

# Stochastic Modeling Approaches and Pricing Techniques for Energy Derivative Contracts

Zur Erlangung des akademischen Grades eines  
Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

bei der Fakultät für  
Wirtschaftswissenschaften  
des Karlsruher Instituts für Technologie

genehmigte

DISSERTATION

von

Dipl.-Wirt.-Math. Nils Unger

Tag der mündlichen Prüfung: 30. Januar 2013

Referentin: Prof. Dr. Marliese Uhrig-Homburg

Korreferent: Prof. Dr. Wolf Fichtner

Karlsruhe 2012

## Danksagung

Ich möchte mich herzlichst bei meiner Doktormutter Frau Prof. Dr. Marliese Uhrig-Homburg für die Möglichkeit der Promotion, die Freiheit bei der Themenwahl, die immer hilfreichen Ratschläge bei der Ausgestaltung der Forschungsprojekte und der Dissertation, den sehr angenehmen persönlichen Umgang und Ihre unendliche Geduld bei der Betreuung meiner Arbeit bedanken.

Mein weiterer Dank gilt Herrn Prof. Dr. Wolf Fichtner für die Zweitbetreuung meiner Arbeit, Herrn Prof. Dr. Oliver Stein für die Übernahme der Rolle als Prüfer in der Prüfungskommission und Frau Prof. Dr. Ute Werner für die Übernahme des Vorsitzes bei meiner mündlichen Prüfung.

Ausserdem möchte ich mich bei dem gesamtem FED-Team (Annika, Claus, Hien, Jasmin, Martin, Philipp, Sorana, Steffen, Timo) für die fachliche Unterstützung und die sehr angenehme Promotionszeit bedanken.

Danke Patrick und Daniel für eure Hilfe jeglicher Art und die unterhaltsame Zeit an und ausserhalb der Universität.

Mein ganz besonderer Dank gilt meiner Familie und meiner Freundin für die bedingungslose und uneingeschränkte Unterstützung in allen Gemütslagen vor und während meiner Promotion.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Stochastic Models for Single Futures Price Dynamics</b>	<b>6</b>
2.1	Literature Review . . . . .	7
2.2	Stochastic Modeling Approaches and Derivatives Pricing and Hedging . . .	9
2.2.1	Stochastic Models . . . . .	9
2.2.2	Option Pricing . . . . .	14
2.2.3	Hedging Strategies . . . . .	15
2.2.4	Variance Swap Contracts . . . . .	16
2.3	Estimation Method . . . . .	19
2.3.1	Bayesian Statistics and the MCMC Algorithm . . . . .	19
2.3.2	Specification of the MCMC Algorithm . . . . .	25
2.4	Empirical Study . . . . .	30
2.4.1	Market Data . . . . .	30
2.4.2	Estimation Results . . . . .	33
2.4.3	Empirical Tests . . . . .	43
<b>3</b>	<b>Stochastic Term Structure Modeling Framework</b>	<b>52</b>
3.1	Trade-off between Tractability and Completeness . . . . .	53
3.2	The Model Framework . . . . .	56
3.2.1	Real Futures Contracts . . . . .	56

<i>CONTENTS</i>	iii
3.2.2 Theoretical Spot and Futures Contracts . . . . .	57
3.3 Implementation and Empirical Results . . . . .	60
3.3.1 Choice of Market Contracts . . . . .	61
3.3.2 Specification and Estimation of the Market Models . . . . .	63
3.3.3 Completing the Model through the Futures Price Curve . . . . .	77
<b>4 The Valuation of Storage Contracts</b>	<b>80</b>
4.1 Storage Contracts . . . . .	81
4.2 Static vs. Dynamic Storage Strategies . . . . .	82
4.3 Stochastic Optimization Problem . . . . .	84
4.4 Theoretical Results . . . . .	86
4.5 Numerical Example . . . . .	94
4.6 Natural Gas Storage Valuation . . . . .	96
<b>5 Conclusion and Outlook</b>	<b>104</b>
<b>A Appendix</b>	<b>107</b>
A.1 Option Pricing Functions . . . . .	107
A.2 MCMC Algorithm for Single Futures Price Dynamics . . . . .	109
A.3 Smooth Futures Price Curve . . . . .	119
A.4 Normal and Log-Normal Market Models . . . . .	121
A.5 MCMC Algorithm for Term Structure Models . . . . .	122
A.6 Normalization of Storage Contracts . . . . .	128
<b>Bibliography</b>	<b>129</b>

# List of Figures

2.1	Black-implied volatility and market-implied volatility smile . . . . .	11
2.2	implied volatility smiles in the stochastic volatility model . . . . .	12
2.3	implied volatility smiles in the jump diffusion model . . . . .	13
2.4	empirical posterior distribution of the mean parameter . . . . .	22
2.5	standard errors of filtered variance states . . . . .	29
2.6	time series of variance swap rates during 2000-2010 . . . . .	31
2.7	time series of futures prices during 1985-2010 . . . . .	32
2.8	time series of absolute futures price returns during 1985-2010 . . . . .	33
2.9	filtered volatility states in the SV and SVJ models . . . . .	34
2.10	filtered jump probabilities in the JD and SVJ models . . . . .	38
2.11	simulated price paths for the GB, JD, SV, and SVJ models . . . . .	39
2.12	quantile-quantile-plots . . . . .	44
2.13	time series of variance swap pricing errors during 2000-2010 . . . . .	46
3.1	day-ahead prices, futures prices, and futures price curve . . . . .	62
3.2	time series of crude oil futures prices and log-returns . . . . .	65
3.3	factor loadings . . . . .	66
3.4	filtered regime process . . . . .	68
3.5	quantile-quantile-plots . . . . .	70
3.6	time series of natural gas futures prices and log-returns . . . . .	72

3.7	factor loadings . . . . .	73
3.8	filtered jump times and jump sizes . . . . .	74
3.9	quantile-quantile-plots . . . . .	76
3.10	weighting functions . . . . .	78
3.11	simulated price spreads . . . . .	79
4.1	trading and delivery periods of spot and futures contracts . . . . .	81
4.2	average futures price curve . . . . .	83
4.3	impact of the summer-winter spread on the storage value . . . . .	85
4.4	spot price dynamics . . . . .	95
4.5	weighting function of the summer-winter spread . . . . .	98
4.6	quantile-quantile-plots . . . . .	100
4.7	initial futures price curve and average volume in storage . . . . .	101
4.8	continuation value function . . . . .	103
A.1	splines without the maximum smoothness criterion . . . . .	120

# List of Tables

2.1	summary statistics . . . . .	33
2.2	model parameter estimates for the time period 1985-2010 . . . . .	35
2.3	model parameter estimates for the time periods 1985-1999 and 2000-2010 . . . . .	36
2.4	option pricing errors during 2000-2010 . . . . .	48
2.5	option pricing errors during 01/2000-09/2008 and 09/2008-09/2009 . . . . .	48
2.6	hedging errors during 2000-2010 and 09/2008-09/2009 . . . . .	50
2.7	expected shortfalls for hedge portfolios . . . . .	51
3.1	model parameter estimates for the crude oil market model . . . . .	68
3.2	model parameter estimates for the natural gas market model . . . . .	74
A.1	prior distributions for the GB, JD, SV, and SVJ models . . . . .	110
A.2	prior distributions for the crude oil and natural gas market models . . . . .	124

# Chapter 1

## Introduction

In the last couple of years, market participants have been increasingly trading energy derivative contracts. This shows up in a strong growth in trading volumes of major energy benchmark contracts from less than 500 million traded contracts in 2007 to more than 800 million traded contracts in 2011.<sup>1</sup> The trading motive differs significantly between physical (e.g., producers or utilities) and non-physical traders (e.g., banks or hedge funds).<sup>2</sup> Physical traders, on the one hand, hold large and complex portfolios consisting of physical assets as well as capital intensive investment projects. These traders have a natural incentive to maximize the profits of their physical assets through market-based operation strategies and to actively manage their energy price exposures in order to reduce their funding costs and default risks (see Acharya, Lochstoer, and Ramadorai (2011)). This makes it necessary for them to determine market-based values for embedded real options in physical assets, to quantify risk-return profiles, and to deduce appropriate hedging strategies.

The implementation of trading and hedging strategies depends crucially on the underlying model specification, where omitted risk factors can have a strong impact on the pricing results, hedge ratios, and risk measures of hedged and unhedged trading positions. It turns out, however, that it is quite difficult to develop a suitable model for energy price dynamics that accounts for the specific features of physical energy trading (e.g., storage costs and limited delivery rates), that appropriately captures the time series properties of the underlying price process (e.g., clustered large returns and price jumps), and that remains tractable from an application point of view. The well-known collapse of Metallgesellschaft and the near bankruptcy of Semgroup are two prominent examples that reveal

---

<sup>1</sup>See [www.futuresindustry.com](http://www.futuresindustry.com) for details.

<sup>2</sup>In the following, market participants are denoted as non-physical traders if physical energy trading is not their core business, even if they are invested in physical assets as large banks are.

the possible implications of underestimating or mismanaging energy price risk.

Non-physical traders, on the other hand, seek for investment opportunities in order to diversify their asset portfolios or to generate excess returns in a low interest rates environment.<sup>3</sup> These non-physical traders satisfy physical traders' hedging demands and try to profit from risk premia in futures and option markets that are paid by risk-averse physical traders to hedge their natural trading positions. For instance, Trolle and Schwartz (2010) find that short selling suitable option portfolios, which can be used as hedging instruments against an increase in price uncertainty, led to sizable Sharpe ratios during the years 1996-2006. The risk premia estimates, however, strongly depend on the underlying model specification, where missing risk factors can lead to spurious estimation results (see Branger and Schlag (2008)).

In summary, physical and non-physical traders (i) require an in-depth understanding of the role of different risk factors and their associated risk premia in energy markets, (ii) have to be able to set up a modeling approach that captures the specific features of physical energy trading, and (iii) are in need of applicable tools for pricing and efficiently managing financial and physical trading positions.

This thesis addresses these challenges and makes a contribution to each of the three issues. Chapter 2 facilitates a distinct understanding of how to distinguish different risk factors. Chapter 3 provides a new modeling approach that links physical and financial energy markets. Chapter 4 offers insights on optimal operation strategies for storage facilities. While we can rely on stochastic modeling approaches for a single futures contract to address the role of risk premia in Chapter 2, the common price behavior of multiple futures contracts and their linkage through energy market frictions becomes important in Chapter 3. In addition to these stochastic modeling approaches, dynamic stochastic optimization tools become critical in Chapter 4, when it comes to exploiting the flexibility inherent in physical assets through market-based operation strategies.

We first consider stochastic modeling approaches for a single futures contract (see Hain, Uhrig-Homburg, and Unger (2012)). Here, an appropriate price process still satisfies standard no-arbitrage conditions (e.g., martingale property under the pricing measure), and the specific features of physical energy trading only influence the model choice indirectly through their impact on the return distribution of the underlying energy commodity. In the energy finance literature, volatility and price jump risks, along with the fundamental diffusive price risk, are the most prominent risk factors. The pricing and hedging performances of stochastic volatility models with jumps are only investigated in very few empirical studies (see, for example, Trolle and Schwartz (2009)) due to the similar impact that volatility and jump risk have on the risk-neutral return distribution. This makes it

---

<sup>3</sup>For example, commodity index investors invested about 190 billion USD in early 2012 (see Falkowski (2011)).

difficult to robustly separate both risk factors, where the robustness of estimation results can be increased if both return and option market data are considered in a statistical filtering approach. However, complex option pricing formulas make it difficult to exploit available option market data in a computationally economic manner. We overcome this estimation problem by incorporating option market information through a suitably weighted option portfolio, instead of multiple single option prices. This allows to significantly reduce filtering errors of latent volatility states without increasing computational time considerably.

In Chapter 2, we apply this estimation approach to obtain a refined picture of volatility and jump risk for the crude oil market from different perspectives. Namely, we estimate a stochastic volatility model with jumps and its nested model specifications based on a comprehensive data set of short-dated crude oil futures and option contracts from 1985 to 2010 in order to answer the following important questions:

- (i) Is volatility and/or jump risk priced in the crude oil option market?
- (ii) If so, what are the risk premia for taking over volatility and jump risk?
- (iii) What is the role of volatility risk for hedging strategies?
- (iv) How risky are hedge portfolios if the underlying risk factors are actively managed?

Our empirical results show that a stochastic volatility component is required to capture strongly fluctuating implied volatility levels over time, but volatility risk alone is not able to reflect pronounced implied volatility smiles of short-dated option contracts. This is because implied volatility smiles in stochastic volatility models flatten, whereas market smiles become more pronounced when approaching maturity. This suggests that another temporary risk factor is priced in the option market. In a stochastic volatility model with jumps, the jump component is able to reproduce pronounced implied volatility smiles of short-dated option contracts, which reduces option pricing errors substantially compared to pure stochastic volatility models. This indicates that both jump and volatility risk are reflected in crude oil option prices.

In addition, we examine the market price of volatility and jump risk in our data set. This is particularly important for physical traders who want to quantify the expected costs of active risk management and for non-physical traders looking for information about the risk-return profile of a potential investment in the crude oil market. Our findings on risk premia show that current empirical results should be reconsidered (see, for example, Doran and Ronn (2008)). It turns out that jump risk is priced with a significant premium, while no significant premium is paid for taking over volatility risk in the crude oil market. Further, we investigate the hedging performance of the different model specifications.

Our hedging results show that hedging errors can be reduced if both price and volatility risk (delta-vega hedging strategy), and not just price risk (delta hedging strategy), are actively managed. In addition, we find that pure stochastic volatility models drastically underestimate the risk of hedge portfolios, while jump diffusion models are much better able to capture the risk inherent in hedge portfolios.

These empirical results can then be used to specify a suitable modeling approach for the pricing and hedging of derivative instruments that are written on a single underlying. It is however not possible to value important real option contracts based on our tested modeling approaches. The reason is that the market value of important physical assets and flexible delivery contracts depends on the common stochastic behavior of multiple futures prices instead of a single one. This makes it necessary to set up a consistent stochastic model for the entire futures price curve that accounts for the specific features of physical energy markets such as futures contracts having delivery periods instead of delivery dates. In the third chapter, we develop a novel modeling framework for the common stochastic price behavior of futures contracts with arbitrary delivery periods (see Uhrig-Homburg and Unger (2012)). The model is able to take specific market frictions into account and can be easily calibrated to market data. The core idea behind obtaining a consistent modeling framework is to capture the stochastic behavior of traded futures contracts with fixed non-overlapping delivery periods through a standard market model and to price all other instruments relative to them based on a smooth interpolation approach. This allows one to specify the underlying price dynamics based on observable market data instead of on latent factors, which simplifies the estimation problem significantly. The interpolation function completes our modeling approach, whilst providing the link between theoretical futures prices that refer to delivery dates and real futures prices that refer to delivery periods. This constructive approach results in a multi-factor spot price process. The spot price dynamics is fully specified by traded instruments instead of latent factors and is applicable for the pricing and hedging of all important energy derivative contracts.

We first introduce the formal model framework and then give some general remarks on its implementation for the U.S. crude oil and natural gas markets. In addition, we discuss a concrete practical application of our modeling approach for valuing a natural storage contract. This requires to first consider the contract design of energy storage contracts. In most cases, standard storage contracts give their holders the right to inject, store, and withdraw natural gas subject to limited injection and withdrawal rates and a limited total storage capacity. These storage options can then be exploited in various traded physical delivery contracts, such as day-ahead and month-ahead futures contracts. The optimal trading strategy is path-dependent due to limited injection, withdrawal, and storage volumes. It depends on multiple market information and, at least to the best of our knowledge, cannot be derived analytically for standard stochastic price dynamics. This

makes it necessary to set up a modeling framework for the entire futures price curve and to discretize the underlying stochastic optimization problem in order to apply numerical valuation algorithms.

In the fourth chapter, we first analyze the storage valuation problem from a theoretical point of view and then use our (theoretical) results to set up a suitable pricing approach. We show that (i) rational storage operators only have to trade in a (small) subset of all delivery contracts and that (ii) optimal trading times and volumes can both be restricted to a finite number. These results allow us to formulate a low-dimensional discrete stochastic optimization problem for arbitrary price dynamics.

Then, we specify the underlying price dynamics via our modeling framework. This has the advantage that model parameters can be directly calibrated to observable market data instead of latent factors and that arbitrage-free price dynamics of arbitrary delivery contracts can be derived. Further, specific futures price movements can be more easily incorporated than in standard spot price models. We implement a three-factor market model component. The first factor accounts for temporary price risk, which largely determines the short-term optimization potential of flexible storage contracts. The second factor is used to capture a stochastic summer-winter spread, which largely determines risk-free trading profits of storage contracts that can be obtained by trading in futures contracts at the beginning of the contract period. The third factor accounts for parallel shifts, which largely explain the common price behavior of futures contracts and can be incorporated without increasing the complexity of the optimization problem. Then, we apply the Longstaff and Schwartz (2001) algorithm to determine the storage value and its dependency on the underlying risk factors. It turns out that a dynamic trading strategy in the day-ahead market can significantly increase the storage value compared to the optimal static trading strategy in futures contracts only and that a convex relation between the summer-winter spread and the storage value exists.

## Chapter 2

# Stochastic Models for Single Futures Price Dynamics

In this chapter, we analyze the role of volatility and jump risk for the pricing and hedging of crude oil option contracts in a unified modeling framework.

In the current energy finance literature, volatility and jump risks have mostly been analyzed separately, even though these risk factors have different implications for risk management and investment decisions. It is possible, for example, to manage volatility risk efficiently through trading in option contracts, while jump risk cannot be efficiently hedged by only a few market contracts. This induces very different optimal hedging strategies and hedging errors for pure stochastic volatility and pure jump diffusion models. In addition, estimation results on risk premia strongly depend on the underlying model specification. Thus, current findings on risk premia estimates for modeling approaches without stochastic volatility or price jumps must be reexamined if both risk factors are considered in option prices.

The following chapter sheds light on these issues and is organized as follows: we start with a short overview of the related literature and then introduce the stochastic modeling approaches that are empirically tested. We then present a novel estimation method based on return and aggregated option market data that allows to filter out latent variance states more robustly. Finally, the different model specifications are tested with regard to their distributional properties and their pricing and hedging performances using a comprehensive data set of crude oil futures and option contracts from 1985 to 2010.

