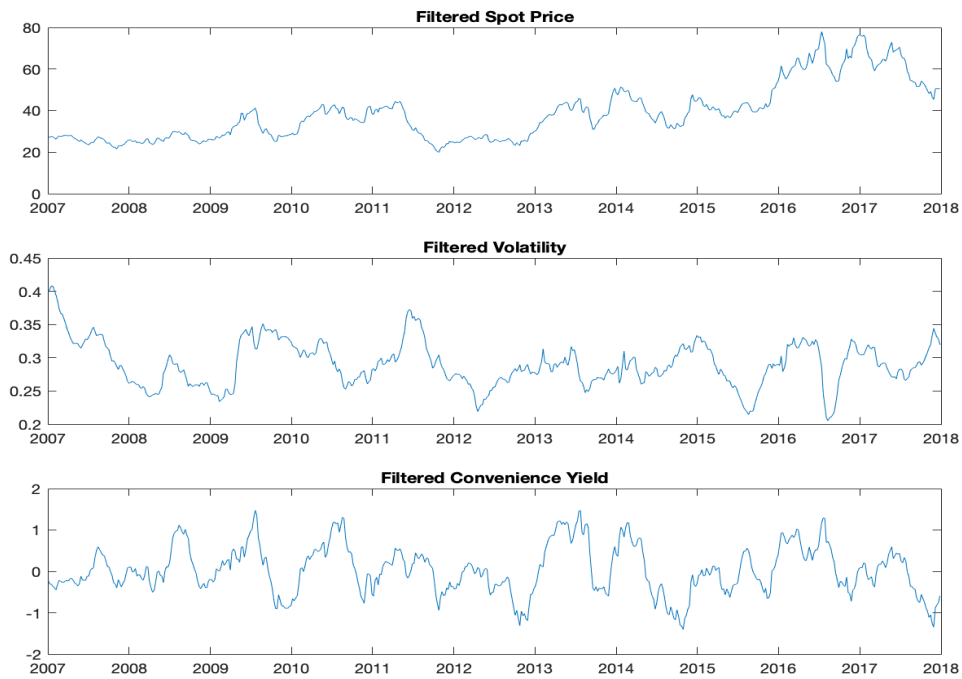
Results without Jump Component in Spot Prices

We estimate parameters from the short-term maturity sample (FPI, F1, F3, F5, F7, F9) and the long-term maturity sample (FPI, F7, F11, F13, F17, F25) separately. The model estimates are presented in Table 2.3. The mean-absolute-errors (MAE) and root-mean-square-error (RMSE) that measure the differences between filtered prices and real prices evaluate the performance of our model. The small MAE and RMSE in Table 2.4 demonstrate that our

[2] If we were to use the built in ODE solver package in MATLAB, it would take about two weeks to calibrate our model.

model fits the real data well. The best fit is obtained for spot prices and futures contracts with an average maturity of 257 days in the month of expiry. We also obtain the dynamics of filtered state variables (spot price, volatility and convenience yield), and present them in Figure 2.3. The volatility and convenience yield are revealed to be stochastic and mean-reverting. Further, we employ the filtered state variables and the pricing equation for futures to construct filtered forward curves for salmon. This is presented in Figure 2.4. Compared to the actual forward curves, the model generated (filtered) forward curves show a generally better prediction for the short-term sample, which can also be observed from Table 2.4. This finding is in accordance with Ewald & Ouyang (2017), who argue that the shapes of forward curves for the long-term panel are more difficult to capture than those for the short-term panel.

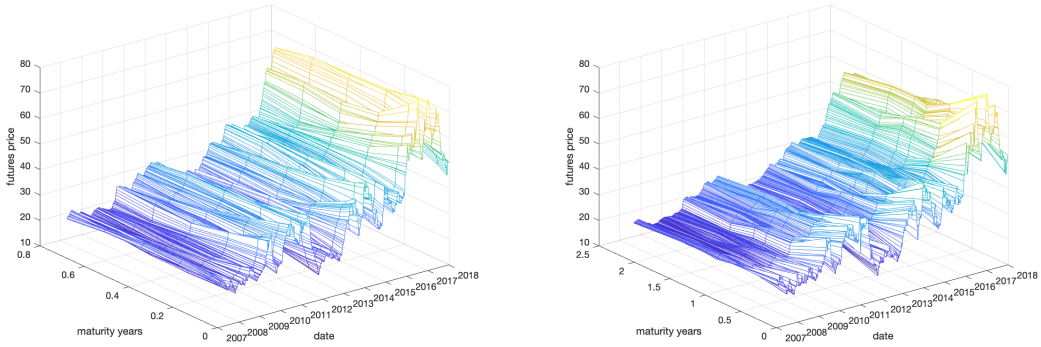
Figure 2.3: Filtered State Variables for LQJD Model without Jumps



Note: The figure shows the filtered state variables obtained through the Extended Kalman Filter, as they are not directly observable.

The variance term $e^{2\alpha^*(t)}\sigma^2(t)$ consists of a linear-quadratic class of state variable and a seasonality term. Hence $\sigma(t)$ can be called the seasonality adjusted volatility, of which the mean reversion speed is 2.219 for the short sample estimates. That means the half-life

Figure 2.4: Term Structure of Salmon Futures - Filtered Forward Curves



(a) Short-term Maturity Sample (b) Long-term Maturity Sample

Note: These three dimensional figures show the salmon futures filtered forward curves from 2007 to 2017.

Table 2.3: MAE and RMSE of Filtered Logarithmic Prices for LQJD Model without Jumps

Short-term Sample						
	FPI	F1	F3	F5	F7	F9
MAE	0.0340	0.0106	0.0295	0.0232	0.0079	0.0270
RMSE	0.0466	0.0140	0.0374	0.0297	0.0107	0.0348
Long-term Sample						
	FPI	F7	F11	F13	F17	F25
MAE	0.0128	0.0386	0.0198	0.0137	0.0236	0.0372
RMSE	0.0167	0.0472	0.0245	0.0179	0.0293	0.0466

Note: Mean-absolute-errors (MAE) and root-mean-square-errors (RMSE) are used to evaluate the differences between filtered prices and real prices (in logarithm).

(equals to $\log(2)/\kappa$) of the shocks to the seasonality adjusted volatility is approximately 0.312 years (about 4 months). The long-run mean of seasonality adjusted volatility (which is equal to μ/κ as $\sigma(t)$ follows an OU process) is 60.57% for the short-term sample.[9] The volatility of seasonality adjusted volatility for the short-term sample is 0.184, and it is highly significant. This confirms the stochastic nature of volatility in salmon markets. The corresponding parameter for the long-term sample is less significant, which implies that the volatility is less sensitive to shocks. We know that the noise in the salmon market originates from three kind of Brownian shocks within our framework. Thus, it can be further inferred that more stochastic behavior of salmon markets can be explained by the volatility with the maturity date approaching, as the coefficient of volatility induced Brownian shocks of the long-term sample is less significant. The long-run mean of seasonality adjusted volatility

[9]The actual long-run mean volatility could be lower than 60.57%, as we remove the seasonal component $e^{2\alpha^*(t)}$ here, which is often less than 1 according to our results.

Table 2.4: Estimated Parameters for LQJD Model without Jump Component in Spot Prices

Parameter	Short-term Sample Estimates (Standard Error)	Long-term Sample Estimates (Standard Error)
μ	1.344 (0.233)***	7.077 (2.308)***
κ	2.219 (0.305)***	18.791 (7.805)***
ν_σ	0.184 (0.017)***	1.209 (0.958)*
β	4.147 (0.226)***	3.270 (0.574)***
ν_δ	1.262 (0.078)***	1.138 (0.250)***
ϕ_1	-0.051 (0.031)*	-0.033 (0.024)
ϕ_1^*	-0.082 (0.032)***	-0.033 (0.021)
ϕ_2	-0.039 (0.031)	-0.036 (0.040)
ϕ_2^*	-0.019 (0.031)	-0.001 (0.017)
α_0	0.008 (0.068)	0.163 (0.102)
ψ_1	-1.004 (0.086)***	-0.803 (0.094)***
ψ_1^*	1.219 (0.062)***	0.752 (0.051)***
ψ_2	2.222 (0.062)***	2.040 (0.200)***
ψ_2^*	1.508 (0.141)***	0.180 (0.171)
λ	-	-
μ_y	-	-
σ_y	-	-
$\rho_{s\sigma}$	0.042 (0.076)	0.065 (0.056)
$\rho_{s\delta}$	0.925 (0.009)***	0.976 (0.006)***
$\rho_{\sigma\delta}$	-0.196 (0.077)**	-0.032 (0.019)*
λ_1	6.187 (2.008)***	9.993 (8.104)
λ_2	6.210 (1.916)***	-9.992 (7.568)
λ_3	-1.027 (0.353)***	-0.077 (0.064)
λ_1^c	-4.610 (0.999)***	-6.787 (2.455)***
λ_2^c	2.255 (0.909)**	3.861 (3.545)
λ_3^c	-0.053 (0.282)	0.139 (0.258)
μ_y'	-	-
ξ_1	0.049 (0.000)***	0.031 (0.000)***
ξ_2	0.022 (0.000)***	0.048 (0.000)***
ξ_3	0.038 (0.000)***	0.025 (0.000)***
ξ_4	0.031 (0.000)***	0.019 (0.000)***
ξ_5	0.013 (0.000)***	0.029 (0.000)***
ξ_6	0.034 (0.000)***	0.047 (0.000)***
Log-likelihood	9147.96	8981.34

Note: We divide our sample into two parts, one of which includes the shorter maturities (FPI, F1, F3, F5, F7, F9), and the other includes longer maturities (FPI, F7, F11, F13, F17, F25). [***] stands for significant at 1% level, [**] significant at 5% level and [*] significant at 10% level.

for the long-term sample is 41.02%, which comes with a stronger mean reversion speed of 18.791. This provides evidence for the so-called Samuelson effect[10] for salmon futures. To

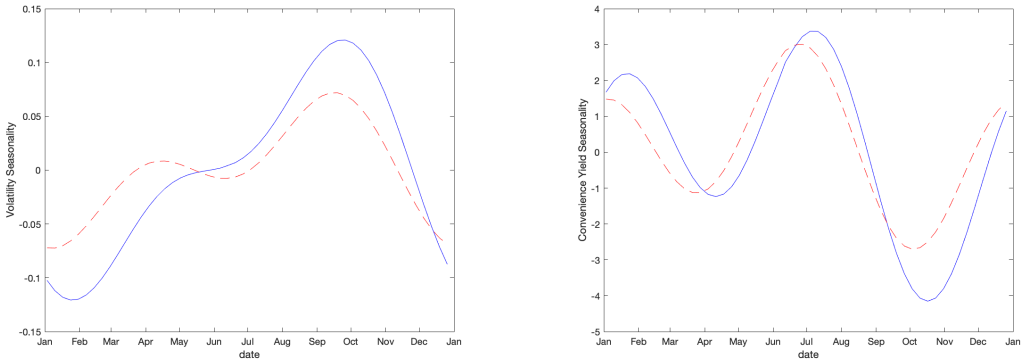
[10]The Samuelson effect refers to the hypothesis by Samuelson (1965a) that "the variations of distant maturity futures are lower than nearby futures prices", which means that the volatility of futures will tend to increase with the expiration date approaching.

further confirm this would require more experiments and is beyond the scope of this chapter. The half-life for the volatility for the long-term sample is about 0.037 years. This means that it takes about 2 weeks for the shocks to volatility to die out. The mean reversion speed of volatility under the pricing measure for the short-term sample (i.e. $\kappa+\lambda_2\sqrt{1-\rho_{\sigma\delta}^2}$) is 8.309, which is much higher than that under the physical measure. The mean reversion speed of volatility under the pricing measure for the long-term sample is 8.804, lower than that under the physical measure. The mean reversion speeds for different samples are close under the measure Q, but the long-run level of volatility under the measure Q is more complicated, since it partially depends on the diffusive risk premium for the volatility process $\nu_\sigma\Lambda_2(t)$, and $\nu_\sigma\Lambda_2(t)$ monotonically depends on the instantaneous convenience yield $\delta(t)$.

All the parameters attached to the stochastic convenience yield are highly significant, which means the convenience yield modeled in form of an OU process with seasonality adjustments fits the salmon data well. This is in accordance with the findings of Ewald & Ouyang (2017). The mean reversion speed parameter β equals 4.147 for the short-term sample, about twice the mean-reverting speed of the volatility process, and 3.270 for the long-term sample. While the mean-reverting speed of convenience yield is 4.342 in Ewald et al. (2016) for the short-term sample, in which they use the Schwartz multi-factor model. The half-life for the unexpected shocks to the convenience yield is 0.167 (about 2 months) for the short-term sample and 0.212 (about 77 days) for the long-term sample. The volatility of the convenience yield for the long-term sample is 1.138, which is slightly lower than that for the short-term sample (1.262). The high significance of both strengthens the importance of a stochastic convenience yield. The mean-reverting speed under the pricing measure, $\beta+\lambda_3\nu_\delta$, is 2.851 for the short-term sample. This is lower than under the physical measure. The mean-reversion for the long-term sample under the pricing measure is close to that under the physical measure (3.182). Particularly, the convenience yield reverts to a level that depends on the season, characterized by $\alpha(t)$. The seasonality adjusted long-run mean of the convenience yield α_0/β for the short-term sample is 0.002, which is similar to that for the long-term sample 0.05. Both of them are close to 0, and this can be confirmed by observing the filtered dynamics of the convenience yield in Figure 2.3.

The high significance of seasonal parameters in the convenience yield process, as ob-

Figure 2.5: Seasonal Changes of Volatility and Convenience Yield



(a) Seasonal Changes of Volatility (b) Seasonal Changes of Convenience Yield

Note: The blue line depicts seasonal factors for the short-term sample, and the red dashed line for the long-term sample.

served in Table 2.4, confirms the seasonality in convenience yield. The seasonal parameters attached to the volatility are less significant for the short-term sample, and not significant for the long-term sample. This indicates that the volatility for the longer maturity futures is not affected by seasonality but for the shorter terms it likely is. There appears to be an averaging effect in the longer term contracts, which at this point we cannot fully capture in the model.

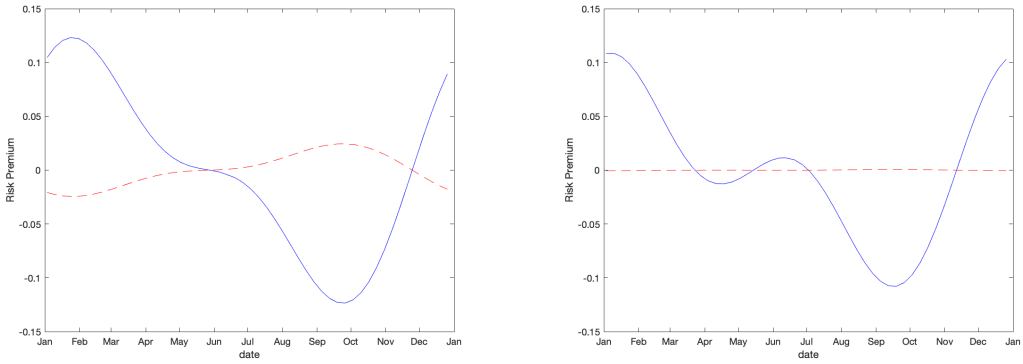
We plot the seasonal functions of volatility $\alpha^*(t)$ and convenience yield $\alpha(t)$ in Figure 2.5, correspondingly. The short-term sample and the long-term sample seasonality reflect similar patterns, except for different magnitudes. $\alpha(t)$ reaches its global maximum in July, and approaches the global minimum in October and November. This reveals the dynamics of the seasonal long-run level of convenience yield. While the volatility shows almost converse seasonal dynamics: it reaches the global minimum in January and February, and global maximum in October. This is consistent with the theory of storage. According to Working (1933) and Kaldor (1939), the futures prices tend to be in contango when the commodity supplies are high. Then the inventories will decrease, which gives a rise to the premium of futures prices up to the full storage cost and decreases the volatility of the spot and futures prices. The convenience yield and volatility will evolve in the opposite direction when the commodity supplies are tight. This is precisely the case for the salmon market, since it is known that salmon has a higher growth rate in warmer sea water (especially in summertime), which increases farmed fish supplies, resulting in a lower convenience yield and a higher volatility in November.

The correlation between spot price and volatility is 0.042 for the short-term sample, and 0.065 for the long-term sample, but neither of them is significant. A positive relationship between price and volatility is also found in Solibakke (2012). Moreover, Bloznelis (2016) finds that the correlation is close to zero for small fish. The correlation between spot price and convenience yield is 0.925 for the short-term sample and 0.976 for the long-term sample. The high correlation between price and convenience yield is also observed in Ewald & Ouyang (2017), in which the correlation is 0.855 for the short-term sample and 0.908 for the long-term sample. Both of the correlations between volatility and convenience yield for the short and long-term sample are negative. The positive correlation between spot price and volatility, which is called the "inverse leverage effect", is often observed for commodities especially agricultural commodities (see for example Richter & Sørensen (2002)). This is opposite to the classical leverage effect which depicts the negative correlation between equity spot price and volatility, which has a prevalent explanation based on a hypothesis by Black (1976): as the prices of an equity decline, the underlying company becomes mechanically more leveraged since the relative value of their debt rises relative to that of their equity, then their stock is likely to become riskier, hence more volatile. While one possible explanation for the inverse leverage effect is as follows: Assuming that hedging activities reduce the volatility and that speculative activities increase the volatility, the majority of traders in fish market are hedgers. Since the spot price is mainly affected by the supply of salmon, a fall in the spot price is often attributed to an increase of supply, indicating a pessimistic market sentiment, then causing more hedging activities, which finally decreases the volatility.

Seasonal Risk Premium

Risky assets need to generate excess returns over the risk-less rate in order for investors to invest. These excess returns are generally referred to as risk premia. The risk premium ΔRP_t

Figure 2.6: Risk Premium Dynamics



(a) Risk Premium for the Short-term Sample (b) Risk Premium for the Long-term Sample

Note: The blue line depicts the overall risk premium and the red dashed line the contribution originating from the convenience yield.

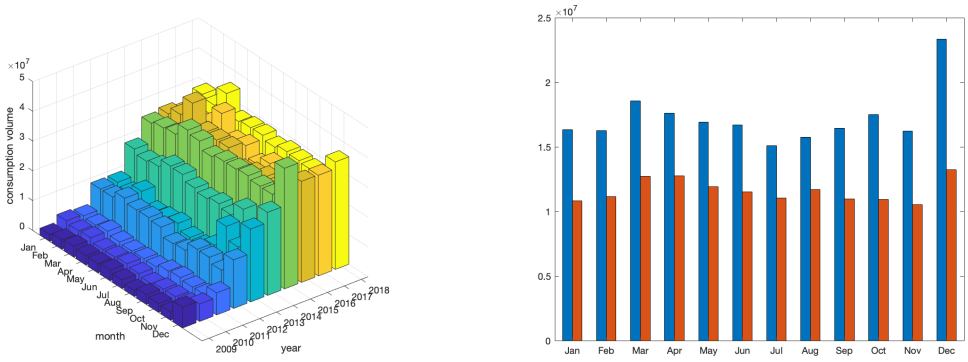
for a fixed time interval Δt can be defined as,

$$\begin{aligned}
\Delta RP_t &= \mathbb{E}^P(\Delta \log(S_t)) - \mathbb{E}^Q(\Delta \log(S_t)), \\
&= \Lambda_1(t) e^{\alpha^*(t)} \sigma(t) \Delta t, \\
&= \Bigg(\sqrt{1 - \rho_{s\sigma}^2 - \Big(\frac{\rho_{s\sigma} - \rho_{s\delta}\rho_{\sigma\delta}}{\sqrt{1-\rho_{\sigma\delta}^2}}\Big)^2} \cdot (\lambda_1 \sigma(t) + \lambda_1^c) + \frac{\rho_{s\sigma} - \rho_{s\delta}\rho_{\sigma\delta}}{\sqrt{1-\rho_{\sigma\delta}^2}} \cdot (\lambda_2 \sigma(t) + \lambda_2^c) \\
&+ \rho_{s\delta}(\lambda_3 \delta(t) + \lambda_3^c) \Bigg) e^{\alpha^*(t)} \sigma(t) \Delta t. \qquad (2.35)
\end{aligned}$$

As before we choose $\Delta t = 7/365$, i.e. one week. As can be observed, there are two trigonometric functions in the formula for the risk premium, $\alpha^*(t)$ appears explicitly and $\alpha(t)$ implicitly through affecting $\delta(t)$. In conclusion the risk premium exhibits seasonal patterns too. It should be further noted that the risk premium is state dependent, i.e. it depends on the volatility and convenience yield level. This is a difference between our model and other affine class models. We decompose the total risk premium into the compensation for the volatility and the compensation for the convenience yield, which are $\big(\Lambda_1(t) - \rho_{s\delta}\lambda_3\delta(t)\big)e^{\alpha^*(t)}\sigma(t)\Delta t$ and $\rho_{s\delta}\lambda_3\delta(t)e^{\alpha^*(t)}\sigma(t)\Delta t$ respectively. The total seasonal risk premium and its component for the convenience yield for the short-term sample and long-term sample are plotted in Figure 2.6.

From Figure 2.6, we can see that the contribution of the convenience yield risk premium to the overall risk premium is fairly small and close to zero for the long-term sample. The total risk premium reaches a local maximum in the end of each year and a local minimum

Figure 2.7: EU Salmon Consumption Chart



(a) Demand between 2008-2018 (b) Monthly Average and Standard Deviation

Note: (a) shows monthly salmon consumption (in kilogram) details from 2008-2018 for the EU. The blue bars in (b) are monthly averages of salmon consumption (in kilogram) from 2008-2018, and the red bars present monthly standard deviations.

during the summertime. The seasonal changes in the risk premium have a higher magnitude for the short-term sample than for the long-term sample. One possible reason is that the excess return shrinks in the middle of each year, during which there are excess supplies of salmon. While in winter especially around Christmas and Easter holidays, there is a significant increase in consumption of salmon, resulting in an increase of uncertainty (risk) in the demands. The increasing uncertainty drives the risk premium to rise since the market asks for higher returns for bearing higher risk. Bessembinder & Lemmon (2002) provide an equilibrium model stating that the risk premium of commodities is positively related to the variance of demand. This phenomenon is confirmed in electricity market by Kolos & Ronn (2008) and natural gas market by Shao et al. (2015). They argue that a large risk premium implied through the futures and a high demand volatility are usually accompanied by a large positive market price of demand risk. Thus, it will induce a positive correlation between risk premium and demand volatility. We conduct a similar analysis as in Shao et al. (2015) for the salmon futures market.

We obtain the monthly data for salmon consumption within the EU from the European Market Observatory for Fisheries and Aquaculture (EUMOFA). Due to data availability, we only extract data from 2008 to 2018. The summary charts are presented in 2.7. A large increase in the past 10 years can be observed in Figure 2.7 (a). We then compute the monthly mean and standard deviation of salmon consumption. Two peaks around the Christmas and

Easter holidays are found in Figure 2.7 (b). This is similar to the risk premium's patterns in Figure 2.6. We obtain the correlation between mid-month risk premium and demand volatility and run a regression for the mid-month risk premium on the demand volatility. However, the regression coefficient is not statistically significant.[11] This is most probably due to the limited number of data points in the regression. Nevertheless, the correlation between mid-month risk premium and demand uncertainty (standard deviation), 0.3421, supports the former assertion and provides economic significance.

2.4.3 Comparison with Two-factor Model

How much better is the new model than Ewald and Ouyang (2017) ? To answer this question we use the same data as before, but this time we calibrate the two-factor model from Ewald & Ouyang (2017), which is a seasonal extension of the two-factor Schwartz (1997) model. We provide details of the model under the measure P as following[12],

$$\mathrm{d}S(t) = (\mu - \delta(t))S(t)\mathrm{d}t + \sigma_1 S(t)\mathrm{d}Z_1(t), \tag{2.36}$$

$$\mathrm{d}\delta(t) = \kappa(\alpha(t) - \delta(t))\mathrm{d}t + \sigma_2 \mathrm{d}Z_2(t), \tag{2.37}$$

where $Z_1(t)$ and $Z_2(t)$ are two Brownian motions with correlation coefficient ρ, and

$$\alpha(t) = \alpha_0 + \sum_{k=1}^{2} \Big(\gamma_k cos(2\pi kt) + \gamma_k^* sin(2\pi kt)\Big).$$

Ewald & Ouyang (2017) specify the following form for the market price of risk and change of measure

$$\mathrm{d}\tilde{Z}_1(t) = \mathrm{d}Z_1(t) + \frac{\mu - r}{\sigma_1}\mathrm{d}t,$$
$$\mathrm{d}\tilde{Z}_2(t) = \mathrm{d}Z_2(t) + \frac{\lambda}{\sigma_2}\mathrm{d}t.$$

The calibration results[13] from using the Kalman filter (KF) and Maximum likelihood (ML) are presented in Table 2.6. The mean-reverting speed for both samples (4.635 for the

[11] The coefficient is 2.273×10^{-10}, and its 95% confidence interval is $[-0.419 \times 10^{-8}, 0.464 \times 10^{-8}]$.

[12] Some of the symbols here are the same as in our previous notations, as we try to stay consistent in notations as in Ewald & Ouyang (2017). We hope this will not confuse readers.

[13] The necessary details for the calibration such as futures pricing equation and moment conditions are provided in Ewald & Ouyang (2017), so we do not state them here again.

Table 2.5: MAE and RMSE of Filtered Logarithmic Prices for the Two-Factor Model

Short-term Sample						
	FPI	F1	F3	F5	F7	F9
MAE	0.0342	0.0096	0.0332	0.0268	0.0042	0.0318
RMSE	0.0470	0.0125	0.0430	0.0364	0.0055	0.0437
Long-term Sample						
	FPI	F7	F11	F13	F17	F25
MAE	0.0043	0.0480	0.0246	0.0202	0.0211	0.0316
RMSE	0.0057	0.0615	0.0313	0.0263	0.0277	0.0409

Note: Mean-absolute-errors (MAE) and root-mean-square-errors (RMSE) are used to evaluate the differences between filtered prices and real prices (in logarithm).

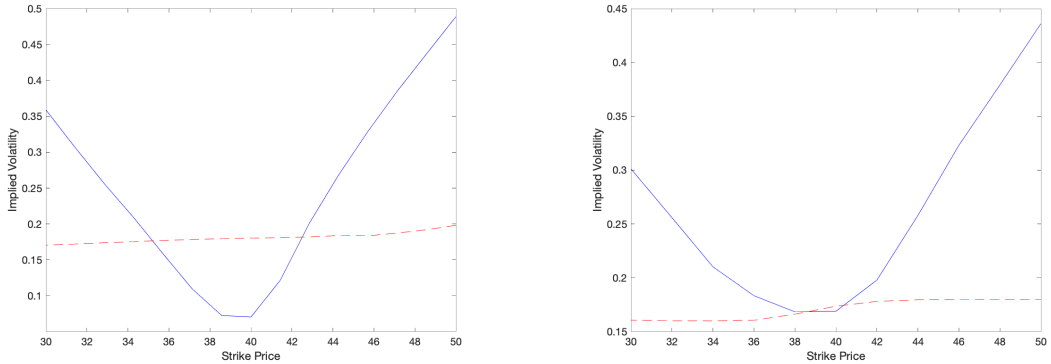
Table 2.6: Estimated Parameters for the Two-factor Model

Parameter	Short-term Sample Estimates (Standard Error)	Long-term Sample Estimates (Standard Error)
μ	0.021 (0.078)	0.108 (0.028)***
κ	4.635 (0.150)***	3.500 (0.160)***
α_0	-0.090 (0.084)	-0.055 (0.647)
σ_1	0.307 (0.010)***	0.434 (0.039)***
σ_2	1.568 (0.066)***	1.426 (0.144)***
ρ	0.897 (0.011)***	0.982 (0.005)***
λ	0.562 (0.370)	0.041 (0.078)
γ_1	0.056 (0.013)***	-0.022 (0.471)
γ_2	0.212 (0.020)***	0.034 (0.412)
γ_1^*	0.165 (0.011)***	0.009 (0.307)
γ_2^*	0.022 (0.009)**	0.002 (0.261)
ξ_1	0.050 (0.000)***	0.019 (0.001)***
ξ_2	0.021 (0.000)***	0.063 (0.000)***
ξ_3	0.043 (0.000)***	0.032 (0.000)***
ξ_4	0.037 (0.000)***	0.027 (0.000)***
ξ_5	0.009 (0.000)***	0.028 (0.000)***
ξ_6	0.044 (0.000)***	0.043 (0.000)***
Log-likelihood	8851.52	8615.97

Note: We divide our sample into two parts, of which one has short-term maturities (FPI, F1, F3, F5, F7, F9), and the other has long-term maturities (FPI, F7, F11, F13, F17, F25). [***] stands for significant at 1% level, [**] significant at 5% level and [*] significant at 10% level.

short-term sample and 3.500 for the long-term sample) are close to our LQJD model without jumps. Beyond that, the volatility of the convenience yield and correlation between spot price and convenience yield are also similar to those from our LQJD model. We present the filtering performance of the two-factor model in Table 2.5 by Mean-absolute-errors (MAE) and root-mean-square-errors (RMSE). The LQJD model without jumps is observed to be better in capturing the short term structure of salmon futures, but nearly the same for the

Figure 2.8: Implied Volatilities from European Options



(a) European Option with one-Month Maturity (b) European Option with three-Month Maturity

Note: The blue line depicts the implied volatilities of options priced by the LQJD model and the red dashed line the implied volatilities by the two-factor model. The underlying one-year-maturity futures price is 40, and the risk-free interest rate is 0.01.

long term structure. Hence, we argue that the LQJD model without jumps fits the short term sample better, not only because of the lower MAE and RMSE, but also because of the statistical significance of the parameters attached to the stochastic volatility.