## 2.1 Literature Review

In equity markets, a broad range of risk factors and their associated risk premia have been investigated in various empirical studies. Bakshi, Cao, and Chen (1997), Bates (2000), Chernov and Ghysels (2000), Pan (2002), Bakshi and Kapadia (2003), Jones (2003), Eraker (2004), Broadie, Chernov, and Johannes (2007), and Carr and Wu (2009) find clear empirical evidence for volatility and jump risk in return and option price data, but obtain varying results for market prices of volatility and jump risk.<sup>1</sup> For instance, Chernov and Ghysels (2000) and Jones (2003) find a significant negative market price of volatility risk, while Pan (2002) and Eraker (2004) obtain an insignificant one. Furthermore, Pan (2002) estimates a significant jump risk premium in contrast to Eraker (2004), who finds no clear evidence for a jump risk premium in his empirical study. The different estimation results can be partially explained through different model specifications, different underlying time periods, different data sets (return and/or option price data), and different estimation methods. For instance, Branger and Schlag (2008) show that discrete hedging and model specification can have such a large impact on risk premia estimates that standard tests can lead to unreliable results. In summary, it can be said that volatility and jump risk seem to exist in equity markets, but the quantification of their associated risk premia remains a challenging task that requires to take model specification issues very seriously.

In our empirical study, we analyze the role of stochastic volatility and price jumps for short-dated crude oil futures contracts and options on them. The role of stochastic volatility has been far less thoroughly investigated for commodity markets than for stock markets. Trolle and Schwartz (2009) test term structure models with different stochastic volatility specifications in the crude oil futures market between 1990 and 2006. They show that both a temporary volatility risk factor with a high mean reversion rate as well as a persistent volatility risk factor with a low mean reversion rate are required to obtain a good pricing performance for option contracts with different maturities. However, large option pricing errors arise for short maturities in all of their tested model specifications. This can be traced back to flattening implied volatility smiles in stochastic volatility models when time to maturity reaches zero, while market smiles become more pronounced. In addition, the authors find that volatility risk is largely unspanned by

---

<sup>1</sup>In more recent studies, empirical tests indicate that a jump component in the volatility process is important for capturing crisis periods with clustered large returns induced by unexpected economic events (e.g., collapse of Lehman Brothers). Furthermore, it is generally not possible to derive the sign of certain risk premia in an equilibrium model based on standard assumptions (e.g., risk-averse market makers). For instance, Liu and Pan (2003) show, in a partial equilibrium framework, that the market price of volatility risk can be positive or negative depending on market participants' relative risk-aversion.

price risk.<sup>2</sup> Consequently, traders can reduce hedging errors for a single option contract if they actively hedge volatility risk by trading in other option contracts. Indeed, Trolle and Schwartz (2009) confirm that a delta-vega hedging strategy in futures and option markets significantly reduces mean hedging errors compared to a delta hedging strategy in futures markets only.<sup>3</sup>

The estimation results on the market price of variance risk differ among empirical studies. Trolle and Schwartz (2009) obtain insignificant market prices of variance risk in their model specifications. In contrast, Doran and Ronn (2008) find a significant negative market price of variance risk in their single factor volatility model based on at-the-money option contracts between 1994 and 2004. In a model-free approach, Kang and Pan (2011) estimate a negative overall variance risk premium for different maturities.<sup>4</sup>

The role of jumps for crude oil futures price dynamics has only been investigated in a very few studies, such as those of Dempster, Medova, and Tang (2010) or Brooks and Prokopczuk (2011). Dempster, Medova, and Tang (2010) consider a two-factor price model with a temporary and a persistent jump component. They show that jump events can clearly be linked to unexpected political events and find that jump diffusion models are able to capture the distributional properties of crude oil futures price returns rather well between 2000 and 2006. In addition, they visually examine model-implied volatilities and state that jump risk is potentially able to explain pronounced market-implied volatility smiles. In a current study, Brooks and Prokopczuk (2011) analyze a stochastic volatility model with a price and volatility jump component and its nested model specifications using crude oil spot price data during 1985-2010. They show that price jumps improve the time series properties measured by the Deviance Information Criterion (DIC) (see Spiegelhalter, Best, Carlin, and van der Linde (2002)) compared to pure stochastic volatility models. They also find weak evidence for jumps in the volatility process. Interestingly, Brooks and Prokopczuk (2011) estimate a four times lower jump intensity than in Dempster, Medova, and Tang (2010) and find a clearly larger negative correlation parameter between price and volatility innovations than in Trolle and Schwartz (2009). The lower jump intensity can be explained by their additional stochastic volatility component, while a possible explanation for the larger negative correlation parameter could be the longer underlying time period, which includes the Financial Crisis, during which

---

<sup>2</sup>The estimated correlation parameters between futures price and volatility innovations for all model specifications are between -0.15 and 0.15.

<sup>3</sup>In their paper, Trolle and Schwartz (2009) only present mean hedging errors without providing standard deviations or other quantiles. This information would give a more precise picture (see Branger, Krautheim, Schlag, and Seeger (2012)).

<sup>4</sup>In short, Kang and Pan (2011) find that historical squared log-returns are on average lower than option-implied variance levels. This can be traced back on a negative market price of variance risk or on a non-zero jump risk premia.

time crude oil prices collapsed and the volatility level strongly increased.

All of the above-mentioned empirical studies on commodity markets analyze volatility and jump risk separately under the pricing measure, even though strong empirical evidence for both risk factors exists under the physical measure (see Brooks and Prokopczuk (2011)). This naturally leads to the following two questions. How can volatility and jump risk be robustly separated based on available market information? What is the impact of both risk factors on pricing and hedging errors as well as on risk premia estimates? The answers to these questions are essential for active risk management, suitable risk measures of hedge portfolios, and efficient investment decisions.

## 2.2 Stochastic Modeling Approaches and Derivatives Pricing and Hedging

In this section, we specify the stochastic volatility model with jumps (SVJ) and its nested model specifications for the futures price dynamics under the physical and risk-neutral measure (see, for example, Broadie, Chernov, and Johannes (2007)). Further, we provide pricing and hedging formulas for European option contracts and introduce variance swap contracts.

### 2.2.1 Stochastic Models

We first present a formal description of the SVJ model and then discuss the role of each model component in greater detail by considering nested modeling approaches. The superscripts  $\mathbb{P}$  and  $\mathbb{Q}$  are used to display model parameters that can differ among the physical and the risk-neutral measure, whereas model parameters without a superscript have to be the same under both measures. In the SVJ model, the futures price dynamics under the physical measure is given by<sup>5</sup>

$$df_t = (\alpha_t^{\mathbb{P}} - \lambda_z \bar{\mu}_z^{\mathbb{P}}) f_{t-} dt + \sqrt{v_t} f_{t-} dw_{f,t}^{\mathbb{P}} + (e^{z_t} - 1) f_{t-} dn_{f,t}, \quad (2.1)$$

$$dv_t = \kappa_v^{\mathbb{P}} (\theta_v^{\mathbb{P}} - v_t) dt + \sigma_v \sqrt{v_t} dw_{v,t}^{\mathbb{P}}, \quad (2.2)$$

where  $w_{f,t}^{\mathbb{P}}$  and  $w_{v,t}^{\mathbb{P}}$  are correlated Wiener processes with  $d[w_{f,t}^{\mathbb{P}}, w_{v,t}^{\mathbb{P}}] = \rho_{f,v} dt$ . The two state variables  $f_t$  and  $v_t$  denote the futures price referring to a fixed maturity date and the latent variance state at time  $t$ . We assume that the market price of diffusion risk is parameterized as  $\eta_f \sqrt{v_t}$  following Broadie, Chernov, and Johannes (2007). The jump

---

<sup>5</sup>In the following,  $f_{t-}$  denotes the left limit of  $f_s$  when  $s \rightarrow t$  and  $s < t$  ( $f_{t-} = \lim_{s \uparrow t} f_s$ ).

component is modeled by a Poisson process  $n_{f,t}$  with constant jump intensity  $\lambda_z$  and (percentage) jump sizes  $z_t$  that are normally distributed with mean  $\mu_z^{\mathbb{P}}$  and standard deviation  $\sigma_z^{\mathbb{P}}$ . We allow the mean jump size and the jump size variance to differ among both measures and restrict the jump intensity to be the same under  $\mathbb{P}$  and  $\mathbb{Q}$ .<sup>6</sup> The drift parameter  $\alpha_t^{\mathbb{P}}$  is equal to the expected excess return (futures price risk premium) of the underlying futures price dynamics. It is given by

$$\alpha_t^{\mathbb{P}} = \underbrace{\lambda_z \bar{\mu}_z^{\mathbb{P}} - \lambda_z \bar{\mu}_z^{\mathbb{Q}}}_{\text{jump risk component}} + \underbrace{\eta_f v_t}_{\text{diffusion component}}, \quad (2.3)$$

where  $-\lambda_z \bar{\mu}_z^{\mathbb{P}} = -\lambda_z(e^{\mu_z^{\mathbb{P}}+0.5(\sigma_z^{\mathbb{P}})^2} - 1)$  and  $-\lambda_z \bar{\mu}_z^{\mathbb{Q}} = -\lambda_z(e^{\mu_z^{\mathbb{Q}}+0.5(\sigma_z^{\mathbb{Q}})^2} - 1)$  are the jump compensators under the physical and risk-neutral measure. The variance process follows a classical square-root diffusion process with a constant long-term variance level  $\theta_v^{\mathbb{P}}$ , mean reversion rate  $\kappa_v^{\mathbb{P}}$ , and volatility of volatility parameter  $\sigma_v$ .<sup>7</sup> We parameterize the market price of variance risk as  $\eta_v \sigma_v^{-1} \sqrt{v_t}$  (see, for example, Broadie, Chernov, and Johannes (2007)). It follows then that the risk-neutral futures price dynamics is given by

$$df_t = -\lambda_z \bar{\mu}_z^{\mathbb{Q}} f_{t-} dt + \sqrt{v_t} f_{t-} dw_{f,t}^{\mathbb{Q}} + (e^{z_t} - 1) f_{t-} dn_{f,t}, \quad (2.4)$$

$$dv_t = \kappa_v^{\mathbb{Q}} (\theta_v^{\mathbb{Q}} - v_t) dt + \sigma_v \sqrt{v_t} dw_{v,t}^{\mathbb{Q}}, \quad (2.5)$$

where the risk-neutral model parameters of the variance process are given by  $\kappa_v^{\mathbb{Q}} = \kappa_v^{\mathbb{P}} + \eta_v$  and  $\theta_v^{\mathbb{Q}} = \frac{\kappa_v^{\mathbb{P}}}{\kappa_v^{\mathbb{Q}}} \theta_v^{\mathbb{P}}$ .

In the next step, we discuss the role of each risk factor for modeling, pricing, and hedging purposes in greater detail.

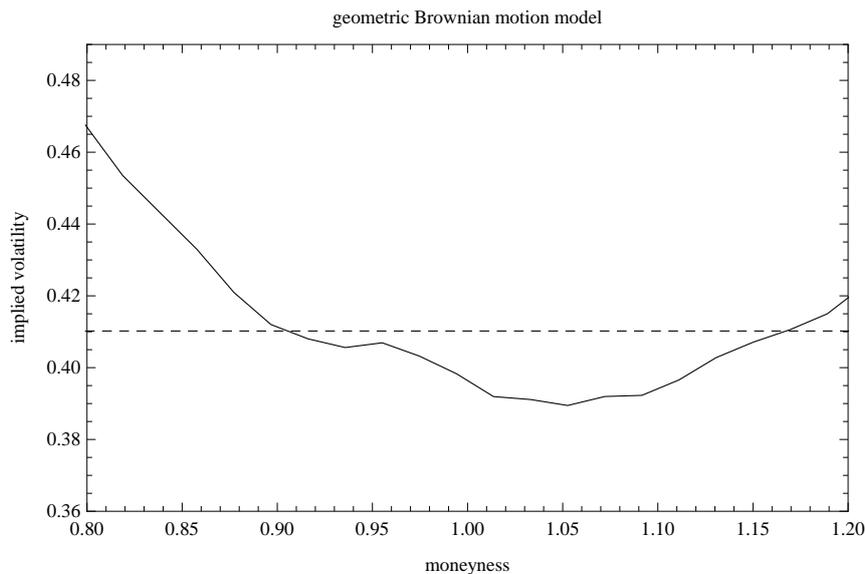
## The Geometric Brownian Motion Model (The Black Model)

The geometric Brownian motion (GB) model provides the most important benchmark approach for modeling futures price dynamics. It assumes that the underlying futures price process follows a geometric Brownian motion with a constant volatility level over time (see Black (1976)). The GB model can be obtained from the SVJ model if the jump component is equal to zero  $n_{f,t} \equiv 0$  ( $\lambda_z = 0$ ) and the variance process is constant over time  $v_t \equiv \sigma_f^2$  ( $v_0 = \theta_v^{\mathbb{P}} = \sigma_f^2$  and  $\sigma_v = 0$ ).<sup>8</sup>

<sup>6</sup>In general, the measure change for a jump process is more flexible and only requires that both jump distributions are predictable and have the same support (see Cont and Tankov (2002)).

<sup>7</sup>In the following,  $v_t$  and  $\sqrt{v_t}$  denote the variance state and the volatility state at time  $t$ , respectively.

<sup>8</sup>For completeness, we set  $\eta_v \sigma_v^{-1} = 0$ ,  $\mu_z^{\mathbb{P}} = \mu_z^{\mathbb{Q}} = 0$ , and  $\sigma_z^{\mathbb{P}} = \sigma_z^{\mathbb{Q}} = 1$ .



**Figure 2.1: Black-implied volatility and market-implied volatility smile**

The solid line shows the market-implied volatility smile based on option contracts on the front-month West Texas Intermediate (WTI) crude oil futures contract from June 18, 2002. The dashed line shows the fitted Black-implied volatility level. Each moneyness category is defined as strike price divided by futures price.

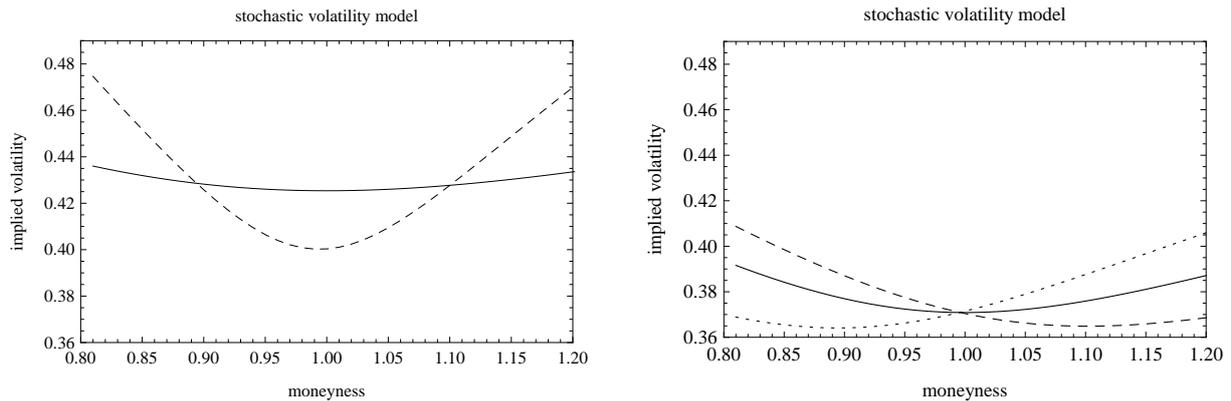
It is thus given by

$$df_t = \mu_f^{\mathbb{P}} f_t dt + \sigma_f f_t dw_{f,t}^{\mathbb{P}}, \quad (2.6)$$

where  $\mu_f^{\mathbb{P}} = \eta_f \sigma_f^2$ . In the GB model, log-returns are normally distributed and closed-form European option price formulas exist. This allows one to estimate the unknown model parameters, i.e., the constant drift  $\mu_f^{\mathbb{P}}$  and volatility  $\sigma_f$  parameters, on historical return and/or option price data by standard techniques. However, the GB model is not able to capture heavy-tailed return distributions, clusters in large returns, and implied volatility smiles or skews. This implies that the GB model provides little flexibility to capture tail risk and market-implied volatilities. Figure 2.1 shows, for example, that the Black model systematically undervalues out-of-the-money option contracts compared to at-the-money option contracts in the crude oil market.

### The Stochastic Volatility Model (The Heston Model)

The square-root stochastic volatility (SV) model is a widely used extension of the Black model. It captures clusters in small, medium, and large price returns by a stochastic variance process (see Heston (1993)) and fits into the SVJ model framework by setting



**Figure 2.2: implied volatility smiles in the stochastic volatility model**

The left graph shows the implied volatility smile for a small ( $\sigma_v = 0.8$ , solid) and large ( $\sigma_v = 2.6$ , dashed) volatility of volatility parameter in the stochastic volatility model (2.7). The right graph shows the impact of the correlation parameter on the shape of the implied volatility smile. The symmetric smile form (solid line) corresponds to a correlation parameter of zero ( $\rho_{f,v} = 0$ ), the skew (dashed line) is obtained for a negative correlation parameter  $\rho_{f,v} = -0.2$ , and the reverse skew (dotted line) follows from a positive correlation parameter  $\rho_{f,v} = 0.2$ .