Additionally, we conduct an experiment to highlight the importance of incorporating stochastic volatility in our LQJD model. With the calibration results from Table 2.4 and 2.6, we price hypothetical European options written on salmon futures using the LQJD model and the two-factor model, and back out the corresponding implied volatilities of the Black & Scholes model. We assume that the risk-free interest rate is 0.01 and two European options with one-month maturity and three-month maturity are traded in the fish market. The underlying futures price with one-year maturity is 40, and the state variables of volatility and convenience yield are 0.2 and 0.3, correspondingly. We obtain the implied volatilities with respect to different strike prices and present them in Figure 2.8. Evidently, a volatility smile can be observed for the LQJD model, while the implied volatilities for the two-factor model is more or less invariant with respect to the different strike prices. Therefore, including stochastic volatility enables the LQJD model to be more capable of capturing the volatility smile. It may also shed some light on commodity markets such as crude oil markets where the volatility smile is well recorded (Trolle & Schwartz (2009)).

2.4.4 Calibration: Full Model with Jumps

For calibrating the LQJD model with jumps in spot prices, we adopt an implied-state quasi-maximum likelihood (IS-QML) method[14], introduced by Santa-Clara & Yan (2010). Taking advantage of the linear form of futures and options pricing formulas, they introduce the IS-QML method by extending the implied-state general method of moments (Pan, 2002) compatible with the maximum likelihood estimator.

We illustrate the method briefly. For estimation, we need to assume that prices of some of the futures contracts are observed without errors, depending on the number of latent state variables (here it is three), while the other contracts are observed with i.i.d. Gaussian errors. We begin by assuming a set of initial parameters. From the analytic futures pricing formula (2.12), jointly with the ODE system (2.13), (2.14) and (2.15), we can back out the three latent state variables. We denote these as follows

$$\begin{pmatrix} Y_1(t,T_1) \\ Y_2(t,T_2) \end{pmatrix} = \begin{pmatrix} F_1(X_t,t,T_1) \\ F_2(X_t,t,T_2) \end{pmatrix} + \begin{pmatrix} 0 \\ \xi_t \end{pmatrix}, \tag{2.38}$$

$$X_t = F_1^{-1}(t, T_1, Y_1), \tag{2.39}$$

where Y_1 is a three-dimensional vector of prices of futures contracts that are observed without errors, and Y_2 are the remaining contracts. Above, T_1 and T_2 denote corresponding vectors of maturities. The functions $F_1(\cdot)$ and $F_2(\cdot)$ are the futures pricing formulas in (2.12) and we have ξ_t a vector of i.i.d. Gaussian errors. The vector of latent variables X_t is then used to obtain prices of futures contracts that are observed with errors $Y_2(t, T_2)$. After that the estimate for ξ_t is computed. Therefore, we can acquire the value of the likelihood function, which is of the following form:

$$\tilde{\mathscr{L}}(\theta) = \log f_X(X_t^\theta | X_{t-1}^\theta) + \log f_\xi(\xi_t^\theta), \tag{2.40}$$

where $f_\xi(\cdot)$ is the Gaussian density function for the observation error ξ_t^θ, and $f_X(\cdot)$ is the conditional density of the vector of state variables X_t^θ. The discretization method is expressed in (2.29), (2.30) and (2.31), with a truncation jump frequency M ($M \in \mathbb{N}^+$) of the related

[14] We also tried the EKF with QML introduced by Chang & Kim (2001), which takes Poisson noise into consideration. However, the convergence has been poor in this case.

truncated Poisson process B_t. Since the function $f_X(\cdot)$ is approximated by the density function of a truncated Poisson-normal mixture distribution, we need to set up M as a prior in practice. Here we fix the truncation jump times M to be 1 so that there are at most 1 jump per week. It is possible to set a greater value for M, as our model compatibility with reality increases with greater M. Yet it will also come with a significant increase in computational time. Then, we have

$$f_X(X_t|X_{t-1}) = (1-\lambda)\Phi(\mu_{X_t|X_{t-1}}, \sigma^2_{X_t|X_{t-1}}) + \lambda\Phi(\mu_{X_t|X_{t-1}}, \sigma^2_{X_t|X_{t-1}} + \sigma^2_y),$$

where $\mu_{X_t|X_{t-1}}$ and $\sigma^2_{X_t|X_{t-1}}$ are mean and variance of X_t conditioned on X_{t-1}, respectively, λ is the jump intensity, σ^2_y is the variance of jump magnitude, and

$$\Phi(\mu_{X_t|X_{t-1}}, \sigma^2_{X_t|X_{t-1}}) = (2\pi\sigma^2_{X_t|X_{t-1}})^{-1/2} \exp\big(-(X_t - \mu_{X_t|X_{t-1}})^2/2\sigma^2_{X_t|X_{t-1}}\big).$$

Through iteration we are able to obtain the optimal set of parameters by maximizing the log likelihood function (2.40). We use the same sets of samples to calibrate our LQJD with jumps model. It should be noted here that we assume that the observation errors are the same across different maturities at each time t in order to be parsimonious in parameters. The results are presented in Table 2.7. We observe that the parameters attached to volatility and stochastic convenience yield do not deviate by too much from our previous results in Table 2.4. Moreover, there is also an "inverse leverage effect" (positive correlation between spot price and volatility), although they are not statistically significant. The spot price and convenience yield is still tightly correlated (correlations are 0.975 for the short-term sample and 0.977 for the long-term sample). Of all those findings, the most important is that the jump intensity is small enough to be ignored and insignificant for both samples. For this reason, we stick to the findings of our previous results for the LQJD model without jumps in spot prices.[15]

Given the better performance in fitting futures data than other models, it is natural to expect the LQJD model without jumps will be more feasible to hedge long-term futures/forward contracts with existing short-term futures contracts. Let us explain the procedure concisely, as this might be left for future research. Following Schwartz (1997), after

[15]The likelihood values are much higher than previous results, but they are not comparable since we use a different likelihood function for the model with jumps.

Table 2.7: Estimated Parameters for the LQJD Model with Jump Component in Spot Prices

Parameter	Short-term Sample Estimates (Standard Error)	Long-term Sample Estimates (Standard Error)
μ	1.157 (0.216)***	3.399 (0.447)***
κ	3.507 (0.432)***	10.923 (1.578)***
ν_σ	0.258 (0.069)***	0.026 (0.016)
β	4.312 (0.309)***	4.487 (0.205)***
ν_δ	0.798 (0.101)***	1.409 (0.064)***
ϕ_1	0.553 (0.331)*	-0.051 (0.026)**
ϕ_1^*	-0.333 (0.358)	-0.032 (0.029)
ϕ_2	0.378 (0.121)***	0.016 (0.027)
ϕ_2^*	-0.243 (0.051)***	-0.018 (0.030)
α_0	-0.045 (0.078)	0.201 (0.029)***
ψ_1	-1.280 (0.055)***	-1.102 (0.082)***
ψ_1^*	1.065 (0.237)***	1.207 (0.059)***
ψ_2	0.844 (0.281)***	1.994 (0.142)***
ψ_2^*	1.127 (0.176)***	1.400 (0.151)***
λ	0.000 (0.099)	0.000 (0.001)
μ_y	-2.722 (0.941)***	-1.878 (0.731)***
σ_y	1.402 (0.218)***	0.004 (0.001)***
$\rho_{s\sigma}$	0.107 (0.565)	0.056 (0.920)
$\rho_{s\delta}$	0.975 (0.017)***	0.977 (0.004)***
$\rho_{\sigma\delta}$	-0.097 (0.097)	-0.038 (1.359)
λ_1	7.652 (1.523)***	7.272 (1.160)
λ_2	9.930 (3.473)***	-3.217 (5.340)
λ_3	0.733 (0.937)	-2.176 (0.437)***
λ_1^c	-9.960 (5.107)*	3.567 (1.624)**
λ_2^c	6.138 (1.880)***	6.794 (4.242)*
λ_3^c	2.241 (2.314)	1.671 (1.650)
μ_y'	2.165 (0.942)**	1.877 (0.731)**
ξ	0.034 (0.000)***	0.071 (0.000)***
Log-likelihood	20446.75	19306.08

Note: We divide our sample into two parts, of which one has short-term maturities (FPI, F1, F3, F5, F7, F9), and the other has long-term maturities (FPI, F7, F11, F13, F17, F25). [***] stands for significant at 1% level, [**] significant at 5% level and [*] significant at 10% level.

calibrate our model using historic data, the sensitivity of the target futures/forward contract with respect to the spot price and each state variable should equal to that of the proposed portfolio of short term futures contract with respect to the same factors. In this way, all sources of risks in our model can be eliminated. Hence, the number of the hedging futures contracts should be equal to the number of risks/factors in our model. As we do not assume an interest risk, the futures contracts here are of forward type, and there should be three different futures contracts to fully hedge the underlying long-term futures contract.

2.5 Concluding Remarks

We have developed a general jump-diffusion model for describing the dynamics of commodity prices in the presence of stochastic seasonal volatility and convenience yield. Analytical formulas for futures prices and prices of European options written on futures are derived using the inverse Fourier transform. Our model is connected to a number of other models that recently appeared in the Operations Research literature, but is new and unique in the way that it includes seasonal stochastic volatility and convenience yield. Both are fundamentally important for operational models of many commodity markets. Using futures contracts from the fish pool market we use the extended Kalman filter and quasi-maximum likelihood estimator to fit our model. The results reveal that our model fits the data well. We find that the volatility factor has higher persistence and lower volatility, whereas the convenience yield factor is much more volatile. Positive correlation between spot prices and volatility is established and significant seasonality is detected in the convenience yield for both samples, as well as in the volatility of spot prices. Further, we find that the risk premium in the price process shows seasonal patterns, depending on volatility and convenience yield. A positive relationship between the seasonal risk premium and the uncertainty of EU salmon demand is established. By comparing to the two factor model in Ewald & Ouyang (2017), we highlight the better performance of our LQJD model without jumps in pricing futures contracts, as the MAE and the RMSE of our model are reported to be less than the two factor model. In addition, the LQJD model can capture the volatility smile that makes it unique to those term structure models, and is capable to be utilized to more general commodity markets with option contracts. Then we add jumps in the LQJD model, and the results also reveal that LQJD model with jumps cannot beat the one without jumps. Because when incorporating jumps in spot prices and calibrating the corresponding model using an implied-state quasi-maximum likelihood estimator we find that the intensity of the jump component is close to 0 and insignificant, meaning there is no evidence supporting jumps in the salmon spot prices.

Chapter 3

Linear Quadratic Jump Diffusion Models with Co-Jumps in Volatility

3.1 Introduction

The increasing availability of high frequency observations of financial asset prices presents opportunities in modeling asset prices, but it also demands more robust and effective estimation methods. As in Chapter 2, we continue to employ the general linear quadratic jump diffusion (LQJD) class models introduced by P. Cheng & Scaillet (2007). In line with the settings in P. Cheng & Scaillet (2007), suppose an m-dimensional state vector X_t driven by an n-dimensional Brownian motion ($n \leq m$) and a pure jump process N such that

$$\mathrm{d}X_t = \mu(X_t, t)\mathrm{d}t + \sigma(X_t, t)\mathrm{d}W_t + \mathrm{d}N_t. \tag{3.1}$$

For X_t to be a LQJD model, it is required that the drift coefficients matrix $\mu(X_t, t)$, the covariance coefficients matrix $\Omega(X_t, t) = \sigma(X_t, t)\sigma(X_t, t)^\intercal$, and the jump intensity $\lambda(X_t, t)$ are linear quadratic with respect to X_t, i.e.

$$\varkappa(X_t, t) = \frac{1}{2}X^\intercal\Lambda(t)X + b(t)^\intercal X + c(t), \tag{3.2}$$

where $\intercal$ denotes transposition, and $\Lambda(t)$, $b(t)$ and $c(t)$ are time dependent deterministic functions. Conventional continuous stochastic volatility models have been rendered deficient in capturing abrupt changes in asset prices (Bakshi et al., 1997; Eraker et al., 2003), which in turn has triggered numerous studies on the impacts of including jumps in the price process. Further Pan (2002), Eraker et al. (2003) and Aït-Sahalia & Jacod (2009b) indicate that incorporating jumps in the volatility process is as important in accounting for sudden changes in

realized volatilities. Using the methodology of power variations and multi-power variations adopted for high frequency returns, we calibrate a linear quadratic volatility jump diffusion model, in which the continuous part of the variance process $\sigma(X_t, t)^2$ belongs to the LQJD class. Our setup so far is semi-parametric. However, we later investigate a fully parametric specification and assess whether the assumption of normal distributed jump sizes is suitable for our model and data.

The variance process in our linear quadratic volatility jump diffusion model is a superposition of the continuous linear quadratic part and the discontinuous part, denoted as

$$\sigma^2(t) = \sigma_c^2(t) + V_d(t). \tag{3.3}$$

We use the general Gaussian Ornstein-Uhlenbeck (hereafter OU) process as the continuous part of the volatility process, i.e.

$$\mathrm{d}\sigma_c(t) = \theta(\mu - \sigma_c(t))\mathrm{d}t + \nu \mathrm{d}W(t),$$

where θ, μ and ν are parameters, and $W(t)$ is a Brownian motion. This causes the continuous part of variance to be linear quadratic. One may doubt the validity of OU process for modelling volatility, since the OU process does not guarantee the volatility process to be positive, which is the key advantage of the CIR process. However, the CIR process is used to model the variance process(the square of the volatility process), and the squared OU process here we used for the variance process is non-negative. We model the discontinuous part $V_d(t)$ as a moving average of the past jumps. The linear quadratic feature is distinct from the Cox-Ingersoll-Ross process (hereafter CIR process), which belongs to the AJD class, in allowing the co variance structure of shocks to state variables to be unrestricted, according to Santa-Clara & Yan (2010) and Christoffersen et al. (2012). The discontinuous part of the variance is incorporating analytically tractable jumps in variance while preserving the mean-reverting feature of variance (Todorov, 2011). The superposition structure in our model is close to Todorov (2009a), but we use a linear quadratic volatility structure instead of a CIR process.

While modelling stock price using jump processes is motivated by observing large and sudden changes in historic prices, modelling stochastic volatility by jump processes is not so

directly observable, but is out of statistical findings. Continuous diffusion processes driven by Brownian motions are employed to depict volatility process as they can capture the so-called "volatility clustering", i.e. large price moves tend to be followed by large moves and small price moves tend to be followed by small moves. However, there could be counterexamples, where the discontinuous or jump processes can be used. Bates (1996) finds that the real conditional transition distribution of volatility is far more leptokurtotic than hypothesized using diffusion processes, observing large increases in volatility.

Bates (2000) and Pan (2002) confirm the existence of jumps in volatility of S&P 500 index. To see it more directly, we plot a figure (3.1) of realized variance of S&P 500 index, which we will elaborate later. The red line can be understood as the continuous volatility part and the black line as the jump part. Jumps in returns can capture large movements in returns, but the impact of jumps is transient. While volatility driven by diffusion processes or Brownian motions is persistent but can only increase gradually as this is the feature of Brownian motion, which evolves as a sequence of small Gaussian increments. Eraker et al. (2003) argue that the jumps in volatility can provide a rapidly moving yet persistent factor that is in line with statistical observations. Hence, with those high frequency modelling literature we mentioned previously, there is no doubt about the necessity of jumps in volatility.

Notably that the availability of intra-daily data of equity prices enables us to use econometric methods (see for example Barndorff-Nielsen & Shephard (2003b) and Barndorff-Nielsen & Shephard (2004)) to calibrate continuous parameters as well as jump parameters in models. The general idea is that the bi-power variation provides a non-parametric approximation of the continuous part of the realized variance, which can be estimated using the sum of intra-daily squared returns. Taking GMM as an example, by specifying the market model, the separated continuous parameters and jump parameters can be calibrated by matching analytic moments of integrated variance and quadratic variation. We will elaborate it explicitly later.

In the empirical part, we assess the quality of fit of our model to high-frequency stock data and compare it with the classical affine jump diffusion model, also called Bates model, and the linear quadratic volatility model with jumps in price. We estimate these by matching moments of daily power variations and bi-power variations within a semi-parametric setting.

The setting is semi-parametric, as we do not explicitly specify a parametric form for the distribution of jumps in the price as well as volatility process. Our empirical analysis shows that our model fits the data well. It is capable to capture the abrupt changes in volatility processes, and it performs outstandingly better than those models which only feature jumps in the price processes. This is similar to the result of Todorov (2009a). We further show that our semi parametric setup is superior to any fully parametric setup in which the two jump sizes are specified as being normal distributed. To do this, we repeat our estimation procedure within a fully parametric setting of our model, and find that while the result rejects over-identification, the coefficients of the normal distribution for the jump size are not significant. This may be due to the infeasibility of the normality assumption or problems within the estimation methodology. To exclude the latter, we conduct a Monte Carlo analysis, which shows that our estimation methodology would indeed return siginicant parameters for the normal distributions in the two jump sizes, were the jump sizes indeed normally distributed. In conclusion, they are not.

Bates (1996) introduced jumps into the price process within the well known Heston model. Following this, there have been two main trends in the jump diffusion literature, either assuming that the compensator of the price jumps (the jump intensity in the compound Poisson jumps case) is stochastic or allowing jumps in the volatility process. Duffie et al. (2000) analyzed the general AJD class and assumes that the intensity of jumps is a deterministic function of volatility. Time-changed Lévy processes have also been discussed in the literature (for example see Carr et al. (2003) and Carr & Wu (2004)). Santa-Clara & Yan (2010) employed correlated linear quadratic processes in both variance and compensator of jumps and estimated it under both physical and risk-neutral measure. Barndorff-Nielsen & Shephard (2001) proposed Lévy-driven OU processes, i.e. the moving average of realized Lévy processes, to capture sudden changes in volatilities. Later Barndorff-Nielsen & Shephard (2003a) explicitly investigated the distributional properties of those processes. Brockwell (2001) and Brockwell & Lindner (2012) explored different weighting structures of moving averages and generalized this to Lévy-driven continuous autoregressive moving average (CARMA) models. The setting of simultaneous jumps in price process and volatility process was also statistically confirmed by Jacod & Todorov (2010) and Jacod et al. (2017).

Further, Jacod et al. (2017) tested for non-correlation between jumps in price and volatility, rejecting the non-correlation hypothesis.

Both, the empirical part as well as the Monte Carlo experiment, are inspired by the idea of matching closed form moment conditions of returns. In the low frequency (i.e. daily frequency) case, T. G. Andersen & Sørensen (1996) adopted a Monte Carlo study to test the feasibility of inference in matching different moments when estimating stochastic autoregressive volatility (SARV) models using the general method of moments (GMM). Pan (2002) used joint moments inference of both returns and volatility to estimate affine diffusion processes. Turning to high frequency data, it was T. G. Andersen et al. (2003) who introduced a nonparametric estimator for realized volatility, which is simply the sum of squared returns. Jacod (1994) provided a general description for the asymptotic behavior of that estimator for the case of Brownian semimartingales. Bollerslev & Zhou (2002) used realized volatility as a proxy of integrated volatility, and estimated jump diffusion models by matching sample moments of integrated volatility. Further, Aït-Sahalia (2004) analyzed the practicality of using moments of realized volatility to estimate stochastic volatility with jumps models. Barndorff-Nielsen & Shephard (2004) introduced a robust method named bipower variation to disentangle jumps from the continuous part of returns. Following this Barndorff-Nielsen et al. (2006) derived asymptotic properties by imposing the Central Limit Theorem (hereafter CLT) for multipower variations. Todorov (2009b) implemented Monte Carlo analysis to justify the effectiveness of using joint moments inference of realized power variation and multipower variations when estimating jump diffusion processes. In addition, Mancini (2009) estimated coefficients in jump diffusion models using a threshold truncation method called truncated variation, which had been introduced by Mancini (2001). A rough fractional stochastic volatility (RFSV) model, adopted by Gatheral et al. (2018), depicted the logarithmic volatility as a fractional Brownian motion, and was proved to be remarkably consistent with high frequency financial time series.

The remainder of this chapter is structured as follows. Sections 2 and 3 provide the specifications of our model and provide parameters and proofs for essential moments which are used in the estimation. Section 4 presents the results of the estimation process based on high frequency data. In this section we also provide several robustness test results that

support our methodology. Section 5 provides the details of the Monte Carlo experiment that supports our semi parametric methodology. Finally, section 6 provides some concluding remarks.

3.2 Model Specification

We consider a linear quadratic volatility jump diffusion model in which the underlying asset price is affected by two types of risk: a continuous diffusive risk, denoted by a Brownian Motion, and discontinuous risk, denoted by a general jump process, both possibly multi-dimensional. The linear quadratic structure is similar to Santa-Clara & Yan (2010), where it is assumed that diffusive volatility and the square root jump intensity follow a Gaussian Ornstein-Uhlenbeck (OU) process. We set our model within a probability space $(\Omega,\ \mathcal{F},\ \{\mathcal{F}_t\}_{t\geqslant 0},\ P)$ equipped with the natural filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$. We denote with $P(t)$ the underlying asset price and let $R(t)$ denote its logarithmic returns. We the specify our model as follows:

$$\mathrm{d}R(t) = \alpha(t)\mathrm{d}t + \sigma(t-)\mathrm{d}W_1(t) + \int_{\mathbb{R}} J(x)\tilde{\mu}(\mathrm{d}t, \mathrm{d}x), \tag{3.4}$$

$$\sigma^2(t) = V_c(t) + V_d(t), \tag{3.5}$$

$$V_c(t) = \sigma_c^2(t), \text{ where } \mathrm{d}\sigma_c(t) = \theta(\mu - \sigma_c(t))\mathrm{d}t + \nu\mathrm{d}W_2(t), \tag{3.6}$$

$$\mathrm{d}V_d(t) = -\kappa V_d(t)\mathrm{d}t + \int_{\mathbb{R}} Q(x)\mu(\mathrm{d}t, \mathrm{d}x), \tag{3.7}$$

where $W_1(t)$ and $W_2(t)$ are two independent standard Brownian motions. The measure $\mu(\mathrm{d}t, \mathrm{d}x)$ is a time-homogeneous Poisson random measure, and $\tilde{\mu}(\mathrm{d}t, \mathrm{d}x)$ its compensated version, with compensator $\nu(\mathrm{d}t, \mathrm{d}x) = \mathrm{d}tG(\mathrm{d}x)$ with $G : \mathbb{R} \to \mathbb{R}^+$. In the special case of a compound Poisson process, the compensator simplifies to $\lambda f(x)\mathrm{d}t$, where λ is the intensity of Poisson jumps and $f(x)$ is the probability density function of the jump size. Here we assume possibly simultaneous jumps in the return process $R(t)$ and the discontinuous part of volatility process $V_d(t)$. $J(x)$ and $Q(x)$ are jump sizes, with $J : \mathbb{R} \to \mathbb{R}$ and $Q : \mathbb{R} \to \mathbb{R}^+$. We assume that J and Q are deterministic functions. The fact that jumps in the returns and volatility are driven by the same random source, the measure $\mu(\mathrm{d}t, \mathrm{d}x)$, results in possibly

simultaneous jumps, i.e. co-jumps, a feature that has recently been studied in the context of COGARCH in Klüppelberg et al. (2004) and in the context of jump driven stochastic volatility in Todorov (2011) for example. As indicated previously, we set our model semi-parametrically, i.e. we do not restrict the distribution of the jump parts. This feature, also used in Todorov (2011), takes on some advantages of GMM in contrast to the Maximum Likelihood Estimator (MLE). Note that because of the use of high-frequency, the drift term $\alpha(t)$ is practically zero and we can ignore it in the following it, compare Bollerslev & Zhou (2002)).

By imposing a superposition assumption on the coefficients of the diffusive part in the return process, we form a stochastic volatility process with jumps similar as in Todorov (2009a). The continuous part of the volatility process is assumed to follow an OU process. By employing Itô's formula, it is not difficult to find that the drift part features a linear quadratic structure with respect to the volatility $\sigma_c(t)$ if the variance follows a squared OU process. We find that the variance process satisfies

$$\begin{aligned} V_c(t) &= \sigma_c^2(t), \text{ where } \mathrm{d}\sigma_c(t) = \theta(\mu - \sigma_c(t))\mathrm{d}t + \nu \mathrm{d}W_2(t), \\ \mathrm{d}V_c(t) &= (2\theta\mu\sigma_c(t) + \nu^2 - 2\theta V_c(t))\mathrm{d}t + 2\nu\sigma_c(t)\mathrm{d}W_2(t). \end{aligned} \tag{3.8}$$

The discontinuous part of the volatility $V_d(t)$ follows a Non-Gaussian OU process, which is a special case of a Lévy-Driven processes (see Barndorff-Nielsen & Shephard (2001) and Brockwell & Lindner (2012)). The discontinuous part of the volatility is a CARMA(1,0) process, and it can be represented as a weighted sum of past jumps (assuming the initial value to be zero),

$$V_d(t) = \int_{-\infty}^{t} \int_{\mathbb{R}} e^{\kappa(s-t)} Q\mu(\mathrm{d}s, \mathrm{d}x) \tag{3.9}$$

In order to be consistent with moment conditions used in the estimation part, we then denote the *Quadratic Variation* (QV) process of the underlying return process during period $(t, t+a]$as,

$$[R, R]_{(t,t+a]} = \int_{t}^{t+a} \sigma^2(s)\mathrm{d}s + \int_{t}^{t+a} \int_{\mathbb{R}} J^2 \mu(\mathrm{d}s, \mathrm{d}x). \tag{3.10}$$

In addition, the *integrated variance* (IV) during the interval $(t, t+a]$ can be separated into

continuous part and discontinuous part, denoted as

$$IV_{(t,t+a]} = \int_t^{t+a} \sigma^2(s)\mathrm{d}s = \int_t^{t+a} V_c(s)\mathrm{d}s + \int_t^{t+a} V_d(s)\mathrm{d}s. \tag{3.11}$$

Let us note at this point that we do not restrict our jump process to be a compound Poisson process. We only estimate the cumulants of jumps, as Aït-Sahalia & Jacod (2009a) claimed evidence of small infinite activities of jumps. That enables us to make assumptions only on the integrability of those cumulants.

The classical leverage effect is implicitly specified in our model. Most often the leverage effect is realized by assuming negative correlation between the two Brownian Motions in return and volatility processes. However, the leverage effect can also be captured by assuming a negative co-variance in the jumps of returns and volatility. Precisely, the approach generally assumes a negative correlation between jump sizes in return process and volatility process, which in turn makes the correlation between return process and volatility process negative. In our model this can be realized by appropriately choosing the functions J and Q. The empirical analysis makes this evident through the estimation of the mixed moments of jump sizes in return and volatility processes. It is worth noting that our setting of correlated jumps can be seen as a special case of a sort of discontinuous leverage effect, which has indeed been investigated (jointly with the continuous leverage effect) by Aït-Sahalia et al. (2017). They also obtained the central limit theorems of the estimators for each type of leverage effect. It may be of great interest to explore the joint inference of the power and multi-power variations and the leverage estimator for existing models by using high frequency data.

3.3 Moment Conditions

As the GMM esimator we employed here are mainly composed by moments of the *quadratic variation* and the *integrated variance* that is defined in Section 3.4.1, it is worthwhile to investigate the tractability of such moment conditions. We start by the basic components of those moment conditions. The following Lemma 3.3.1 serves as the moments of the integrand of continuous part in the the *quadratic variation* and the *integrated variance*.

Lemma 3.3.1 (Moments of the Linear Quadratic Volatility Process)**.** *Given the initial value*

σ_0, *for a Linear Quadratic Volatility process of the type* $V(t) = \sigma^2(t)$ *with* $\mathrm{d}\sigma(t) = \theta(\mu - \sigma(t))\mathrm{d}t + \nu \mathrm{d}W(t)$ *(*$\theta > 0, \mu > 0, \nu > 0$ *and* $W(t)$ *is a Wiener Process), we have the following moments for* $\sigma(t)$*.*[1]

$$\mathbb{E}(\sigma_T|\mathcal{F}_t) = e^{-\theta(T-t)}\sigma_t + \mu(1 - e^{-\theta(T-t)}),\ \forall t \leqslant T, \tag{3.12}$$

$$\mathbb{E}(\sigma_T^2|\mathcal{F}_t) = e^{-2\theta(T-t)}\sigma_t^2 + (2\mu e^{-\theta(T-t)} - 2\mu e^{-2\theta(T-t)})\sigma_t + 2\mu^2(1 - e^{-\theta(T-t)}) + \frac{\nu^2 - 2\theta\mu^2}{2\theta}(1 - e^{-2\theta(T-t)}), \tag{3.13}$$

$$\lim_{t\to\infty} \mathbb{E}(\sigma_t) = \mu, \tag{3.14}$$

$$\lim_{t\to\infty} \mathbb{E}(\sigma_t^2) = \frac{\nu^2}{2\theta} + \mu^2, \tag{3.15}$$

$$\mathbb{E}(\sigma_s^2\sigma_u^2) = \left(\mu^2 + \frac{\nu^2}{2\theta}\right)^2 + \frac{2\mu^2\nu^2}{\theta}e^{-\theta(s-u)} + \frac{\nu^4}{2\theta^2}e^{-2\theta(s-u)},\ \forall u \leqslant s. \tag{3.16}$$

Proof. To obtain equation (3.12), the first order (uncentered) conditional moment of the OU process, we multiply σ_t by $e^{\theta t}$, through Itô's lemma we get,

$$e^{\theta T}\sigma_T = e^{\theta t}\sigma_t + \mu(e^{\theta T} - e^{\theta t}) + \int_t^T \nu e^{\theta t}\mathrm{d}W_s. \tag{3.17}$$

Taking the expectation on both sides conditional on $\mathcal{F}_t$, we get

$$\mathbb{E}(\sigma_T|\mathcal{F}_t) = e^{-\theta(T-t)}\sigma_t + \mu(1 - e^{-\theta(T-t)}).$$

For the second order conditional moment, we apply Itô's lemma to $e^{2\theta t}\sigma_t^2 - 2\mu e^{2\theta t}\sigma_t$, jointly with equation (3.17), we have,

$$\sigma_T^2 = 2\mu\Big(e^{-\theta(T-t)}\sigma_t + \mu(1 - e^{-\theta(T-t)})\Big) + e^{-2\theta(T-t)}\sigma_t^2 - 2\mu e^{-2\theta(T-t)}\sigma_t + \frac{\nu^2 - 2\theta\mu^2}{2\theta}(1 - e^{-2\theta(T-t)}) + \int_t^T \Big(2\nu e^{2\theta(s-T)}(\sigma_s - \mu) + 2\mu\nu e^{\theta(s-T)}\Big)\mathrm{d}W_s, \tag{3.18}$$

and by taking expectations on both sides conditional on $\mathcal{F}_t$ we get equation (3.13). Particularly at $t = 0$,

$$\mathbb{E}(\sigma_T) = e^{-\theta T}\mathbb{E}(\sigma_0) + \mu(1 - e^{-\theta T}).$$

The stationary unconditional moments can be obtained by simply taking the limit of time to infinity, so we have equation (3.14) and (3.15) (substituting T with t). Additionally, we

[1] Here we identify $\mathbb{E}(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_0)$

have unconditional third and fourth moments, $\lim_{t\to\infty}\mathbb{E}(\sigma_t^3) = \mu^3 + \frac{3\mu\nu^2}{2\theta}$ and $\lim_{t\to\infty}\mathbb{E}(\sigma_t^4) = \mu^4 + \frac{3\mu^2\nu^2}{2\theta} + \frac{3\nu^4}{4\theta^2}$, since the stationary distribution of σ_t is Gaussian.

The mixed moment (3.16), provided σ_t is covariance stationary, can be easily obtained by utilizing equation (3.13), (3.14), (3.15) and third and fourth moment through the *law of iterated expectations* $\mathbb{E}\Big(\sigma_u^2\mathbb{E}(\sigma_s^2|\mathcal{F}_u)\Big) = \mathbb{E}(\sigma_s^2\sigma_u^2)$, $\forall u \leqslant s$. □

We then obtain the moment conditions of the *integrated variance* by decomposing the it into continuous part and discontinuous part, as the variance process is defined to be the superposition of a linear-quadratic process and a non-Gaussian Lévy driven OU process. Given the previous lemma, we find those moments still tractable, though a little bit tedious. We present results in Theorem 3.3.1. Moment conditions for the *quadratic variation* are introduced in Theorem 3.3.2.

Theorem 3.3.1 (Moments of the Integrated Variance)**.** *For the integrated variance we defined as* $IV_{(t,t+a]} = \int_t^{t+a}\sigma^2(\tau)\mathrm{d}\tau = \int_t^{t+a}V_c(\tau)\mathrm{d}\tau + \int_t^{t+a}V_d(\tau)\mathrm{d}\tau$, $\forall a \in \mathbb{R}^+$, *where* $V_c(t)$ *follows a linear quadratic diffusion process and* $V_d(t)$ *follows a Non-Gaussian Lévy driven OU process, we have the following moments,*

$$\mathbb{E}\left(\int_t^{t+a}V_d(\tau)\mathrm{d}\tau|\mathcal{F}_s\right) = \int_{-\infty}^{s}\int_{\mathbb{R}}\frac{e^{\kappa(u-t)} - e^{\kappa(u-t-a)}}{\kappa}Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x) + a\frac{\int_{\mathbb{R}}QG(\mathrm{d}x)}{\kappa}, \tag{3.19}$$

$$\lim_{t\to\infty}\mathbb{E}(IV_{(t,t+a]}) = a\left(\frac{\nu^2}{2\theta} + \mu^2 + \frac{\int_{\mathbb{R}}QG(\mathrm{d}x)}{\kappa}\right), \tag{3.20}$$

$$\begin{aligned}\lim_{t\to\infty}Var(IV_{(t,t+a]}) &= \frac{4\mu^2\nu^2(e^{-\theta a} + \theta a - 1)}{\theta^3} + \frac{\nu^4(e^{-2\theta a} + 2\theta a - 1)}{4\theta^3} \\ &+ \frac{e^{2\kappa a} - 2e^{\kappa a} - 2e^{-2\kappa a} + 6e^{-\kappa a} + 2\kappa a - 3}{2\kappa^3}\int_{\mathbb{R}}Q^2G(\mathrm{d}x),\end{aligned} \tag{3.21}$$

$$\begin{aligned}\lim_{t\to\infty}Cov(IV_{(t,t+a]}, IV_{(t+h,t+h+a]}) &= a\left(\frac{\nu^2}{2\theta} + \mu^2\right)\frac{(\nu^2 - 2\theta\nu^2)(e^{-2\theta h} - e^{-2\theta(h-a)})}{4\theta^2} \\ &+ \left(a\left(\frac{\nu^2}{2\theta} + \mu^2\right)^2 + \frac{2\mu^2\nu^2(1 - e^{-\theta a})}{\theta^2} + \frac{\nu^4(1 - e^{-2\theta a})}{4\theta^3}\right)\cdot\left(\frac{e^{-2\theta(h-a)} - e^{-\theta h}}{2\theta}\right) \\ &+ \left(\frac{1 - e^{-\theta a}}{\theta}\left(\mu^3 + \frac{3\mu\nu^2}{2\theta}\right) + \mu\left(a - \frac{1 - e^{-\theta a}}{\theta}\right)\left(\frac{\nu^2}{2\theta} + \mu^2\right)\right) \\ &\cdot\left(\frac{e^{-\theta(h-a)} - e^{-\theta h}}{\theta} - \frac{e^{-2\theta(h-a)} - e^{-2\theta h}}{2\theta}\right)2\mu \\ &+ \frac{(e^{\kappa a} - 1)e^{-\kappa(2a+h)}}{2\kappa^3}\int_{\mathbb{R}}Q^2G(\mathrm{d}x), \textit{ for } h = ia,\ i \in \mathbb{N}^+.\end{aligned} \tag{3.22}$$

Proof. We start by proving equation (3.19). We use *Fubini's Theorem* and obtain[2]

$$\begin{aligned}\int_t^{t+a} V_d(\tau)\mathrm{d}\tau &= \int_t^{t+a}\int_{-\infty}^{\tau}\int_{\mathbb{R}} e^{\kappa(u-\tau)}Q\mu(\mathrm{d}u,\mathrm{d}x)\mathrm{d}\tau \\ &= \int_t^{t+a}\int_u^{t+a} e^{\kappa(u-\tau)}\mathrm{d}\tau\int_{\mathbb{R}} Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x) \\ &+ \int_{-\infty}^{t}\int_t^{t+a} e^{\kappa(u-\tau)}\mathrm{d}\tau\int_{\mathbb{R}} Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x) + a\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}. \end{aligned} \tag{3.23}$$

Taking conditional expectations with regards to $\mathcal{F}_s$ on both sides we obtain equation (3.19).

For equation (3.20), by the *law of iterated expectations* we know that

$$\begin{aligned}\lim_{t\to\infty}\mathbb{E}(IV_{(t,t+a]}) &= \lim_{t\to\infty}\int_t^{t+a}\mathbb{E}(V_c(\tau))\mathrm{d}\tau + \lim_{t\to\infty}\int_t^{t+a}\mathbb{E}(V_d(\tau))\mathrm{d}\tau \\ &= a\left(\frac{\nu^2}{2\theta}+\mu^2+\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}\right).\end{aligned}$$

To prove equation (3.21), we first calculate $\lim_{t\to\infty}\mathbb{E}(IV^2_{(t,t+a]})$,

$$\begin{aligned}\lim_{t\to\infty}\mathbb{E}(IV^2_{(t,t+a]}) &= \lim_{t\to\infty}\mathbb{E}\Big(\Big(\int_t^{t+a}V_c(\tau)\mathrm{d}\tau+\int_t^{t+a}V_d(\tau)\mathrm{d}\tau\Big)^2\Big) \\ &= \lim_{t\to\infty}\mathbb{E}\Big(\Big(\int_t^{t+a}V_c(\tau)\mathrm{d}\tau\Big)^2\Big) + \lim_{t\to\infty}\mathbb{E}\Big(\Big(\int_t^{t+a}V_d(\tau)\mathrm{d}\tau\Big)^2\Big) \\ &+ 2a^2\left(\frac{\nu^2}{2\theta}+\mu^2\right)\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}.\end{aligned}$$

To compute these expressions, we first derive the mean of the squared integrated continuous variance with the result of equation (3.16),

$$\begin{aligned}\lim_{t\to\infty}\mathbb{E}\Big(\Big(\int_t^{t+a}V_c(\tau)\mathrm{d}\tau\Big)^2\Big) &= \int_0^a\int_0^a\mathbb{E}\Big(\sigma_c^2(s)\sigma_c^2(u)\Big)\mathrm{d}s\mathrm{d}u \\ &= a^2\left(\mu^2+\frac{\nu^2}{2\theta}\right)^2 + \frac{4\mu^2\nu^2(e^{-\theta a}+\theta a-1)}{\theta^3} + \frac{\nu^4(e^{-2\theta a}+2\theta a-1)}{4\theta^3},\end{aligned} \tag{3.24}$$

[2]We swap the order of integration of the jump part of the integrated variance, in order to obtain the compensated Poisson measure and further the expectation of it.

then we deal with the squared integrated discontinuous variance of equation (3.23),

$$
\begin{aligned}
\mathbb{E}\Big(\Big(\int_t^{t+a} V_d(\tau)\mathrm{d}\tau\Big)^2\Big) &= \mathbb{E}\Big(\Big(\int_t^{t+a}\int_u^{t+a} e^{\kappa(u-\tau)}\mathrm{d}\tau \int_{\mathbb{R}} Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x) \\
&+ \int_{-\infty}^{t}\int_t^{t+a} e^{\kappa(u-\tau)}\mathrm{d}\tau \int_{\mathbb{R}} Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x) + a\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}\Big)^2\Big) \\
&= \mathbb{E}\Big(\Big(\int_t^{t+a}\int_u^{t+a} e^{\kappa(u-\tau)}\mathrm{d}\tau \int_{\mathbb{R}} Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x)\Big)^2\Big) \\
&+ \mathbb{E}\Big(\Big(\int_{-\infty}^{t}\int_t^{t+a} e^{\kappa(u-\tau)}\mathrm{d}\tau \int_{\mathbb{R}} Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x)\Big)^2\Big) + a^2\left(\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}\right)^2 \\
&= \mathbb{E}\Big(\int_t^{t+a}(\int_u^{t+a} e^{\kappa(u-\tau)}\mathrm{d}\tau)^2 \int_{\mathbb{R}} Q^2\mu(\mathrm{d}u,\mathrm{d}x)\Big) \\
&+ \mathbb{E}\Big(\int_{-\infty}^{t}(\int_t^{t+a} e^{\kappa(u-\tau)}\mathrm{d}\tau)^2 \int_{\mathbb{R}} Q^2\mu(\mathrm{d}u,\mathrm{d}x)\Big) + a^2\left(\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}\right)^2 \\
&= \frac{e^{2\kappa a} - 2e^{\kappa a} - 2e^{-2\kappa a} + 6e^{-\kappa a} + 2\kappa a - 3}{2\kappa^3}\int_{\mathbb{R}} Q^2G(\mathrm{d}x) + a^2\left(\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}\right)^2,
\end{aligned}
\tag{3.25}
$$

where we use the time-homogeneity of the jump Process and the isometry formula.[3]

Through equations (3.24) and (3.25) we calculate $\lim_{t\to\infty}\mathbb{E}(IV^2_{(t,t+a]})$.

$$
\begin{aligned}
\lim_{t\to\infty}\mathbb{E}(IV^2_{(t,t+a]}) &= a^2\left(\mu^2 + \frac{\nu^2}{2\theta}\right)^2 + \frac{4\mu^2\nu^2(e^{-\theta a} + \theta a - 1)}{\theta^3} + \frac{\nu^4(e^{-2\theta a} + 2\theta a - 1)}{4\theta^3} \\
&+ \frac{e^{2\kappa a} - 2e^{\kappa a} - 2e^{-2\kappa a} + 6e^{-\kappa a} + 2\kappa a - 3}{2\kappa^3}\int_{\mathbb{R}} Q^2G(\mathrm{d}x) + a^2\left(\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}\right)^2 \\
&+ 2a^2\left(\frac{\nu^2}{2\theta} + \mu^2\right)\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}.
\end{aligned}
\tag{3.26}
$$

We know that $\lim_{t\to\infty} Var(IV_{(t,t+a]}) = \lim_{t\to\infty}\mathbb{E}(IV^2_{(t,t+a]}) - \lim_{t\to\infty}\mathbb{E}(IV_{(t,t+a]})^2$, which shows (3.21).

As for the covariance of IV, we have,

$$
\lim_{t\to\infty} Cov(IV_{(t,t+a]}, IV_{(t+h,t+h+a]}) = \mathbb{E}(IV_{(t,t+a]}IV_{(t+h,t+h+a]}) - \mathbb{E}(IV_{(t,t+a]})^2,
$$

for $h = ia,\ i \in \mathbb{N}^+$. We obtain,

$$
\begin{aligned}
\lim_{t\to\infty}\mathbb{E}(IV_{(t,t+a]}IV_{(t+h,t+h+a]}) &= \lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a} V_c(\tau)\mathrm{d}\tau\int_{t+h}^{t+h+a} V_c(\tau)\mathrm{d}\tau\Big) \\
&+ \lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a} V_d(\tau)\mathrm{d}\tau\int_{t+h}^{t+h+a} V_d(\tau)\mathrm{d}\tau\Big) + \lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a} V_c(\tau)\mathrm{d}\tau\int_{t+h}^{t+h+a} V_d(\tau)\mathrm{d}\tau\Big) \\
&+ \lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a} V_d(\tau)\mathrm{d}\tau\int_{t+h}^{t+h+a} V_c(\tau)\mathrm{d}\tau\Big),
\end{aligned}
$$

[3] See Proposition 8.7 in Cont & Tankov (2003) for the isometry formula for general martingales.

where the third term $\lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a} V_c(\tau)\mathrm{d}\tau \int_{t+h}^{t+h+a} V_d(\tau)\mathrm{d}\tau\Big)$ can be expressed as

$$\begin{aligned}
&\lim_{t\to\infty}\mathbb{E}\left(\int_t^{t+a} V_c(\tau)\mathrm{d}\tau \int_{t+h}^{t+h+a} V_d(\tau)\mathrm{d}\tau\right)\\
&\quad= \lim_{t\to\infty}\mathbb{E}\left(\int_t^{t+a} V_c(\tau)\mathrm{d}\tau\left(\int_t^{t+a}\int_u^{t+a} e^{\kappa(u-\tau)}\,\mathrm{d}\tau\int_{\mathbb{R}} Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x)\right.\right.\\
&\quad\left.\left.+\int_{-\infty}^{t}\int_t^{t+a} e^{\kappa(u-\tau)}\mathrm{d}\tau\int_{\mathbb{R}} Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x)+a\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}\right)\right)\\
&\quad= a\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}\lim_{t\to\infty}\mathbb{E}\left(\int_t^{t+a} V_c(\tau)\mathrm{d}\tau\right)\\
&\quad= a^2\left(\frac{\nu^2}{2\theta}+\mu^2\right)\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa},
\end{aligned}\tag{3.27}$$

using the independence of the continuous and discontinuous parts of the integrated variance. The fourth term $\lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a} V_d(\tau)\mathrm{d}\tau \int_{t+h}^{t+h+a} V_c(\tau)\mathrm{d}\tau\Big)$ also equals to $a^2\left(\frac{\nu^2}{2\theta}+\mu^2\right)\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}$.

For other terms we have,

$$\begin{aligned}
&\lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a} V_c(\tau)\mathrm{d}\tau \int_{t+h}^{t+h+a} V_c(\tau)\mathrm{d}\tau\Big)\\
&\quad= \lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a} V_c(\tau)\mathrm{d}\tau\mathbb{E}\big(\int_{t+h}^{t+h+a} V_c(\tau)\mathrm{d}\tau|\mathcal{F}_{t+a}\big)\Big)\\
&\quad= \left(\left(a-\frac{e^{-2\theta h}-e^{-2\theta(h-a)}}{-2\theta}\right)\frac{\nu^2-2\theta\mu^2}{2\theta}+2\mu^2 a\right)\lim_{t\to\infty}\mathbb{E}\left(\int_t^{t+a}\sigma_c^2(\tau)\mathrm{d}\tau\right)\\
&\quad+\frac{e^{-2\theta h}-e^{-2\theta(h-a)}}{-2\theta}\lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a}\sigma_c^2(\tau)\sigma_c^2(t+a)\mathrm{d}\tau\Big)+2\mu\Big(\frac{e^{-\theta h}-e^{-\theta(h-a)}}{-\theta}\\
&\quad-\frac{e^{-2\theta h}-e^{-2\theta(h-a)}}{-2\theta}\Big)\lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a}\sigma_c^2(\tau)\sigma_c(t+a)\mathrm{d}\tau\Big).
\end{aligned}$$

Using the *law of iterated expectations* jointly with equation (3.16) we get,

$$\begin{aligned}
\lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a}\sigma_c^2(\tau)\sigma_c^2(t+a)\mathrm{d}\tau\Big) &= \lim_{t\to\infty}\int_t^{t+a}\mathbb{E}\Big(\sigma_c^2(\tau)\mathbb{E}\big(\sigma_c^2(t+a)|\mathcal{F}_\tau\big)\Big)\mathrm{d}\tau\\
&= a\left(\frac{\nu^2}{2\theta}+\mu^2\right)^2+\frac{2\mu^2\nu^2}{\theta^2}\left(1-e^{-\theta a}\right)+\frac{\nu^4}{4\theta^3}\left(1-e^{-2\theta a}\right),
\end{aligned}$$

$$\begin{aligned}
\lim_{t\to\infty}\mathbb{E}\Big(\sigma_c^2(\tau)\sigma_c(t+a)\mathrm{d}\tau\Big) &= \lim_{t\to\infty}\int_t^{t+a}\mathbb{E}\Big(\sigma_c^2(\tau)\mathbb{E}\big(\sigma_c(t+a)|\mathcal{F}_\tau\big)\Big)\mathrm{d}\tau\\
&= \frac{1-e^{-\theta a}}{\theta}\left(\mu^3+\frac{3\mu\nu^2}{2\theta}\right)+\mu\left(a-\frac{1-e^{-\theta a}}{\theta}\right)\left(\frac{\nu^2}{2\theta}+\mu^2\right).
\end{aligned}$$

We already know that the integrated continuous variance satisfies $\lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a}\sigma_c^2(\tau)\mathrm{d}\tau\Big) =$

$a\left(\frac{\nu^2}{2\theta}+\mu^2\right)$. Therefore, we have,

$$\begin{aligned}
&\lim_{t\to\infty}\mathbb{E}\Big(\int_t^{t+a}V_c(\tau)\mathrm{d}\tau\int_{t+h}^{t+h+a}V_c(\tau)\mathrm{d}\tau\Big)\\
&=a\left(\frac{\nu^2}{2\theta}+\mu^2\right)\Bigg(\frac{(\nu^2-2\theta\nu^2)(e^{-2\theta h}-e^{-2\theta(h-a)})}{4\theta^2}\\
&+\frac{(\nu^2+2\theta\mu^2)a}{2\theta}\Bigg)+\Bigg(a\left(\frac{\nu^2}{2\theta}+\mu^2\right)^2+\frac{2\mu^2\nu^2(1-e^{-\theta a})}{\theta^2}+\frac{\nu^4(1-e^{-2\theta a})}{4\theta^3}\Bigg)\\
&\cdot\left(\frac{e^{-2\theta(h-a)}-e^{-\theta h}}{2\theta}\right)+\Bigg(\frac{1-e^{-\theta a}}{\theta}\left(\mu^3+\frac{3\mu\nu^2}{2\theta}\right)+\mu\left(a-\frac{1-e^{-\theta a}}{\theta}\right)\Bigg(\frac{\nu^2}{2\theta}\\
&+\mu^2\Bigg)\Bigg)\cdot\left(\frac{e^{-\theta(h-a)}-e^{-\theta h}}{\theta}-\frac{e^{-2\theta(h-a)}-e^{-2\theta h}}{2\theta}\right)2\mu,
\end{aligned}\tag{3.28}$$

$$\begin{aligned}
&\mathbb{E}\Big(\int_t^{t+a}V_d(\tau)\mathrm{d}\tau\int_{t+h}^{t+h+a}V_d(\tau)\mathrm{d}\tau\Big)=\mathbb{E}\Bigg(\Bigg(\int_t^{t+a}\int_u^{t+a}e^{\kappa(u-\tau)}\mathrm{d}\tau\int_{\mathbb{R}}Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x)\\
&+\int_{-\infty}^{t}\int_t^{t+a}e^{\kappa(u-\tau)}\mathrm{d}\tau\int_{\mathbb{R}}Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x)+a\frac{\int_{\mathbb{R}}QG(\mathrm{d}x)}{\kappa}\Bigg)\\
&\cdot\Bigg(\int_{t+h}^{t+h+a}\int_u^{t+h+a}e^{\kappa(u-\tau)}\mathrm{d}\tau\int_{\mathbb{R}}Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x)\\
&+\int_{-\infty}^{t+h}\int_{t+h}^{t+h+a}e^{\kappa(u-\tau)}\mathrm{d}\tau\int_{\mathbb{R}}Q\tilde{\mu}(\mathrm{d}u,\mathrm{d}x)+a\frac{\int_{\mathbb{R}}QG(\mathrm{d}x)}{\kappa}\Bigg)\Bigg)\\
&=\mathbb{E}\Big(\int_{-\infty}^{t}\Big(\int_t^{t+a}e^{\kappa(u-\tau)}\mathrm{d}\tau\int_{t+h}^{t+h+a}e^{\kappa(u-\tau)}\mathrm{d}\tau\Big)Q^2\mu(\mathrm{d}u,\mathrm{d}x)\Big)\\
&+a^2\left(\frac{\int_{\mathbb{R}}QG(\mathrm{d}x)}{\kappa}\right)^2\\
&=\frac{(e^{\kappa a}-1)e^{-\kappa(2a+h)}}{2\kappa^3}\int_{\mathbb{R}}Q^2G(\mathrm{d}x)+a^2\left(\frac{\int_{\mathbb{R}}QG(\mathrm{d}x)}{\kappa}\right)^2.
\end{aligned}\tag{3.29}$$

Thus with equation (3.27), (3.28) and (3.29) we have,

$$\begin{aligned}
&\lim_{t\to\infty}\mathbb{E}(IV_{t,t+a}IV_{t+h,t+h+a})=a\left(\frac{\nu^2}{2\theta}+\mu^2\right)\Bigg(\frac{(\nu^2-2\theta\nu^2)(e^{-2\theta h}-e^{-2\theta(h-a)})}{4\theta^2}\\
&+\frac{(\nu^2+2\theta\mu^2)a}{2\theta}\Bigg)+\Bigg(a\left(\frac{\nu^2}{2\theta}+\mu^2\right)^2+\frac{2\mu^2\nu^2(1-e^{-\theta a})}{\theta^2}+\frac{\nu^4(1-e^{-2\theta a})}{4\theta^3}\Bigg)\\
&\cdot\left(\frac{e^{-2\theta(h-a)}-e^{-\theta h}}{2\theta}\right)+\Bigg(\frac{1-e^{-\theta a}}{\theta}(\mu^3+\frac{3\mu\nu^2}{2\theta})+\mu\left(a-\frac{1-e^{-\theta a}}{\theta}\right)\left(\frac{\nu^2}{2\theta}+\mu^2\right)\Bigg)\\
&\cdot\left(\frac{e^{-\theta(h-a)}-e^{-\theta h}}{\theta}-\frac{e^{-2\theta(h-a)}-e^{-2\theta h}}{2\theta}\right)2\mu+\frac{(e^{\kappa a}-1)e^{-\kappa(2a+h)}}{2\kappa^3}\int_{\mathbb{R}}Q^2G(\mathrm{d}x)\\
&+a^2\left(\frac{\int_{\mathbb{R}}QG(\mathrm{d}x)}{\kappa}\right)^2+2a^2\left(\frac{\nu^2}{2\theta}+\mu^2\right)\frac{\int_{\mathbb{R}}QG(\mathrm{d}x)}{\kappa}.
\end{aligned}\tag{3.30}$$

With equation (3.30) we then obtain the covariance of IV (3.22). □

Theorem 3.3.2 (Moments of the Quadratic Variation)**.** *For the Quadratic Variation defined as* $QV_{(t,t+a]} = \int_t^{t+a} \sigma^2(s)\mathrm{d}s + \int_t^{t+a}\int_{\mathbb{R}} J^2\mu(\mathrm{d}s,\mathrm{d}x)$*, we find the following moments,*

$$\lim_{t\to\infty}\mathbb{E}(QV_{(t,t+a]}) = a\left(\frac{\nu^2}{2\theta}+\mu^2+\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}\right)+a\int_{\mathbb{R}} J^2G(\mathrm{d}x), \tag{3.31}$$

$$\begin{aligned}\lim_{t\to\infty} Var(QV_{(t,t+a]}) &= \frac{4\mu^2\nu^2(e^{-\theta a}+\theta a-1)}{\theta^3}+\frac{\nu^4(e^{-2\theta a}+2\theta a-1)}{4\theta^3}\\ &+\frac{e^{2\kappa a}-2e^{\kappa a}-2e^{-2\kappa a}+6e^{-\kappa a}+2\kappa a-3}{2\kappa^3}\int_{\mathbb{R}} Q^2G(\mathrm{d}x)\\ &+a\frac{2(\kappa a+e^{-\kappa a}-1)}{\kappa^2}\frac{\int_{\mathbb{R}} QJ^2G(\mathrm{d}x)}{\kappa}+a\int_{\mathbb{R}} J^4G(\mathrm{d}x).\end{aligned} \tag{3.32}$$

Proof. We omit the derivation of the mean of QV here as it is straightforward. As for the variance of QV, we start from deriving $\lim_{t\to\infty}\mathbb{E}(QV^2_{(t,t+a]})$:

$$\begin{aligned}\lim_{t\to\infty}\mathbb{E}(QV^2_{(t,t+a]}) &= \mathbb{E}\Big(\Big(\int_t^{t+a} V_c(\tau)\mathrm{d}\tau+\int_t^{t+a} V_d(\tau)\mathrm{d}\tau\\ &+\int_t^{t+a}\int_{\mathbb{R}} J^2\tilde{\mu}(\mathrm{d}u,\mathrm{d}x)+a\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}\Big)^2\Big)\\ &=\lim_{t\to\infty}\mathbb{E}(IV^2_{(t,t+a]})+\lim_{t\to\infty}2a^2\mathbb{E}(V_c(t))\int_{\mathbb{R}} J^2G(\mathrm{d}x)\\ &+a\frac{2(\kappa a+e^{-\kappa a}-1)}{\kappa^2}\int_{\mathbb{R}} QJ^2G(\mathrm{d}x)\\ &+2a^2\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}\int_{\mathbb{R}} J^2G(\mathrm{d}x)+a\int_{\mathbb{R}} J^4G(\mathrm{d}x)+a^2(\int_{\mathbb{R}} J^2G(\mathrm{d}x))^2\\ &=a^2(\mu^2+\frac{\nu^2}{2\theta})^2+\frac{4\mu^2\nu^2(e^{-\theta a}+\theta a-1)}{\theta^3}+\frac{\nu^4(e^{-2\theta a}+2\theta a-1)}{4\theta^3}\\ &+\frac{e^{2\kappa a}-2e^{\kappa a}-2e^{-2\kappa a}+6e^{-\kappa a}+2\kappa a-3}{2\kappa^3}\int_{\mathbb{R}} Q^2G(\mathrm{d}x)+a^2(\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa})^2\\ &+2a^2(\frac{\nu^2}{2\theta}+\mu^2)\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa}+2a^2(\frac{\nu^2}{2\theta}+\mu^2+\frac{\int_{\mathbb{R}} QG(\mathrm{d}x)}{\kappa})\int_{\mathbb{R}} J^2G(\mathrm{d}x)\\ &+a\frac{2(\kappa a+e^{-\kappa a}-1)}{\kappa^2}\int_{\mathbb{R}} QJ^2G(\mathrm{d}x)+a\int_{\mathbb{R}} J^4G(\mathrm{d}x)+a^2(\int_{\mathbb{R}} J^2G(\mathrm{d}x))^2.\end{aligned}$$

Therefore we can obtain equation (3.32) by $\lim_{t\to\infty} Var(QV_{(t,t+a]}) = \lim_{t\to\infty}\mathbb{E}(QV^2_{(t,t+a]}) - \lim_{t\to\infty}\mathbb{E}(QV_{(t,t+a]})^2$. □

Note that in Theorem 3.3.1 and Theorem 3.3.2 we express the moments of jumps in the form of cumulants, as we do not impose additional assumptions on jumps. The cumulants can be easily transferred into analytic forms when specifying the distribution of jumps.

3.4 Details of Estimation

3.4.1 Theoretic Foundations

In this section we introduce some general theory on asymptotic properties of power and multipower variations, as well as techniques for disentangling jumps from observed asset returns. Our methodology is based on using moment conditions for power variations and multipower variations to obtain a general method of moments (GMM) estimator for high frequency data.