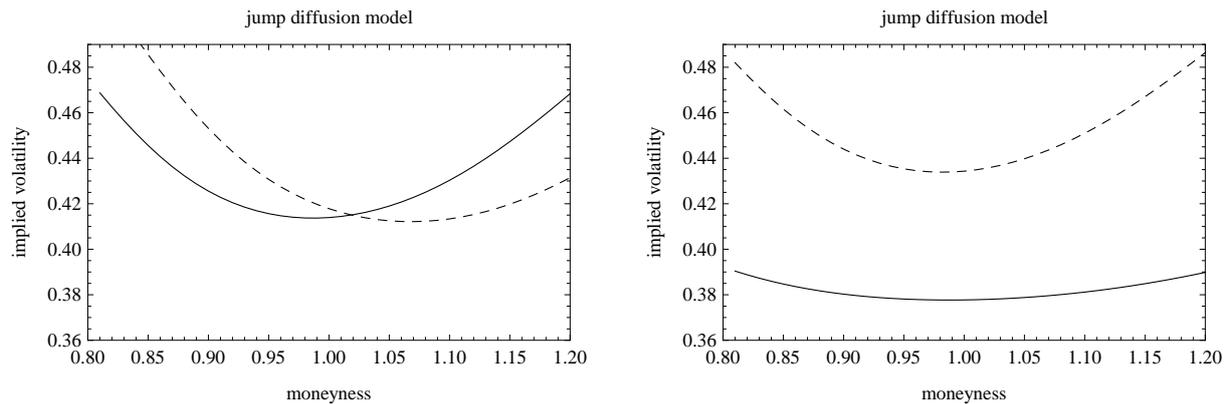
the jump component equal to zero  $n_{f,t} \equiv 0$  ( $\lambda_z = 0$ ).<sup>9</sup> The SV model is given by

$$df_t = \eta_f v_t f_t dt + \sqrt{v_t} f_t dw_{f,t}^{\mathbb{P}}, \quad (2.7)$$

$$dv_t = \kappa_v^{\mathbb{P}} (\theta_v^{\mathbb{P}} - v_t) dt + \sigma_v \sqrt{v_t} dw_{v,t}^{\mathbb{P}}. \quad (2.8)$$

In contrast to the Black model, the long-term variance level can differ under the physical and risk-neutral measure. This allows one to capture potential differences between average squared log-returns (variance level under the physical measure) and option-implied volatility levels (variance level under the risk-neutral measure). The volatility of volatility parameter  $\sigma_v$  mainly determines the excess kurtosis of futures price returns relative to the Black model. Basically, a higher  $\sigma_v$  increases out-of-the-money option prices and reduces at-the-money option prices without impacting the average implied volatility level among moneyness categories (see Figure 2.2 (left)). In addition, the correlation  $\rho_{f,v}$  between futures price and variance innovations influences the skewness of the return distribution, where a negative (positive) correlation parameter  $\rho_{f,v}$  induces a (reverse) skew in the model-implied volatility smile. This is illustrated in Figure 2.2 (right).

<sup>9</sup>For completeness, we set  $\mu_z^{\mathbb{P}} = \mu_z^{\mathbb{Q}} = 0$  and  $\sigma_z^{\mathbb{P}} = \sigma_z^{\mathbb{Q}} = 1$ .



**Figure 2.3: implied volatility smiles in the jump diffusion model**

The graphs illustrate the impact of the mean jump size parameter  $\mu_z^{\mathbb{Q}}$  (left) and the jump size volatility parameter  $\sigma_z^{\mathbb{Q}}$  (right) on the smile form in the jump diffusion model (2.9). The left graph shows the implied volatility smile for  $\mu_z^{\mathbb{Q}} = 0$  (solid) and  $\mu_z^{\mathbb{Q}} = -0.1$  (dashed). The right graph shows the implied volatility smile for  $\sigma_z^{\mathbb{Q}} = 0.1$  (solid) as well as  $\sigma_z^{\mathbb{Q}} = 0.2$  (dashed).

### The Jump Diffusion Model (The Merton Model)

The jump diffusion (JD) model adds jumps to the Black model (see Merton (1973)). The intention of including a jump component is to capture rare large price returns that can arise, for example, due to an unexpected political decision or a terrorist attack. These extreme price movements are modeled by a Poisson process with normally distributed (percentage) jump sizes. The JD model can be obtained from the SVJ model by setting  $\sigma_v$  equal to zero and the current variance state  $v_0$  at its long-term level  $\theta_v^{\mathbb{P}} = \sigma_f^2$

$$df_t = (\tilde{\alpha}^{\mathbb{P}} - \lambda_z \bar{\mu}_z^{\mathbb{P}}) f_t dt + \sigma_f f_t dw_{f,t}^{\mathbb{P}} + (e^{z_t} - 1) f_t dn_{f,t}, \quad (2.9)$$

where both the futures price risk premium  $\tilde{\alpha}^{\mathbb{P}} = \lambda_z \bar{\mu}_z^{\mathbb{P}} - \lambda_z \bar{\mu}_z^{\mathbb{Q}} + \eta_f \sigma_f^2$  and the volatility process  $\sqrt{v_t} = \sigma_f$  are constant over time.<sup>10</sup> The jump process leads to a heavy-tailed return distribution, where its skewness can be determined by the mean jump size and its kurtosis by the jump intensity and jump size volatility. The impact of the jump size parameters on option prices is as follows: (i) the mean jump size influences the low point of the implied volatility curve (see Figure 2.3 (left)), and (ii) the jump intensity and jump size volatility decide on the difference between implied volatilities for at-the-money and out-of-the-money option contracts (see Figure 2.3 (right)). In contrast to the SV model, a more pronounced implied volatility smile is directly linked to the variation of the futures price dynamics, since more jumps or jumps with a larger variation also increase the overall volatility level (see Figure 2.3 (right)).

<sup>10</sup>For completeness, we set  $\eta_v \sigma_v^{-1} = 0$ .

### The Stochastic Volatility Model with Jumps

The SVJ modeling approach adds a stochastic volatility and a price jump component to the Black model. This allows one to model volatility clusters as well as single extreme price movements in return data and provides more flexibility to capture option market information. The disadvantage is that it is rather difficult to separate the stochastic volatility component from the jump component when only a single data source (return data or option price data) is used. For instance, stochastic volatility and price jumps can both be used to capture skewness and excess kurtosis in return distributions and are both able to reflect implied volatility smiles or skews. Therefore, we undertake different cross-sectional empirical tests in order to separate both components. The separation is important, since stochastic volatility and price jumps imply different optimal hedging strategies, which we turn to in Subsection 2.2.3.

#### 2.2.2 Option Pricing

Next, we present, without proofs, pricing and hedging formulas for European option contracts under all model specifications. The theoretical underpinning and a comprehensive explanation of the underlying methodology can be found in Bakshi and Madan (2000). In what follows, we drop the  $t$ -subscripts from both state variables  $f$  and  $v$ , where reasonable, in order to simplify notation. In addition, although it is easy to incorporate an affine-linear stochastic interest rate process, we do not account for interest rate uncertainty due to its minor impact on market prices of short-dated option contracts (see, for example, Casassus and Collin-Dufresne (2005) or Trolle and Schwartz (2009)).

**Lemma 1** (*European Option Price Formula*)

*In the GB, JD, SV, and SVJ models, the market value of a European call option with maturity date  $\tau$  and strike price  $k$  on a futures contract is given by*

$$c_t(k, \tau) = e^{-r(\tau-t)} (f\pi_t^{(1)}(\tau) - k\pi_t^{(2)}(\tau)), \quad (2.10)$$

where

$$\pi_t^{(j)}(\tau) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[ \frac{e^{-i\phi \ln[k]} h_t^{(j)}(\tau, f, v, \phi)}{i\phi} \right] d\phi, \quad j \in \{1, 2\}. \quad (2.11)$$

In Appendix A.1, we provide the concrete functional forms of  $h_t^{(1)}(\cdot)$  and  $h_t^{(2)}(\cdot)$ . In contrast to a naive pricing approach that integrates out each underlying risk factor, the above

option pricing formula only requires the numerical calculation of a one-dimensional integral term. This speeds up computational time considerably and increases the robustness of numerical results. Nevertheless, it is very time consuming to incorporate option market information directly through the above option pricing formula in an estimation approach. This is especially problematic if the pricing formula must be evaluated numerous times as in simulation-based estimation methods (e.g., Markov chain Monte Carlo algorithm).<sup>11</sup> In this case, the resulting computational effort can become unmanageable. We solve this problem later on by using aggregated option market information instead of multiple individual option prices.

### 2.2.3 Hedging Strategies

Next, we briefly present the delta and delta-vega hedging strategy for the different model specifications (see, for example, Bakshi, Cao, and Chen (1997)). In the delta hedging strategy, we choose a hedge portfolio that is locally immune to infinitesimally small price changes. The hedge portfolio consists of (i) a short position in the underlying “target” call option contract with strike  $k$  and maturity  $\tau$  and (ii) a position of  $\delta_{fut}$  futures long. The  $\delta_{fut}$ -position is given by the first derivative of the option price formula with respect to the current futures price

$$\begin{aligned}\delta_{fut} &= \frac{\partial c_t}{\partial f}(k, \tau) \\ &= e^{-r(\tau-t)} \pi_t^{(1)}(\tau) \geq 0.\end{aligned}\tag{2.12}$$

This ensures that the target call option contract and  $\delta_{fut}$  futures contracts have the same sensitivity to small price movements.

In the delta-vega hedging strategy, we choose a hedge portfolio that is locally immune to infinitesimally small movements in both the underlying price and variance processes. The hedge portfolio consists of (i) a short position in the “target” call option contract with strike  $k_1$  and maturity  $\tau$ , (ii) a long position of  $\delta_{opt}$  in a “hedge” call option contract with strike  $k_2$  and maturity  $\tau$ , and (iii) a position of  $\delta_{fut}$  futures long. The additional hedge position is required to control for variance risk. It can be derived independently of the  $\delta_{fut}$ -position, because the underlying futures price is unaffected by movements in the variance process. The  $\delta_{opt}$ -position is equal to the following ratio

$$\delta_{opt} = \frac{\frac{\partial c_t}{\partial v}(k_1, \tau)}{\frac{\partial c_t}{\partial v}(k_2, \tau)}.$$

---

<sup>11</sup>In our case, we would have to evaluate the option pricing formula about 2.5 billion times.

The ratio is chosen so that one “target” option contract and  $\delta_{opt}$  “hedge” option contracts have the same sensitivity to infinitesimally small changes in the variance process. Then, we can calculate the  $\delta_{fut}$ -position in the underlying futures contract. We have to adjust the  $\delta_{fut}$ -position (2.12), since futures price movements influence the model price of the hedge option contract. The  $\delta_{fut}$ -position is given by

$$\delta_{fut} = \delta^{(1)} - \delta_{opt}\delta^{(2)},$$

where

$$\delta^{(1)} = \frac{\partial c_t}{\partial f}(k_1, \tau) \text{ and } \delta^{(2)} = \frac{\partial c_t}{\partial f}(k_2, \tau).$$

These two hedging positions ensure that the hedge portfolio is locally immune to price and variance risk in the respective model framework.

## 2.2.4 Variance Swap Contracts

In our estimation approach, we have to fit unobservable variance states, jump events, and jump sizes to market data, where their latent nature makes it difficult to obtain robust estimates based on return data only. It is generally possible to obtain more precise estimation results if option market information is incorporated in an estimation approach. However, as already mentioned, using option market data directly is computationally intensive (see Broadie, Chernov, and Johannes (2007)). In addition, multiple option prices referring to different strikes have to be weighted “suitably” in order to filter out the single variance state at any point in time. For that reason, we use a novel estimation approach that overcomes both problems by using “variance swap rates” instead of multiple option prices. The variance swap rate  $vs_{t,\tau}$  is simply defined as the “expected average annualized quadratic variation” of the underlying risk-neutral futures price process in the time period  $[t, \tau]$

$$vs_{t,\tau} = \frac{1}{\tau - t} \mathbb{E}_t^{\mathbb{Q}} [(\sigma_{t,\tau})^2]. \quad (2.13)$$

It can be calculated using two approaches: (i) a model-based approach based on the underlying risk-neutral price process and (ii) a market-based approach based on cross-sectional option prices (see Carr and Wu (2009)). Next, we show that an affine-linear relation between latent variance states and variance swap rates exists in the SV and SVJ models. This allows us to “filter out” latent variance states by solving simple linear equations based on variance swap rates instead of using highly non-linear option price

formulas directly. First, we consider in greater detail the model-based approach for calculating variance swap rates. In the SVJ model, variance swap rates depend on the latent variance and jump process as follows (see Carr and Wu (2009)):

$$\begin{aligned}
 v_{St,\tau} &= \frac{1}{\tau-t} \mathbb{E}_t^{\mathbb{Q}} \left[ \left( \int_t^\tau v_s ds \right) \right] + \lambda_z \int_{\mathbb{R}^0} x^2 g_{nd}(x) dx \\
 &= \underbrace{\theta_v^{\mathbb{Q}} + \frac{1 - e^{-\kappa_v^{\mathbb{Q}}(\tau-t)}}{\kappa_v^{\mathbb{Q}}(\tau-t)} (v_t - \theta_v^{\mathbb{Q}})}_{\text{variance component}} + \underbrace{\lambda_z ((\mu_z^{\mathbb{Q}})^2 + (\sigma_z^{\mathbb{Q}})^2)}_{\text{jump component}}, \tag{2.14}
 \end{aligned}$$

where  $g_{nd}$  denotes the density function of a normal distribution with mean  $\mu_z^{\mathbb{Q}}$  and standard deviation  $\sigma_z^{\mathbb{Q}}$ . Second, we introduce the market-based approach to calculate variance swap rates based on European option contracts with a continuum of strike prices. Breeden and Litzenberger (1978) provide an intuitive approach to derive the risk-neutral return distribution based on the second derivative of the call option price formula with respect to the strike price  $\frac{\partial^2 c_t}{\partial^2 k}(k, \tau)$ . Especially, their approach can be used to extract variance swap rates based on European option contracts with a continuum of strike prices. However, in order to apply this method, it is necessary to interpolate and extrapolate missing option prices, which can be challenging, since suitable interpolation and extrapolation functions for the second derivative of the option price curve are hard to find. For that reason, Carr and Wu (2009) develop a more robust approach for deriving variance swap rates. It is based on option prices directly and does not require the second derivative of the option price function. They show the following relation between variance swap rates and out-of-the-money European option prices:<sup>12</sup>

$$v_{St,\tau} = \frac{2}{\tau-t} \int_0^\infty \frac{o_t(k, \tau)}{e^{-r(\tau-t)} k^2} dk + \varepsilon_{vsr}, \tag{2.15}$$

where  $o_t(k, \tau)$  is the market price of an European out-of-the-money option contract with strike  $k$  and maturity  $\tau$  and  $\varepsilon_{vsr}$  is the approximation error in the presence of price jumps.<sup>13</sup>

---

<sup>12</sup>We skip the technical details for deriving the option-based measure and refer to Carr and Wu (2009) for details.

<sup>13</sup>In (2.15), the variance swap rate is calculated based on call option prices for strikes above the current futures price and put option prices for strikes below the current futures price.

The error term  $\varepsilon_{vsr}$  is equal to<sup>14</sup>

$$\varepsilon_{vsr} = \begin{cases} 0, & \text{GB and SV models} \\ -2\lambda_z \left( e^{\mu_z^{\mathbb{Q}} + 0.5(\sigma_z^{\mathbb{Q}})^2} - 1 - \mu_z^{\mathbb{Q}} - 0.5((\mu_z^{\mathbb{Q}})^2 + (\sigma_z^{\mathbb{Q}})^2) \right), & \text{JD and SVJ models} \end{cases}.$$

This approach for calculating variance swap rates based on option prices has become a widely used market standard. Up to this point, we have shown that variance swap rates can be calculated in the model framework or based on European option contracts with a continuum of strike prices. If the underlying price dynamics is “correct”, the model-based and market-based expressions (2.14) and (2.15) are equal. This can then be exploited to filter out latent variance states if the remaining model parameters are known.

**Lemma 2** (*Variance Swap Rates*)

*There exists an affine-linear relationship between the variance swap rate and the latent variance state in the SV and SVJ models. The affine-linear relation is given by*

$$\frac{2}{\tau - t} \int_0^\infty \frac{o_t(k, \tau)}{e^{-r(\tau-t)} k^2} dk = \theta_v^{\mathbb{Q}} + \frac{1 - e^{-\kappa_v^{\mathbb{Q}}(\tau-t)}}{\kappa_v^{\mathbb{Q}}(\tau - t)} (v_t - \theta_v^{\mathbb{Q}}) + c_z^{\mathbb{Q}}, \quad (2.16)$$

where

$$c_z^{\mathbb{Q}} = \begin{cases} 0, & \text{SV model} \\ 2\lambda_z \left( e^{\mu_z^{\mathbb{Q}} + 0.5(\sigma_z^{\mathbb{Q}})^2} - 1 - \mu_z^{\mathbb{Q}} \right), & \text{SVJ model} \end{cases}.$$

It is important to keep in mind that the left-hand side of (2.16) must be approximated, since only a finite number of option contracts are actively traded at the market. The resulting approximation error can be accounted for by assuming that variance swap rates are observed with noise.

In addition to variance swap rates, we also consider the so called “variance risk premium” that is defined as the difference between the expected average annualized quadratic variation under the physical and risk-neutral measure

$$\frac{1}{\tau - t} \left( \mathbb{E}_t^{\mathbb{P}} [(\sigma_{t,\tau})^2] - \mathbb{E}_t^{\mathbb{Q}} [(\sigma_{t,\tau})^2] \right) = \frac{1}{\tau - t} \mathbb{E}_t^{\mathbb{P}} [(\sigma_{t,\tau})^2] - vs_{t,\tau}. \quad (2.17)$$

In the SVJ model, the variance risk premium depends on the market price of variance risk and both risk-neutral jump size parameters. If no concrete model specification is considered, it is common to approximate the variance risk premium by using variance

---

<sup>14</sup>See Carr and Wu (2009) for details.

swap rates for  $\frac{1}{\tau-t}\mathbb{E}_t^{\mathbb{Q}}[(\sigma_{t,\tau})^2]$  and squared log-returns for  $\frac{1}{\tau-t}\mathbb{E}_t^{\mathbb{P}}[(\sigma_{t,\tau})^2]$  over the time period in question (see, for example, Carr and Wu (2009)).

## 2.3 Estimation Method

In this section, we introduce our unified estimation framework for the GB, JD, SV, and SVJ models. We first offer a short introduction to the Bayesian statistical inference approach for estimating model parameters and state variables from return and variance swap data. Then, we concretely specify our estimation approach for all model specifications.