The discussion below is valid for a return process which can be represented via an arbitrary semimartingale X_t of type $X_t = X_0 + \int_0^t \sigma_s \mathrm{d}W_s + Y_t$, where the volatility process σ_t ($\sigma > 0$) is adapted to the same filtration $\mathcal{F}_t$, while W_t is a Brownian Motion and Y_t is a pure jump process. We let Δ_n denote the width of the sampling interval, and the process X_t is observed at equally spaced times $i\Delta_n$ for $i = 0, 1, ..., \lfloor T/\Delta_n \rfloor$. We denote the observed return as,

$$\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n},$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x for any real number x. We use $n = \lfloor 1/\Delta_n \rfloor$ to represent the observation frequency during a time interval of unit length. In conclusion there will be nT observations during the period $[0, T]$.

Power Variations

For any $p > 0$, we define the *realized power variation* of order p for X_t as

$$B(p, \Delta_n)_t = \sum_{i=1}^{n} |\Delta_i^n X|^p. \tag{3.33}$$

This corresponds to daily realized variance (hereafter RV) for $p = 2$ and one observation period is one day. This has been extensively exploited in financial econometrics, compare T. G. Andersen et al. (2009). We have

$$RV_t = B(2, \Delta_n) = \sum_{i=1}^{n} |\Delta_i^n X|^2,$$

and for $\Delta_n \to 0$ (the observation frequency n goes to infinity), RV converges to *realized quadratic variation* in probability (see Barndorff-Nielsen & Shephard (2003b)), i.e.

$$RV_t \xrightarrow{P} [X, X]_t, \tag{3.34}$$

where $[X, X]_t$ is the *quadratic variation* (hereafter QV) process of X_t (in this case $[X, X]_t = \int_0^t \sigma_s^2 \mathrm{d}s + \sum_{0<s\leqslant t} Y_t^2$). This result is crucial in order to match moment conditions when implementing the general method of moments (GMM). One may refer to Jacod & Protter (2012) for the related central limit theorem (CLM) for the convergence rate and the limiting distribution, in order to make inference about the estimation statistics.

Another special case is when $p = 4$, the *realized fourth variation* (hereafter FV) is able to eliminate the continuous part of returns, and it converges to the sum of jumps raised to power four during one single period[4], i.e.

$$FV_t = B(4, \Delta_n) = \sum_{i=1}^{n} |\Delta_i^n X|^4 \xrightarrow{P} \sum_{0<s\leqslant t} Y_t^4. \tag{3.35}$$

Bipower Variation

There are several ways to distangle the various sources from QV, one common and jump-robust estimator for the *integrated variance* (IV) is multipower variations

$$IV_t = \int_0^t \sigma_s \mathrm{d}s.$$

Following Barndorff-Nielsen & Shephard (2004), we define the *realized bipower variation* (BV) as

$$BV_t = \frac{\pi}{2} \sum_{i=2}^{n} |\Delta_i^n X||\Delta_{i-1}^n X|. \tag{3.36}$$

By defining BV, we are able to express RV as the sum of BV and *jump variation* JV.

3.4.2 GMM Estimator

We construct a GMM type estimator by using moment conditions of the power variations and the bipower varations. In our Linear quadratic volatility with co-jumps model, the parameter

[4] Todorov (2009b) and Todorov (2011) also used this as one of moment conditions.

set is given as

$$\xi = \left(\theta, \mu, \nu, \kappa, \int_{\mathbb{R}} J^2 G(\mathrm{d}x), \int_{\mathbb{R}} J^4 G(\mathrm{d}x), \int_{\mathbb{R}} QG(\mathrm{d}x), \int_{\mathbb{R}} Q^2 G(\mathrm{d}x), \int_{\mathbb{R}} QJ^2 G(\mathrm{d}x)\right)'.$$

We employ the following assumptions in order to guarantee the existence and consistency of the asymptotic distribution of our estimator.

Assumption 3.4.1. (i) *$\alpha(t)$ is a locally bounded predictable process, and $\int_{\mathbb{R}} J^2 G(\mathrm{d}x) < \infty$.*
(ii) *$\theta > 0$, $\mu > 0$, and $\nu > 0$.*
(iii) *$\kappa > 0$, $\int_{\mathbb{R}} QG(\mathrm{d}x) < \infty$ and $\int_{\mathbb{R}} Q^2 G(\mathrm{d}x) < \infty$.*
(iv) *$\int_{\mathbb{R}} QJ^2 G(\mathrm{d}x) < \infty$ and $\int_{\mathbb{R}} J^4 G(\mathrm{d}x) < \infty$.*

Parts (i) and (iii) of the assumption guarantee the local existence of the quadratic variation process of the underlying return process in (3.10), which is obvious. Part (ii) guarantees the stationarity of the OU process $\sigma_c(t)$[5], and Part (iii) implies that the jump part of the volatility $V_d(t)$ is weakly stationary and square-integrable. Part (iv) is used to guarantee the existence of the second moment of the QV (see equation (3.32)).

In the following we use the notion of an infeasible estimator as introduced in Todorov (2009b).

Assumption 3.4.2. (i) *The infeasible estimator $\hat{\xi}_{nR}$ of ξ for the return process $R(t)$ is defined as*

$$\hat{\xi}_{nR} = argmin\ g(\xi)'\hat{W}g(\xi),$$

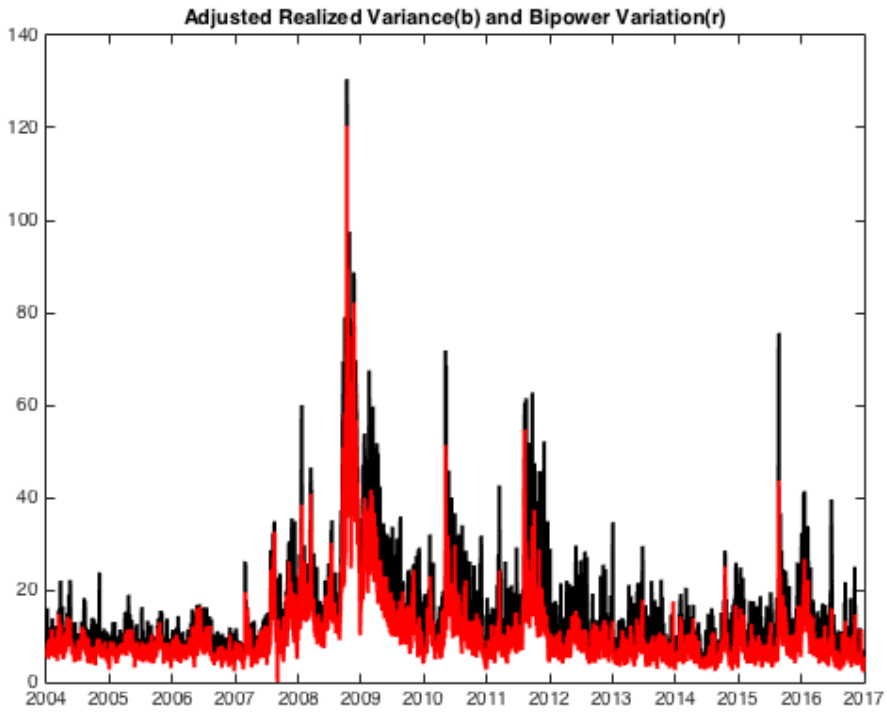
where $g(\xi)$ is the mean of sample moments, and $\hat{W}$ converges in probability to the positive semi-definite variance-covariance matrix W. The infeasible elements of the data set are IV and QV, since they are unobservable.
(ii) *Replacing QV with RV and IV with BV in the data set, the feasible estimator $\hat{\xi}_R$ of ξ for the return process $R(t)$ is defined as*

$$\hat{\xi}_R = argmin\ g(\xi)'\hat{W}g(\xi).$$

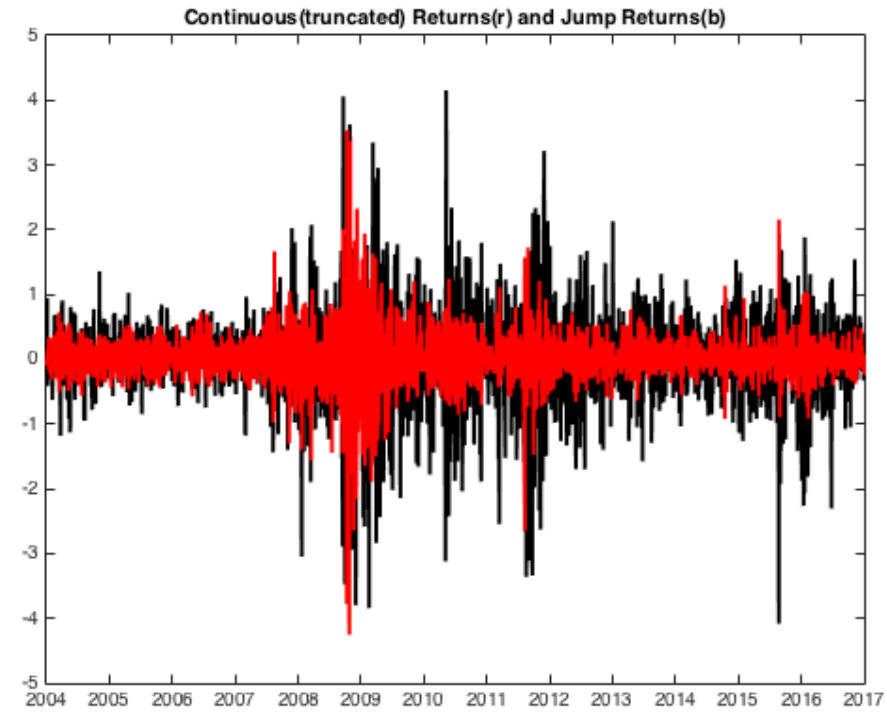
(iii) *Both estimators $\hat{\xi}_{nR}$ and $\hat{\xi}_R$ are consistent and asymptotically normal.*

[5] Compare Doob (1942)

Figure 3.1: The Power&Bipower Variations and Truncated Returns



The daily adjusted Realized Variance (black) and the Bipower Variation (red) (The Realized Variance is adjusted to be no less than the Bipower Variation)



Continuous returns (red) and jump returns (black) disentangled by truncation method ($\alpha = 4.5$)

Note that the assumptions here are in fact close to those in Todorov (2009b). The reason

for infeasibility is that IV and QV are not observable, so econometricians use proxies for them, here we use RV for QV and BV for IV. In the robustness test we also use the realized tripower variation (RTV) as a proxy for IV, for which the performance has been tested by Mancini (2001).

As for the exact moment conditions we have used for matching, these are

1. Mean and Variance of QV
2. Mean, Variance and Auto-correlation of IV
3. Mean of FV

With respect to the auto-correlation of IV, we use auto-correlation for lag 1, lag 2, lag 3, lag 5, lag 6, lag 7, lag 9, and the average of auto-correlation for lag 11-20, lag 21-30 and lag 31-40. Overall we have 15 moments, hence we have 6 degrees of freedom for the Linear Quadratic Stochastic Volatility with Co-Jumps model. Using more than 15 moments did not improve the result, while slowing down the computational process.

Regarding the truncation level of the discrete return observations, we use $\alpha = 4.5$ to truncate the return process. The RV, BV and truncated returns of the sample are presented in Figure 3.1. As for the optimal weighting matrix, we use a Heteroskedasticity and Autocorrelation Consistent (HAC) covariance matrix W, specifying a Parzen kernel[6] with a lag length of 80.

3.4.3 Affine Jump Diffusion Models

We assess our model against the conventional affine jump diffusion model, also referred to as the Bates model:

$$\mathrm{d}R(t) = \alpha(t)\mathrm{d}t + \sqrt{V(t)}\mathrm{d}W_1(t) + \int_{\mathbb{R}} J\tilde{\mu}(\mathrm{d}t, \mathrm{d}x), \tag{3.37}$$

$$\mathrm{d}V(t) = \kappa(\theta - V(t))\mathrm{d}t + \sigma\sqrt{V(t)}\mathrm{d}W_2(t). \tag{3.38}$$

Notice, that this model does not feature jumps in the volatility process. However we make similar assumptions as before on the square integrability of jumps in the return process,

[6] See Chapter 6 of Gallant (2009)

in order to guarantee the existence of moment conditions. The moment conditions for the general affine jump diffusion model have been provided in Todorov (2011). Additionally, we apply the reparameterizations $\sigma_v = \sigma\sqrt{\frac{\theta}{2\kappa}}$ to avoid identification problems. The variance follows a CIR process, hence we impose the Feller condition $\sigma_v < \theta$ and $\kappa > 0$ to guarantee stationarity and positivity.

3.4.4 Data and Estimation Outputs

In our empirical analysis, we use 5-min returns from the S&P 500 index from Jan 1, 2004 to Dec 31, 2016, acquired from Tick Data Database. Excluding weekends, holidays and non-trading days, we have 3250 days of trading data in total. Due to the limitation of the database, we only use high frequency data from 09.00 - 15.00 each day, which consists of 73 5-min observations in total.[7] Estimation results for different models are reported in Table 3.1 and Table 3.2.

Panel A in Table 3.1 shows that our linear quadratic volatility with jumps model works well. Evidence for this is the significance of parameters and the small J-statistics in the overidentification test. Moreover, the two (arguably) most important parameters, those that reflect the two mean-reverting speeds in our model, indicate two different half-life periods, i.e. the discontinuous part of the volatility dies out much more quickly than the continuous one. This interesting result is in line with Todorov (2009a). Another interesting observation is that there is evidence for non-zero correlation between the jump sizes in the return and volatility processes from the estimation results of cumulants in the same panel.

On the other hand, the Bates model (which exlcudes jumps in the volatility process) gets rejected by the J-test, as shown in Panel A of Table 3.1. This highlights the importance of incorporating jumps in the volatility process. Table 3.2 provides further evidence of the superiority of our model as compared with the traditional Bates model.

In the following discussion we estimate our model by using different proxies for integrated volatility, in order to verify the robustness of our results. We use the Realized

[7] The theory of consistency of estimating realized variance hinges on the finer sampled observation over an interval. Yet the market microstructure frictions might bring the estimator more noise and hence less accurate. Here we follow the ad hoc approach and choose the 5-min sampling frequency. See more discussions in Aït-sahalia et al. (2005) and Bollerslev et al. (2011).

Table 3.1: Estimation for Linear Quadratic Volatility with Jump Diffusion Models

$$\mathrm{d}R(t) = \alpha(t)\mathrm{d}t + \sigma(t-)\mathrm{d}W_1(t) + \int_{\mathbb{R}} J\tilde{\mu}(\mathrm{d}t, \mathrm{d}x)$$
$$\sigma^2(t) = V_c(t) + V_d(t)$$
$$V_c(t) = \sigma_c^2(t), \text{ where } \mathrm{d}\sigma_c(t) = \theta(\mu - \sigma_c(t))\mathrm{d}t + \nu \mathrm{d}W_2(t)$$
$$dV_d(t) = -\kappa V_d(t)\mathrm{d}t + \int_{\mathbb{R}} Q\mu(\mathrm{d}t, \mathrm{d}x)$$

Panel A: The LQJD Model (By (Multi)Power Variations)		
	With Price and Volatility Jumps	With Only Price Jumps
θ	0.7032 (0.0661)***	0.445 (0.042)***
μ	5.181e-3 (2.998e-3)*	6.573e-3 (1.358e-3)***
ν	6.289e-3 (1.79e-3)***	5.202e-3 (1.103e-3)***
κ	2.257 (1.57e-3)***	
$\int_{\mathbb{R}} J^2 G(\mathrm{d}x)$	2e-6 (3e-6)	6e-6 (3.5e-5)
$\int_{\mathbb{R}} J^4 G(\mathrm{d}x)$	6.982e-3 (1e-6)***	0.102 (5.96e-3)
$\int_{\mathbb{R}} QG(\mathrm{d}x)$	1.03e-4 (2.9e-5)***	
$\int_{\mathbb{R}} Q^2 G(\mathrm{d}x)$	2.69e-3 (4e-6)***	
$\int_{\mathbb{R}} QJ^2 G(\mathrm{d}x)$	0.016 (1e-5)***	
Overidentification Test		
Test Statistics	0.0031	25.4494
Degree of Freedom	6	8
P-Value	1.0000	0.0013

Panel B: The LQJD Model (Robustness Tests)		
	By Tripower Variation	By Truncated Variation
θ	0.7986 (0.0518)***	0.5741 (0.1014)***
μ	4.126e-3 (1.247e-3)***	3.251e-3 (1.086e-3)***
ν	7.489e-3 (1.146e-3)***	6.816e-3 (1.493e-3)***
κ	2.3815 (1.85e-4)***	2.3539 (0.0392)***
$\int_{\mathbb{R}} J^2 G(\mathrm{d}x)$	2e-6 (9e-6)	4.5e-5 (3e-6)***
$\int_{\mathbb{R}} J^4 G(\mathrm{d}x)$	7.457e-3 (1e-6)***	1.905e-3 (1e-6)***
$\int_{\mathbb{R}} QG(\mathrm{d}x)$	1e-4 (2.2e-5)***	1.08e-4 (1.2e-5)***
$\int_{\mathbb{R}} Q^2 G(\mathrm{d}x)$	4.025e-3 (1e-6)***	2.7e-5 (1e-6)***
$\int_{\mathbb{R}} QJ^2 G(\mathrm{d}x)$	0.0179 (2e-6)***	1.194e-3 (4.6e-5)***
Overidentification Test		
Test Statistics	0.0061	0.0090
Degree of Freedom	6	6
P-Value	1.0000	1.0000

Note: Standard errors in parentheses. [***] Significant at 1% level; [**] Significant at 5% level; [*] Significant at 10% level. In Sagan's J test, the null hypothesis is that the over-identifying restrictions are valid.

Table 3.2: Estimation for the Affine Jump Diffusion Model

$$\mathrm{d}R(t) = \alpha(t)\mathrm{d}t + \sqrt{V(t)}\mathrm{d}W_1(t) + \int_{\mathbb{R}} J\tilde{\mu}(\mathrm{d}t, \mathrm{d}x)$$

$$\mathrm{d}V(t) = \kappa(\theta - V(t))\mathrm{d}t + \sigma\sqrt{V(t)}\mathrm{d}W_2(t)$$

Affine Jump Diffusion Model	
	With Only Price Jumps
θ	0.0346 (2.075e-3)***
κ	0.3541 (0.0292)***
σ	0.0346 (2.288e-3)***
$\int_{\mathbb{R}} J^2 G(\mathrm{d}x)$	0.0216 (1.5e-3)***
$\int_{\mathbb{R}} J^4 G(\mathrm{d}x)$	2.462e-3 (4.32e-4)***
Overidentification Test	
Test Statistics	699.6899
Degree of Freedom	8
P-Value	0.0000

Tripower Variation (hereafter RTV) and Truncated Variation (TV) to replace the BV used in the previous part. These were also used in Todorov (2009b) and Mancini (2009) correspondingly. The estimation results obtained from using these two proxies are presented in panel B of Table 3.1. We do not discuss these further as they are similar to the previous part. The RTV and TV are defined as follows.

$$RTV_t = \mu_{2/3}^{-3} \sum_{i=3}^{n} |\Delta_i^n X|^{2/3} |\Delta_{i-1}^n X|^{2/3} |\Delta_{i-2}^n X|^{2/3},$$

$$TV_t = \sum_{i=1}^{n} |r_{t,i}^c|^2,$$

where $\mu_a = \mathbb{E}(|u|^a)$ and $u \sim \mathcal{N}(0,1)$, and $r_{t,i}^c$ is the continuous part of the discretely observed return. By defining the discretely observed return,

$$r_{t,i} = \Delta_{(t-1)n+i}^n X, \; i = 1, 2, ..., n,$$

and a truncation threshold

$$\mathrm{CUT}_{t,i} = \alpha \Delta_n^{0.49} \sqrt{\tau_i BV_{t-1}},$$

$r_{t,i}^c$ is denoted as

$$r_{t,i}^c = r_{t,i} \mathbb{1}_{|r_{t,i}| \leqslant \mathrm{CUT}_{t,i}},$$

Table 3.3: Estimation for Linear Quadratic Volatility Model with Normal Jumps

$$\mathrm{d}R(t) = \alpha(t)\mathrm{d}t + \sigma(t-)\mathrm{d}W_1(t) + \int_{\mathbb{R}} J\tilde{\mu}(\mathrm{d}t, \mathrm{d}x)$$
$$\sigma^2(t) = V_c(t) + V_d(t)$$
$$V_c(t) = \sigma_c^2(t), \text{ where } \mathrm{d}\sigma_c(t) = \theta(\mu - \sigma_c(t))\mathrm{d}t + \nu \mathrm{d}W_2(t)$$
$$\mathrm{d}V_d(t) = -\kappa V_d(t)\mathrm{d}t + \int_{\mathbb{R}} Q\mu(\mathrm{d}t, \mathrm{d}x)$$

With Normal Jumps in Price and Volatility	
θ	0.6145 (0.2215)***
μ	0.0003 (0.0006)
ν	0.0060 (0.0010)***
κ	0.0216 (1.5e-3)
λ	4e-5 (101.377)
μ_J	0.5619 (96.454)
σ_J	0.5913 (63.715)
μ_Q	0.3920 (90.203)
σ_Q	0.5406 (21.047)
ρ	0.0110 (10.487)
Overidentification Test	
Test Statistics	4.3168
Degree of Freedom	5
P-Value	0.5048

where $\mathbb{1}_E(x)$ is an indicator function, and α is suggested to lie in the interval $[3.5, 4.5]$.(Aït-Sahalia & Jacod, 2009a; T. G. Andersen et al., 2011)

Finally, we run our methodology through a fully parametrized version of our model. We follow the conventional assumption that jump sizes in both return and volatility process are normally distributed ($J(x) \sim \mathcal{N}(\mu_J, \sigma_J^2)$ and $Q(x) \sim \mathcal{N}(\mu_Q, \sigma_Q^2)$), and further that these two normal distribution may be correlated with correlation coefficient ρ. We observe that the parameters that relate to these distributions become non-significant, although the J-test does not reject the fully parameterized model, see 3.3. In conclusion, we want to know whether the distribution assumption affects our result. If not, the assumption of normally distributed jump-sizes for both return and volatility maybe fundamentally flawed, as the previous results clearly indicate the existence of jumps. To follow up on this analysis, we provide a the following Monte Carlo experiment.

$$\mathrm{d}R(t) = \alpha(t)\mathrm{d}t + \sigma(t-)\mathrm{d}W_1(t) + \int_{\mathbb{R}} J\tilde{\mu}(\mathrm{d}t, \mathrm{d}x)$$

$$\sigma^2(t) = V_c(t) + V_d(t)$$

$$V_c(t) = \sigma_c^2(t), \text{ where } \mathrm{d}\sigma_c(t) = \theta(\mu - \sigma_c(t))\mathrm{d}t + \nu \mathrm{d}W_2(t)$$

$$\mathrm{d}V_d(t) = -\kappa V_d(t)\mathrm{d}t + \int_{\mathbb{R}} Q\mu(\mathrm{d}t, \mathrm{d}x)$$

Table 3.4: Details for Monte Carlo Simulations

Case	Parameters									
	θ	μ	ν	κ	λ	μ_J	σ_J	μ_Q	σ_Q	ρ
low frequency	0.6	5e-3	6e-3	2.5	0.1	1e-4	3.2e-3	1e-3	0.01	-0.7
high frequency	0.6	5e-3	6e-3	2.5	0.3	1e-4	3.2e-3	1e-3	0.01	-0.7
large mean and variance	0.6	5e-3	6e-3	2.5	0.1	5e-4	0.01	5e-3	0.02	-0.7

Note: We simulate the jump process as a compound Poisson type here for the sake of parsimony, while other types are also feasible. In addition we simulate the jump sizes to be correlated normal distributions, and the number of iteration is 100 for each simulation.

3.4.5 Monte Carlo Experiment

In order to back up our result on the non-normality of jump sizes in our model, we conduct a small panel Monte Carlo experiment. The idea is to run a simulation based on our fully parameterized model, and estimate it again to see whether the coefficients of the jump distribution are significant or not. The simulated data replace the high frequency data from the S&P 500 previously used. Details of the data generating process are stated in Table 3.4. The discretization method used in the Monte Carlo simulation is Euler-Maruyama. Since the randomness of jump processes originates from two sources, i.e. the jump time and the jump size, we follow the method provided by Kloeden & Platen (2013) when simulating compound Poisson processes. We model the high frequency observations by determining that the data are observed in 1-min frequency, and the number of trading days is 1000 with 8 trading hours in each day. Again the RV and BV are summed up in 5-min frequency. The results are presented in Table 3.5, Table 3.6 and Table 3.7.

We summarize the results of our Monte Carlo experiement as follows:

- Our estimation based on the simulated data with distribution assumption shows satisfactory performance, which means that our estimator is efficient when the actual distribution of jump size is normal. Therefore, we can say that the assumption of normal distributed jump sizes is not feasible for our S&P 500 high frequency dataset.
- The similarity of results with and without distribution specifications, implies that our estimator is indeed consistent with parameters fully specified or with only cumulants estimated.
- The mean-reversion speed of the discontinuous part of volatility is slightly upward biased. That makes the stationary volatility to be downward biased, which was indicated by T. G. Andersen & Sørensen (1996) and Todorov (2011).

3.5 Conclusion

In this chapter, we construct a linear quadratic volatility with jumps model, in which the variance process is a superposition of two separate parts, i.e. the continuous part (a squared

Gaussian OU process) and the discontinuous part (a Lévy-driven OU process). We derive the analytic moment conditions for the quadratic variation and the integrated variance of our model and then calibrate our model by matching sample moments of realized power variations and realized bipower variations. Our results indicate that our model fits the market data well and shows superior performance compared with the conventional affine model, i.e. Bates model, which does not incorporate jumps in the volatility process. We also conduct two robustness tests, i.e. use other types of estimator (realized tripower variation and truncated variation) to consolidate our basic results.

While the distribution of both jump sizes of our model in the context are left unrestricted, i.e. we estimate them in cumulants, it is interesting that we get insignificant parameters when we specify the jump size distribution to be normal. That leads us to implement a Monte Carlo experiment in order to investigate the identification problem in the distribution of the jump sizes. The Monte Carlo experiments confirms our view that the traditional normal distribution may not be suitable for S&P 500 high frequency data.