### 2.3.1 Bayesian Statistics and the MCMC Algorithm

In Bayesian statistics, we always start with ex-ante beliefs about a vector of interest  $u \in \mathbb{R}^{dim}$ .<sup>15</sup> The ex-ante beliefs reflect our intuition about the vector of interest based on information beyond the underlying data set. This ex-ante information is captured through a so called prior distribution, which generally has a large standard deviation if little additional information exists. Then, we use the information in the underlying data set  $d$  to change our prior beliefs to a consistent posterior distribution. The “updating” of our beliefs is conducted through the Bayes theorem, which states that the density of the posterior distribution  $p(u|d)$  is proportional to the likelihood function  $p(d|u)$  times the density of the prior distribution  $p(u)$ .<sup>16</sup>

$$\begin{aligned} p(u|d) &= \frac{p(d|u)p(u)}{\int p(d|u)p(u)du} \\ &\propto p(d|u)p(u). \end{aligned}$$

It is now our objective to determine the posterior distribution of the vector of interest  $u$ . The posterior distribution reflects our (subjective) beliefs about all model parameters and state variables subject to our prior beliefs and the underlying data set. It turns out, however, that the posterior distribution is highly complex for the JD, SV, and SVJ models and cannot be calculated in closed-form.

In such cases, the Markov chain Monte Carlo (MCMC) algorithm can be used to approximate intractable posterior distributions using a simulation-based approach. We first illustrate the general procedure of the MCMC algorithm through some simple examples

---

<sup>15</sup>In the following, we assume that the vector of interest contains model parameters and state variables.

<sup>16</sup>The expression  $p(u|d) \propto p(d|u)p(u)$  means that the left-hand side  $p(u|d)$  is proportional to the right-hand side  $p(d|u)p(u)$ .

before we discuss in greater detail its implementation for the different model specifications.

In the MCMC algorithm, we construct a Markov chain  $\{u^{(g)}\}_{g=0}^{\infty}$  with a tractable transition kernel that has the posterior distribution as its limiting distribution<sup>17</sup>

$$\frac{1}{G} \sum_{g=0}^{G-1} \mathbf{1}_{\{u^{(g)} \leq \tilde{u}\}} \xrightarrow{G \rightarrow \infty} F_u^{post}(u \leq \tilde{u} | d), \text{ with probability 1.} \quad (2.18)$$

The cumulative posterior distribution function  $F_u^{post}(\cdot)$  is then approximated through a sufficiently long path of the underlying Markov chain. At first glance, it seems rather complicated to implement such an indirect sampling approach. However, it turns out that the MCMC algorithm provides a unified sampling method for all posterior distributions of our model specifications.

We start with two simple one-dimensional estimation problems to illustrate the main components of the Markov chain Monte Carlo approach – the Metropolis-Hastings and the Gibbs Sampling algorithm.

**Example:** (Black Model)

In this example, we estimate the (unknown) drift parameter of the Black model based on  $n$  log-returns  $d = \{\ln f_{t_{i+1}} - \ln f_{t_i}\}_{i=1}^n$ . The log-returns are generated through the data-generating process<sup>18</sup>

$$\ln f_{t_{i+1}} - \ln f_{t_i} = \mu_f \delta t + \sigma_f \sqrt{\delta t} \varepsilon_{f,t_i}, \quad (2.19)$$

where  $\delta t$  is equal to  $1/252$  and  $\varepsilon_{f,t_i}$  is normally distributed with a mean of zero and a standard deviation of one for  $i = 1, \dots, n$ . In addition to the log-return data, it is assumed that prior information about the unknown drift parameter exists and that the variance parameter is known. The prior distribution of  $\mu_f$  is a normal distribution with mean  $\mu_0$  and variance  $\sigma_0^2$ .

Now, our objective is to calculate the posterior distribution of the mean parameter conditional on the observed log-return data and the prior information. In such a case, we can apply the Bayes theorem, which states that the unnormalized density of the posterior distribution is simply given by the likelihood function times the density of the prior

---

<sup>17</sup>The indicator function  $\mathbf{1}_{\{u^{(g)} \leq \tilde{u}\}}$  is equal to one if all components of  $u^{(g)}$  are smaller than or equal to  $\tilde{u}$ .

<sup>18</sup>We skip the superscripts in order to reduce notational burden.

distribution<sup>19</sup>

$$p(\mu_f | \sigma_f^2, d) \propto \underbrace{\left( \frac{1}{2\pi\sigma_f^2} \right)^{n/2} e^{-\frac{\sum_{i=1}^n (d_{t_i} - \mu_f)^2}{2\sigma_f^2}}}_{\text{likelihood function}} \underbrace{\frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(\mu_f - \mu_0)^2}{2\sigma_0^2}}}_{\text{prior density}} \propto e^{-\frac{(\mu_f - \mu_{post})^2}{2\sigma_{post}^2}},$$

where

$$\mu_{post} = \frac{\sigma_f^2 \mu_0 + \sigma_0^2 \sum_{i=1}^n d_{t_i}}{\sigma_f^2 + n\sigma_0^2} \quad \text{and} \quad \sigma_{post}^2 = \frac{\sigma_f^2 \sigma_0^2}{\sigma_f^2 + n\sigma_0^2}.$$

In this case, we can determine the normalization constant easily, since the posterior distribution is equal to a normal distribution, i.e.,

$$p(\mu_f | \sigma_f^2, d) = \underbrace{\frac{1}{\sqrt{2\pi}\sigma_{post}}}_{\text{normalization constant}} e^{-\frac{(\mu_f - \mu_{post})^2}{2\sigma_{post}^2}}.$$

As a result, we do not have to apply the MCMC algorithm, since the posterior distribution is of well-known form. This analytical updating of prior information only works for specific data-generating processes and prior distributions. If we consider again, for example, the data-generating process (2.19) with a log-normal prior distribution with parameters  $\mu_0$  and  $\sigma_0^2$ , we obtain the following density of the posterior distribution

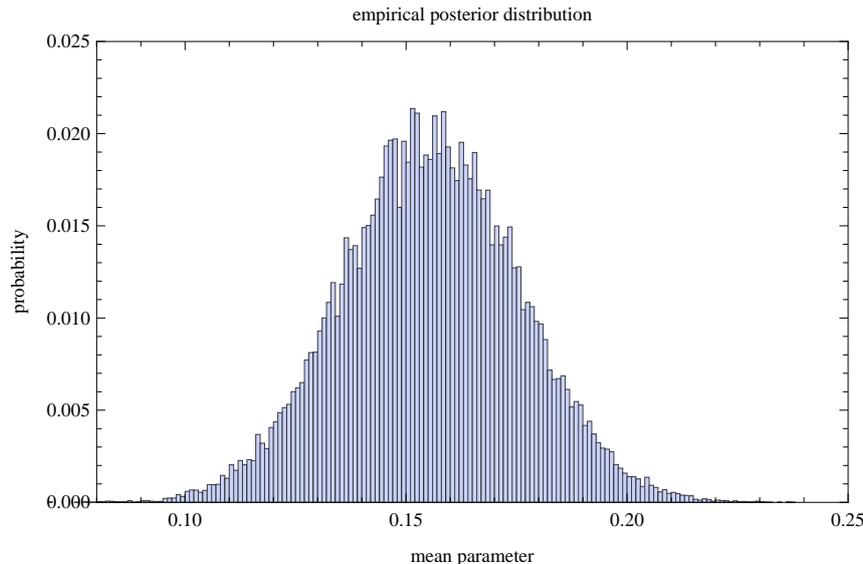
$$p(\mu_f | \sigma_f^2, d) \propto \underbrace{\left( \frac{1}{2\pi\sigma_f^2} \right)^{n/2} e^{-\frac{\sum_{i=1}^n (d_{t_i} - \mu_f)^2}{2\sigma_f^2}}}_{\text{likelihood function}} \underbrace{\frac{1}{\sqrt{2\pi}\sigma_0\mu_f} e^{-\frac{(\ln \mu_f - \mu_0)^2}{2\sigma_0^2}}}_{\text{prior density}}, \quad \mu_f > 0. \quad (2.20)$$

In this case, the posterior distribution is of non-standard form and an approximation through a standard Monte Carlo simulation is not possible.

We can, however, choose the MCMC sampling algorithm to approximate such an intractable distribution. This requires us to specify a tractable transition kernel that ensures that the Markov chain converges to the posterior distribution asymptotically. Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller (1953) introduced and Hastings (1970) refined an accept-reject approach to specify the transition kernel that satisfies the limiting condition (2.18). The Metropolis-Hastings (M-H) algorithm can briefly be explained as follows: we choose an (arbitrary) starting value  $u^{(0)}$  and draw a “candidate value”  $w^{(1)}$  based on an arbitrary tractable proposal density  $\tilde{p}(u^{(0)}, w^{(1)})$  with the same support as the conditional posterior distribution.

---

<sup>19</sup>The updating rule can be found in DeGroot (1990) or Tsay (2005).



**Figure 2.4: empirical posterior distribution of the mean parameter**

This figure shows the empirical posterior distribution of the mean parameter  $\mu_f$ . The underlying data set consists of one hundred simulated log-returns ( $n=100$ ), which are generated through (2.19). The model parameters are given by  $\mu_f = 0.15$  (unknown) and  $\sigma_f = 0.2$  (known). The prior distribution of the mean parameter is log-normal with parameters  $\mu_0 = -3$  and  $\sigma_0 = 1.5$ . The proposal density is log-normal with parameters  $\mu_{prop}^{(g)} = -0.5$  and  $\sigma_{prop}^{(g)} = 1$ . The empirical posterior distribution is obtained from the Metropolis-Hastings algorithm based on 500.000 iteration steps.

The candidate value is then accepted with probability  $a(u^{(0)}, w^{(1)})$ , where

$$a(u^{(0)}, w^{(1)}) = \min\left\{\frac{p(w^{(1)}|d) \tilde{p}(w^{(1)}, u^{(0)})}{p(u^{(0)}|d) \tilde{p}(u^{(0)}, w^{(1)})}, 1\right\}, \quad (2.21)$$

and rejected with probability  $1 - a(u^{(0)}, w^{(1)})$ . If the candidate value is accepted,  $u^{(1)}$  is set equal to  $w^{(1)}$ . Otherwise,  $u$  remains unchanged, i.e.,  $u^{(1)} = u^{(0)}$ . Then, we once again sample a candidate value based on  $\tilde{p}(u^{(1)}, w^{(2)})$  that is accepted with probability  $a(u^{(1)}, w^{(2)})$ . This procedure is repeated several times until a certain abort criterion is satisfied (see, for example, Geyer (1992)). In Figure 2.4, we show the empirical distribution of a simulation run with 500.000 draws based on a log-normal prior distribution with  $\mu_0 = -3$  and  $\sigma_0 = 1.5$  that approximates the real underlying posterior distribution.

It turns out that the performance of the approximation crucially depends on the proposal density. Basically, if the proposal distribution is close to the posterior distribution, acceptance probabilities are close to one. This reduces the correlation among individual draws  $\{u^{(g)}\}_{g=0}^{\infty}$  and increases convergence rates compared to proposal densities that do not fit the posterior distribution well (see Roberts and Tweedie (2008)).

However, it is often very complicated to obtain a “good” proposal density for high-dimensional posterior distributions. In this case, it is often much easier to divide a single

high-dimensional sampling problem into multiple low-dimensional ones (see, for example, Geman and Geman (1984)). In the Gibbs Sampling approach, this “divide and conquer” principle is picked up as follows. We first divide the vector of interest into  $h$  blocks  $u = (u_1, \dots, u_h)$ . The partition is chosen so that sampling from each lower-dimensional conditional posterior distribution  $p(u_i|u_{-i}, d)$ , where  $u_{-i} = (u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_h)$ , is possible with standard techniques for  $i = 1, \dots, h$ . Then, we construct the Markov chain to approximate the posterior distribution by iteratively drawing from the conditional posterior distributions. In detail, we start with an (arbitrary) initial value  $u^{(0)}$ . In each update step, we sample a new value  $u_i^{(1)}$  based on the conditional posterior distribution  $p(u_i^{(1)}|u_{-i}^{(0)}, d)$ , where  $u_{-i}^{(0)} = (u_1^{(1)}, \dots, u_{i-1}^{(1)}, u_{i+1}^{(0)}, \dots, u_h^{(0)})$ .<sup>20</sup> This means that  $u_j$  for each  $j < i$  is the draw from the current simulation run and  $u_j$  for each  $j > i$  is the draw from the previous simulation run. The transition kernel for one iteration step is then simply given by

$$p(u^{(g-1)}, u^{(g)}) = \prod_{i=1}^h p(u_i^{(g)}|u_{-i}^{(g-1)}, d).$$

When all individual blocks have been updated, the procedure is repeated for  $u^{(1)}$  instead of  $u^{(0)}$ . This is done until a certain abort criterion is satisfied (see, for example, Geyer (1992)).

In some cases, it is useful to mix the Metropolis-Hastings and Gibbs Sampling algorithms. Then, the posterior distribution is partitioned into multiple smaller blocks and the Metropolis-Hastings algorithm is applied to sample from the conditional posterior distribution of one or more blocks.

**Example:** (Black Model)

In this example, we illustrate the Gibbs Sampling and the Metropolis-Hastings algorithm for the Black model based on  $n$  log-returns  $d = \{\ln f_{t_{i+1}} - \ln f_{t_i}\}_{i=1}^n$ . The log-returns are again generated by

$$\ln f_{t_{i+1}} - \ln f_{t_i} = \mu_f \delta t + \sigma_f \sqrt{\delta t} \varepsilon_{f,t_i}, \quad i = 1, \dots, n. \quad (2.22)$$

Here, however, both the drift parameter  $\mu_f$  and the variance parameter  $\sigma_f^2$  are assumed to be unknown. The two-dimensional prior distribution is given by independent standard normal and inverse gamma marginal prior distributions for the mean and the variance

---

<sup>20</sup>We present the MCMC algorithm for a fixed updating sequence. In the empirical part, we update individual blocks in a random order.

parameters

$$p(\mu_f) \sim \mathcal{N}(\mu_0, \sigma_0^2), \quad \mu_0 \in \mathbb{R}, \sigma_0 \in \mathbb{R}^+, \quad (2.23)$$

$$p(\sigma_f^2) \sim \mathcal{IG}(\alpha_0, \beta_0), \quad \alpha_0, \beta_0 \in \mathbb{R}^+. \quad (2.24)$$

The underlying stochastic process and the prior distributions uniquely determine the density of the posterior distribution, which is given by

$$\begin{aligned} p(\mu_f, \sigma_f^2 | d) &\propto p(d | \mu_f, \sigma_f^2) p(\mu_f, \sigma_f^2) \\ &\propto \underbrace{\left(\frac{1}{\sigma_f^2}\right)^{n/2} \exp\left\{-\sum_{i=1}^n \frac{(d_{t_i} - \mu_f)^2}{2\sigma_f^2}\right\}}_{\text{likelihood function}} \underbrace{(\sigma_f^2)^{-\alpha_0-1} e^{-\frac{\beta_0}{\sigma_f^2}}}_{\text{prior density of } \sigma_f^2} \underbrace{\exp\left\{-\frac{(\mu_f - \mu_0)^2}{2\sigma_0^2}\right\}}_{\text{prior density of } \mu_f}. \end{aligned}$$

The unnormalized density of the posterior distribution is of unknown form. Now, we show two different approaches for sampling from such an intractable two-dimensional distribution. First, we consider the Gibbs Sampling approach. The initial points are the two conditional posterior distributions  $p(\mu_f | \sigma_f^2, d)$  and  $p(\sigma_f^2 | \mu_f, d)$  that can be expressed in closed-form due to the tractable prior distributions (2.23) and (2.24). They are given by

$$p(\mu_f | \sigma_f^2, d) \sim \mathcal{N}(\mu_{post}, \sigma_{post}^2), \quad (2.25)$$

$$p(\sigma_f^2 | \mu_f, d) \sim \mathcal{IG}(\alpha_{post}, \beta_{post}), \quad (2.26)$$

where

$$\begin{aligned} \mu_{post} &= \frac{\sigma_f^2 \mu_0 + \sigma_0^2 \sum_{i=1}^n d_{t_i}}{\sigma_f^2 + n\sigma_0^2}, & \alpha_{post} &= \alpha_0 + 0.5n, \\ \sigma_{post}^2 &= \frac{\sigma_f^2 \sigma_0^2}{\sigma_f^2 + n\sigma_0^2}, & \beta_{post} &= \beta_0 + 0.5 \sum_{i=1}^n (d_{t_i} - \mu_f)^2. \end{aligned}$$

Then, we choose a starting value for the variance parameter  $(\sigma_f^2)^{(0)}$  and sequentially draw from both conditional posterior distributions  $p(\mu_f^{(g-1)} | (\sigma_f^2)^{(g-1)}, d)$  and  $p((\sigma_f^2)^{(g)} | \mu_f^{(g-1)}, d)$  for  $g = 1, \dots, G$ . The resulting path  $\{(\mu_f^{(g)}, (\sigma_f^2)^{(g)})\}_{g=0}^G$  is used to approximate the common posterior distribution.<sup>21</sup> Second, we consider the Metropolis-Hastings algorithm. In the M-H algorithm, we use a two-dimensional proposal density to simulate both model parameters simultaneously through an accept-reject approach. It is, as mentioned, a non-trivial task to find a proposal distribution that is close to the true unknown posterior distribution. In the case of the Black model, both conditional posterior distributions can

---

<sup>21</sup>It is common to ignore the first  $n_{burn}$  draws (e.g.,  $n_{burn} = 5000$ ) to lower the impact of the starting value on the outcome.

be used to specify an “appropriate” proposal density, which is given by

$$\tilde{p}((\mu_f^{(g-1)}, (\sigma_f^2)^{(g-1)}), (\mu_f^{(g)}, (\sigma_f^2)^{(g)})) = \prod_{j=1}^2 \tilde{p}_j((\mu_f^{(g-1)}, (\sigma_f^2)^{(g-1)}), (\mu_f^{(g)}, (\sigma_f^2)^{(g)})) \quad (2.27)$$

with

$$\tilde{p}_1((\mu_f^{(g-1)}, (\sigma_f^2)^{(g-1)}), (\mu_f^{(g)}, (\sigma_f^2)^{(g)})) \sim \mathcal{N}(\mu_{pro}, \sigma_{pro}^2), \quad (2.28)$$

$$\tilde{p}_2((\mu_f^{(g-1)}, (\sigma_f^2)^{(g-1)}), (\mu_f^{(g)}, (\sigma_f^2)^{(g)})) \sim \mathcal{IG}(\alpha_{pro}, \beta_{pro}). \quad (2.29)$$

The model parameters of (2.28) and (2.29) are given by

$$\begin{aligned} \mu_{pro} &= \frac{(\sigma_f^2)^{(g-1)}\mu_0 + \sigma_0^2 \sum_{i=1}^n d_{t_i}}{(\sigma_f^2)^{(g-1)} + n\sigma_0^2}, & \alpha_{pro} &= \alpha_0 + 0.5n, \\ \sigma_{pro}^2 &= \frac{(\sigma_f^2)^{(g-1)}\sigma_0^2}{(\sigma_f^2)^{(g-1)} + n\sigma_0^2}, & \beta_{pro} &= \beta_0 + 0.5 \sum_{i=1}^n (d_{t_i} - \mu_f^{(g-1)})^2. \end{aligned}$$

In each simulation step, we draw a candidate value based on (2.27) and calculate the acceptance probability by inserting the candidate and the current value in (2.21).