Table 3.5: Monte Carlo Results with Low Jump Frequency

Parameter	True Value	Mean	RMSE	5th percentile	Median	95th percentile
Panel A: low frequency without distribution assumption						
θ	0.6	0.4995	0.1005	0.4989	0.4995	0.4999
μ	5e-3	4.893e-3	1.184e-4	4.807e-3	4.888e-3	4.981e-3
ν	6e-3	4.825e-3	1.18e-3	4.741e-3	4.82e-3	4.914e-3
κ	2.5	2.469	0.0314	2.465	2.469	2.472
$\int_{\mathbb{R}} J^2 G(\mathrm{d}x)$	1e-6	9.9997e-7	3.060e-11	9.9996e-7	9.9997e-7	9.9997e-7
$\int_{\mathbb{R}} J^4 G(\mathrm{d}x)$	3e-11	2.5e-8	2.497e-8	2.5e-8	2.5e-8	2.5e-8
$\int_{\mathbb{R}} QG(\mathrm{d}x)$	1e-4	7.953e-5	2.045e-5	7.948e-5	7.955e-5	7.962e-5
$\int_{\mathbb{R}} Q^2 G(\mathrm{d}x)$	1.01e-5	1.546e-5	5.364e-6	1.504e-5	1.548e-5	1.586e-5
$\int_{\mathbb{R}} QJ^2 G(\mathrm{d}x)$	1.44e-9	1.5e-8	1.356e-8	1.5e-8	1.5e-8	1.5e-8
Panel B: low frequency with distribution assumption						
θ	0.6	0.50104	0.099	0.49991	0.50099	0.50206
μ	5e-3	4.933e-3	1.46e-4	4.463e-3	4.969e-3	4.984e-3
ν	6e-3	4.925e-3	1.083e-3	4.454e-3	4.959e-3	4.974e-3
κ	2.5	2.4992	3.084e-3	2.4886	2.4999	2.5004
λ	0.1	0.1239	0.0239	0.12169	0.12410	0.12413
μ_J	1e-4	2.5e-4	1.5e-4	2.49e-4	2.5e-4	2.5e-4
σ_J	3.16e-3	4.97e-3	1.81e-3	4.58e-3	4.997e-3	5.001e-3
μ_Q	1e-3	9.83e-4	3.3e-5	8.82e-4	9.91e-4	9.92e-4
σ_Q	0.01	9.96e-3	6.1e-5	9.91e-3	9.95e-3	1.01e-2
ρ	-0.7	-0.7513	0.0513	-0.7528	-0.7512	-0.7509

Table 3.6: Monte Carlo Results with High Jump Frequency

Parameter	True Value	Mean	RMSE	5th percentile	Median	95th percentile
Panel A: high frequency without distribution assumption						
θ	0.6	0.5444	0.056	0.5405	0.5415	0.5493
μ	5e-3	5.55e-3	8.52e-4	4.671e-3	6.023e-3	6.127e-3
ν	6e-3	5.385e-3	8.65e-4	4.652e-3	5.844e-3	5.943e-3
κ	2.5	2.492	0.029	2.458	2.514	2.519
$\int_{\mathbb{R}} J^2 G(\mathrm{d}x)$	3e-6	2.0003e-6	1e-6	2e-6	2.0004e-6	2.0004e-6
$\int_{\mathbb{R}} J^4 G(\mathrm{d}x)$	9e-11	2.5e-8	2.491e-8	2.5e-8	2.5e-8	2.5e-8
$\int_{\mathbb{R}} QG(\mathrm{d}x)$	3e-4	2.6878e-4	3.29e-5	2.5615e-4	2.7656e-4	2.7836e-4
$\int_{\mathbb{R}} Q^2 G(\mathrm{d}x)$	3.03e-5	3.2319e-5	5.24e-6	2.6331e-5	3.6203e-5	3.6543e-5
$\int_{\mathbb{R}} QJ^2 G(\mathrm{d}x)$	4.32e-9	2.5e-8	2.068e-8	2.5e-8	2.5e-8	2.5e-8
Panel B: high frequency with distribution assumption						
θ	0.6	0.501	0.099	0.4945	0.4949	0.5149
μ	5e-3	5.103e-3	1.04e-4	5.071e-3	5.105e-3	5.129e-3
ν	6e-3	5.088e-3	9.13e-4	5.054e-3	5.093e-3	5.111e-3
κ	2.5	2.5064	6.92e-3	2.5040	2.5051	2.5115
λ	0.3	0.2997	4.54e-4	0.2992	0.2996	0.3003
μ_J	1e-4	2.5e-4	1.5e-4	2.49e-4	2.5e-4	2.5e-4
σ_J	3.16e-3	5.007e-3	1.85e-3	5.001e-3	5.008e-3	5.011e-3
μ_Q	1e-3	1.041e-3	4.1e-5	1.035e-3	1.04e-3	1.048e-3
σ_Q	0.01	1.086e-2	6.3e-5	1.066e-2	1.089e-2	1.099e-2
ρ	-0.7	-0.7505	0.0505	-0.7509	-0.7506	-0.7501

Table 3.7: Monte Carlo Results with Large Mean and Variance

Parameter	True Value	Mean	RMSE	5th percentile	Median	95th percentile
Panel A: large mean and variance without distribution assumption						
θ	0.6	0.550898	0.0491	0.550768	0.550891	0.551026
μ	5e-3	4.303e-3	6.97e-4	4.267e-3	4.309e-3	4.337e-3
ν	6e-3	4.300e-3	1.7e-3	4.273e-3	4.300e-3	4.333e-3
κ	2.5	2.5001	3.73e-3	2.4943	2.4998	2.5073
$\int_{\mathbb{R}} J^2 G(\mathrm{d}x)$	1.0025e-5	1.4972e-5	4.95e-6	1.4970e-5	1.4972e-5	1.4973e-5
$\int_{\mathbb{R}} J^4 G(\mathrm{d}x)$	3.02e-9	2.5e-8	2.198e-8	2.5e-8	2.5e-8	2.5e-8
$\int_{\mathbb{R}} QG(\mathrm{d}x)$	5e-4	3.2415e-4	1.7587e-4	3.1964e-4	3.2434e-4	3.2846e-4
$\int_{\mathbb{R}} Q^2 G(\mathrm{d}x)$	4.25e-5	5.2358e-5	9.87e-6	5.1549e-5	5.2346e-5	5.3247e-5
$\int_{\mathbb{R}} QJ^2 G(\mathrm{d}x)$	6.388e-8	4e-8	2.39e-8	4e-8	4e-8	4e-8
Panel B: large mean and variance with distribution assumption						
θ	0.6	0.5061	0.096	0.4750	0.5028	0.5383
μ	0.005	4.863e-3	1.37e-4	4.851e-3	4.862e-3	4.872e-3
ν	0.006	4.861e-3	1.14e-3	4.841e-3	4.864e-3	4.872e-3
κ	2.5	2.4749	0.026	2.4576	2.4751	2.4862
λ	0.1	0.1159	0.016	0.1136	0.1156	0.1178
μ_J	5e-4	3.999e-4	1e-4	3.999e-4	3.999e-4	3.999e-4
σ_J	0.01	0.014769	4.77e-3	0.014553	0.014771	0.015024
μ_Q	5e-3	3.546e-3	1.45e-3	3.521e-3	3.546e-3	3.571e-3
σ_Q	0.02	0.01870	1.31e-3	0.01846	0.01867	0.01898
ρ	-0.7	-0.7512	0.0513	-0.7554	-0.7518	-0.7462

Chapter 4

American Real Options under Stochastic Volatility and Model Uncertainty

4.1 Introduction

Classic articles on derivatives pricing often assume that a pricing measure exists and specific events occur with a certain probability distribution under that measure. However, the probability distribution is not usually known a priori in reality, which is named uncertainty or ambiguity. Formally, Ellsberg (1961), following Knight (1921), defines the random variables with known probability distribution *certain*, and those with unknown probability distribution *uncertain*. The uncertain parameters of a model will be characterized by a family of probability measures, and the feature that the probability distribution is not fully fixed is called *ambiguity* here. We are interested in numerically evaluating financial American options and real life investments i.e. real options of American style under specific models that accommodate ambiguity, especially in multivariate settings.

Generally, the theory of model uncertainty tends to overthrow the assumption that the distribution of underlying asset prices is fixed, due to the uncertainty of the market. Thus, the orthodox theory of option pricing cannot be directly used either in financial markets or real life investment valuations and may need amendments. Most financial markets and real investments studies assumes the future markets or situations are characterized by a certain probability measures over states of nature. This amounts to that the agent or firm is definitely certain that future conditions are depicted by this special probability measure. However, this may be seriously biased and the agent may not so sure about the future uncertainty and may

think other probability measures, as revealed by the Ellsberg paradox. Hence, the model uncertainty approach seems unusual but rather realistic and necessary.

There are two main directions of theoretical settings for describing ambiguity and corresponding optimal strategies. One is to postulate a set of equivalent probability measures, which is known as multiple priors or drift ambiguity, given the existence of a dominating probability measure. Gilboa & Schmeidler (1989) axiomatize the agent's robust decision and introduce the maxmin method under drift ambiguity. It describes that the mother nature drives the agent to evaluate the payoff or utility function with the worst belief, and then seek for the optimal strategy to maximize it. The other direction is to assume the absence of a dominating probability measure, which means that all priors are not necessarily absolutely continuous to a reference measure. This leads to the capability of involving sources of uncertainty on the standard deviation of a distribution, i.e. volatility ambiguity. One can find the first attempts to attack the pricing problem in such a context in Avellaneda et al. (1995) and Lyons (1995). Specifically, this chapter considers valuation of American options under stochastic volatility model and optimal fish harvesting decision under stochastic convenience yield model, both within drift ambiguity framework.

The models used in this chapter are related to two main streams of previous literature. First, the feature of stochastic volatility is a generalization of the famous constant-volatility model in Black & Scholes (1973). As the stochastic volatility models are capable of better capturing the implied volatility smile and the leverage effect, there are substantial classes of models that explicitly parameterize the stochastic volatility, for example, models introduced by Stein & Stein (1991), Heston (1993), Bates (1996), etc. Nevertheless, all of them concentrate on European options. It is natural but challenging to extend to the American case, because there is no analytical solution for the prices of American options, even under the simplest Black-Scholes model. Second, the advantages of stochastic convenience yield[1] model over the Black-Scholes model in pricing commodity derivatives are initially indicated by Gibson & Schwartz (1990) and now well known (refer to Schwartz & Smith (2000), Ewald et al. (2019) and Moreno et al. (2019) for instance). It is widely applied to the field

[1] Gibson & Schwartz (1990) describe the convenience yield as "the flow of services accruing to the holder of the spot commodity but not to the owner of a futures contract".

of real options, which can be understood as an adaption of methods of traditional financial options to real life investments (see e.g. Dixit & Pindyck (1994)). Cortazar et al. (2008) numerically value the copper mine by using a stochastic convenience yield model. Ewald et al. (2016) investigate the optimal fish harvesting problem and evaluate the fish farm under stochastic convenience yield model numerically.

As for the specifications of drift ambiguity, we employ the κ-*ignorance* introduced by Chen & Epstein (2002) in one dimensional case, so that the density generator θ lies in an interval that centers at origin with radius κ. Nishimura & Ozaki (2007) and Trojanowska & Kort (2010) adopt such a setting in optimal investment decision in real options. They highlight that an increase of drift ambiguity leads to a lower value of investment, which is in sharp contrast of the effect of an increase in risk. X. Cheng & Riedel (2013) and Vorbrink (2011) address the optimal stopping problem of exotic options under the Black-Scholes model. For the multivariate case, we consider an ellipsoid shape of uncertainty set, which is initially proposed by Goldfarb & Iyengar (2003) on the portfolio optimization problems. Similar approaches have been taken by Cohen & Tegnér (2017) for European options pricing and Balter & Pelsser (2020) for hedging strategy. Note that the ellipsoid uncertaity set reduces to an ellipse in two dimensions. Thus, we name it the *elliptical ambiguity*. To the best of our knowledge, our work is the first attempt to valuing American options under stochastic volatility or real options of American style under stochastic convenience yield with drift ambiguity.

We investigate the formulation of the optimal value of American option and American real option, for which we take a single-rotation fish farm valuation as an example. Without ambiguity, it is well known that the value is essentially a supermartingale and has a dual representation of a reflected backward stochastic differential equation(RBSDE), according to El Karoui, Pardoux, & Quenez (1997) and El Karoui, Kapoudjian, Pardoux, Peng, & Quenez (1997). We link the multivariate case with the *elliptical ambiguity* to the solutions of RBSDEs and prove the existence and uniqueness, provided the foundations of one dimensional case by X. Cheng & Riedel (2013) and Vorbrink (2011).

The question remains how we numerically evaluate the optimization function given the correct formulation. We propose two possible algorithms to conduct numerical im-

plementations. The first one, Stratified Regression One-step Forward Dynamic Programming(SRODP), is constructed through the "Max method" for RBSDEs by Gobet & Lemor (2008), joint with a general stratification approach by Gobet et al. (2016). Algorithms for RBSDEs are based on the convergence results by Ma & Zhang (2005) and Bouchard & Chassagneux (2008). The general stratification approach is a sampling method to approximate the object conditional expectation function by local polynomials, which differs from the conventional Monte Carlo methods (see Tsitsiklis & Van Roy (2001), Longstaff & Schwartz (2001) and Glasserman (2013) to name a few) in not depending on the starting point and starting value. Gobet et al. (2016) implement it in multivariate European option case and underline its superiority in conserving computational memory and efficiency, especially when enabling GPU computing. The idea behind the general stratification approach also contributes to the second algorithm, Stratified Least Square Monte Carlo(SLSM), which is a combination of the classic dynamic programming principle and that approach.

We conduct numerical experiments in non-ambiguous cases to show the convergence of algorithms to the *exact value* of European and American options. In cases with ambiguity, results by the two proposed algorithms are close to the benchmark by Least Square Monte Carlo (LSM) algorithm(Longstaff & Schwartz, 2001) under Black-Scholes model. Moreover, it is observed that the American option value increases when shrinking the uncertainty interval. This is in accordance with theoretical analysis in X. Cheng & Riedel (2013) and Vorbrink (2011). However, we do not have an *exact value* under multivariate settings as we argue in Section 4.3.4. Results for values of American options in that case show that the SRODP algorithm has superior efficiency than the SLSM algorithm, which is confirmed by the fish farm valuation case. In spite of it, the SLSM algorithm still provides possible solution in scenarios when the optimal generator of the RBSDE cannot be solved explicitly.

The remainder of this chapter is constructed as follows. Section 4.2 formulate the American option value under Heston's model to the solution of an RBSDE, obtain the optimal driver within the *elliptical ambiguity* framework and prove the existence and uniqueness of the solution. Section 4.3 introduces the two algorithms. Section 4.4 presents numerical results of financial options. Section 4.5 demonstrates the application of theoretical results and algorithms in optimal fish harvesting problem. The last section concludes.

4.2 Theoretical Framework

4.2.1 American Option Pricing

We characterize the probability space by the triplet $(\Omega, \mathcal{F}, P)$, and let $(\mathcal{F}_t)_{0\leq t\leq T}$ be the augmented filtration generated by a standard two dimensional Brownian motion $W = (W^1, W^2)^*$. Without uncertainty in probability measures, the dynamics of the price process S_t and the variance process V_t in Heston (1993) under the objective measure P are denoted as,

$$\mathrm{d}S_t/S_t = \mu_t \mathrm{d}t + \sqrt{V_t}\big(\rho \mathrm{d}W_t^1 + \sqrt{1-\rho^2}\mathrm{d}W_t^2\big), \tag{4.1}$$

$$\mathrm{d}V_t = \alpha(\beta - V_t)\mathrm{d}t + \sigma\sqrt{V_t}\mathrm{d}W_t^1. \tag{4.2}$$

There exists a money market account γ with a risk-free rate r, and γ evolves according to,

$$\mathrm{d}\gamma_t = r\gamma_t \mathrm{d}t, \;\; \gamma_0 = 1. \tag{4.3}$$

The payoff function of an option is

$$H_t := H(t, X_t) = \Phi(X_t), \tag{4.4}$$

if we denotes $X_t = (S_t, V_t)^*$. To price an option, we need to choose an equivalent martingale measure, under which the discounted option price $J_t\gamma_t^{-1}$ is the smallest supermartingale dominating the discounted payoff.(see for example Chapter 5 of Pham (2009)) Denote a stopping time τ as the exercise time, the option price J_t is denoted as,

$$J_t := J(t, X_t) = \operatorname*{ess\,sup}_{\tau\in\mathscr{T}_t} \mathbb{E}^Q[H_\tau \gamma_{\tau-t}^{-1}|\mathcal{F}_t],$$

where $\mathscr{T}$ is the set of all stopping times dominated by T, and $\mathscr{T}_t = \{\tau \in \mathscr{T}; t \leq \tau \leq T\}$. The conditional Radon-Nikodym derivative required for the change of measure is defined as,

$$\frac{\mathrm{d}Q|_{\mathcal{F}_t}}{\mathrm{d}P|_{\mathcal{F}_t}} = \mathcal{E}\big(-\int_0^{\cdot} \lambda_s^* \mathrm{d}W_s\big)_t,$$

where the stochastic exponential $\mathcal{E}(\cdot)_t$ of $-\int_0^{\cdot} \lambda_s^* \mathrm{d}W_s$ is a P-martingale, given by,

$$\mathcal{E}\big(-\int_0^{\cdot} \lambda_s^* \mathrm{d}W_s\big)_t = \exp\bigg(-\int_0^t \lambda_s^* \mathrm{d}W_s - \frac{1}{2}\int_0^t \lambda_s^* \lambda_s \mathrm{d}s\bigg).$$

In addition, we still need the stochastic exponential to satisfy technical conditions such as progressively measurablity and the Novikov condition to have Q as an equivalent measure to P (see for example Chapter 3 of Karatzas & Shreve (1998)).

Therefore, the price process and the variance process under the risk-neutral measure Q are given as,

$$\begin{aligned} \mathrm{d}S_t/S_t &= r\mathrm{d}t + \sqrt{V_t}\big(\rho \mathrm{d}\tilde{W}_t^1 + \sqrt{1-\rho^2}\mathrm{d}\tilde{W}_t^2\big), \\ \mathrm{d}V_t &= \tilde{\alpha}(\tilde{\beta} - V_t)\mathrm{d}t + \sigma\sqrt{V_t}\mathrm{d}\tilde{W}_t^1. \end{aligned}$$

By the Girsanov theorem, a vector of market price of risk process $\lambda_t = (\lambda_t^1, \lambda_t^2)^*$ is given as,

$$\begin{aligned} \mathrm{d}\tilde{W}_t^1 &= \lambda_t^1 \mathrm{d}t + \mathrm{d}W_t^1, \\ \mathrm{d}\tilde{W}_t^2 &= \lambda_t^2 \mathrm{d}t + \mathrm{d}W_t^2, \\ \frac{\mu_t - r}{\sqrt{V_t}} &= \rho\lambda_t^1 + \sqrt{1-\rho^2}\lambda_t^2. \end{aligned}$$

Note that here the Feller condition should be satisfied to guarantee the strict positivity of variance, that is, $2\alpha\beta > \sigma^2$. In order to keep the variance process within the affine class under Q measure, we employ an essential affine structure of market price specification $\lambda_t^1 = a\sqrt{V_t} + b$ for some constant a, which is introduced by Duffee (2002). Thus, we have,

$$\begin{aligned} \tilde{\alpha} &= \alpha + a\sigma, \\ \tilde{\beta} &= \frac{\alpha\beta - b\sigma}{\alpha + a\sigma}. \end{aligned}$$

In the Markovian framework of the underlying processes, the American option price J_t can be viewed as a unique viscosity solution to the following Hamilton-Jacobi-Bellman (hereafter HJB) variational inequality:

$$\begin{aligned} \min\left\{rJ(t,x) - \frac{\partial J}{\partial t} - \mathcal{L}J(t,x), J(t,x) - H(t,x)\right\} &= 0, \ \forall t \in [0,T), \\ J(T,x) &= H(T,x), \end{aligned}$$

where $\mathcal{L}$ is the infinitesimal generator of X with,

$$\mathcal{L}J(t,x) = \frac{1}{2}(S_t^2 V_t \frac{\partial^2 J}{\partial S^2} + \sigma^2 V_t \frac{\partial^2 J}{\partial V^2} + \rho\sigma S_t V_t \frac{\partial^2 J}{\partial S \partial V}) + rS_t \frac{\partial J}{\partial S} + \tilde{\alpha}(\tilde{\beta} - V_t)\frac{\partial J}{\partial V}.$$

This is a well-known result by general Itô's lemma and Feynman-Kac representation formula.(see for example Theorem 9.4.7 and 9.4.8 of Pascucci (2011))

4.2.2 Set of Multiple Priors

We use the previous setting of probability space, in which P is a reference measure here. This means that it functions only to fix sets with measure zero (P-null sets) in order to define a set of equivalent probability measures, and it does not have to be the real world objective measure.

Following Chen & Epstein (2002), we start constructing a set of multiple priors $\mathscr{P}$ by defining a density generator (Girsanov kernel) $\theta_t : [0,T] \times \Omega \to \mathbb{R}^d$.(for Heston's model $d = 2$) Note that θ_t is allowed to be dynamic and stochastic by definition. θ_t is defined in a such way that the conditional Radon-Nikodym derivative M_t is a P-martingale with $M_0 = 1$, where

$$\frac{\mathrm{d}Q^\theta|_{\mathcal{F}_t}}{\mathrm{d}P|_{\mathcal{F}_t}} = M_t = \mathcal{E}\big(-\int_0^{\cdot} \theta_s^* \mathrm{d}W_s\big)_t, \tag{4.5}$$
$$\mathcal{E}\big(-\int_0^{\cdot} \theta_s^* \mathrm{d}W_s\big)_t = \exp\Big(-\int_0^t \theta_s^* \mathrm{d}W_s - \frac{1}{2}\int_0^t \theta_s^* \theta_s \mathrm{d}s\Big).$$

The probability measure Q^θ of certain set in the filtration $\mathcal{F}_T$ is given by,

$$Q^\theta(A) = \mathbb{E}^P\big[\mathbb{1}_A \mathcal{E}(-\int_0^{\cdot} \theta_s^* \mathrm{d}W_s)_T\big], \forall A \in \mathcal{F}_T. \tag{4.6}$$

The set of multiple priors is defined by,

$$\mathscr{P}^\Theta := \{Q^\theta : \theta_t \in \Theta \text{ and } Q^\theta \text{ is defined in (4.6)}\}.$$

It is straightforward that Q^θ will be identical to be reference measure P if we let $\theta = 0$. Additionally, we will have $Q = Q^\theta$ if $\theta_t = \lambda_t$ for all $t \in [0,T]$. The parameter in charge of the drift model uncertainty is the so-called market price of risk, which is a byproduct of the Girsanov theorem. In fact, as the range of market price of risk increase, the value of the option on the ambiguity-averse agent's perspective tends to lower. This fact is proved theoretically by Nishimura & Ozaki (2007) and confirmed here in Figure 4.1 by numerical results. We will discuss the impact of change the range of uncertainty in detail in Section 4.4.3. The last step is to define a set of density generators Θ. For now we have two options:

- κ-*ignorance*: $\Theta = \{(\theta_t) : |\theta_t^i| \leq \kappa_i,$ for $t \in [0,T]$ and $i = 1, ..., d\}$, for a fixed positive constant $\kappa = (\kappa_1, ..., \kappa_d)$ in $\mathbb{R}_+^d$. This is introduced by Chen & Epstein (2002). In

addition, it can be naturally extended to a case where κ is time-varying, which is called *IID ambiguity*.

- *elliptical ambiguity*: $\Theta = \{(\theta_t) : \theta_t^* \Sigma^{-1} \theta_t \leq \chi, t \in [0, T]\}$ for some fixed positive semi-definite matrix Σ and positive constant χ. This is similar to the elliptical uncertainty sets used in Cohen & Tegnér (2017). Note that elliptical ambiguity set implies and elliptical set of priors, while *κ-ignorance* implies a rectangular set of priors. Elliptical ambiguity set nests circular ambiguity set that restricts the Euclidean norm of density generators within a compact and convex set.

4.2.3 Financial Markets with Multiple Priors

We characterize that the agent in the real world is facing various probabilities with respect to specific events, which is known as *Knightian uncertainty*. That means the agent cannot be absolutely certain about the drift coefficients μ_t, α and β, in Heston's stochastic volatility model (4.1)-(4.2). With partial information of specified events, the agent may act with a prior probability measure, and this can be seen as ambiguity. We define a set of equivalent probability measures $\mathscr{P}^{\Theta}$ that the agent may refer to by a set of density generators Θ.

Then we proceed to the evaluation rule of financial claims and the dynamics of underlying processes when there is ambiguity. According to Gilboa & Schmeidler (1989), the behaviour agent with multiple priors will be uncertainty-averse, or ambiguity-averse if the agent acts in accordance with some certain sensible axioms. That means the agent will use a probability measure corresponding to the "worst" case scenario when evaluating a claim. In the American option case, the desired price is denoted as,

$$v_t := v(t, X_t) = \operatorname*{ess\,sup}_{\tau \in \mathscr{T}_t} \operatorname*{ess\,inf}_{Q^\theta \in \mathscr{P}^\Theta} \mathbb{E}^{Q^\theta}[H_\tau \gamma_{\tau-t}^{-1} | \mathcal{F}_t],\ t \in [0, T], \tag{4.7}$$

and we will use $v_t := \operatorname{ess\,sup}_{\tau \in \mathscr{T}_t} \operatorname{ess\,inf}_{\theta \in \Theta} \mathbb{E}^{\theta}[H_\tau \gamma_{\tau-t}^{-1} | \mathcal{F}_t]$ for simplicity.

A direct result by the Girsanov theorem reveals the dynamics of underlying processes with ambiguity,

$$\begin{aligned} \mathrm{d}S_t/S_t &= \mu_t \mathrm{d}t - \sqrt{V_t}\big(\rho\theta_t^1 + \sqrt{1-\rho^2}\theta_t^2\big)\mathrm{d}t + \sqrt{V_t}\big(\rho \mathrm{d}W_t^{\theta 1} + \sqrt{1-\rho^2}\mathrm{d}W_t^{\theta 2}\big), \\ \mathrm{d}V_t &= \alpha(\beta - V_t)\mathrm{d}t - \sigma\theta_t^1\sqrt{V_t}\mathrm{d}t + \sigma\sqrt{V_t}\mathrm{d}W_t^{\theta 1}, \end{aligned}$$

where $W^\theta = (W^{\theta 1}, W^{\theta 2})^*$ is a standard two dimensional Brownian motion under measure Q^θ, and

$$\mathrm{d}W_t^{\theta 1} = \theta_t^1 \mathrm{d}t + \mathrm{d}W_t^1,$$
$$\mathrm{d}W_t^{\theta 2} = \theta_t^2 \mathrm{d}t + \mathrm{d}W_t^2.$$

For the evaluation purpose, we change the reference measure directly to the risk-neutral measure Q, as the agent does not know the real world probability measure and it is equivalent to P. This approach is also adopted by X. Cheng & Riedel (2013), Vorbrink (2011) and Cohen & Tegnér (2017). Thus we are considering under probability space $(\Omega, \mathcal{F}, Q)$, and the filtration $(\mathcal{F}_t)_{0\le t\le T}$ is generated by $\tilde{W} = (\tilde{W}^1, \tilde{W}^2)^*$. To avoid confusion and stay in the previous multiple priors framework, we let $Q \in \mathscr{P}^\Theta$, and we have equivalently,

$$\mathrm{d}S_t/S_t = r\mathrm{d}t - \sqrt{V_t}\big(\rho\theta_t^1 + \sqrt{1-\rho^2}\theta_t^2\big)\mathrm{d}t + \sqrt{V_t}\big(\rho \mathrm{d}\tilde{W}_t^{\theta 1} + \sqrt{1-\rho^2}\mathrm{d}\tilde{W}_t^{\theta 2}\big),$$
$$\mathrm{d}V_t = \alpha(\beta - V_t)\mathrm{d}t - \sigma\theta_t^1\sqrt{V_t}\mathrm{d}t + \sigma\sqrt{V_t}\mathrm{d}\tilde{W}_t^{\theta 1},$$

where

$$\mathrm{d}\tilde{W}_t^{\theta 1} = \theta_t^1 \mathrm{d}t + \mathrm{d}\tilde{W}_t^1, \tag{4.8}$$
$$\mathrm{d}\tilde{W}_t^{\theta 2} = \theta_t^2 \mathrm{d}t + \mathrm{d}\tilde{W}_t^2. \tag{4.9}$$

4.2.4 Relation to Reflected Backward Stochastic Differential Equations

Following El Karoui, Kapoudjian, et al. (1997), we introduce some notations for a proper definition of the reflected backward stochastic differential equation.

$$\mathbb{L}^2 = \Big\{\xi \text{ is an } \mathcal{F}_T\text{-measurable random variable s.t. } \mathbb{E}[|\xi|^2] < +\infty\Big\}, \tag{4.10}$$
$$\mathbb{H}^2 = \Big\{\{\phi_t,\ 0 < t < T\} \text{ is a predictable process s.t. } \mathbb{E}[\int_0^T |\phi_t|^2 \mathrm{d}t] < +\infty\Big\}, \tag{4.11}$$
$$\mathbb{S}^2 = \Big\{\{\phi_t,\ 0 < t < T\} \text{ is a predictable process s.t. } \mathbb{E}[\sup_{0\le t\le T} |\phi_t|^2] < +\infty\Big\}. \tag{4.12}$$

Definition 4.2.1. *Assume conditions such as the terminal condition $\xi \in \mathbb{L}^2$, the continuous reflection bound $L \in \mathbb{S}^2$, a uniform Lipschitz generator $f(t, y, z)$ and $f(\cdot, y, z) \in \mathbb{H}^2$, $\forall (y, z) \in \mathbb{R} \times \mathbb{R}^2$ are met, a triplet $\{(Y_t, Z_t, K_t),\ t \in [0, T]\}$ of $\mathcal{F}_t$ progressively measurable processes taking values in $\mathbb{R}$, $\mathbb{R}^2$ and $\mathbb{R}_+$ is the unique solution of the following reflected backward stochastic differential equation (RBSDE) satisfying:*

(i) $Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\mathrm{d}s + K_T - K_t - \int_t^T Z_s^* \mathrm{d}B_s$,

(ii) $Y_t \geq L_t$ *and* $(Y, Z, K) \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{L}^2$,

(iii) K_t *is a continuous and increasing process,* $\int_0^T (Y_s - L_t)\mathrm{d}K_t = 0$ *and* $K_0 = 0$,

where B *is a* d*-dimensional standard Brownian motion under reference measure* Q*,* L_t *is called an "obstacle" or reflection bound, and* K_t *is a process that "pushes* Y_t *upwards" minimally in the sense of condition (iii).* $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ *is the generator.*

The existence and uniqueness of the solution of above RBSDE are proved in Section 6 of El Karoui, Kapoudjian, et al. (1997). Followed by Proposition 2.3 of El Karoui, Kapoudjian, et al. (1997), Y_t satisfies for each $t \in [0, T]$:

$$Y_t = \operatorname*{ess\,sup}_{\tau \in \mathscr{T}_t} \mathbb{E}^Q[\int_t^\tau f(s, Y_s, Z_s)\mathrm{d}s + L_\tau \mathbb{1}_{\{\tau < T\}} + \xi \mathbb{1}_{\{\tau = T\}} | \mathcal{F}_t].$$

where $\mathscr{T}$ is the set of all stopping times dominated by T, and $\mathscr{T}_t = \{\tau \in \mathscr{T}; t \leq \tau \leq T\}$.

Without uncertainty, which means we take $\theta_t = 0$ for all $t \in [0, T]$, the American option price J_t with payoff H_t at time t, defined in equation (4.4), will have a dual representation of the solution of an RBSDE (Y_t). Hence $J_t = Y_t$, and the RBSDE is denoted in differential form as,

$$-\mathrm{d}Y_t = -rY_t\mathrm{d}t + \mathrm{d}K_t - Z_t^* \mathrm{d}\tilde{W}_t,\ Y_T = H_T, \tag{4.13}$$

with obstacle $L_t = H_t\ \forall t \in [0, T]$. This is obtained by taking a linear generator $f(t, y, z) = -ry$ and let $B_t = \tilde{W}_t$, according to Proposition 7.1 of El Karoui, Kapoudjian, et al. (1997). Further, this result can be extended to the case of multiple priors.