### 2.3.2 Specification of the MCMC Algorithm

In the next step, we specify the Markov chain Monte Carlo algorithm for the GB, JD, SV, and SVJ models. This requires us to choose the market data that is to be considered in the estimation approach, prior distributions, the partition of the vector of interest (model parameters and state variables), and each sampling approach.

In a first step, we use a log-transformation on the underlying stochastic process and then discretize the log futures price process through the quasi Monte Carlo method (see Eraker (2004) or Broadie, Chernov, and Johannes (2007)).<sup>22</sup> It follows then that

$$y_{t_i} = \ln f_{t_{i+1}} - \ln f_{t_i} = (-\lambda_z \bar{\mu}_z^Q + \bar{\eta}_f v_{t_i}) \delta t + \sqrt{v_{t_i}} \delta t \varepsilon_{f,t_i} + z_{t_i} \delta n_{f,t_i}, \quad (2.30)$$

$$v_{t_{i+1}} - v_{t_i} = \kappa_v^{\mathbb{P}} (\theta_v^{\mathbb{P}} - v_{t_i}) \delta t + \sigma_v \sqrt{v_{t_i}} \delta t \varepsilon_{v,t_i}, \quad (2.31)$$

---

<sup>22</sup>In the special case of the Black model, the log-transformation eliminates all discretization errors. However, if we consider more complex return distributions (e.g., JD, SV, and SVJ model), it is generally not possible to find a suitable transformation such that the modified process has normally distributed returns. Then, a discretization error arises, since, for instance, the time-continuous Poisson process is approximated through a Bernoulli random variable and/or variance returns are assumed to be normally distributed instead of chi-squared distributed.

where  $\bar{\eta}_f = \eta_f - 0.5$  and the time distance between two observations of the futures price process  $\delta t$  is set equal to one.<sup>23</sup> In (2.30) and (2.31),  $\varepsilon_{f,t_i}$  and  $\varepsilon_{v,t_i}$  are normally distributed random variables with zero means, standard deviations of one, and correlation parameter  $\rho_{f,v}$ . Further,  $\delta n_{f,t_i}$  is a Bernoulli distributed random variable with jump probability  $\lambda_z$ , and  $z_{t_i}$  is normally distributed with mean  $\mu_z^{\mathbb{P}}$  and standard deviation  $\sigma_z^{\mathbb{P}}$ .

In our estimation approach, we generally choose uninformative prior distributions for all model parameters and state variables. The only exceptions are the jump intensity and the jump size variance parameter, where prior distributions capture our intuition that jumps are rare events that induce large returns. The concrete prior distributions are given in Table A.1 in Appendix A.2.

In the MCMC algorithm, we also must decide whether to sample each single model parameter and state variable sequentially or to group several ones and update them simultaneously. Liu, Wong, and Kong (1994) point out that sampling multiple highly correlated model parameters or state variables at once can potentially increase convergence rates. However, posterior distributions of multiple parameters are often highly complex and of unknown form, which means that such blocks have to be updated through the Metropolis-Hastings algorithm. Unfortunately, it is difficult to find adequate proposal densities for such high-dimensional conditional posterior distributions. For that reason, we prefer a sequential sampling approach, which is the favored method in most empirical studies using comparable price dynamics (see Eraker, Johannes, and Polson (2003), Asgharian and Bengtsson (2006), Brooks and Prokopczuk (2011), and Larsson and Nossman (2011)).

In the next step, we briefly introduce our individual update steps given that only return data  $d = \{y_{t_i}\}_{i=1}^n$  is considered in the estimation approach. Here, we obtain tractable conditional posterior distributions for the drift parameters of the futures price and variance processes, the jump intensity, the mean jump size, the jump size variance, as well as for jump times and jump sizes (see Asgharian and Bengtsson (2006)). In addition, the volatility of volatility parameter  $\sigma_v$  is updated through an inverse gamma distribution, even though the conditional posterior distribution is only inverse gamma distributed for  $\rho_{f,v} = 0$  (see Eraker, Johannes, and Polson (2003)).<sup>24</sup> The Metropolis-Hastings algorithm is only used to update latent variance states and the correlation parameter (see Appendix A.2).<sup>25</sup>

---

<sup>23</sup>This means that the underlying return distributions depend on business days instead of calendar days.

<sup>24</sup>In a simulation study, we tested approximation errors under various parameter constellations. We found that approximation errors are negligible, even if  $\rho_{f,v}$  is not close to zero.

<sup>25</sup>In several other empirical studies,  $\sigma_v$  and  $\rho_{f,v}$  are updated simultaneously by a suitable reparameterization (see Jacquier, Polson, and Rossi (2004) or Brooks and Prokopczuk (2011)). However, we found that such an updating step is numerically unstable in our case.

Now, we turn to the update steps given that log-returns and variance swap rates  $d = (\{y_{t_i}\}_{i=1}^n, \{vs_{t_i, \tau_i}\}_{i=1}^n)$  are incorporated in the estimation approach. The additional market information changes the conditional posterior distribution of the latent variance states, while all other conditional posterior distributions are unaffected. The reason is that option market information only has an indirect impact on physical model parameters and is completely uninformative for jump times and sizes. Thus, we only have to discuss the impact of variance swap rate data on the update step of the latent variance state in greater detail.

In short, the conditional posterior distribution for each variance state  $v_{t_i}$  can be expressed by using the Bayes theorem as follows:

$$\begin{aligned} p(v_{t_i} | u_{-v_{t_i}}, vs, y) &= \frac{p(v_{t_i}, vs, y | u_{-v_{t_i}})}{p(vs, y | u_{-v_{t_i}})} \\ &\propto p(v_{t_i}, vs, y | u_{-v_{t_i}}), \end{aligned}$$

where  $u_{-v_{t_i}}$  corresponds to the vector of model parameters and state variables excluding  $v_{t_i}$ ,  $vs$  denotes the vector of variance swap rates  $vs = \{vs_{t_i, \tau_i}\}_{i=1}^n$ , and  $y$  is equal to the vector of log-return data  $y = \{y_{t_i}\}_{i=1}^n$ . In order to more easily grasp the impact of the different model components on the filtering approach, we split the density function of the conditional posterior distribution  $p(v_{t_i} | u_{-v_{t_i}}, vs, y)$  into two analytically tractable components

$$\begin{aligned} p(v_{t_i} | u_{-v_{t_i}}, vs, y) &\propto p(v_{t_i}, vs, y | u_{-v_{t_i}}) \\ &\propto p(v_{t_i}, y | u_{-v_{t_i}}) p(vs | u_{-v_{t_i}}, v_{t_i}, y). \end{aligned} \quad (2.32)$$

In (2.32),  $p(v_{t_i}, y | u_{-v_{t_i}})$  corresponds to the joint density function of the current variance state and log-returns of the futures price process, and  $p(vs | u_{-v_{t_i}}, v_{t_i}, y)$  is the likelihood function of variance swap rates conditional on all model parameters, state variables, and return data. These functions can be further simplified by integrating out all terms that do not depend on  $v_{t_i}$ . It follows then that

$$p(v_{t_i} | u_{-v_{t_i}}, vs, y) \propto p(v_{t_i}, y | u_{-v_{t_i}}) p(vs_{t_i, \tau_i} | u_{-v_{t_i}}, v_{t_i}),$$

where  $p(vs_{t_i, \tau_i} | u_{-v_{t_i}}, v_{t_i})$  is equal to one if no variance swap rate is available at  $t_i$  for  $i = 1, \dots, n$ . The function  $p(v_{t_i}, y | u_{-v_{t_i}})$  provides the link between the current latent variance state to preceding and succeeding variance states and the preceding and current futures price log-returns.

The functional form of  $p(v_{t_i}, y|u_{-v_{t_i}})$  is given by (see Brooks and Prokopczuk (2011))

$$p(v_{t_i}, y|u_{-v_{t_i}}) \propto v_{t_i}^{-1} \exp(-\omega_1) \exp(-(\omega_2 + \omega_3)) \quad (2.33)$$

with

$$\begin{aligned} \omega_1 &= \frac{(y_{t_i} - (-\lambda_z \bar{\mu}^{\mathbb{Q}} + \bar{\eta}_f v_{t_i}) - z_{t_i} \delta n_{f,t_i})^2}{2v_{t_i}}, \\ \omega_2 &= \frac{(v_{t_i} - (v_{t_{i-1}} + \kappa_v^{\mathbb{P}}(\theta_v^{\mathbb{P}} - v_{t_{i-1}})) - \rho_{f,v} \sigma_v (y_{t_{i-1}} - (-\lambda_z \bar{\mu}^{\mathbb{Q}} + \bar{\eta}_f v_{t_{i-1}}) - z_{t_{i-1}} \delta n_{f,t_{i-1}}))^2}{2(1 - \rho_{f,v}^2) \sigma_v^2 v_{t_{i-1}}}, \\ \omega_3 &= \frac{(v_{t_{i+1}} - (v_{t_i} + \kappa_v^{\mathbb{P}}(\theta_v^{\mathbb{P}} - v_{t_i})) - \rho_{f,v} \sigma_v (y_{t_i} - (-\lambda_z \bar{\mu}^{\mathbb{Q}} + \bar{\eta}_f v_{t_i}) - z_{t_i} \delta n_{f,t_i}))^2}{2(1 - \rho_{f,v}^2) \sigma_v^2 v_{t_i}}. \end{aligned}$$

The first component  $\exp(-\omega_1)$  puts more mass on large variance states  $v_{t_i}$  of the conditional posterior distribution when large positive or negative diffusive returns  $\varepsilon_{f,t_i} = (y_{t_i} - (-\lambda_z \bar{\mu}^{\mathbb{Q}} + \bar{\eta}_f v_{t_i}) - z_{t_i} \delta n_{f,t_i})$  are extracted from the log-return data. The second component  $\exp(-(\omega_2 + \omega_3))$  captures the time series properties of the variance process and the dependency structure between future price and variance innovations. The relative impact of both components on the overall conditional posterior distribution mainly depends on the volatility of volatility parameter  $\sigma_v$ , where the preceding and succeeding variance states become more important for smaller volatility of volatility parameters.

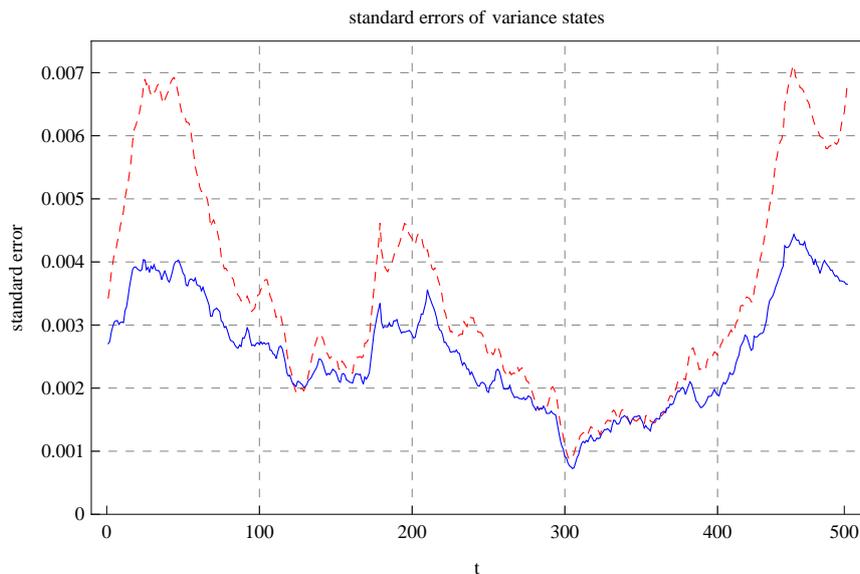
The novel part of our estimation approach is the additional component  $p(v_{t_i, \tau_i} | u_{-v_{t_i}}, v_{t_i})$  that incorporates “forward-looking” market expectations about average variance levels into the filtering method. It is given by

$$p(v_{t_i, \tau_i} | u_{-v_{t_i}}, v_{t_i}) = \frac{1}{\sqrt{2\pi} \sigma_e} \exp\left(-\frac{(v_{t_i, \tau_i} - v_{t_i, \tau_i}^{mod})^2}{2\sigma_e^2}\right), \quad (2.34)$$

where

$$v_{t_i, \tau_i}^{mod} = \theta_v^{\mathbb{Q}} + \frac{1 - e^{-\kappa_v^{\mathbb{Q}}(\tau_i - t_i)}}{\kappa_v^{\mathbb{Q}}(\tau_i - t_i)} (v_{t_i} - \theta_v^{\mathbb{Q}}) + \lambda_z ((\mu_z^{\mathbb{Q}})^2 + (\sigma_z^{\mathbb{Q}})^2). \quad (2.35)$$

It is derived from the affine-linear relation between the current variance state and the variance swap rate (see (2.14)) under the assumption that variance swap rates are observed with independent normally distributed error terms having zero means and standard deviations of  $\sigma_e$ . Otherwise, if we assume that variance swap rates are observed without any noise, we obtain a singular (maximal informative) conditional posterior distribution. Further, the incorporation of variance swap rates in the MCMC algorithm allows us to estimate the risk-neutral jump parameters  $\mu_z^{\mathbb{Q}}$  and  $\sigma_z^{\mathbb{Q}}$  as well as the risk-neutral parameters  $\kappa_v^{\mathbb{Q}}$  and  $\theta_v^{\mathbb{Q}}$  of the variance process simultaneously with the physical parameters. These



**Figure 2.5: standard errors of filtered variance states**

The dashed red line shows the standard errors of latent variance states given that only return data is used in the estimation method. The solid blue line shows the standard errors of latent variance states given that return and variance swap data is used in the estimation approach.

model parameters can then be used to extract different types of risk premia from variance swap data.

In summary, our estimation approach allows us to link unobservable latent variance states to observable market data. This should improve the robustness of the estimation results and makes it possible to bring different sources of market information together. This is particularly important for obtaining good hedging results in a real market environment, since the hedging performance is highly dependent on the ability to capture the common stochastic behavior of futures and option prices over time. Moreover, it reduces potential inconsistencies between historical and implied parameter estimates that can lead to spurious risk premia estimates and have a strong impact on the option pricing performance in two-stage estimation methods.

It seems obvious that incorporating an additional data source in an estimation approach should lower estimation errors, but the magnitude of improvement is unclear. For that reason, we conduct a simulation study for the SVJ model. The model parameters are selected close to those obtained by Larsson and Nossman (2011). The concrete parameter values are as follows:<sup>26</sup>  $\lambda_z = 6.3$ ,  $\mu_z^{\mathbb{P}} = -0.02$ ,  $\sigma_z^{\mathbb{P}} = 0.08$ ,  $\mu_z^{\mathbb{Q}} = -0.02$ ,  $\sigma_z^{\mathbb{Q}} = 0.16$ ,  $\rho_{f,v} = 0$ ,  $\theta_v^{\mathbb{P}} = 0.126$ ,  $\kappa_v^{\mathbb{P}} = 3.78$ ,  $\sigma_v = 0.756$ ,  $\eta_v = 0$ , and  $\sigma_e = 0.001$ .<sup>27</sup> Based on this parameter

<sup>26</sup>The parameter values correspond to annual decimals.

<sup>27</sup>In addition, we test various parameter constellations (e.g., positive/negative correlations and/or positive/negative mean jump sizes) and obtain similar results under all scenarios.

setup, we simulate 50 data sets of log-returns and variance swap rates consisting of 500 observations. Then, we perform two separate estimation runs for each simulated data set, one which makes use of return data only, and another one which uses both return and variance swap data.

The simulation results confirm the positive impact of using variance swap rates on estimating latent variance states. We find that incorporating variance swap rates reduces the root mean squared error between the filtered and the true variance process by about 20 percent. In addition, the standard deviation of the posterior distribution of variance states is reduced by about 20 percent. This can also be seen in Figure 2.5, which shows the standard errors of the latent variance states for one representative data set. Overall, our results confirm that using variance swap data can help us to produce more robust estimates of latent variance states without increasing computational time considerably (about 10 percent on average in our case).

## 2.4 Empirical Study

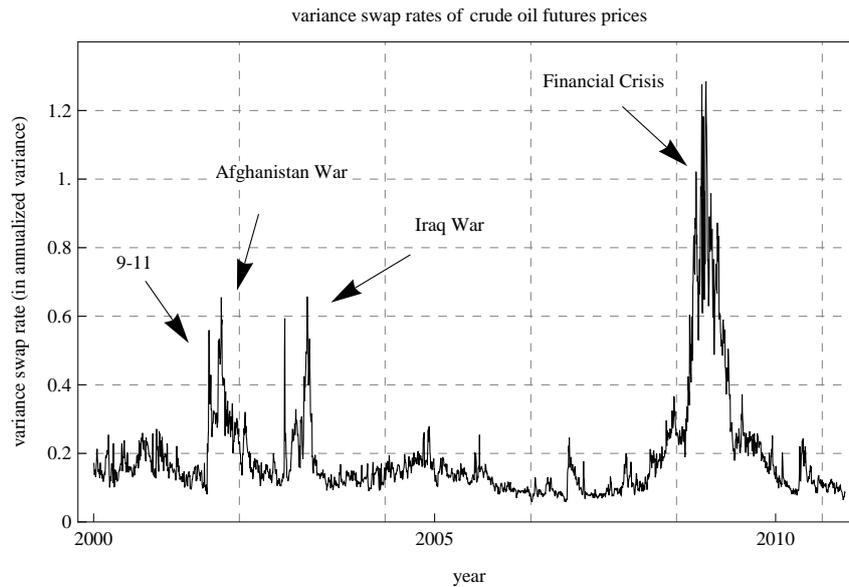
In the following two subsections, we first present the underlying data set and discuss our estimation results. Then, we test the different model specifications with regard to their distributional properties as well as their pricing and hedging performances.