Suppose conditions needed for the existence and uniqueness are satisfied, $(Y_t^\theta, Z_t^\theta, K_t^\theta), t \in [0, T]$ are the solution to the following linear RBSDE

$$-\mathrm{d}Y_t^\theta = f(t, Y_t^\theta, Z_t^\theta)\mathrm{d}t + \mathrm{d}K_t^\theta - Z_t^{\theta *} \mathrm{d}\tilde{W}_t^\theta,\ Y_T^\theta = H_T, \tag{4.14}$$

with generator $f(t, Y_t^\theta, Z_t^\theta) = -rY_t^\theta$ and obstacle $L_t = H_t$, $\forall t \in [0, T]$, where $\tilde{W}^\theta = (\tilde{W}^{\theta 1}, \tilde{W}^{\theta 2})^*$, defined in (4.8)-(4.9), is a standard two dimensional Brownian motion under probability measure Q^θ,

Proposition 4.2.1. *Let u_t be defined as,*

$$u_t := \operatorname*{ess\,sup}_{\tau \in \mathscr{T}_t} \mathbb{E}^{\theta}[H_\tau \gamma_{\tau-t}^{-1} | \mathcal{F}_t], \tag{4.15}$$

where γ_t is defined in (4.3). Then u_t has a dual representation of the solution of an RBSDE in equation (4.14), which means $u_t = Y_t^\theta$ for every t in $[0, T]$. Additionally, u_t is the unique viscosity solution of the following variational inequality:

$$\min\Big\{ -f(t, Y_t^\theta, Z_t^\theta) - \frac{\partial u}{\partial t} - \mathcal{L}u(t,x), u(t,x) - H(t,x) \Big\} = 0, \ (t,x) \in [0,T) \times \mathbb{R}^2, \tag{4.16}$$

$$u(T,x) = H(T,x), \tag{4.17}$$

with the generator explicitly denoted by

$$f(t, Y_t^\theta, Z_t^\theta) = -rY_t^\theta - \theta_t^1 Z_t^{\theta 1} - \theta_t^2 Z_t^{\theta 2}. \tag{4.18}$$

Proof. By the Girsanov theorem, (4.14) is equivalent to

$$-\mathrm{d}Y_t^\theta = -(rY_t^\theta + \theta_t^* Z_t^\theta)\mathrm{d}t + \mathrm{d}K_t^\theta - Z_t^{\theta *}\mathrm{d}\tilde{W}_t, \tag{4.19}$$

then by Proposition 7.1 of El Karoui, Kapoudjian, et al. (1997), $\{(Y_t^\theta, Z_t^\theta, K_t^\theta), \ t \in [0,T]\}$ satisfies,

$$\Gamma_t Y_t^\theta = \operatorname*{ess\,sup}_{\tau \in \mathscr{T}_t} \mathbb{E}^Q[\Gamma_\tau H_T \mathbb{1}_{\{\tau = T\}} + \Gamma_\tau H_\tau \mathbb{1}_{\{\tau < T\}}],$$
$$\mathrm{d}\Gamma_t = \Gamma_t\big(r\mathrm{d}t + \theta_t^* \mathrm{d}\tilde{W}_t\big), \ \Gamma_0 = 1.$$

We know $\Gamma_{\tau-t} = \gamma_{\tau-t}^{-1} M_{\tau-t}$, where M is defined in (4.5), and we have (4.15) by definition of conditional expectation under measure Q^θ. The second part is a direct computation from Theorem 8.6 of El Karoui, Kapoudjian, et al. (1997). □

Proposition 4.2.2. *Under the κ-ignorance and the elliptical ambiguity framework as in the section 4.2.2, the American option price v_t under the worst case belief as defined in (4.7) is the value of a minimax (optimal) control problem such that,*

$$v_t = \operatorname*{ess\,inf}_{\theta \in \Theta} Y_t^\theta, \tag{4.20}$$

and we have,

$$\operatorname*{ess\,sup}_{\tau\in\mathscr{T}_t}\operatorname*{ess\,inf}_{\theta\in\Theta}\mathbb{E}^{\theta}[H_\tau\gamma_{\tau-t}^{-1}|\mathcal{F}_t]=\operatorname*{ess\,inf}_{\theta\in\Theta}\operatorname*{ess\,sup}_{\tau\in\mathscr{T}_t}\mathbb{E}^{\theta}[H_\tau\gamma_{\tau-t}^{-1}|\mathcal{F}_t],\ \forall t\in[0,T], \tag{4.21}$$

where Y_t^θ is an element of the solution to the RBSDE (4.14), and H_t is the payoff function defined in (4.4). Moreover, there exists a pair $(\tilde{\tau},\tilde{\theta}_t)$, $\tilde{\tau}\in[0,T]$, $\tilde{\theta}\in\Theta$ such that it reaches the optimal (saddle) point $(\tilde{\tau},\tilde{\theta}_t)$ when the generator given by (4.18) satisfies $f(t,Y_t^{\tilde{\theta}},Z_t^{\tilde{\theta}})=$ $\operatorname{ess\,inf}_{\theta\in\Theta}f(t,Y_t^\theta,Z_t^\theta)$, $\forall t\in[0,T]$.

Proof. Equation (4.20) holds if we take the essential infimum of the generator of (4.19) while remain other parameters unchanged, according to the comparison theorem, Theorem 4.1 of El Karoui, Kapoudjian, et al. (1997). Then equation (4.21) follows directly from Theorem 7.2 of El Karoui, Kapoudjian, et al. (1997). The only thing we need is the existence of the minimum of the generator $f(t,Y_t^\theta,Z_t^\theta)$.

From Proposition 4.2.1 we know that $f(t,Y_t^\theta,Z_t^\theta)=-rY_t^\theta-\theta_t^*Z_t^\theta$, then the minimum of $f(t,Y_t^\theta,Z_t^\theta)$ corresponds to the maximum of $\theta_t^*Z_t^\theta$. The existence of $\max_{\theta\in\Theta}\theta_t^*Z_t^\theta$ under *κ-ignorance* and *IID ambiguity* is proved in Section 2.4 of Chen & Epstein (2002). The existence of maximum point of generator under elliptical ambiguity follows similarly because the set of generators Θ is also compact and convex valued in $\mathbb{R}^2$. □

As the solution of the RBSDE (4.14) stands for the option's value under the ambiguity probability measure Q^θ governed by the uncertainty parameter θ, we are able to obtain the options value ν_t on the ambiguity-averse agent's perspective if we can the lower bound of the solution of the RBSDE (4.14).

Let us consider the case of *elliptical ambiguity* $\Theta=\{\theta:\theta\Sigma^{-1}\theta^*\leq\chi\}$, as defined in section 4.2.2. To obtain the minimum of the generator, we can solve the following optimization problem,

$$\begin{aligned}&f(t,Y_t^{\tilde{\theta}},Z_t^{\tilde{\theta}})=\min_{\theta\in\Theta}(-rY_t^\theta-\theta_t^*Z_t^\theta),\\&\text{subject to }\ \theta\Sigma^{-1}\theta^*=\chi,\end{aligned}$$

where we have the equality in the constraint, since there is no internal stationary points of the elliptical set due to the affine structure inside the minimum, this argument is similar to that

in Cohen & Tegnér (2017). Solving the Lagrangian, the solution of the above optimization problem is given by,

$$f(t, Y_t^{\tilde{\theta}}, Z_t^{\tilde{\theta}}) = -rY_t^{\tilde{\theta}} - \sqrt{Z_t^{\tilde{\theta}^*}\Sigma^* Z_t^{\tilde{\theta}}\chi}, \tag{4.22}$$

$$\text{with } \tilde{\theta} = -\frac{Z_t^{\tilde{\theta}^*}\Sigma^*}{\sqrt{Z_t^{\tilde{\theta}^*}\Sigma^* Z_t^{\tilde{\theta}}/\chi}}. \tag{4.23}$$

It is notable that in the case of κ-ignorance, in which the uncertainty set is rectangular (convex and compact), the optimal solution is located in one of four vertexes of the rectangle.

Remark 4.2.1. *Given the above optimal generator, we still need to check if the uniqueness of the solution of the RBSDE (4.19) still holds. To prove that, is sufficient to show the generator is uniform Lipschitz, i.e.* $|f(t, y, z) - f(t, y', z')| \leq$ *Constant* $\cdot\, (|y - y'| + ||z - z'||)\, \forall y, y' \in \mathbb{R},\ z, z' \in \mathbb{R}^d$. *Note that*

$$\begin{aligned}
|f(t, y, z) - f(t, y', z')| &= |(ry - ry') + (\sqrt{z^*\Sigma z\chi} - \sqrt{z^*\Sigma z\chi})|, \\
&\leq r|y - y'| + \sqrt{\chi}(||\Sigma^{1/2}z|| - ||\Sigma^{1/2}z'||), \\
&\leq r|y - y'| + \sqrt{\chi}\Sigma^{1/2}||z - z'||,
\end{aligned}$$

where the first equality holds by the definition of the optimal generator; the second inequality holds by the triangle inequality and the last inequality holds by the reverse triangle inequality. Then the optiomal generator is uniform Lipschitz because Σ *is a positive semi-definite matrix and* χ *is a positive constant by definition.*

4.3 Numerical Methods

In this section we briefly introduce the conventional numerical methods for reflected backward stochastic differential equations (RBSDEs) and our proposed amended method for RBSDEs, and investigate the feasibility of a new numerical dynamic programming method without using the theory of RBSDE.

Firstly we split the time interval $[0, T]$ in to N equal parts with each part to be Δt. Hence we have a time grid $\pi : 0 = t_0 < ... < t_N = T$ with $(i + 1)$-th time step $t_{i+1} - t_i$ denoted as Δ_i, and the $(i + 1)$-th Brownian motion increment under measure Q: $\tilde{W}_{t_{i+1}} - \tilde{W}_{t_i}$ is defined

by $\Delta\tilde{W}_i$. The conditional expectation $\mathbb{E}^Q[\cdot|\mathcal{F}_{t_i}]$ under measure Q is denoted as $\mathbb{E}_i^Q[\cdot]$. Then we simulate the dynamics of X_t under Q measure in M different paths by only one time step Δt, using the Euler scheme,

$$X_{i+1}^\pi = X_i^\pi + b(t_i, X_i^\pi)\Delta t + \sigma(t_i, X_i^\pi)^* \Delta\tilde{W}_i. \tag{4.24}$$

4.3.1 Numerical Methods for BSDEs

We begin with numerical methods for solutions of discrete backward stochastic differential equations (BSDEs), since methods for RBSDEs are based on those for BSDEs.

- One-step Forward Dynamic Programming (ODP) scheme: Based on the one-step scheme introduced by Bouchard & Touzi (2004), Lemor et al. (2006) propose a modified scheme as such,

$$Z_i^\pi = \frac{1}{\Delta_i}\mathbb{E}_i^Q[Y_{i+1}^\pi \Delta\tilde{W}_i],\ Y_N^\pi = H(X_N^\pi), \tag{4.25}$$

$$Y_i^\pi = \mathbb{E}_i^Q[Y_{i+1}^\pi + f(X_i^\pi, Y_{i+1}^\pi, Z_i^\pi)\Delta_i]. \tag{4.26}$$

  This Markovian representation of a BSDE is obtained by taking conditional expectation on both sides of the following discrete BSDE:

$$Y_{i+1}^\pi - Y_i^\pi = -f(X_i^\pi, Y_{i+1}^\pi, Z_i^\pi)\Delta_i + Z_i^\pi \Delta\tilde{W}_i.$$

- Multi-step Forward Dynamic Programming (MDP) scheme: Bender & Denk (2007) introduced a more stable multi-step scheme:

$$Z_i^\pi = \frac{1}{\Delta_i}\mathbb{E}_i^Q[\big(Y_N^\pi + \sum_{k=i}^{N-1} f(X_k^\pi, Y_{k+1}^\pi, Z_k^\pi)\Delta_k\big)\Delta\tilde{W}_i],\ Y_N^\pi = H(X_N^\pi), \tag{4.27}$$

$$Y_i^\pi = \mathbb{E}_i^Q[Y_N^\pi + \sum_{k=i}^{N-1} f(X_k^\pi, Y_{k+1}^\pi, Z_k^\pi)\Delta_k],\ \text{for } i = 0, 1, ..., N-1. \tag{4.28}$$

  The backward induction (4.27)-(4.28) are obtained by substituting subsequent terms of Y_i^π into (4.25)-(4.26) and applying the tower law of conditional expectation. According to Gobet & Turkedjiev (2016), the advantage of the MDP over the ODP is that the error of approximating conditional expectation is the average rather than the sum of local error terms, thus the result is tighter in a sense.

4.3.2 Numerical Methods for RBSDEs

"Max method": Introduced in Gobet & Lemor (2008), it is a method to approximate (X_t, Y_t) by a discrete-time doublet (X^π, Y^π) for a discrete RBSDE:

$$Z_i^\pi = \frac{1}{\Delta_i}\mathbb{E}_i^Q[Y_{i+1}^\pi \Delta\tilde{W}_i],\ Y_N^\pi = H(X_N^\pi), \tag{4.29}$$

$$\tilde{Y}_i^\pi = \mathbb{E}_i^Q[Y_{i+1}^\pi + f(X_i^\pi, Y_{i+1}^\pi, Z_i^\pi)\Delta_i], \tag{4.30}$$

$$Y_i^\pi = \tilde{Y}_i^\pi \vee H(X_i^\pi). \tag{4.31}$$

The "Max method" (4.29), (4.30) and (4.31) is nothing different from the ODP for the discrete BSDE except the last step (4.31). The convergence of solutions is proved by Ma & Zhang (2005) when the obstacle $H(\cdot)$ is of class $C^{1,2}$. Bouchard & Chassagneux (2008) extend the case to general Bermudan/American case and prove that the rate of convergence is at least $N^{-1/4}$.[2]

In this chapter we use the "Max method", along with the stratification regression method introduced by Gobet et al. (2016) to approximate solutions of discrete RBSDEs. We use the name Stratified Regression One-step Forward Dynamic Programming (SRODP) for this scheme hereafter.

4.3.3 Approximating Conditional Expectation by Stratified Regression

As for practically operating the backward induction, one can choose to approximate the conditional expectation $\mathbb{E}_i^Q[\cdot]$ with $\hat{\mathbb{E}}^Q[\cdot|X_i]$ by solving a least-square optimization problem

$$\hat{\mathbb{E}}^Q[\cdot|X_i] := \arg\inf_\phi \frac{1}{M}\sum_{m=1}^{M} |\phi(X_i^{(m)}) - \mathbb{E}^Q[\cdot|X_i^{(m)}]|^2, \tag{4.32}$$

where generally $\phi(\cdot)$ is defined in $L^2(\Omega, \mathcal{F}, Q)$, and $X_i^{(m)}$ is the (m)-th path of all the paths by Monte Carlo simulation. To make it suitable for a search policy, one can fix $\phi(\cdot)$ in a finite linear space $\mathscr{K}$. Then it is possible to choose a polynomial basis $p_1(x), p_2(x), ..., p_K(x)$, (possibly using the Legendre, Laguerre or simple power polynomials, according to Longstaff

[2]The MDP scheme will not work for the RBSDE case, since the maximization operators in (4.31) does not get cancelled by tower law when substituting subsequent terms of Y_i^π into (4.29)-(4.30).

& Schwartz (2001)) and let $y := [\mathbb{E}^Q[\cdot|X_i^{(m)}]]_m$, $P := [p_k(X_i^{(m)})]_{m,k}$. The general least-square optimization (4.32) becomes solving a linear regression:

$$\inf_{\mathscr{K}} \frac{1}{M} \sum_{m=1}^{M} |\phi(X_i^{(m)}) - \mathbb{E}^Q[\cdot|X_i^{(m)}]|^2 = \inf_{\beta \in \mathbb{R}^K} \frac{1}{M} |P\beta - y|_2^2. \tag{4.33}$$

Conventional methods of using all the simulated paths (Tsitsiklis & Van Roy, 2001) or the in-the-money paths (Longstaff & Schwartz, 2001) to run the regression and globally approximate the objective function are popular in finance. Yet it yields a higher number of polynomials when dealing with multi-dimensional cases. Gobet et al. (2016) propose a general stratified regression method to approximate conditional expectations locally in partitions (hypercubes) of simulated sample paths. The main benefits of it is one can minimize the memory consumed by simulations and regression coefficients, since one only needs to generate samples on one hypercube at a time and the number of polynomials required is much less than the traditional method. We briefly state it here, as it is also an essential part of our proposed numerical method without using the theory of RBSDEs.

Firstly we choose a domain of discretized state space $D \subset \mathbb{R}^d$ centered on $\mathscr{O}$ ($D = \prod_{k=1}^d]\mathscr{O}_k - R, \mathscr{O}_k + R]$) with radius R, and partition it into small hypercubes with the same length of edge δ. Then we generate certain randomly distributed points of X^π and simulate the dynamics. Approximation of conditional expectation by least-square regression is implemented independently on different hypercubes. The general stratification in Gobet et al. (2016) differs from the traditional stratified sampling in that it does not need the explicit distribution of X_i^π, since the conditional expectation is determined by the transition equation of X after t_i, and simulations of X^π can start from an random variable at arbitrary time rather than a fixed point at time 0. This feature enables us to resimulate at each time point and save memory for storing simulated paths. It also sheds some light to another method without using the theory of RBSDEs. We present it in the next subsection.

4.3.4 Stratified Least-square Monte Carlo Method

To solve (4.7), one natural conjecture is to split the set of density generators Θ in (4.6) into, say, L discrete points, where each single point stands for an equivalent probability measure $Q^{\theta_i}(i = 0, 1, ..., L)$. Then we simulate X_t under each measure Q^{θ_i}, evaluate the American

option price under Q^{θ_i} using the Least-square Monte Carlo (LSM) algorithm, and at last select the smallest one after comparing L option prices. This method seems to work fine due to the duality argument in (4.21). However, it is not actually always correct. Since we are fixing the density generator θ_i to be a constant in order to implement Monte Carlo simulation, we cannot guarantee that the essential minimum of the generator of the RBSDE (4.19) will be obtained since Z_t is dynamic and stochastic. The only case this method will work is when the sign of Z_t is certain, and Z_t in (4.18) is directly related to the Greeks Delta and Vega.

A possible alternative will be to simulate the density generator processes M_t^θ and X_t simultaneously. The difficulty is that we will have L different choices of θ at each discretized time point, which makes the simulation of forward process M_t^θ to be a L-nomial tree. This will finally consume explosive memory of storing simulated paths when increasing the number of discretization time points N, for example, it will have to store L^N simulations for the last step. Thus, this alternative is not realistic.

We raise another potential way to solve the minimax problem and, in the mean time, avoid using the theory of RBSDEs. We take the name Stratified Least-square Monte Carlo (SLSM) method hereafter. The intuitive is to take advantage of the Girsanov theorem, in order to take a family of probability measures into account. Because the Girsanov theorem relies heavily on the stochastic exponential to transform measures, it might be possible to use the stochastic exponential to transform ambiguity probability measures to the reference measure and hence solve the optimization problem. The idea, using stratification and resimulation at each time point after discretization, can be seen as an extension of regression based method. However, the method we propose is very different from the LSM method, in which the underlying process X_t and the stochastic exponential M_t defined in (4.5) must be simulated from a fixed point at time 0. Instead, we simulate the underlying processes at t_i by only one step forward to t_{i+1}, assuming that X_{t_i} and M_{t_i} follow some certain distributions, for example, logistic distribution (Gobet et al., 2016) or Laplace distribution. This means we take M_t as an extra dimension of underlying state variables. Then we evaluate the cash flow at time t_i by comparing the immediate payoff with the continuation value of American option, which is approximated by basis regression method. This implies that we only need to store one-step simulation and coefficients in the basis regression. We state our idea as follows:

(i) Firstly we discretize the stochastic exponential M_t defined in (4.5), and select L discrete points from the set of density generator Θ. Assuming X_{t_i} and M_{t_i} are i.i.d. conditional logistic random variables, we simulate the dynamics of X_t and M_t under Q measure in M different paths by only one time step Δt, either by the Euler dynamics,,

$$X_{i+1}^{\pi} = X_i^{\pi} + b(t_i, X_i^{\pi})\Delta t + \sigma(t_i, X_i^{\pi})^* \Delta \tilde{W}_i,$$
$$M_{i+1}^{\pi,\theta_j} = M_i^{\pi} \exp\left\{ -\frac{1}{2}\theta_{t_i,j}^2 \Delta t - \theta_{t_i,j}^* \Delta \tilde{W}_i \right\},\ j = 1, 2, ..., L.$$

(ii) We define the one-step back continuation value of the American option under measure Q^{θ_j} at time t_i,

$$h(S_i^{\pi}, M_i^{\pi}, \theta_j) := \mathbb{E}_i^Q[\frac{M_{i+1}^{\pi,\theta_j} Y_{i+1}^{\pi,\tilde{\theta}} \gamma_{\Delta t}^{-1}}{M_i^{\pi}}] = \mathbb{E}_i^{\theta_j}[Y_{i+1}^{\pi,\tilde{\theta}} \gamma_{\Delta t}^{-1}],\ Y_N^{\pi,\tilde{\theta}} = H(S_N^{\pi}),$$
$$h(S_i^{\pi}, M_i^{\pi}, , \tilde{\theta}) = \operatorname*{ess\,inf}_{\theta_j \in \Theta} \mathbb{E}_i^Q[\frac{M_{i+1}^{\pi,\theta_j} Y_{i+1}^{\pi,\tilde{\theta}} \gamma_{\Delta t}^{-1}}{M_i^{\pi}}],$$
$$Y_i^{\pi,\tilde{\theta}} = h(S_i^{\pi}, M_i^{\pi}, \tilde{\theta}) \vee H(S_i^{\pi}).$$

It should be noted that we will have L different realizations of M_{i+1}^{π,θ_j} as for each M_i^{π}, since we split Θ into L discrete points. This should be differentiated from M paths of underlying processes.

(iii) The approximation of $h(S_i^{\pi}, M_i^{\pi,\theta})$ is then obtained by solving the linear least-square regression (4.33),

$$\hat{h}(S_i^{\pi}, M_i^{\pi}, \theta_j) = \sum_{k=1}^{K} c_k^j p_k^j(X_i^{\pi}, M_i^{\pi}),$$
$$\hat{h}(S_i^{\pi}, M_i^{\pi}, \tilde{\theta}) = \operatorname*{ess\,inf}_{\theta_j \in \Theta} \hat{h}(S_i^{\pi}, M_i^{\pi}, \theta_j),$$

where $p_k^j(X_i^{\pi}, M_i^{\pi})$ is the basis functions defined in (4.33) such as Hermite or Legendre polynomials, and c_k^j is the constant coefficient of each polynomial. In practice, we take the approximated American option value $\hat{Y}_i^{\pi,\tilde{\theta}}$ as $\hat{Y}_i^{\pi,\tilde{\theta}} = \hat{h}(S_i^{\pi}, M_i^{\pi}, \tilde{\theta}) \vee H(S_i^{\pi})$.

(iv) We take the above steps from time t_{N-1} to t_1 recursively, and re-simulate the trajectories of $X_{t_i}^{\pi}$ and $M_{t_i}^{\pi}$ at each time point in order to conserve memory. The difference of our method to the LSM algorithm is that we store the regression coefficients $c_k,\ k = 1, 2..., K$ rather than the continuation value $\hat{h}(S_i^{\pi}, M_i^{\pi,\theta})$ at each time point,

which means we will use new sample paths to evaluation the continuation value. Stratification method is also employed, i.e. to stratify $X_{t_i}^{\pi}$ and $M_{t_i}^{\pi}$ to several different strata (hypercubes) in order to approximate the objective function locally. Stratification also enables us to use highly parallelized computation to decrease computation time.

4.4 Numerical Results for Financial Options

We will start this part by presenting results for European and American put options using SRODP within the Black-Scholes (hereafter BS) and Heston's framework.[3] (Black & Scholes, 1973; Heston, 1993) The main purpose is to show that our proposed methods converge and results are close enough to *the exact value*, since we have closed form European option prices and American option prices (by numerical PDE methods).

4.4.1 European Put Option Prices

We use Laguerre local polynomials up to first order (LP1) as the basis functions when implementing SRODP scheme. We fix the space domain for the logarithmic stock price to be $D = [-6.5, 6.5]$. We use 1000 hypercubes and 2000 simulations for each hypercube. Furthermore, we launch 50 times the SRODP algorithm, and collect each time the result, denoted as $(Y_0^i)_{1\leq i\leq 50}$. The SRODP European put option price is then the empirical mean as following,

$$\bar{Y}_0 = \frac{1}{50}\sum_{t=1}^{50} Y_0^i,$$

we also calculate the empirical standard deviation,

$$\sigma_0 = \frac{1}{49}\sqrt{\sum_{t=1}^{50} |Y_0^i - \bar{Y}_0|^2},$$

then the standard error of the mean is $\frac{\sigma_0}{\sqrt{50}}$ and the 95 percent confidence interval is $[\bar{Y}_0 - \frac{1.96\sigma_0}{\sqrt{50}}, \bar{Y}_0 + \frac{1.96\sigma_0}{\sqrt{50}}]$. We compare the SRODP and the BS results, and present the standard deviation and the 95% confidence interval of SRODP in the following Table 4.1.

[3] It should be noted that the proposed SLSM method is identical to the SRODP method in cases without ambiguity.

Table 4.1: European Put Option Prices

S_0	BS	SRODP	std	95% C.I.
36	3.844	3.8379	0.0412	[3.8264, 3.8493]
38	2.852	2.8447	0.0349	[2.8351, 2.8544]
40	2.066	2.0675	0.0331	[2.0583, 2.0767]
42	1.465	1.4625	0.0254	[1.4554, 1.4695]
44	1.017	1.0205	0.0220	[1.0144, 1.0266]

Note: The strike price is 40, volatility is 0.2, risk free rate is 0.06, and time to maturity is 1. We use 5 time steps, 1000 hypercubes and 2000 simulations for each hypercube. Each 50 times SRODP algorithm takes about 120 seconds.

4.4.2 American Put Option Prices without Ambiguity

One Dimensional Case

We use LP1 as the basis functions when implementing SRODP scheme. We fix the space domain for the logarithmic stock price to be $D = [-6.5, 6.5]$. We use 1000 hypercubes and 2000 simulations for each hypercube, and we launch 50 times the SRODP algorithm. The PDE and LSM results are extracted from Longstaff & Schwartz (2001), and PDE results are taken as *the exact value*. The option is exercisable 50 times per year, so it is of American style. The results are in Table 4.2.

Table 4.2: One Dimensional American Put Option Prices

S_0	BS	PDE	LSM (std)	SRODP	std	95% C.I.
36	3.844	4.478	4.472 (0.010)	4.4825	0.0193	[4.4772, 4.4879]
38	2.852	3.250	3.244 (0.009)	3.2570	0.0169	[3.2523, 3.2617]
40	2.066	2.314	2.313 (0.009)	2.3203	0.0158	[2.3160, 2.3247]
42	1.465	1.617	1.617 (0.007)	1.6224	0.0123	[1.6190, 1.6258]
44	1.017	1.110	1.118 (0.007)	1.1125	0.0108	[1.1095, 1.1155]

Note: The strike price is 40, volatility is 0.2, risk free rate is 0.06, and time to maturity is 1. We use 50 time steps, 1000 hypercubes and 2000 simulations for each hypercube. The LSM results are from Longstaff & Schwartz (2001). Each 50 times SRODP algorithm takes about 2400 seconds.

Two Dimensional Case

We extend the one dimensional American option prices to the case under Heston's model. The American option price under Heston's model can be obtained via standard numerical PDE methods and are exploited by previous researches (see for example Ikonen & Toivanen (2008)). Thus, the PDE results from Ikonen & Toivanen (2008) are taken as *the exact value*. Meanwhile, we employ the same parameters for the dynamics of stock price and volatility

as in Ikonen & Toivanen (2008) for simplicity. We fix the space domain for the logarithmic stock price to be $[-6.5, 6.5]$ and the space domain for the volatility to be $[0, 10]$. We launch 10 times for the LSM algorithm and the SRODP algorithm. We also implement the standard "Max method" for RBSDE for 10 times, i.e. without stratification. We conduct several experiments to show the convergence of the SRODP method with the number of hypercubes and simulations increasing. LP2 stands for regressions with local polynomials up to the second order.

It can be observed that both LP1 and LP2 regression SRODP scheme converges when increase the number of hypercubes (from Table 4.3-4.7) or the number of simulated paths per hypercube (from Table 4.5-4.6). However, the number of hypercubes required for the LP2 scheme to converge is much less than that for the LP1 scheme, although higher order of local polynomials come with more computational burden. It should be noted that the standard RBSDE algorithm converge as well, since the standard deviations shrink when the number of simulation paths increase from $10,000$ to $80,000$. The standard deviations are as small as that for the LSM algorithm. The reason is when applying the stopping rule and approximating the stopping time in the backward induction, both the standard RBSDE and LSM algorithms use the original paths of state variables for simulation(Glasserman, 2013). One can certainly resimulate the whole sample paths when applying the stopping rule to evaluate the option as this is the more realistic case, yet it will come with larger standard deviations. Moreover in Table 4.7, either the SRODP algorithm with LP1 or LP2 scheme is close enough to *the exact value* and the LSM results, whereas the pricing bias by the standard RBSDE algorithm is not negligible, especially for at-the-money and out-of-money options.

Table 4.3: Two Dimensional American Put Option Prices (a)

S_0	PDE	LSM (std)	RBSDE (std)	LP1:SRODP(std)	Time(s)	LP2:SRODP(std)	Time(s)
8	2	1.9929 (0.0027)	1.9909 (0.0019)	2.5287 (0.0163)	21.81	2.1777 (0.0299)	51.63
9	1.1076	1.1125 (0.0074)	1.1334 (0.0058)	1.9240 (0.0198)	22.17	1.4866 (0.0412)	63.38
10	0.52	0.5376 (0.0067)	0.5660 (0.0063)	1.3622 (0.0379)	23.34	0.9549 (0.0355)	52.39
11	0.2137	0.2252 (0.0040)	0.2675 (0.0050)	0.8788 (0.0276)	22.54	0.5523 (0.0283)	53.03
12	0.082	0.0843 (0.0029)	0.1059 (0.0038)	0.4247 (0.0320)	22.15	0.2406 (0.0336)	52.90

Note: The strike price is 10, initial volatility V_0 is 0.0625, risk free rate r is 0.1, and time to maturity is 0.25. $\tilde{\alpha} = 5$, $\tilde{\beta} = 0.16$, $\sigma = 0.9$ and $\rho = 0.1$. We use 25 time steps, 10 hypercubes and 3000 simulations for each hypercube. For each single run of the LSM algorithm and standard RBSDE algorithm, we use 10,000 simulations.