### 2.4.1 Market Data

We start with a short description of our data set, which is obtained from the Bloomberg database. The data set consists of daily settlement prices of WTI crude oil futures and option contracts traded at the Chicago Mercantile Exchange (CME) Group. We have access to front-month futures prices from January 1, 1985 to December 31, 2010 and to option market data from January 1, 2000 to December 31, 2010.<sup>28</sup> The front-month futures contract is rolled over eight days before its expiry date in order to avoid maturity effects. In addition, we skip futures price returns at rolling days from our data set in order to avoid predictable price movements. The option price data set consists of, on average, 18 option contracts with different strike prices on every business day, where option prices below 0.05 USD are eliminated as in Trolle and Schwartz (2009). Further, we choose the three-month Treasury bill rate as the risk-free and constant interest rate on every business day.

---

<sup>28</sup>In detail, we take settlement prices that are determined by a “Settlement Price Committee” at the end of regular trading hours (currently 2:30 p.m. EST) as in Trolle and Schwartz (2009).



**Figure 2.6: time series of variance swap rates during 2000-2010**

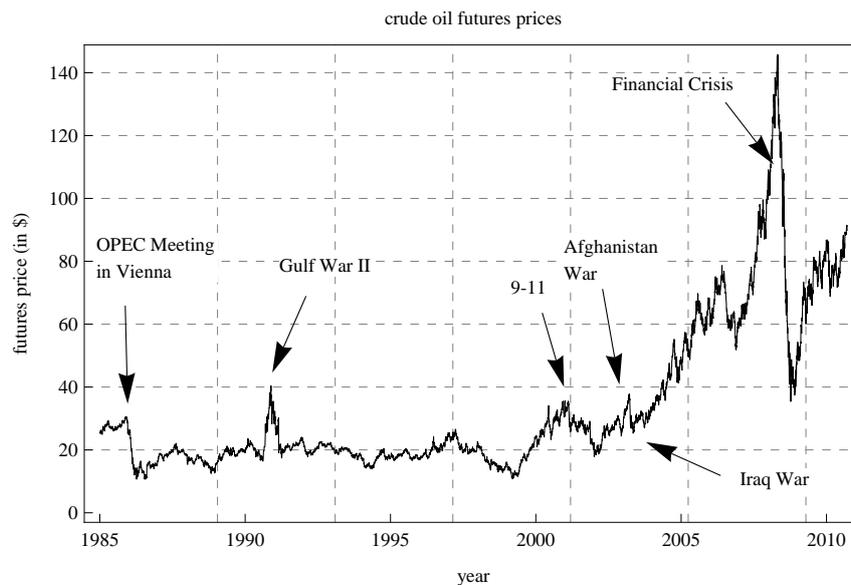
This figure shows the time series of variance swap rates based on WTI front-month crude oil futures options between 2000 and 2010.

The variance swap rates are calculated in three steps. First, we use an approximation approach that was introduced by Barone-Adesi and Whaley (1987) to derive corresponding European option prices from American option prices.<sup>29</sup> Then, we calculate Black-implied volatilities for each traded option contract and interpolate and extrapolate implied volatilities for missing strike prices based on cubic splines. Third, variance swap rates are calculated based on (2.15) ignoring the negligible approximation error term  $\varepsilon_{vsr}$ .<sup>30</sup> The time series of variance swap rates for our data set is shown in Figure 2.6.

In Figure 2.7, we plot the historical time series of the futures price process during 1985-2010. There are three conspicuous peaks and drops. In early 1986, OPEC (Organization of the Petroleum Exporting Countries) members failed to agree on a production limit at a Meeting in Vienna. This resulted in a price drop of more than 40 percent over the following couple of months. The Gulf War II led to a strong decline in crude oil prices during 1991. In September 2008, the front-month crude oil futures price collapsed in less than one year to a third of its previously reached highest level. Figure 2.8 shows absolute futures price returns that are clearly clustered in 1986 (OPEC Meeting in Vienna), 1991 (Gulf War II), and 2008 (Financial Crisis). Further, variance swap rates exhibit a large

<sup>29</sup>In general, the differences between American and European option prices are rather small for short-dated option contracts. Therefore, potential option pricing errors that arise in the Barone-Adesi and Whaley approximation approach are not large.

<sup>30</sup>In the estimation step, we explicitly account for the approximation error term in the presence of jumps.



**Figure 2.7: time series of futures prices during 1985-2010**

This figure shows the time series of WTI front-month crude oil futures prices between 1985 and 2010.

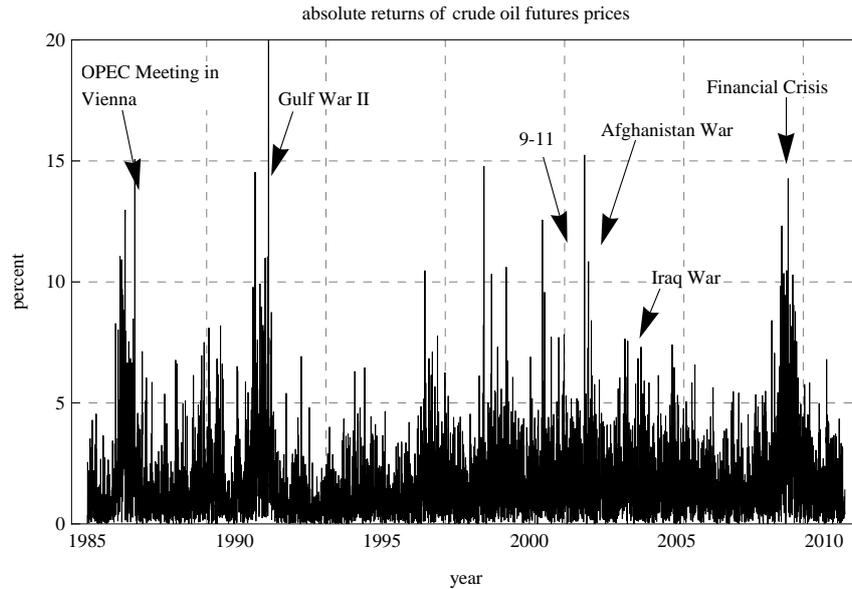
peak in the Financial Crisis and two smaller peaks in 2001 (Afghanistan War) and 2003 (Iraq War) (see Figure 2.6).

In Table 2.1, we provide the summary statistics of log-return data for the complete time period and two subsamples. The first four moments are relatively stable over time and show a clear non-normal behavior. In detail, log-returns are moderately left-skewed, where the skewness is less pronounced in the more recent time period from 2000-2010.<sup>31</sup> In addition, log-returns exhibit significant excess kurtosis in all time periods indicating a return distribution with fat-tails. At a first glance, we find no clear evidence for a positive or negative correlation between future prices changes and volatility movements. Instead volatility peaks coincide with both strong futures price increases (e.g., 2007/2008 (Oil Price Rally)) and declines (e.g., 2008 (Financial Crisis)) in our sample.

To get a rough intuition about the risk-neutral return distribution, we compute the average implied volatility smile for our data sample. We find a mostly symmetric smile form with the lowest implied volatilities for moneyness levels slightly larger than one.<sup>32</sup> This suggests that a stochastic process is required that is able to capture excess kurtosis, but no positive or negative skewness compared to a simple geometric Brownian motion model.

<sup>31</sup>If we compare the skewness of crude oil and equity log-returns, we find that crude oil log-returns are moderately left-skewed ( $> -1.20$ ) compared to equities (e.g.,  $\approx -2.00$  for S&P 500 (see Asgharian and Bengtsson (2006))).

<sup>32</sup>The shape of the implied volatility smile is similar to the market-implied volatility smile on June 18, 2002 (see Figure 2.1).



**Figure 2.8: time series of absolute futures price returns during 1985-2010**

This figure shows the time series of the absolute price returns of the WTI front-month crude oil futures contract during 1985-2010.

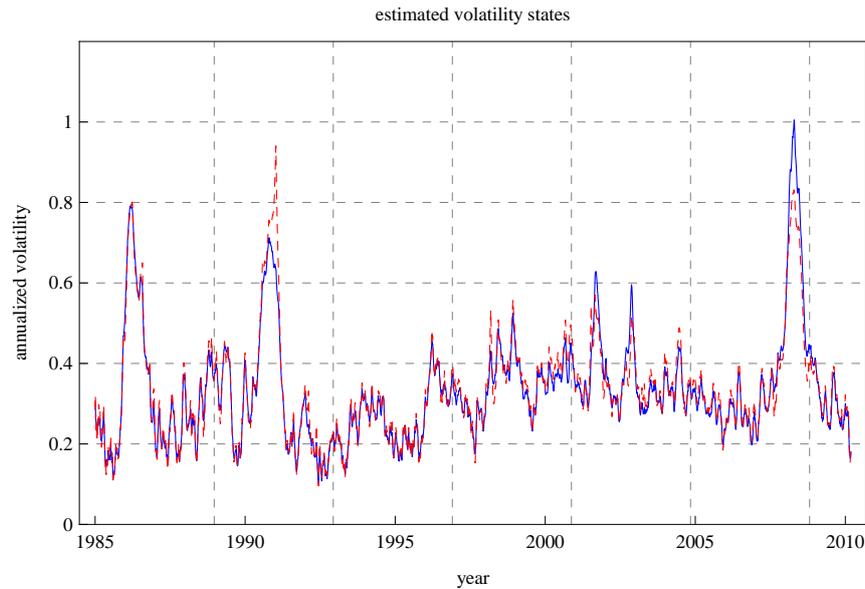
	Mean	Standard Deviation	Skewness	Kurtosis	Min	Max
01/1985-12/2010	0.0001	0.0233	-0.7930	18.0154	-0.3841	0.1403
01/1985-12/1999	-0.0001	0.0228	-1.1815	28.2070	-0.3841	0.1403
01/2000-12/2010	0.0002	0.0241	-0.3102	5.9084	-0.1654	0.1334

**Table 2.1: summary statistics**

The table reports means, standard deviations, skewness, and kurtosis for daily log-returns of WTI front-month crude oil futures prices between 1985-2010, 1985-1999, and 2000-2010.

## 2.4.2 Estimation Results

In the next step, we separately estimate model parameters and state variables for the GB, JD, SV, and SVJ models based on our MCMC algorithm. The model parameter estimates refer to the mean of the posterior distribution. We fit all model specifications to market data from three different time periods (1985-2000, 2000-2010, and 1985-2010) in order to test the robustness of the estimation results. Note that option price data only exists since the year 2000 due to data restrictions. We first discuss our estimation results for physical model parameters and state variables and then present our risk premia estimates.



**Figure 2.9: filtered volatility states in the SV and SVJ models**

This figure shows the estimated volatility states for the stochastic volatility model (2.7) (dashed red) and the stochastic volatility model with jumps (2.1) (solid blue) for the years 2000-2010.

## Model Parameters and State Variables

Figure 2.9 shows the filtered volatility processes for the SV and SVJ model. The volatility processes are quite similar. The only two exceptions are the result of a single extreme negative price jump of more than 35 percent in 1991 (Gulf War II) and large variance swap rates in 2008 (Financial Crisis). These two exceptions can be explained as follows: (i) an extreme price movement of more than 30 percent on a single day cannot be filtered out through a jump event in the SV model and (ii) large variance swap rates increase latent variance states more strongly in the SVJ than in the SV model, since variance swap rates react less sensitively to changes in variance states in the SV than in the SVJ model.<sup>33</sup> In both models, the volatility process attains values between 10 and 100 percent, increases strongly in 1986 (OPEC Meeting in Vienna), 1991 (Gulf War II), and 2008 (Financial Crisis), and reverts to its long-term level in about 100 business days after each of these peaks (see Figure 2.9).

Table 2.2 show that the long-term volatility level  $\sqrt{\theta_v^{\mathbb{P}}}$  is 36 percent and that the volatility process is moderately persistent with an estimated (daily) mean reversion rate ( $\kappa_v^{\mathbb{P}}/252$ ) of about 1.5 percent. Furthermore, the volatility of volatility parameter is more than double the size of estimated values for equities (see, for example, Pan (2002), Eraker (2004), and

<sup>33</sup>This will be explained in greater detail when we discuss our estimation results regarding the market prices of risk.

Broadie, Chernov, and Johannes (2007)). The posterior distribution of the correlation coefficient  $\rho_{f,v}$  between price and volatility innovations has a slightly negative mean. This result coincides roughly with those obtained in Trolle and Schwartz (2010), who also find weak evidence for a small, negative dependence between price and volatility innovations. The estimated variance parameters remain relatively stable in the two distinct subsamples with a slightly lower long-term volatility level in the first than in the second subsample due to the Financial Crisis (34 % for 1985-2000 vs. 39 % for 2000-2010) (see Table 2.3).

	1985-2010				
	GB	JD	SV	SVJ	SVJ0
$\lambda_z$	-	34.1738 (4.4043)	-	1.0951 (0.6305)	1.3112 (0.6293)
$\mu_z^{\mathbb{P}}$	-	[25.4361,44.2773]	-	[0.2303,3.1924]	[0.3276,3.2133]
$\sigma_z^{\mathbb{P}}$	-	-0.0027 (0.0019)	-	-0.0241 (0.0316)	-0.0201 (0.0247)
$(\sigma_f^{\mathbb{P}})^2$	0.1368 (0.0025)	0.0446 (0.0022)	-	[0.1144,0.0374]	[0.1065,0.0346]
$\rho_{f,v}$	[0.1312,0.1425]	[0.0401,0.0503]	-	[0.0605,0.1498]	[0.0586,0.1591]
$\kappa_v^{\mathbb{P}}$	-	0.0712 (0.0029)	-0.2111 (0.0931)	-0.1756 (0.0634)	-0.1694 (0.0509)
$\theta_v^{\mathbb{P}}$	-	-	[3.2950,7.1251]	[1.9408,4.8491]	[1.8174,4.6671]
$\sigma_v$	-	-	5.0598 (0.8213)	3.3141 (0.6233)	3.3727 (0.6170)
$\kappa_v^{\mathbb{Q}}$	-	-	[0.1094,0.1693]	[0.1038,0.1848]	[0.1007,0.1807]
$\theta_v^{\mathbb{Q}}$	-	-	0.8324 (0.0467)	0.6832 (0.0379)	0.6543 (0.0438)
$\eta_v$	-	-	[0.7317,0.9509]	[0.6046,0.7812]	[0.5750,0.7433]
$c_z^{\mathbb{P}} - c_z^{\mathbb{Q}}$	-	-	0.6701 (0.6797)	5.0857 (1.8559)	-
	-	-	[0.1271,3.4107]	[1.0381,9.5901]	-
	-	-	1.8749 (1.2589)	0.1062 (0.0813)	-
	-	-	[0.1848,4.8563]	[0.0421,0.4137]	-
	-	-	-4.3898 (1.1214)	1.7716 (1.9460)	-
	-	-	[-6.7268,-1.0413]	[-2.4946,6.4024]	-
	-	-0.0493 (0.0045)	-	-0.0332 (0.0091)	-0.0283 (0.0071)
	-	[-0.0590,-0.0384]	-	[-0.0539,-0.0117,]	[-0.04437,-0.0109,]

**Table 2.2: model parameter estimates for the time period 1985-2010**

This table reports posterior means, standard deviations (in parenthesis), and 1% to 99% credibility intervals (in square brackets) for the GB, JD, SV, and SVJ (SVJ0) models. The model parameters are estimated based on the complete time period from 1985 to 2010. The market price of diffusion risk is set to zero in all model specifications ( $\eta_f = 0$ ). The market price of variance risk is estimated in the SVJ model, whereas it is set to zero in the SVJ0 model. The parameter values correspond to annual decimals.

	1985-1999			2000-2010				
	GB	JD	SV	SVJ0	GB	JD	SV	SVJ0
$\lambda_z$	-	39.5994 (4.5145)	-	2.5374 (1.0386)	-	18.44516 (4.1527)	-	0.8675 (0.5380)
$\mu_z^P$	-	[29.8179,50.6602]	-	[0.7505,5.5148]	-	[10.3734,29.5747]	-	[0.1191,2.6151]
$\sigma_z^P$	-	-0.0019 (0.0022)	-	0.0011 (0.0174)	-	-0.0084 (0.0052)	-	-0.0382 (0.0387)
$(\sigma_I^P)^2$	-	[-0.0070,0.0033]	-	[-0.0380,0.0472]	-	[-0.0220,0.0033]	-	[-0.1420,0.0585]
$\rho_{f,v}$	-	0.0451 (0.0023)	-	0.0839 (0.0153)	-	0.0508 (0.0046)	-	0.0803 (0.0178)
$\kappa_v^P$	-	[0.0403,0.0508]	-	[0.0556,0.1262]	-	[0.0414,0.0630]	-	[0.0508,0.1395]
$\theta_v^P$	-	0.1307 (0.0031)	-	-	0.1388 (0.0037)	0.1035 (0.0047)	-	-
$\sigma_v$	-	[0.1238,0.1378]	-	-	[0.1303,0.1478]	[0.0926,0.1146]	-	-
$\kappa_v^Q$	-	-	-0.1350 (0.0709)	-0.1068 (0.0611)	-	-	-0.1980 (0.0839)	-0.1811 (0.0813)
$\theta_v^Q$	-	-	[-0.3011,0.0425]	[-0.2506,0.0279]	-	-	[-0.4281,-0.0373]	[-0.4298,-0.0525]
$\eta_v$	-	-	8.6227 (1.1843)	4.5191 (1.0623)	-	3.6107 (1.0651)	-	3.7738 (1.0784)
$c_z^P - c_z^Q$	-	-	[6.0278,11.5287]	[2.2270,7.0970]	-	-	[1.2772,6.3062]	[1.3951,6.4947]
	-	-	0.1119 (0.0109)	0.1146 (0.0177)	-	0.17046 (0.0701)	0.1574 (0.0303)	0.1574 (0.0303)
	-	-	[0.0899,0.1415]	[0.0842,0.1705]	-	[0.1205,0.2749]	[0.1106,0.2528]	[0.1106,0.2528]
	-	-	1.0610 (0.0576)	0.7606 (0.0551)	-	0.6530 (0.0513)	0.6718 (0.0541)	0.6718 (0.0541)
	-	-	[0.9313,1.2065]	[0.6325,0.8819]	-	[0.5480,0.7822]	[0.5608,0.8139]	[0.5608,0.8139]
	-	-	-	-	-	0.9617 (0.7336)	-	-
	-	-	-	-	-	[0.0378,3.3208]	-	-
	-	-	-	-	-	1.5492 (2.4415)	-	-
	-	-	-	-	-	[0.1654,13.8295]	-	-
	-	-	-	-	-	-2.6489 (1.244)	-	-
	-	-	-	-	-	[-5.5631,0.3352]	-	-
	-	-	-	-	-	-0.0370 (0.0071)	-	-0.0286 (0.0076)
	-	-	-	-	-	[-0.0520,-0.0186]	-	[-0.04540,-0.0098]

**Table 2.3: model parameter estimates for the time periods 1985-1999 and 2000-2010**

This table reports posterior means, standard deviations (in parenthesis), and 1% to 99% credibility intervals (in square brackets) for the GB, JD, SV, and SVJ0 models. The model parameters are estimated based on two subsample periods from 1985-1999 and 2000-2010. The market price of diffusion risk and variance risk are set equal to zero in all model specifications ( $\eta_f = 0$  and  $\eta_v = 0$ ). The parameter values correspond to annual decimals.