Table 4.4: Two Dimensional American Put Option Prices (b)

S_0	PDE	LSM (std)	RBSDE (std)	LP1:SRODP(std)	Time(s)	LP2:SRODP(std)	Time(s)
8	2	1.9910 (0.0013)	1.9909 (0.0013)	2.1090 (0.0155)	213.90	1.9966 (0.0087)	485.55
9	1.1076	1.1071 (0.0040)	1.1331 (0.0035)	1.3375 (0.0127)	215.86	1.1257 (0.0277)	488.99
10	0.52	0.5326 (0.0036)	0.5680 (0.0048)	0.9581 (0.0139)	218.83	0.5441 (0.0126)	457.90
11	0.2137	0.2227 (0.0021)	0.2665 (0.0028)	0.6346 (0.0091)	231.67	0.2315 (0.0160)	453.64
12	0.082	0.0816 (0.0014)	0.1059 (0.0026)	0.3393 (0.0054)	230.50	0.0778 (0.0141)	451.28

Note: The strike price is 10, initial volatility V_0 is 0.0625, risk free rate r is 0.1, and time to maturity is 0.25. $\tilde{\alpha} = 5$, $\tilde{\beta} = 0.16$, $\sigma = 0.9$ and $\rho = 0.1$. We use 25 time steps, 30 hypercubes and 3000 simulations for each hypercube. For each single run of the LSM algorithm and standard RBSDE algorithm, we use 30,000 simulations.

Table 4.5: Two Dimensional American Put Option Prices (c)

S_0	PDE	LSM (std)	RBSDE (std)	LP1:SRODP(std)	Time(s)	LP2:SRODP(std)	Time(s)
8	2	1.9909 (0.0009)	1.9909 (0.0010)	2.0142 (0.0116)	165.05	2.0052 (0.0121)	500.76
9	1.1076	1.1055 (0.0032)	1.1332 (0.0029)	1.2060 (0.0219)	167.71	1.1226 (0.0362)	500.71
10	0.52	0.5326 (0.0027)	0.5682 (0.0036)	0.5237 (0.0307)	169.05	0.5501 (0.0499)	499.92
11	0.2137	0.2215 (0.0019)	0.2672 (0.0024)	0.2201 (0.0143)	168.17	0.2273 (0.0106)	500.16
12	0.082	0.0808 (0.0012)	0.1054 (0.0015)	0.1254 (0.0071)	167.26	0.0828 (0.0113)	500.31

Note: The strike price is 10, initial volatility V_0 is 0.0625, risk free rate r is 0.1, and time to maturity is 0.25. $\tilde{\alpha} = 5$, $\tilde{\beta} = 0.16$, $\sigma = 0.9$ and $\rho = 0.1$. We use 25 time steps, 50 hypercubes and 1000 simulations for each hypercube. For each single run of the LSM algorithm and standard RBSDE algorithm, we use 50,000 simulations.

Table 4.6: Two Dimensional American Put Option Prices (d)

S_0	PDE	LSM (std)	RBSDE (std)	LP1:SRODP(std)	Time(s)	LP2:SRODP(std)	Time(s)
8	2	1.9909 (0.0009)	1.9909 (0.0010)	2.0110 (0.0054)	652.96	1.9967 (0.0067)	1299.77
9	1.1076	1.1055 (0.0032)	1.1332 (0.0029)	1.1945 (0.0081)	669.21	1.1130 (0.0305)	1359.20
10	0.52	0.5326 (0.0027)	0.5682 (0.0036)	0.5101 (0.0181)	635.60	0.5315 (0.0158)	1296.12
11	0.2137	0.2215 (0.0019)	0.2672 (0.0024)	0.2220 (0.0077)	638.56	0.2241 (0.0080)	1285.82
12	0.082	0.0808 (0.0012)	0.1054 (0.0015)	0.1246 (0.0053)	635.32	0.0828 (0.0036)	1286.07

Note: The strike price is 10, initial volatility V_0 is 0.0625, risk free rate r is 0.1, and time to maturity is 0.25. $\tilde{\alpha} = 5$, $\tilde{\beta} = 0.16$, $\sigma = 0.9$ and $\rho = 0.1$. We use 25 time steps, 50 hypercubes and 3000 simulations for each hypercube. For each single run of the LSM algorithm and standard RBSDE algorithm, we use 50,000 simulations.

Table 4.7: Two Dimensional American Put Option Prices (e)

S_0	PDE	LSM (std)	RBSDE (std)	LP1:SRODP(std)	Time(s)	LP2:SRODP(std)	Time(s)
8	2	1.9908 (0.0006)	1.9908 (0.0009)	1.9930 (0.0042)	1367.99	1.9932 (0.0047)	3855.30
9	1.1076	1.1063 (0.0021)	1.1325 (0.0022)	1.1375 (0.0105)	1325.64	1.1172 (0.0136)	3856.99
10	0.52	0.5320 (0.0021)	0.5692 (0.0030)	0.5387 (0.0129)	1307.86	0.5306 (0.0108)	3866.87
11	0.2137	0.2210 (0.0016)	0.2670 (0.0019)	0.2393 (0.0084)	1308.15	0.2211 (0.0050)	3849.48
12	0.082	0.0801 (0.0009)	0.1058 (0.0015)	0.0868 (0.0027)	1307.39	0.0812 (0.0032)	3851.47

Note: The strike price is 10, initial volatility V_0 is 0.0625, risk free rate r is 0.1, and time to maturity is 0.25. $\tilde{\alpha} = 5$, $\tilde{\beta} = 0.16$, $\sigma = 0.9$ and $\rho = 0.1$. We use 25 time steps, 80 hypercubes and 3000 simulations for each hypercube. For each single run of the LSM algorithm and standard RBSDE algorithm, we use 80,000 simulations.

4.4.3 American Put Option Prices with Ambiguity

One Dimensional Case

For the one dimensional case with ambiguity, for example under κ-*ignorance*, X. Cheng & Riedel (2013) and Vorbrink (2011) argue that the worst case evaluation is achieved when $\theta = -\kappa$. This means that the density generator will stay invariant, enabling us to directly simulate the state variables under the worst case measure and using the LSM results as the benchmark to test the accuracy of our proposed algorithms, as the LSM results are known to be close enough to *the exact value* for simple one dimensional case.

In this part we initially differentiate the SRODP and SLSM algorithm. The implementation of the SLSM algorithm is introduced in Section 4.3.4. As the drift ambiguity is introduced in the dynamics of state variables, it adds an extra dimension (the stochastic exponential M_t) to the state variables. One should choose the number of hypercubes per dimension as appropriate, since the number of total hypercubes grows geometrically when raising dimensions. As we show in Section 4.4.2, regressions with local polynomials up to the second order (LP2) have superior performance compared to LP1. Hence, we present results with LP2 scheme. $\kappa = 0.3$, so the ambiguity interval, i.e. the set Θ for the density generator is $[-0.3, 0.3]$. We run the LSM and SLSM algorithm for 10 times. The result in 4.8 shows that the SLSM algorithm works well in one dimensional case with ambiguity, given the benchmark result by the LSM algorithm. The other parameters are in line with experiments in Section 4.4.2. For the SLSM algorithm, we choose the space domain for the logarithmic stochastic exponential to be $[-6, 6]$.[4]

[4]We have tested several different lengths for the space domain of the logarithmic stochastic exponential, and the results are close to each other. They are available upon request.

Table 4.8: One Dimensional American Put Option Prices with Ambiguity (a)

S_0	LSM (std)	SLSM (std)	95% C.I.	Time(s)
36	3.8243 (0.0134)	3.8214 (0.0531)	[3.7886, 3.8543]	578.87
38	2.6001 (0.0160)	2.5700 (0.0336)	[2.5492, 2.5908]	584.49
40	1.7188 (0.0199)	1.7179 (0.0366)	[1.6952, 1.7406]	583.62
42	1.1169 (0.0101)	1.1178 (0.0273)	[1.1009, 1.1347]	580.02
44	0.7137 (0.0051)	0.7158 (0.0162)	[0.7058, 0.7259]	580.76

Note: The strike price is 40, volatility is 0.2, risk free rate is 0.06, time to maturity is 1, and $\theta_t \in [-0.3, 0.3]$. We use 5 time steps. For the SLSM algorithm, we use 50 hypercubes and 2800 simulations for each hypercube, and 11 uniformly selected points ($L = 11$) for the density generator. For each single run of the LSM algorithm, we use 30,000 simulations.

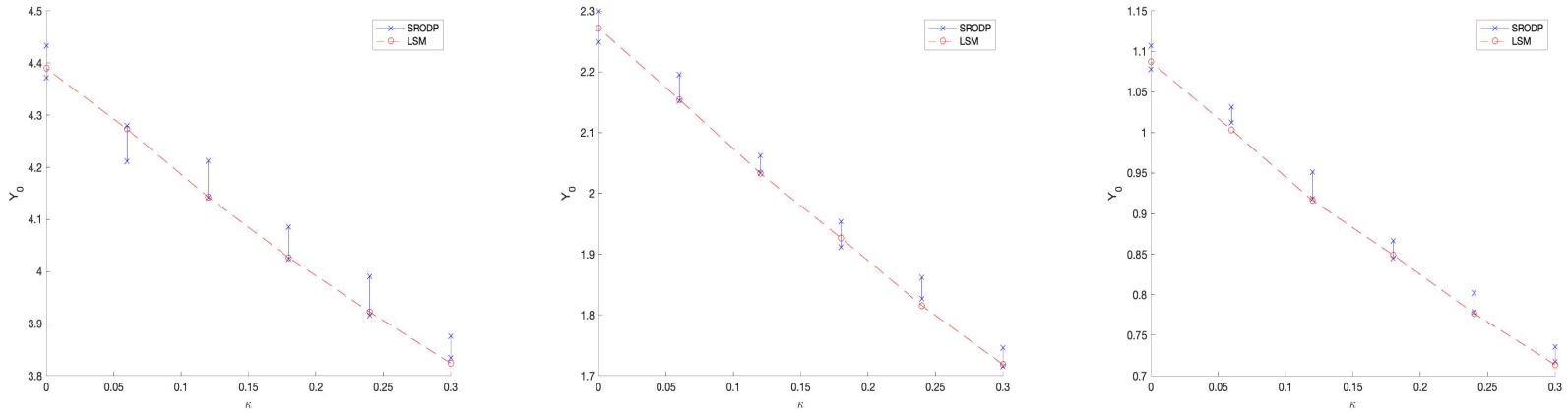
To make the SRODP and the SLSM algorithms directly comparable, we should obtain results for both algorithms, which are computed at approximate levels of computational cost. We conduct several experiments and find that the SRODP algorithm with LP2 scheme and 4000 hypercubes and 2000 simulations per hypercube will have similar computational cost. Table 4.9 summarizes the results for one dimensional American put options prices with drift ambiguity by the SRODP algorithm. Compared with those prices by the SLSM algorithm in Table 4.8, the SRODP algorithm have an evident advantage in reducing standard deviations. Later we will see that advantage is even more remarkable in two dimensional case. Taken the LSM results as our benchmark, we also find that the SRODP results are closer to the benchmark than the standard RBSDE results, which are generally upward biased.

Table 4.9: One Dimensional American Put Option Prices with Ambiguity (b)

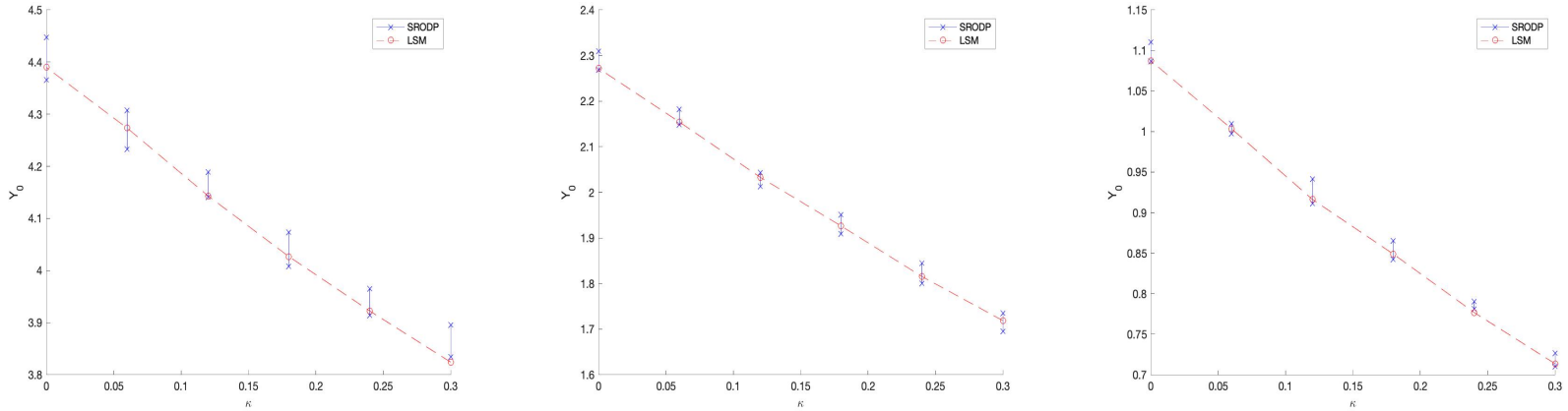
S_0	RBSDE (std)	SRODP (std)	95% C.I.	Time(s)
36	3.8787 (0.0145)	3.8429 (0.0412)	[3.8173, 3.8684]	583.50
38	2.6490 (0.0120)	2.6036 (0.0262)	[2.5874, 2.6199]	582.12
40	1.7747 (0.0127)	1.7218 (0.0307)	[1.7028, 1.7408]	583.25
42	1.1782 (0.0078)	1.1285 (0.0122)	[1.1209, 1.1360]	583.59
44	0.7758 (0.0051)	0.7126 (0.0141)	[0.7039, 0.7213]	589.65

Note: The strike price is 40, volatility is 0.2, risk free rate is 0.06, time to maturity is 1, and $\theta_t \in [-0.3, 0.3]$. We use 5 time steps. For the SRODP algorithm, we use 4000 hypercubes and 3000 simulations for each hypercube. For each single run of the standard RBSDE algorithm, we use 30,000 simulations.

Figure 4.1: One Dimensional American Put Option Prices by LP0 and LP1 Schemes



(a) LP0:$S_0 = 36$ (b) LP0:$S_0 = 40$ (c) LP0:$S_0 = 44$



(a) LP1:$S_0 = 36$ (b) LP1:$S_0 = 40$ (c) LP1:$S_0 = 44$

Note: The y-axis stands for American put option prices with strike price 40, volatility 0.2, risk free rate 0.06 and time to maturity 1. The x-axis stands for the radius (κ) of ambiguity interval. We use 5 time steps, 1000 hypercubes and 2000 simulations for each hypercube for the SRODP algorithm.

Besides, we find that regressions in the SRODP algorithm with only constant local polynomials (LP0) have decent convergence results in one dimensional cases as well. We present results comparison for LP1 and LP0 in Figure 4.1. Almost all the LSM results lie in the 95% confidence interval of the corresponding SRODP results. It is notable that put options prices with ambiguity will increase to prices without ambiguity when ambiguity interval shrinks to 0 length. This finding is also in accordance with theoretical arguments in Nishimura & Ozaki (2007).[5]

Two Dimensional Case

As for Heston's model with ambiguity, the space domain remains identical to the case without ambiguity. We employ the *elliptical ambiguity* as the form of uncertainty set. The reason is that the *elliptical ambiguity* resembles the form of statistical uncertainty given by the Wald's test when implementing maximum likelihood estimator for model calibration, as claimed by Cohen & Tegnér (2017). We run the SLSM algorithm for 10 times, and the space domain for the logarithmic stochastic exponential is $[-6, 6]$. Results in Table 4.10 are obtained by regressions with LP2 scheme. Σ and χ are given as,

$$\Sigma = \begin{bmatrix} 4/9 & 0 \\ 0 & 0.01 \end{bmatrix}, \quad \chi = 9. \tag{4.34}$$

Table 4.10: Two Dimensional American Put Option Prices with Ambiguity (a)

S_0	RBSDE (std)	SLSM (std)	95% C.I.	Time(s)
8	1.9041 (0.0062)	1.9323 (0.0179)	[1.9212, 1.9434]	7867.27
9	0.9977 (0.0064)	1.0697 (0.0107)	[1.0631, 1.0763]	7883.07
10	0.4268 (0.0062)	0.5173 (0.0144)	[0.5084, 0.5262]	7832.34
11	0.1621 (0.0042)	0.2004 (0.0194)	[0.1884, 0.2124]	7878.48
12	0.0521 (0.0023)	0.0445 (0.0105)	[0.0379, 0.0510]	7844.89

Note: The strike price is 10, initial volatility V_0 is 0.0625, risk free rate r is 0.1, and time to maturity is 0.25. $\tilde{\alpha} = 5$, $\tilde{\beta} = 0.16$, $\sigma = 0.9$ and $\rho = 0.1$. We use 5 time steps. For the SLSM algorithm, we use 20 hypercubes and 3000 simulations for each hypercube, and randomly selected 23 discrete points ($L = 23$) in the ellipse. For the standard RBSDE algorithm, we use 10,000 simulations. We launch 10 times of all the algorithms.

For a similar computational budget, we implement the SRODP algorithm with 150 hypercubes each dimension and 3000 simulations per hypercube. The standard RBSDE results

[5]The linear shape of these slopes is directly related to the Greek Rho, which is a constant here.

are also presented in Table 4.10, 4.11 and 4.12, but they are generally upward biased, as we see in one dimensional case. We observe that the SRODP results in 4.11 have much smaller standard deviations than the SLSM results. Further, we implement the SLSM algorithm with 30 hypercubes and 3000 simulations per hypercube, and present the results in Table 4.12. Results in Table 4.11 (SRODP) and in Table 4.12 (SLSM) have close standard deviations, but the SLSM algorithm takes almost 3 times computational time in order to reach the convergence, impling that the SRODP algorithm is more efficient in evaluating the American options. It should be noted that we do not have an *exact value* here, as we argue in the Section 4.3.4. However, results for the SRODP and SLSM algorithm should be closer as we increase the number of hypercubes or the simulated paths per hypercube. We can obtain more accurate results for the SLSM algorithm when increase the number of selected points in the set of density generator, as this algorithm is essentially an optimization in pointwise sense.

Table 4.11: Two Dimensional American Put Option Prices with Ambiguity (b)

S_0	RBSDE (std)	SRODP (std)	95% C.I.	Time(s)
8	1.9044 (0.0049)	1.8866 (0.0060)	[1.8829, 1.8903]	7944.22
9	0.9994 (0.0043)	0.9826 (0.0099)	[0.9765, 0.9886]	7990.26
10	0.4306 (0.0059)	0.4181 (0.0069)	[0.4139, 0.4224]	7959.31
11	0.1627 (0.0032)	0.1522 (0.0052)	[0.1490, 0.1554]	7969.27
12	0.0524 (0.0016)	0.0481 (0.0018)	[0.0470, 0.0492]	7981.40

Note: The strike price is 10, initial volatility V_0 is 0.0625, risk free rate r is 0.1, and time to maturity is 0.25. $\tilde{\alpha} = 5$, $\tilde{\beta} = 0.16$, $\sigma = 0.9$ and $\rho = 0.1$. We use 5 time steps. For the SRODP algorithm, we use 150 hypercubes and 3000 simulations for each hypercube. For the standard RBSDE algorithm, we use 15,000 simulations. We launch 10 times of all the algorithms.

Table 4.12: Two Dimensional American Put Option Prices with Ambiguity (c)

S_0	RBSDE (std)	SLSM (std)	95% C.I.	Time(s)
8	1.9074 (0.0047)	1.9267 (0.0114)	[1.9197, 1.9338]	20576.43
9	0.9974 (0.0040)	1.0394 (0.0083)	[1.0343, 1.0446]	20587.83
10	0.4316 (0.0037)	0.4774 (0.0110)	[0.4706, 0.4843]	20586.73
11	0.1651 (0.0029)	0.1667 (0.0105)	[0.1602, 0.1731]	21042.48
12	0.0531 (0.0014)	0.0422 (0.0047)	[0.0393, 0.0451]	20511.94

Note: The strike price is 10, initial volatility V_0 is 0.0625, risk free rate r is 0.1, and time to maturity is 0.25. $\tilde{\alpha} = 5$, $\tilde{\beta} = 0.16$, $\sigma = 0.9$ and $\rho = 0.1$. We use 5 time steps. For the SLSM algorithm, we use 30 hypercubes and 3000 simulations for each hypercube, and randomly selected 23 discrete points ($L = 23$) in the ellipse. For the standard RBSDE algorithm, we use 20,000 simulations. We launch 10 times of all the algorithms.

4.5 An Application in Real Options: Optimal Fish Harvesting Decision

In this section we discuss the optimal fish harvesting and corresponding fish farm evaluation problem. Generally, the fish farmer, or the farm manager, acts rationally to maximize the future benefits by choosing an optimal harvesting time. Meanwhile, the manager evaluates the value of a single lease of farm, which is the case of single rotation, or the value of the farm ownership, i.e. infinite number of rotations. The evaluation of fish harvesting is of some similarities to the American option pricing. For an explicit description of the problem, one can refer to Ewald et al. (2016) and Asche & Bjorndal (2011).

We are interested in evaluating fish farm with ambiguity in the single-rotation scenario. This means that the manager will earn the revenue of harvesting fish while incurring feeding and harvesting cost, but will have to return the fish farm to the original owner when completing a single harvesting cycle. Essentially, the manager decides whether to harvest or postpone it to future, by weighing the instant harvesting benefits against future expected benefits. In the absence of an actively trading market, the probability measure of expectation used for evaluation is subject to the manager's personal belief. We use the market pricing measure Q as our reference measure, as it has been used by Ewald et al. (2016) for the fish harvesting problem without ambiguity.[6]

We start with the problem without ambiguity. Specifically, we use a two-factor model for the dynamics of state variables under Q measure, following Ewald et al. (2016),

$$\begin{aligned} \mathrm{d}S_t/S_t &= (r-\delta_t)\mathrm{d}t + \sigma_1(\hat{\rho}\mathrm{d}\tilde{W}_t^1 + \sqrt{1-\hat{\rho}^2}\mathrm{d}\tilde{W}_t^2), \\ \mathrm{d}\delta_t &= \hat{\kappa}(\hat{\alpha}-\delta_t-\hat{\lambda})\mathrm{d}t + \sigma_2\mathrm{d}\tilde{W}_t^1, \end{aligned} \tag{4.35}$$

where $r, \sigma_1, \hat{\rho}, \hat{\kappa}, \hat{\alpha}, \hat{\lambda}$ and σ_2 are constants.[7] Further, we assume that the average weight w_t

[6] In this chapter we only consider the single-rotation case in order to demonstrate the applicability of our approaches to the two-dimensional real option problem. It can be naturally extended to the infinite-rotation case in a similar way.

[7] Some of symbols here are the same to our previous notations for Heston's model. We hope this will not confuse readers.

of an individual fish and the total number n_t of fish in a farm follow,

$$w_t = w_\infty(\hat{a} - \hat{b}\exp\{-\hat{c}t\})^3, \tag{4.36}$$

$$\mathrm{d}n_t = -\hat{m}_t n_t \mathrm{d}t, \tag{4.37}$$

where the equation (4.36) is known as the Von Bertalanffy growth function, which is extensively used in fisheries biology(see for example Haddon (2010)), and w_∞ is the asymptotic weight of fish. $\hat{a}$, $\hat{b}$ and $\hat{c}$ are constants. $\hat{m}_t$ represents the mortality rate and is assumed to be constant $\hat{m}$ here for simplicity. By solving an ordinary differential equation (4.37) we have

$$n_t = n_0 \exp\{-\hat{m}t\}. \tag{4.38}$$

Therefore, the biomass $\hat{M}_t$ of the fish farm is given by

$$\hat{M}_t = n_t w_t. \tag{4.39}$$

We assume the fish farmer begins with a certain amount of infant fish, thus there is no release cost. Then the value function $\hat{J}_t$ is naturally defined by maximizing the expected harvesting benefits

$$\hat{J}_t = \operatorname*{ess\,sup}_{\hat{\tau}\in\mathscr{T}_t} \mathbb{E}^Q[e^{-r(\hat{\tau}-t)}(S_{\hat{\tau}}\hat{M}_{\hat{\tau}} - C_1\hat{M}_{\hat{\tau}}) - \int_t^{\hat{\tau}} e^{-r(s-t)}C_2F_sn_s\mathrm{d}s|\mathcal{F}_t], \tag{4.40}$$

where $\mathscr{T}_t$ is defined in Section 4.2.4. C_1 stands for the harvesting cost per kilogram and C_2 is the feeding cost per kilogram per year. $F_t = \hat{f}w_t'$, where w_t' is the first order derivative of w_t representing the fish growth rate, and $\hat{f}$ is the feed conversion ratio, according to Asche & Bjorndal (2011). Applying the dynamic programming principle, we obtain a discrete-time approximation of the Hamilton-Jacobi-Bellman equation after equally discretize the time horizon $[0, T]$ with each part to be Δt.

$$\hat{J}_t = \max\left\{S_t\hat{M}_t - C_1\hat{M}_t,\ e^{-r\Delta t}\mathbb{E}^Q[\hat{J}_{t+\Delta t}|\mathcal{F}_t] - \Delta t C_2F_tn_t\right\}. \tag{4.41}$$

Proof. From (4.40) we have

$$
\begin{aligned}
\hat{J}_t =& \max\Big\{S_t\hat{M}_t - C_1\hat{M}_t, \max_{t+\Delta t\le\hat{\tau}\le T}\mathbb{E}^Q[e^{-r(\hat{\tau}-t)}(S_{\hat{\tau}}\hat{M}_{\hat{\tau}} - C_1\hat{M}_{\hat{\tau}}) \\
&- \int_t^{\hat{\tau}} e^{-r(s-t)}C_2F_sn_s\mathrm{d}s|\mathcal{F}_t]\Big\}, \\
=& \max\Big\{S_t\hat{M}_t - C_1\hat{M}_t, \max_{t+\Delta t\le\hat{\tau}\le T} e^{-r\Delta t}\mathbb{E}^Q[e^{-r(\hat{\tau}-t-\Delta t)}(S_{\hat{\tau}}\hat{M}_{\hat{\tau}} - C_1\hat{M}_{\hat{\tau}}) \\
&- \int_{t+\Delta t}^{\hat{\tau}} e^{-r(s-t-\Delta t)}C_2F_sn_s\mathrm{d}s|\mathcal{F}_t] - \int_t^{t+\Delta t} e^{-r(s-t)}C_2F_sn_s\mathrm{d}s\Big\}, \\
=& \max\Big\{S_t\hat{M}_t - C_1\hat{M}_t, e^{-r\Delta t}\mathbb{E}^Q[\max_{t+\Delta t\le\hat{\tau}\le T}\mathbb{E}^Q[e^{-r(\hat{\tau}-t-\Delta t)}(S_{\hat{\tau}}\hat{M}_{\hat{\tau}} - C_1\hat{M}_{\hat{\tau}}) \\
&- \int_{t+\Delta t}^{\hat{\tau}} e^{-r(s-t-\Delta t)}C_2F_sn_s\mathrm{d}s|\mathcal{F}_{t+\Delta t}]|\mathcal{F}_t] - \int_t^{t+\Delta t} e^{-r(s-t)}C_2F_sn_s\mathrm{d}s\Big\}, \\
=& \max\Big\{S_t\hat{M}_t - C_1\hat{M}_t, e^{-r\Delta t}\mathbb{E}^Q[\hat{J}_{t+\Delta t}|\mathcal{F}_t] - \Delta tC_2F_tn_t\Big\},
\end{aligned}
$$

where the first equality holds by splitting the investment decision (4.40) between harvesting now (at time t) and wait for a short time period and reevaluate the fish farm later (at time $t+\Delta t$); the second equality holds by the fact that $F(\cdot)$ and $n(\cdot)$ are deterministic functions; the third equality holds by the tower law of the conditional expectation; the last equality holds by the definition of value function $\hat{J}_{t+\Delta t}$, and approximating $\int_t^{t+\Delta t} e^{-r(s-t)}C_2F_sn_s\mathrm{d}s$ by $\Delta tC_2F_tn_t$(This is justified when Δt goes to zero and terms higher than Δt are eliminated).

□

Such a discrete-time dynamic programming algorithm (4.41) enables us to use numerical methods to evaluate the fish farm, for example, the LSM algorithm. In the following proposition we show that the value function $\hat{J}_t$ coincides with the solution of an RBSDE, which is similar to (4.13).

Proposition 4.5.1. *Consider the RBSDE*

$$
-\mathrm{d}Y_t = g(t, Y_t, Z_t)\mathrm{d}t + \mathrm{d}K_t - Z_t^*\mathrm{d}\tilde{W}_t,\ Y_T = S_T\hat{M}_T - C_1\hat{M}_T, \tag{4.42}
$$

with the constraints

$$
Y_t \ge L_t,\ \int_0^T (Y_s - L_t)\mathrm{d}K_t = 0 \text{ and } K_0 = 0.
$$

Assume that the generator $g(t, Y_t, Z_t) = -rY_t - C_2F_tn_t$*, the obstacle* $L_t = S_t\hat{M}_t - C_1\hat{M}_t$ *and* K_t *is a continuous and increasing process.*

There exists a unique solution $\{(Y_t, Z_t, K_t),\ t \in [0,T]\}$ *of the above RBSDE, which are* $\mathcal{F}_t$ *progressively measurable processes. The process* Y_t *has the dual representation*

$$Y_t = \operatorname*{ess\,sup}_{\hat{\tau}\in\mathscr{T}_t} \mathbb{E}^Q[e^{-r(\hat{\tau}-t)}(S_{\hat{\tau}}\hat{M}_{\hat{\tau}} - C_1\hat{M}_{\hat{\tau}}) - \int_t^{\hat{\tau}} e^{-r(s-t)}C_2F_sn_s\mathrm{d}s|\mathcal{F}_t],\ t\in[0,T]. \quad (4.43)$$

Proof. The above RBSDE can be taken as a special case in Proposition 2.3 of El Karoui, Kapoudjian, et al. (1997). To show the solution corresponds to the value of an optimal stopping problem, it suffices to show that the RBSDE satisfies the technical conditions in Definition 4.2.1. As $F(\cdot)$ and $n(\cdot)$ are deterministic and continuous functions, the generator is uniform Lipschitz, and belongs to $\mathbb{H}^2$. $\hat{M}_t$ is deterministic and continuous, meaning that it is bounded in $[0,T]$. Hence, it can be seen that L_t is in $\mathbb{S}^2$. The terminal value $Y_T = L_T$ again is in $\mathbb{S}^2$, impling that $Y_T \in \mathbb{L}^2$. Thus, the solution to the above RBSDE has a probabilistic representation form (4.43). Then, the uniqueness follows directly from Theorem 5.2 in El Karoui, Kapoudjian, et al. (1997). □

Without ambiguity, the Proposition 4.5.1 allows us to evaluate a single-rotation fish farm by solving an RBSDE. However, the probability measure used for fish farm evaluation is not restricted to the market pricing measure, as we argue that usually there is no actively trading market for that. There is no need to concern about the arbitrage opportunity. Changing the evaluation measure Q to Q^θ defined in Section 4.2.2, the following linear RBSDE

$$-\mathrm{d}Y_t^\theta = g(t, Y_t^\theta, Z_t^\theta)\mathrm{d}t + \mathrm{d}K_t^\theta - Z_t^{\theta *}\mathrm{d}\tilde{W}_t,\ Y_T^\theta = S_T\hat{M}_T - C_1\hat{M}_T, \quad (4.44)$$

with generator $g(t, Y_t^\theta, Z_t^\theta) = -rY_t^\theta - C_2F_tn_t - \theta_t^* Z_t^\theta$ and obstacle $L_t = S_t\hat{M}_t - C_1\hat{M}_t$ has a unique solution $(Y_t^\theta, Z_t^\theta, K_t^\theta),\ t\in[0,T]$. This claim is analogical to Proposition 4.2.1, so we skip the proof here. In addition, Y^θ has the representation

$$Y_t^\theta = \operatorname*{ess\,sup}_{\hat{\tau}\in\mathscr{T}_t} \mathbb{E}^\theta[e^{-r(\hat{\tau}-t)}(S_{\hat{\tau}}\hat{M}_{\hat{\tau}} - C_1\hat{M}_{\hat{\tau}}) - \int_t^{\hat{\tau}} e^{-r(s-t)}C_2F_sn_s\mathrm{d}s|\mathcal{F}_t]. \quad (4.45)$$

When considering drift ambiguity, of which the ambiguity sets $\mathscr{P}^\Theta$ are convex and compact, we can still solve the worst case evaluation problem via solving an RBSDE. The RBSDE and duality arguments are analogical to that in the American option case. We begin

with defining the value of a fish farm in the worst case

$$\hat{v}_t := \operatorname*{ess\,sup}_{\hat{\tau}\in\mathscr{T}_t} \operatorname*{ess\,inf}_{Q^\theta\in\mathscr{P}^\Theta} \mathbb{E}^{Q^\theta}[e^{-r(\hat{\tau}-t)}(S_{\hat{\tau}}\hat{M}_{\hat{\tau}} - C_1\hat{M}_{\hat{\tau}}) - \int_t^{\hat{\tau}} e^{-r(s-t)}C_2F_sn_s\mathrm{d}s|\mathcal{F}_t],\ t\in[0,T]. \tag{4.46}$$

Proposition 4.5.2. *Under the κ-ignorance and the elliptical ambiguity framework as in the section 4.2.2, the fish farm vale $\hat{v}_t$ under the worst case belief as defined in (4.46) is the value of a minimax (optimal) control problem such that,*

$$\hat{v}_t = \operatorname*{ess\,inf}_{\theta\in\Theta} Y_t^\theta, \tag{4.47}$$

and we have,

$$\operatorname*{ess\,sup}_{\hat{\tau}\in\mathscr{T}_t} \operatorname*{ess\,inf}_{\theta\in\Theta} \mathbb{E}^\theta[\hat{H}_{t,\hat{\tau}}|\mathcal{F}_t] = \operatorname*{ess\,inf}_{\theta\in\Theta} \operatorname*{ess\,sup}_{\hat{\tau}\in\mathscr{T}_t} \mathbb{E}^\theta[\hat{H}_{t,\hat{\tau}}|\mathcal{F}_t],\ \forall t\in[0,T], \tag{4.48}$$

where Y_t^θ is an element of the solution to (4.44), and $\hat{H}_{t,\hat{\tau}} = e^{-r(\hat{\tau}-t)}(S_{\hat{\tau}}\hat{M}_{\hat{\tau}} - C_1\hat{M}_{\hat{\tau}}) - \int_t^{\hat{\tau}} e^{-r(s-t)}C_2F_sn_s\mathrm{d}s$. Moreover, there exists a pair $(\tilde{\tau},\tilde{\theta}_t)$, $\tilde{\tau}\in[0,T]$, $\tilde{\theta}\in\Theta$ such that it reaches the optimal (saddle) point $(\tilde{\tau},\tilde{\theta}_t)$ when the generator $g(t,Y_t^{\tilde{\theta}},Z_t^{\tilde{\theta}}) = \operatorname{ess\,inf}_{\theta\in\Theta} g(t,Y_t^\theta,Z_t^\theta)$, $\forall t\in[0,T]$.

Proof. As the proof is almost identical to that of Proposition 4.2.2 except for the payoff function and generator, we skip it here. □

Specifically, under *elliptical ambiguity* $\Theta = \{\theta:\ \theta\Sigma^{-1}\theta^*\le\chi\}$, the optimal generator is

$$g(t,Y_t^{\tilde{\theta}},Z_t^{\tilde{\theta}}) = -rY_t^{\tilde{\theta}} - \sqrt{Z_t^{\tilde{\theta}^*}\Sigma^*Z_t^{\tilde{\theta}}\chi} - C_2F_tn_t, \tag{4.49}$$

$$\text{with } \tilde{\theta} = -\frac{Z_t^{\tilde{\theta}^*}\Sigma^*}{\sqrt{Z_t^{\tilde{\theta}^*}\Sigma^*Z_t^{\tilde{\theta}}/\chi}}. \tag{4.50}$$

Provided the optimal generator (4.49) and the duality arguments (4.48), it is direct that the RBSDE (4.44) with a such generator has a unique solution, as we have an analogical argument in Remark 4.2.1. Therefore, we can utilize the previous SRODP algorithm(in Section 4.3.2) and SLSM algorithm(in Section 4.3.4) to evaluate the fish farm. Parameters in Table 4.13 are extracted from Ewald et al. (2016) for the numerical experiments. We start by demonstrating examples without ambiguity in Table 4.14, in order to show that results are close under different algorithms.

Table 4.13: Parameters for Fish Farm

Parameter	Value	Parameter	Value	Parameter	Value
$\hat{m}$	10%	C_2	7	$\hat{\alpha}$	1.135
$\hat{f}$	1.1	$\hat{a}$	1.113	$\hat{\lambda}$	1.142
n_0	1	$\hat{b}$	1.097	$\hat{\rho}$	0.736
ω_∞	6	$\hat{c}$	1.43	σ_1	0.153
C_1	3	$\hat{\kappa}$	1.012	σ_2	0.206

Table 4.14: Fish Farm Value without Ambiguity

S_0	LSM (std)	RBSDE (std)	SRODP (std)	Time(s)
8	1.3123 (0.0013)	1.3122 (0.0017)	1.3122 (0.0035)	175.26
20	4.5613 (0.0044)	4.5618 (0.0041)	4.5596 (0.0072)	175.69
32	7.8061 (0.0076)	7.8071 (0.0060)	7.8161 (0.0135)	176.26

Note: The initial convenience yield V_0 is -0.3, risk free rate r is 0.0393, and time to maturity is 0.25. We use 5 time steps. For the SRODP algorithm, we use 50 hypercubes and 3000 simulations for each hypercube. For the LSM and standard RBSDE algorithm, we use 10,000 simulations. We launch 10 times of all the algorithms.

In ambiguous case, we adopt the same parameters for the uncertainty set as we define in (4.34). Still, we can observe from Table 4.15 and Table 4.16 that results by the SRODP algorithm have smaller standard deviations than the SLSM algorithm, given approximate computational budgets. Despite that, the SLSM algorithm can be used in cases when the RBSDE technique is not suitable. For example, the RBSDE technique does not apply when the optimal generator cannot be obtained explicitly, as the existence in Proposition 4.2.2 and Proposition 4.5.2 relies heavily on the compactness of the uncertainty set. It is not possible to use the RBSDE technique when the value function is not the solution of an RBSDE. Yet these are left for future research.

Table 4.15: Fish Farm Value with Ambiguity (a)

S_0	RBSDE (std)	SLSM (std)	95% C.I.	Time(s)
8	1.2129 (0.0090)	1.1832 (0.0216)	[1.1698, 1.1966]	3835.66
20	4.2897 (0.0280)	4.0111 (0.0687)	[3.9685, 4.0536]	4094.91
32	7.3635 (0.0535)	6.8086 (0.1221)	[6.7329, 6.8843]	4099.33

Note: The initial convenience yield V_0 is -0.3, risk free rate r is 0.0393, and time to maturity is 0.25. We use 5 time steps. For the SLSM algorithm, we use 20 hypercubes and 2000 simulations for each hypercube. For the standard RBSDE algorithm, we use 10,000 simulations. We launch 10 times of all the algorithms.

Table 4.16: Fish Farm Value with Ambiguity (b)

S_0	RBSDE (std)	SRODP (std)	95% C.I.	Time(s)
8	1.2188 (0.0039)	1.1884 (0.0104)	[1.1820, 1.1949]	4731.34
20	4.3066 (0.0198)	4.1704 (0.0272)	[4.1536, 4.1873]	4682.14
32	7.4226 (0.0481)	7.1065 (0.0354)	[7.0846, 7.1284]	4725.71

Note: The initial convenience yield V_0 is -0.3, risk free rate r is 0.0393, and time to maturity is 0.25. We use 5 time steps. For the SRODP algorithm, we use 70 hypercubes and 5000 simulations for each hypercube. For the standard RBSDE algorithm, we use 20,000 simulations. We launch 10 times of all the algorithms.

4.6 Concluding Remarks

In this chapter, we explore the evaluation of American options with stochastic volatility and single-rotation fish farms with stochastic convenience yield under drift ambiguity framework. We formulate the value function to the solution of a reflected backward differential equations(RBSDEs) and prove the uniqueness of the solution. Moreover, we propose an algorithm(Stratified Regression One-step Forward Dynamic Programming) to numerically solve RBSDEs, combining the traditional numerical RBSDE method by Gobet & Lemor (2008) with a general stratification approach by Gobet et al. (2016). However, the RBSDE approach relies heavily on the explicit formulation of the generators. We also raise another possible numerical algorithm(Stratified Least Square Monte Carlo) without using the theory of RBSDEs, taking advantage of dynamic programming and the general stratification. We conduct numerical experiments to show the convergence of two algorithms. In one dimensional case, our results are in line with the theoretical arguments in X. Cheng & Riedel (2013) and Vorbrink (2011). Further, the SRODP algorithm exhibits superior efficiency in both one and two dimensional cases.

References

Aït-Sahalia, Y. (2004). Disentangling diffusion from jumps. *Journal of Financial Economics*, *74*(3), 487–528.

Aït-Sahalia, Y., Fan, J., Laeven, R. J., Wang, C. D., & Yang, X. (2017). Estimation of the continuous and discontinuous leverage effects. *Journal of the American Statistical Association*, 1–15.

Aït-Sahalia, Y., & Jacod, J. (2009a). Estimating the degree of activity of jumps in high frequency data. *The Annals of Statistics*, 2202–2244.

Aït-Sahalia, Y., & Jacod, J. (2009b). Testing for jumps in a discretely observed process. *The Annals of Statistics*, 184–222.

Aït-sahalia, Y., Mykland, P. A., & Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies*, *18*(2), 351–416.

Andersen, L. (2010). Markov models for commodity futures: theory and practice. *Quantitative Finance*, *10*(8), 831–854.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2009). Parametric and nonparametric volatility measurement. *Handbook of Financial Econometrics: Tools and Techniques*, *1*, 67.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*(2), 579–625.

Andersen, T. G., Dobrev, D., & Schaumburg, E. (2011). A functional filtering and neighborhood truncation approach to integrated quarticity estimation. *National Bureau of Economic Research*.

Andersen, T. G., & Sørensen, B. E. (1996). Gmm estimation of a stochastic volatility model: a monte carlo study. *Journal of Business & Economic Statistics*, *14*(3), 328–352.

Anderson, B. D., & Moore, J. B. (1979). Optimal filtering. *Englewood Cliffs*, *21*, 22–95.

Arismendi, J. C., Back, J., Prokopczuk, M., Paschke, R., & Rudolf, M. (2016). Seasonal stochastic volatility: Implications for the pricing of commodity options. *Journal of Banking and Finance*, *66*(5), 53–65.

Asche, F., & Bjorndal, T. (2011). *The economics of salmon aquaculture* (Vol. 10). John Wiley & Sons.

Asche, F., Misund, B., & Oglend, A. (2016). Determinants of the futures risk premium in atlantic salmon markets. *Journal of Commodity Markets*, *2*(1), 6–17.

Avellaneda, M., Levy, A., & Parás, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, *2*(2), 73–88.

Bachelier, L. (1900). Théorie de la spéculation. In *Annales scientifiques de l'école normale supérieure* (Vol. 17, pp. 21–86).

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of Finance*, *52*(5), 2003–2049.

Balter, A. G., & Pelsser, A. (2020). Pricing and hedging in incomplete markets with model uncertainty. *European Journal of Operational Research*, *282*(3), 911–925.

Barndorff-Nielsen, O. E., & Shephard, N. (2001). Non-gaussian ornstein–uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(2), 167–241.

Barndorff-Nielsen, O. E., & Shephard, N. (2003a). Integrated ou processes and non-gaussian ou-based stochastic volatility models. *Scandinavian Journal of statistics*, *30*(2), 277–295.

Barndorff-Nielsen, O. E., & Shephard, N. (2003b). Realized power variation and stochastic volatility models. *Bernoulli*, *9*(2), 243–265.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of financial econometrics*, *2*(1), 1–37.

Barndorff-Nielsen, O. E., Shephard, N., & Winkel, M. (2006). Limit theorems for multipower variation in the presence of jumps. *Stochastic processes and their applications*, *116*(5), 796–806.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, *9*(1), 69–107.

Bates, D. S. (2000). Post-'87 crash fears in the s&p 500 futures option market. *Journal of Econometrics*, *94*(1), 181–238.

Bender, C., & Denk, R. (2007). A forward scheme for backward sdes. *Stochastic processes and their applications*, *117*(12), 1793–1812.

Bessembinder, H., & Lemmon, M. L. (2002). Equilibrium pricing and optimal hedging in electricity forward markets. *the Journal of Finance*, *57*(3), 1347–1382.

Black, F. (1976). Studies of stock market volatility changes. *1976 Proceedings of the American Statistical Association Bisiness and Economic Statistics Section*.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637–654.

Bloznelis, D. (2016). Salmon price volatility: A weight-class-specific multivariate approach. *Aquaculture economics & management*, *20*(1), 24–53.

Bollerslev, T., Gibson, M. S., & Zhou, H. (2011). Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. *Journal of Econometrics*, *160*(1), 235–245.

Bollerslev, T., & Zhou, H. (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics*, *109*(1), 33–65.

Bouchard, B., & Chassagneux, J.-F. (2008). Discrete-time approximation for continuously and discretely reflected bsdes. *Stochastic Processes and their Applications*, *118*(12), 2269–2293.

Bouchard, B., & Touzi, N. (2004). Discrete-time approximation and monte-carlo simulation of backward stochastic differential equations. *Stochastic Processes and their applications*, *111*(2), 175–206.

Brennan, M. J., & Schwartz, E. S. (1985). Evaluating natural resource investments. *Journal of Business*, 135–157.

Brockwell, P. (2001). Lévy-driven carma processes. *Annals of the Institute of Statistical Mathematics*, *53*(1), 113–124.

Brockwell, P., & Lindner, A. (2012). Lévy-driven time series models for financial data. *Handbook of Statistics*, *30*, 543–563.

Carr, P., Geman, H., Madan, D. B., & Yor, M. (2003). Stochastic volatility for lévy processes. *Mathematical Finance*, *13*(3), 345–382.

Carr, P., & Madan, D. (1999). Option valuation using the fast Fourier transform. *Journal of computational finance*, *2*(4), 61–73.

Carr, P., & Wu, L. (2004). Time-changed lévy processes and option pricing. *Journal of Financial economics*, *71*(1), 113–141.

Casassus, J., & Collin-Dufresne, P. (2005). Stochastic convenience yield implied from commodity futures and interest rates. *The Journal of Finance*, *60*(5), 2283–2331.

Chang, K.-H., & Kim, M.-J. (2001). Jumps and time-varying correlations in daily foreign exchange rates. *Journal of International Money and Finance*, *20*(5), 611–637.

Chen, Z., & Epstein, L. (2002). Ambiguity, risk, and asset returns in continuous time. *Econometrica*, *70*(4), 1403–1443.

Cheng, P., & Scaillet, O. (2007). Linear-quadratic jump-diffusion modeling. *Mathematical Finance*, *17*(4), 575–598.

Cheng, X., & Riedel, F. (2013). Optimal stopping under ambiguity in continuous time. *Mathematics and Financial Economics*, *7*(1), 29–68.

Chiu, M. C., Wong, H. Y., & Zhao, J. (2015). Commodity derivatives pricing with cointegration and stochastic covariances. *European Journal of Operational Research*, *246*(2), 476–486.

Christoffersen, P., Jacobs, K., & Ornthanalai, C. (2012). Dynamic jump intensities and risk premiums: Evidence from s&p500 returns and options. *Journal of Financial Economics*, *106*(3), 447–472.

Cohen, S. N., & Tegnér, M. (2017). European option pricing with stochastic volatility models under parameter uncertainty. In *International symposium on bsdes* (pp. 123–167).

Cont, R., & Tankov, P. (2003). *Financial modelling with jump processes.* Chapman & Hall.

Cortazar, G., Gravet, M., & Urzua, J. (2008). The valuation of multidimensional american real options using the lsm simulation method. *Computers & Operations Research*, *35*(1), 113–129.

Cortazar, G., & Schwartz, E. S. (2003). Implementing a stochastic model for oil futures prices. *Energy Economics*, *25*(3), 215–238.

Date, P., & Islyaev, S. (2016). A fast calibrating volatility model for option pricing. *European Journal of Operational Research*, *243*(2), 599–606.

Delbaen, F., & Schachermayer, W. (2006). *The mathematics of arbitrage*. Springer Science & Business Media.

Derman, E., & Kani, I. (1994). Riding on a smile. *Risk*, *7*(2), 32–39.

Dixit, A. K., & Pindyck, R. S. (1994). *Investment under uncertainty*. Princeton university press.

Doob, J. L. (1942). The brownian movement and stochastic equations. *Annals of Mathematics*, 351–369.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *The Journal of Finance*, *57*(1), 405–443.

Duffie, D. (2001). *Dynamic asset pricing theory.* Princeton University Press.

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, *68*(6), 1343–1376.

Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied volatility functions: Empirical tests. *The Journal of Finance*, *53*(6), 2059–2106.

Dupire, B. (1994). Pricing with a smile. *Risk*, *7*(1), 18–20.

El Karoui, N., Kapoudjian, C., Pardoux, É., Peng, S., & Quenez, M.-C. (1997). Reflected solutions of backward sde's, and related obstacle problems for pde's. *the Annals of Probability*, *25*(2), 702–737.

El Karoui, N., Pardoux, É., & Quenez, M. C. (1997). Reflected backward sdes and american options. *Numerical methods in finance*, *13*, 215–231.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 643–669.

Eraker, B., Johannes, M., & Polson, N. (2003). The impact of jumps in volatility and returns. *The Journal of Finance*, *58*(3), 1269–1300.

Ewald, C.-O., & Ouyang, R. (2017). An analysis of the fish pool market in the context of seasonality and stochastic convenience yield. *Marine Resource Economics*, *32*(4), 431–449.

Ewald, C.-O., Ouyang, R., & Siu, T. K. (2016). On the market-consistent valuation of fish farms: using the real option approach and salmon futures. *American Journal of Agricultural Economics*, *99*(1), 207–224.

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*, *282*(1-2), 119–130.

Fouque, J.-P., Papanicolaou, G., & Sircar, K. R. (2000). *Derivatives in financial markets with stochastic volatility*. Cambridge University Press.

Gallant, A. R. (2009). *Nonlinear statistical models* (Vol. 310). John Wiley & Sons.

Gatheral, J. (2011). *The volatility surface: a practitioner's guide* (Vol. 357). John Wiley & Sons.

Gatheral, J., Jaisson, T., & Rosenbaum, M. (2018). Volatility is rough. *Quantitative Finance*, *18*(6), 933–949.

Geman, H., & Nguyen, V.-N. (2005). Soybean inventory and forward curve dynamics. *Management Science*, *51*(7), 1076–1091.

Gibson, R., & Schwartz, E. S. (1990). Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance*, *45*(3), 959–976.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of mathematical economics*, *18*(2), 141–153.

Glasserman, P. (2013). *Monte carlo methods in financial engineering* (Vol. 53). Springer Science & Business Media.

Gobet, E. (2016). *Monte-carlo methods and stochastic processes: from linear to non-linear*. CRC Press.

Gobet, E., & Lemor, J.-P. (2008). Numerical simulation of bsdes using empirical regression methods: theory and practice. *arXiv preprint arXiv:0806.4447*.

Gobet, E., López-Salas, J. G., Turkedjiev, P., & Vázquez, C. (2016). Stratified regression monte-carlo scheme for semilinear pdes and bsdes with large scale parallelization on gpus. *SIAM Journal on Scientific Computing*, *38*(6), C652–C677.

Gobet, E., & Turkedjiev, P. (2016). Linear regression mdp scheme for discrete backward stochastic differential equations under general conditions. *Mathematics of Computation*, *85*(299), 1359–1391.

Goldfarb, D., & Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of operations research*, *28*(1), 1–38.

Haddon, M. (2010). *Modelling and quantitative methods in fisheries*. CRC press.

Hamilton, J. D. (1994). *Time series analysis* (Vol. 2). Princeton university press Princeton, NJ.

Harrison, J. M., & Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications*, *11*(3), 215–260.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, *6*(2), 327–343.

Hull, J., & White, A. (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, *42*(2), 281–300.

Ikonen, S., & Toivanen, J. (2008). Efficient numerical methods for pricing american options under stochastic volatility. *Numerical Methods for Partial Differential Equations: An International Journal*, *24*(1), 104–126.

Islyaev, S., & Date, P. (2015). Electricity futures price models: Calibration and forecasting. *European Journal of Operational Research*, *247*(1), 144–154.

Jacod, J. (1994). Limit of random measures associated with the increments of a brownian semimartingale. *Journal of Financial Econometrics*.

Jacod, J., Klüppelberg, C., & Müller, G. (2017). Testing for non-correlation between price and volatility jumps. *Journal of Econometrics*, *197*(2), 284–297.

Jacod, J., & Protter, P. (2012). *Discretization of processes. stochastic modelling and applied probability 67*. Springer-Verlag.

Jacod, J., & Todorov, V. (2010). Do price and volatility jump together? *The Annals of Applied Probability*, *20*(4), 1425–1469.

Kaldor, N. (1939). Speculation and economic stability. *The Review of Economic Studies*, *7*(1), 1–27.

Karatzas, I., & Shreve, S. (1998). *Brownian motion and stochastic calculus* (Vol. 113). Springer.

Kloeden, P. E., & Platen, E. (2013). *Numerical solution of stochastic differential equations* (Vol. 23). Springer Science & Business Media.

Klüppelberg, C., Lindner, A., & Maller, R. (2004). A continuous-time garch process driven by a lévy process: stationarity and second-order behaviour. *Journal of Applied Probability*, *41*(3), 601–622.

Knight, F. H. (1921). *Risk, uncertainty and profit* (Vol. 31). Houghton Mifflin.

Kolos, S. P., & Ronn, E. I. (2008). Estimating the commodity market price of risk for energy prices. *Energy Economics*, *30*(2), 621–641.

Lemor, J.-P., Gobet, E., & Warin, X. (2006). Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli*, *12*(5), 889–916.

Lewis, A. L. (2000). *Option valuation under stochastic volatility.* Newport Beach, CA: Finance Press.

Longstaff, F. A., & Schwartz, E. S. (2001). Valuing american options by simulation: a simple least-squares approach. *The review of financial studies*, *14*(1), 113–147.

Lyons, T. J. (1995). Uncertain volatility and the risk-free synthesis of derivatives. *Applied mathematical finance*, *2*(2), 117–133.

Ma, J., & Zhang, J. (2005). Representations and regularities for solutions to bsdes with reflections. *Stochastic processes and their applications*, *115*(4), 539–569.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping brownian motion. *Giornale dell'Istituto Italiano degli Attuari*, *64*(19-47), 44.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, *36*(2), 270–296.

McDonald, R., & Siegel, D. (1986). The value of waiting to invest. *The Quarterly Journal of Economics*, *101*(4), 707–727.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 141–183.

Moreno, M., Novales, A., & Platania, F. (2019). Long-term swings and seasonality in energy markets. *European Journal of Operational Research*, *279*(3), 1011–1023.

Mrázek, M., Pospísil, J., & Sobotka, T. (2016). On calibration of stochastic and fractional stochastic volatility models. *European Journal of Operational Research*, *254*(3), 1036–1046.

Nishimura, K. G., & Ozaki, H. (2007). Irreversible investment and knightian uncertainty. *Journal of Economic Theory*, *136*(1), 668–694.

Øksendal, B., & Sulem, A. (2007). *Applied stochastic control of jump diffusions*. Springer Science & Business Media.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics*, *63*(1), 3–50.

Pascucci, A. (2011). *PDE and martingale methods in option pricing*. Springer.

Pham, H. (2009). *Continuous-time stochastic control and optimization with financial applications* (Vol. 61). Springer.

Rambeerich, N., Tangman, D., Lollchund, M., & Bhuruth, M. (2013). A fast calibrating volatility model for option pricing. *European Journal of Operational Research*, *224*(1), 219–226.

Recchioni, M., & Sun, Y. (2016). An explicitly solvable heston model with stochastic interest rate. *European Journal of Operational Research*, *249*(1), 359–377.

Richter, M., & Sørensen, C. (2002). Stochastic volatility and seasonality in commodity futures and options: The case of soybeans. *Journal of Futures Markets*.

Rubinstein, M. (1985). Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active cboe option classes from august 23, 1976 through august 31, 1978. *The Journal of Finance*, *40*(2), 455–480.

Rubinstein, M. (1994). Implied binomial trees. *The Journal of Finance*, *49*(3), 771–818.

Samuelson, P. A. (1965a). Proof that properly anticipated prices fluctuate randomly. *Management Review*, *6*(2).

Samuelson, P. A. (1965b). Rational theory of warrant pricing. *IMR; Industrial Management Review (pre-1986)*, *6*(2), 13.

Santa-Clara, P., & Yan, S. (2010). Crashes, volatility, and the equity premium: Lessons from S&P 500 options. *The Review of Economics and Statistics*, *92*(2), 435–451.

Sato, K.-i. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge university press.

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

Schwartz, E. S., & Smith, J. E. (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science*, *46*(7), 893–911.

Shao, C., Bhar, R., & Colwell, D. B. (2015). A multi-factor model with time-varying and seasonal risk premiums for the natural gas market. *Energy Economics*, *50*, 207–214.

Solibakke, P. B. (2012). Scientific stochastic volatility models for the salmon forward market: forecasting (un-) conditional moments. *Aquaculture Economics & Management*, *16*(3), 222–249.

Stein, E. M., & Stein, J. C. (1991). Stock price distributions with stochastic volatility: an analytic approach. *The Review of Financial Studies*, *4*(4), 727–752.

Todorov, V. (2009a). Variance risk-premium dynamics: The role of jumps. *The Review of Financial Studies*, *23*(1), 345–383.

Todorov, V. (2009b). Estimation of continuous-time stochastic volatility models with jumps using high-frequency data. *Journal of Econometrics*, *148*(2), 131–148.

Todorov, V. (2011). Econometric analysis of jump-driven stochastic volatility models. *Journal of Econometrics*, *160*(1), 12–21.

Trojanowska, M., & Kort, P. M. (2010). The worst case for real options. *Journal of Optimization Theory and Applications*, *146*(3), 709–734.

Trolle, A. B., & Schwartz, E. S. (2009). Unspanned stochastic volatility and the pricing of commodity derivatives. *The Review of Financial Studies*, *22*(11), 4423–4461.

Tsitsiklis, J. N., & Van Roy, B. (2001). Regression methods for pricing complex american-style options. *IEEE Transactions on Neural Networks*, *12*(4), 694–703.

Vorbrink, J. (2011). American options with multiple priors in continuous time. *Working Paper*.

Wong, H. Y., & Lo, Y. W. (2009). Option pricing with mean reversion and stochastic volatility. *European Journal of Operational Research*, *197*(1), 179–187.

Working, H. (1933). Price relations between July and September wheat futures at chicago since 1885. *Wheat Studies*(06).