In the next step, we consider the filtered jump processes for the JD and the SVJ models. It can be seen that filtered jump events differ significantly in the two models (see Figure 2.10). There are significantly more filtered jump events in the JD model than in the SVJ model. Further, jump events are clustered in 1986 (OPEC Meeting in Vienna), 1991 (Gulf War II), and 2008 (Financial Crisis) in the JD model, while no clear jump clusters are found in the SVJ model. In particular, no jumps are filtered out during the Financial Crisis due to an extremely high volatility level in this period of time. In addition, we do not find a relation between filtered jump events and volatility states as suggested in Bates (2000) and Doran and Ronn (2008). The jump sizes range from -33.8 to 12.3 percent and are on average slightly negative in both models (see Figure 2.10). If we look at the daily jump intensity ( $\lambda_z/252$ ), we find a very high value of 13 percent in the JD and a low value of 0.5 percent for the SVJ model for the complete time period (see Table 2.2). The estimated jump intensity for the JD model is similar to the estimated jump intensities for equity markets. For instance, Johannes, Kumar, and Polson (1999) find jump intensities ranging from 5 to 16 percent for various prominent equity indices based on the JD model. In the case of the SVJ model, we obtain a slightly lower jump intensity than for equity markets. For example, Eraker, Johannes, and Polson (2003) estimate a daily jump intensity of 0.8 percent for the S&P 500. Compared to other commodities, our estimated jump intensity is considerably lower. Brooks and Prokopczuk (2011), for example, find a daily jump intensity of 2.5 percent for heating oil under the SVJ model framework.

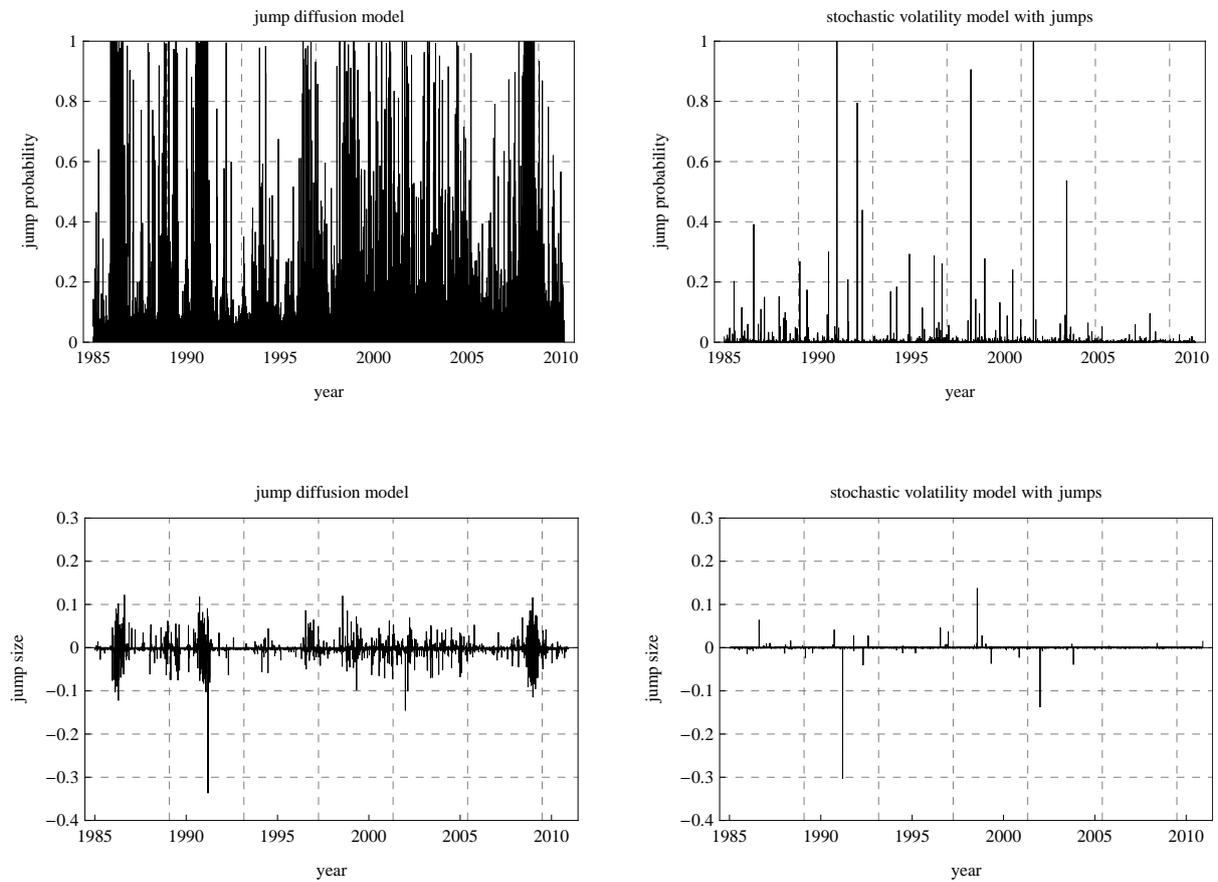
Concerning the results of our two subsamples, we find a higher jump intensity under both model specifications for the subsample from 1985-2000 (15.7 % vs. 6.3 % for the JD model and 1.0 % vs. 0.34 % for the SVJ model). This is consistent with our finding of fewer jump occurrences and a higher average volatility level after 2000.

We also simulate price paths to test whether the respective models are able to capture the time series properties of historical log-returns (see Figure 2.11). The estimated GB model is not able to capture clusters in large returns and price jumps, the JD model results in too many and too small price jumps, and the SV model has difficulty capturing single large absolute returns above 15 percent compared to the historical return data. In contrast, simulated price paths of the SVJ model exhibit similar time series properties as historical log-returns of crude oil futures prices. We have also estimated a stochastic volatility model with a price and volatility jump component. It turns out that robust parameter estimates for the volatility jump component are difficult to obtain and that jumps in the volatility process do not greatly improve the distributional properties as well as the pricing and hedging performances.<sup>34</sup> For this reason, we do not discuss our

---

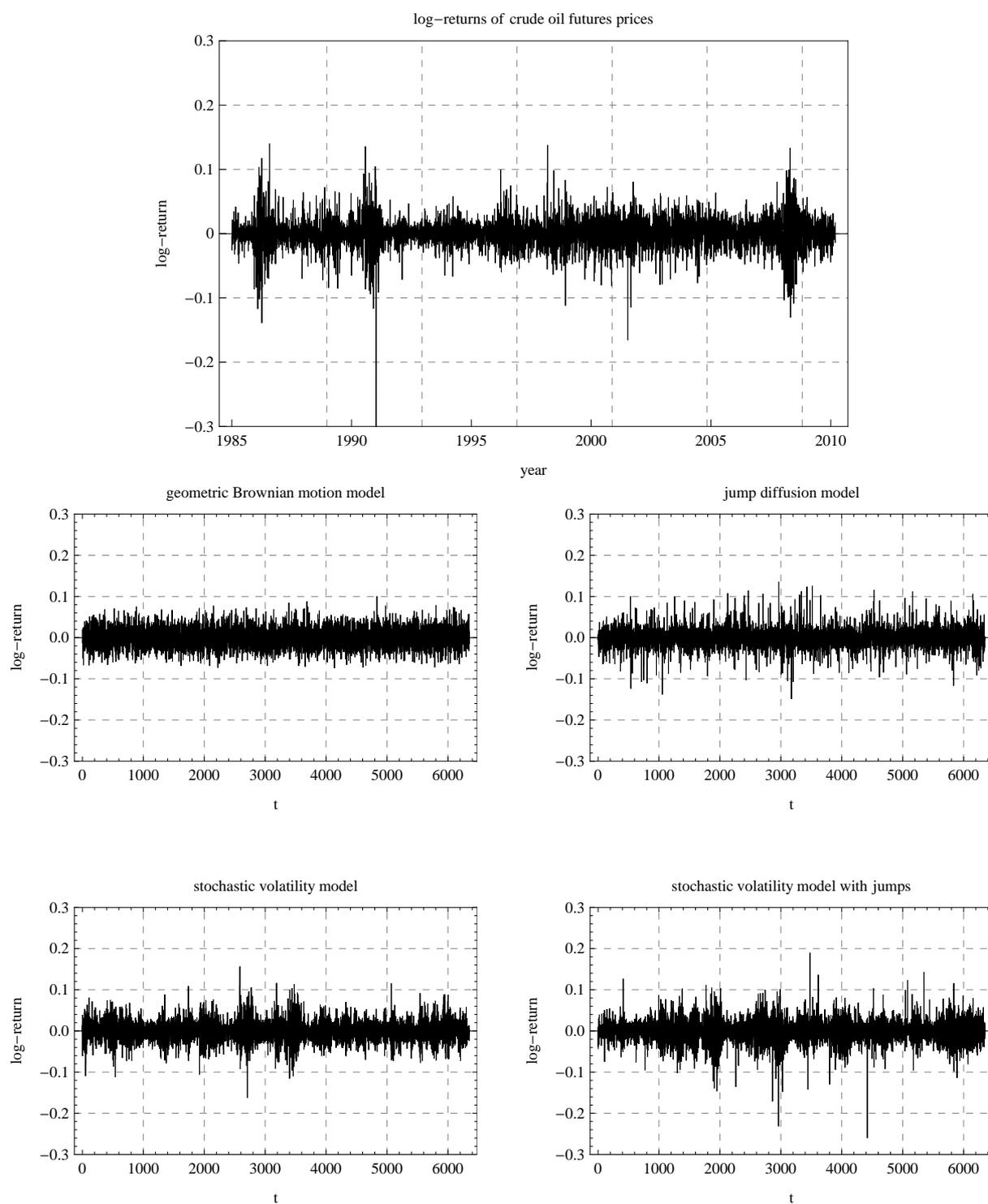
<sup>34</sup>Similar to our results, Brooks and Prokopczuk (2011) only find weak evidence for a jump component in the volatility process based on return data only.

estimation results for this model specification in greater detail.



**Figure 2.10: filtered jump probabilities in the JD and SVJ models**

This figure shows posterior probabilities of jump events (top) and filtered jump sizes (bottom) for the jump diffusion model (2.9) (left panel) and the stochastic volatility model with jumps (2.1) (right panel) at each trading day during 1985-2010.



**Figure 2.11: simulated price paths for the GB, JD, SV, and SVJ models**

This figure shows the time series of real log-returns of the WTI front-month crude oil futures contract during 1985-2010 and one representative simulated price path for each model specification based on the estimated model parameters given in Table 2.2.

## Market Price of Risk

In this paragraph, we present our estimation results for the market price of diffusion, variance, and jump risk. The different types of risk premia can be estimated simultaneously based on historical return and variance swap data in our MCMC algorithm (see Section 2.3), where insignificant risk premia are set equal to zero in order to increase the robustness of our estimation results.

First, we discuss our estimation results for the market price of diffusion risk. The market price of diffusion risk  $\eta_f$  is reflected in the drift components of the physical futures price process (2.1) via the futures price risk premium (2.3), but does not appear in the risk-neutral futures price process (2.4). The reason is that standard no-arbitrage arguments uniquely determine the risk-neutral drift component of traded futures contracts independently of the underlying data set. This means that the market price of diffusion risk can only be estimated based on historical excess return data. The impact of  $\eta_f$  on excess returns differs among our model specifications. It leads to a constant excess return if the variance process is constant over time (GB and JD models) and determines the relation between excess returns and the uncertainty in the market (measured by the current variance state) if the variance process is stochastic (SV and SVJ models).

In all model specifications, we find neither a significant excess return nor a significant relation between variance states and excess returns.<sup>35</sup> This suggests that no premium is paid for taking over diffusion risk in crude oil futures markets. In addition to our MCMC estimates, we simply regress squared log-returns as well as variance swap rates on historical excess returns to test the robustness of our empirical results in a model-free approach. As predicted by our estimation results, neither squared log-returns nor variance swap rates have a significant predictive power for excess returns. For this reason, we assume, henceforth, that the market price of diffusion risk is equal to zero.

In the next step, we consider our estimation results for the variance and jump risk premia. The market prices of variance risk  $\eta_v$ , mean price jump risk premium  $\mu_z^{\mathbb{P}} - \mu_z^{\mathbb{Q}}$ , and volatility of price jumps risk premium  $\sigma_z^{\mathbb{Q}} - \sigma_z^{\mathbb{P}}$  are extracted based on the following relation

---

<sup>35</sup>In an unrestricted MCMC run that explicitly estimates  $\eta_f$ , we obtain the following parameter estimates for the diffusion risk premium  $\eta_f$ : 0.7881 (0.8586) for the GB model and 0.2816 (0.5425) for the SVJ model, where the values in the parenthesis correspond to the standard deviation of the respective posterior distribution. In both cases, the ex-post probability of a positive and a negative market price of diffusion risk is more than 5 percent so that no “significant” diffusion risk premium is found. For the SV model, the mean and standard deviation of the posterior distribution are given by 0.6211 (0.5425), where positive and negative market prices of diffusion risk both have ex-post probabilities of more than 5 percent. In contrast to stochastic volatility models, we are not able to separate the market prices of diffusion and jump risk based on excess return data only in the JD model. The reason is that the aggregated jump risk premium  $\lambda_z \bar{\mu}_z^{\mathbb{P}} - \lambda_z \bar{\mu}_z^{\mathbb{Q}}$  and the market price of diffusion risk lead to a constant excess return (see (2.3)).

between latent variance states and observable out-of-the-money option prices (see (2.16)):

$$\begin{aligned} \frac{2}{\tau-t} \int_0^\infty \frac{o_t(k, \tau)}{e^{-r(\tau-t)} k^2} dk &= \theta_v^{\mathbb{Q}} + \frac{1 - e^{-\kappa_v^{\mathbb{Q}}(\tau-t)}}{\kappa_v^{\mathbb{Q}}(\tau-t)} (v_t - \theta_v^{\mathbb{Q}}) + 2\lambda_z (e^{\mu_z^{\mathbb{Q}} + 0.5(\sigma_z^{\mathbb{Q}})^2} - 1 - \mu_z^{\mathbb{Q}}) \\ &= \frac{\kappa_v^{\mathbb{P}}}{\kappa_v^{\mathbb{P}} + \eta_v} \theta_v^{\mathbb{P}} + \frac{1 - e^{-(\kappa_v^{\mathbb{P}} + \eta_v)(\tau-t)}}{(\kappa_v^{\mathbb{P}} + \eta_v)(\tau-t)} \left( v_t - \frac{\kappa_v^{\mathbb{P}}}{\kappa_v^{\mathbb{P}} + \eta_v} \theta_v^{\mathbb{P}} \right) + c_z^{\mathbb{Q}}, \end{aligned}$$

where  $c_z^{\mathbb{Q}} = 2\lambda_z (e^{\mu_z^{\mathbb{Q}} + 0.5(\sigma_z^{\mathbb{Q}})^2} - 1 - \mu_z^{\mathbb{Q}})$  is denoted as the aggregated variance jump compensator, and the value of the option portfolio  $\frac{2}{\tau-t} \int_0^\infty \frac{o_t(k, \tau)}{e^{-r(\tau-t)} k^2} dk$  is denoted as the non-adjusted variance swap rate. The different risk premia are separately updated by their conditional posterior distributions given that the physical model parameters (e.g.,  $\mu_z^{\mathbb{P}}$  and  $\sigma_z^{\mathbb{P}}$ ) are known. This allows us to consider the impact of the market price of variance risk and the risk-neutral jump distribution parameters  $\mu_z^{\mathbb{Q}}$  and  $\sigma_z^{\mathbb{Q}}$  on non-adjusted variance swap rates directly. The risk-neutral jump size mean  $\mu_z^{\mathbb{Q}}$  and jump size volatility  $\sigma_z^{\mathbb{Q}}$  parameters both have a constant impact on non-adjusted variance swap rates through  $c_z^{\mathbb{Q}}$ . Thus, it is only possible to estimate an aggregated variance jump compensator  $c_z^{\mathbb{Q}}$ , while the individual risk-neutral jump size parameters cannot be estimated in our MCMC algorithm.

Next, we consider the different impact of the market price of variance risk  $\eta_v$  and the aggregated variance jump compensator  $c_z^{\mathbb{Q}}$  on non-adjusted variance swap rates. The main differences are (i) their different impact on the term structure of non-adjusted variance swap rates and (ii) their different impact on the sensitivity between the latent variance process and non-adjusted variance swap rates. In detail, a negative market price of variance risk leads to an increasing variance swap rate in time to maturity, while the aggregated variance jump compensator  $c_z^{\mathbb{Q}}$  has a constant impact on variance swap rates in time to maturity. Furthermore, the sensitivity of variance swap rates to changes in the latent variance process only depends on the market price of variance risk  $\eta_v$  by means of its impact on the risk-neutral mean reversion rate  $\kappa_v^{\mathbb{Q}} = \kappa_v^{\mathbb{P}} + \eta_v$ . These differences allow us to separate both risk premia based on return and non-adjusted variance swap data in the MCMC estimation approach.

Next, we briefly present our estimation results for both risk premia in each model specification. The estimation results show that a significant aggregated market price of jump risk  $c_z^{\mathbb{P}} - c_z^{\mathbb{Q}}$  exists in all jump models, whereas a significant market price of variance risk is only found in the pure stochastic volatility model (see Table 2.2 and 2.3). The estimated large market price of variance risk in the SV model can be explained as follows: we observe a negative variance risk premium  $\frac{1}{\tau-t} (\mathbb{E}_t^{\mathbb{P}} [(\sigma_{t,\tau})^2] - \mathbb{E}_t^{\mathbb{Q}} [(\sigma_{t,\tau})^2])$  between squared log-returns and short-dated variance swap rates. To capture this, a very large market price of variance risk is required, since variance risk diminishes when time to maturity reaches zero. The estimated market price of variance risk in the SV model would then lead

to a strongly increasing (absolute) variance risk premium  $\frac{1}{\tau-t} |\mathbb{E}_t^{\mathbb{P}} [(\sigma_{t,\tau})^2] - \mathbb{E}_t^{\mathbb{Q}} [(\sigma_{t,\tau})^2]|$  in time to maturity  $\tau - t$  that is not empirically observed. For instance, the model-implied variance risk premium is about 10 time larger in absolute terms than its model-free empirical counterpart for a time to maturity of six months (see Kang and Pan (2011)). This indicates that another temporary risk factor, such as jump risk, is priced in the option market.

The robustness of our risk premia estimates is tested by considering empirical model-free variance risk premia for different time to maturities based on market data outside of our initial data set. Here, we compare realized squared log-returns of one-month-, two-month-, three-month-, and six-month-ahead futures contracts and the corresponding variance swap rates. We find that average realized squared log-returns are below their corresponding variance swap rates (negative variance risk premium) and that the absolute variance risk premium is not increasing in time to maturity (see Kang and Pan (2011)). This contradicts the estimated market price of variance risk in the SV model that would imply a strongly increasing absolute variance risk premium in time to maturity. In addition, the negative variance risk premium is largely captured by the aggregated market price of jump risk.

In addition to the above robustness test, we conduct a restricted estimation run with  $\eta_v = 0$  (zero market price of variance risk) for the complete time period for the SVJ model (SVJ0) in order to test the impact of the market price of variance risk on other model parameter estimates. It turns out that all parameter estimates remain largely unchanged, although we obtain slightly lower standard deviations for most of the model parameters in the SVJ0 specification (see Table 2.2).<sup>36</sup>

The MCMC estimation results show that a jump risk premium is paid in the crude oil futures market, but the individual risk-neutral jump parameters  $\mu_z^{\mathbb{Q}}$  and  $\sigma_z^{\mathbb{Q}}$  remain unknown. For that reason, we implement an additional estimation step that exploits the different impact of  $\mu_z^{\mathbb{Q}}$  and  $\sigma_z^{\mathbb{Q}}$  on the smile form of implied volatilities. In detail, we estimate the risk-neutral mean jump size and jump size volatility based on cross-sectional option price data given that the aggregated variance jump compensator is equal to the posterior mean of the MCMC estimate  $\hat{c}_z^{\mathbb{Q}}$ . We first allocate all option contracts to 10 moneyness categories ranging from 0.70 to 1.20 in steps of 0.05 and randomly choose one representative out-of-the-money option contract for each moneyness category for every business day.<sup>37</sup> Then, we fit  $\mu_z^{\mathbb{Q}}$  and  $\sigma_z^{\mathbb{Q}}$  to implied volatilities given that the aggregated

---

<sup>36</sup>For this reason, we set the market price of variance risk equal to zero in the estimation approach for the time period 2000-2010.

<sup>37</sup>We do not consider all available option contracts on any trading day due to the fact that traded option contracts are unequally distributed among different moneyness categories. Furthermore, we choose the moneyness range from 0.70-1.20, because trading volumes outside this interval are very low in our data set.

variance jump compensator is equal to the estimated one:

$$\min_{(\mu_z^{\mathbb{Q}}, \sigma_z^{\mathbb{Q}})} \sum_{i=1}^n \sum_{j=1}^{10} (iv_{t_i}^{mod}(k_j, \tau_i) - iv_{t_i}^{mar}(k_j, \tau_i))^2 \quad (2.36)$$

$$s.t. \quad 2\lambda_z(e^{\mu_z^{\mathbb{Q}}+0.5(\sigma_z^{\mathbb{Q}})^2} - 1 - \mu_z^{\mathbb{Q}}) = \widehat{c}_z^{\mathbb{Q}}, \quad (2.37)$$

where  $iv_{t_i}^{mar}(k_j, \tau_i)$  and  $iv_{t_i}^{mod}(k_j, \tau_i)$  are the market-implied and model-implied volatilities for out-of-the money option contracts with strike price  $k_j$  in the  $j$ -th moneyness category and maturity  $\tau_i$  for  $j = 1, \dots, 10$  and  $i = 1, \dots, n$ .

The estimation results show that the risk-neutral mean jump size is very close to the physical counterpart ( $\mu_z^{\mathbb{P}} = -0.0027$  (0.0019) and  $\mu_z^{\mathbb{Q}} = -0.005$  for the JD model and  $\mu_z^{\mathbb{P}} = -0.0201$  (0.0247) and  $\mu_z^{\mathbb{Q}} = -0.035$  for the SVJ model, whereas the jump size variance is considerable larger than its statistical counterpart ( $\sigma_z^{\mathbb{Q}} = 0.0586$  compared to  $\sigma_z^{\mathbb{P}} = 0.0442$  (0.0022) for the JD model and  $\sigma_z^{\mathbb{Q}} = 0.1743$  compared to  $\sigma_z^{\mathbb{P}} = 0.0921$  (0.0204) for the SVJ model). This indicates a positive volatility of price jumps risk premium and a mean price jump risk premium close to zero. Further, the risk-neutral jump parameters show that better option pricing fits can be obtained by adjusting the jump size variance parameter than by adjusting the mean jump size parameter to option market data.

In summary, we find that jump risk is an important risk factor that is priced with a significant premium in the crude oil market, while no significant premium is found for diffusive price and volatility risk. Our estimation results further indicate that ignoring jump risk, as done in previous studies (see, for example, Doran and Ronn (2008)), leads to an unreliable large market price of variance risk. These results are important when it comes to the pricing of complex bilateral agreements in delivery contracts or other derivative instruments.

### 2.4.3 Empirical Tests

In this subsection, we test the distributional properties, pricing performances, and hedging errors of the different modeling approaches for the underlying time periods (1985-2010, 1985-2000, and 2000-2010).

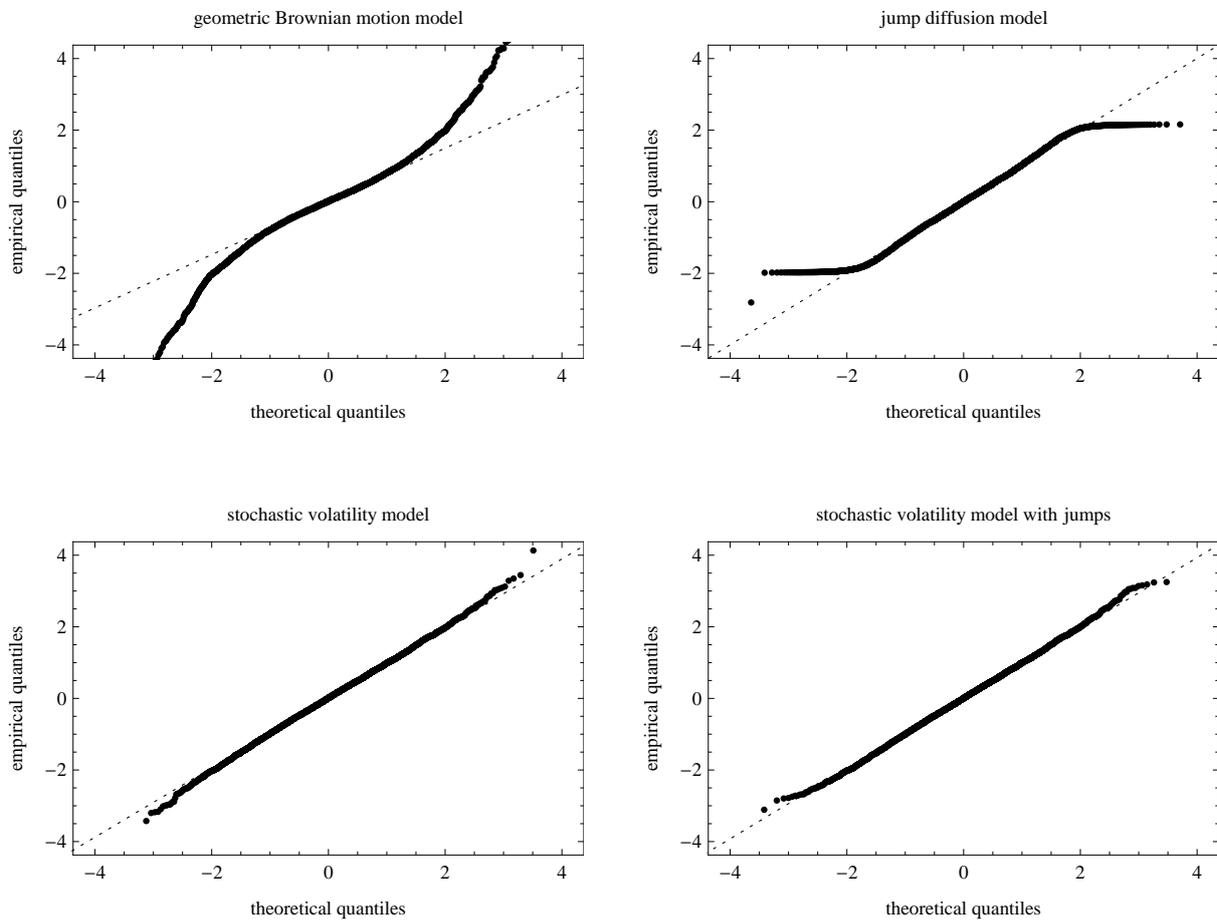
#### Distributional Properties

We use quantile-quantile-plots (QQ-plots) to test the distributional properties of the different model specifications. The residuals are extracted by reformulating the discretized

data-generating process (2.30) as follows:

$$\varepsilon_{f,t_i} = \frac{y_{t_i} - (-\lambda_z \bar{\mu}_z^Q + \bar{\eta}_f v_{t_i}) \delta t - z_{t_i} \delta n_{f,t_i}}{\sqrt{v_{t_i} \delta t}}, \quad i = 1, \dots, n. \quad (2.38)$$

If the underlying modeling approach is “correct”, the residuals are (approximately) normally distributed. The residuals are then tested for normality by simple QQ-plots, which compare theoretical with empirical quantiles. It is important to keep in mind that we have applied a Bayesian estimation approach. Thus, more complex model specifications do not automatically perform better than simpler (nested) model specifications.



**Figure 2.12: quantile-quantile-plots**

This figure shows the quantile-quantile-plots for the GB, JD, SV, and SVJ models based on log-returns of the WTI front-month crude oil futures contract during the years 1985-2010.

Figure 2.12 shows the QQ-plots for all tested models for the complete time period 1985-2010. It can be seen that a stochastic volatility component is required to capture the non-normal behavior of log-returns during 1985-2010. Moreover, the QQ-plot for the jump diffusion model shows that large negative returns are overestimated and large positive returns are underestimated, which is consistent with empirical results from Larsson and Nossman (2011).

We use the Bayesian Deviance Information Criterion (DIC) proposed by Spiegelhalter, Best, Carlin, and van der Linde (2002) to test the distributional properties of the different model specifications. This Bayesian information measure not only accounts for the “goodness of fit” to the data but also penalizes complexity. This makes it suited for model selection problems. The DIC scores are computed by using the simulated posterior distributions obtained from the MCMC algorithm. They are -29,705 for the GB model, -57,205 for the JD model, -131,753 for the SV model, and -140,653 for the SVJ model, where lower values translate into an overall superior model performance. The results confirm the importance of a stochastic volatility component, since the SV and the SVJ model scores are far smaller than the JD model. Moreover, the SVJ model specification performs best although the difference to the simpler SV model is not very large.

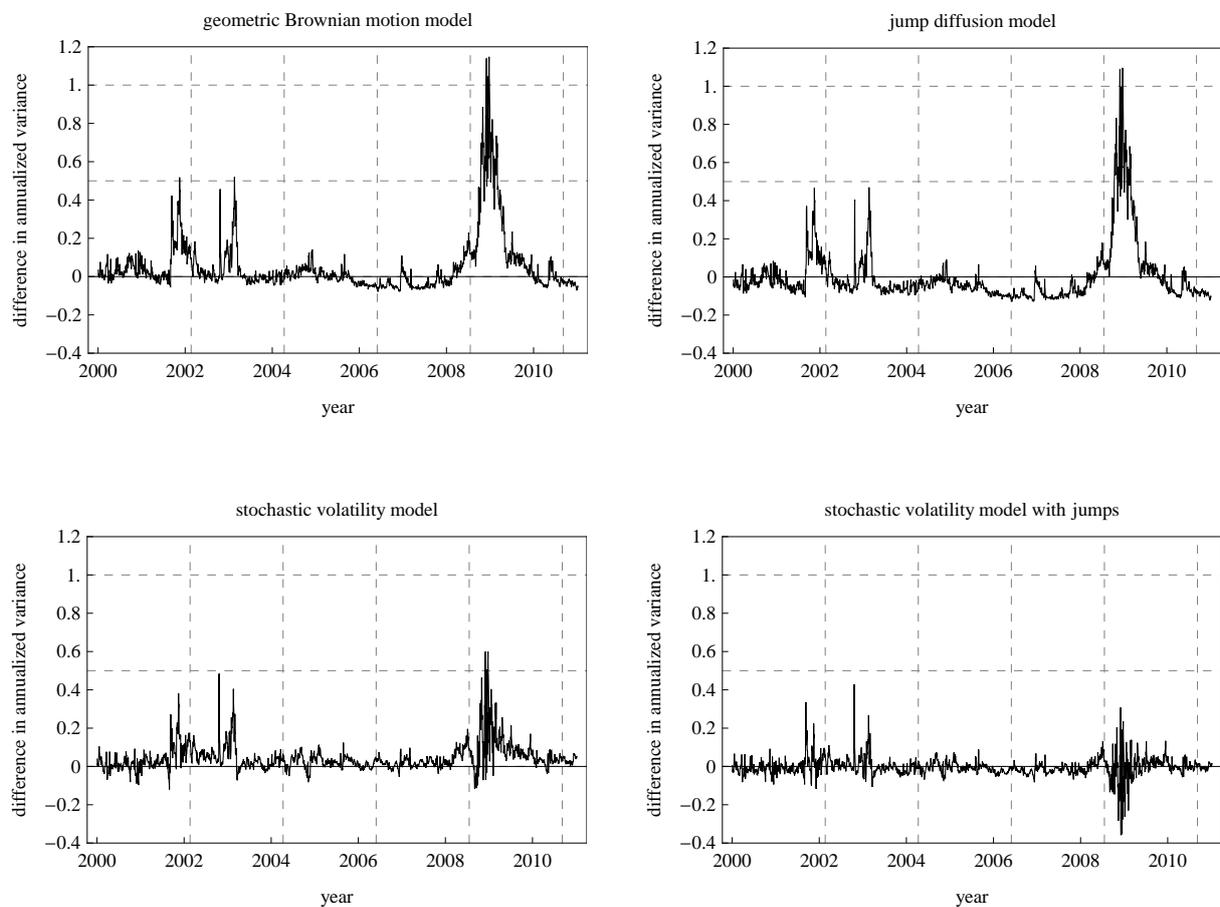
### Option Pricing Performance

In the next step, we compare the different modeling approaches concerning their option pricing performance. This empirical test is particularly useful for traders who want to assess whether their non-linear derivative instruments (e.g., bilateral delivery contracts) are valued consistently to the market for a given modeling approach.

The option pricing performance is measured through two criteria: (i) pricing errors between market-implied and model-implied variance swap rates and (ii) pricing errors between market-implied and model-implied volatilities for different moneyness categories given that the model fits variance swap rates perfectly over time.

The first criterion is used to test whether the underlying model can capture the stochastic behavior of variance swap rates over time. In Figure 2.13, we plot the residuals between the model and market values for variance swap rates. On average, the GB model underestimates the variance swap rates. The reason is that no variance risk exists in the GB model. Thus, the variation of the futures price process has to be the same under the physical and risk-neutral measure. Therefore, the difference between realized squared log-returns and variance swap rates cannot be captured in the GB model. In the JD model, the market-implied and model-implied average variance swap rates coincide through the aggregated market price of jump risk  $c_z^{\mathbb{P}} - c_z^{\mathbb{Q}}$ , but large pricing errors arise between constant model-implied and strongly fluctuating market-implied variance swap rates. The

pricing errors are significantly reduced in the SV and SVJ model due to the stochastic variance process. If we compare both stochastic volatility models, we find that the additional flexibility of the SVJ model allows one to further reduce pricing errors compared to the SV model (see Figure 2.13). In addition to the pricing errors, we calculate the absolute pricing errors (in annualized variance) to assess the overall pricing performance. We obtain the following average absolute pricing errors for the different model specifications: 0.1590 (GB), 0.1501 (JD), 0.0601 (SV), and 0.0461 (SVJ). As expected, absolute pricing errors are at the lowest for the SV and SVJ models.



**Figure 2.13: time series of variance swap pricing errors during 2000-2010**

These graphs show the difference between market-implied and model-implied variance swap rates for the GB model (2.6) (top (left)), the JD model (2.9) (top (right)), the SV model (2.7) (bottom (left)), and the SVJ model (2.1) (bottom (right)). The market-implied variance swap rates are calculated based on option contracts on WTI front-month crude oil futures prices during 2000-2010.

The second criterion is used to test whether the underlying model can reproduce the shape of implied volatilities given that the model-implied variance swap rate is equal to the market-implied variance swap rate on each business day. This side condition is met by recalibrating the constant variance parameter  $\sigma_f^2$  (GB and JD models) or latent variance states  $\{v_{t_i}\}_{i=1}^n$  (SV and SVJ models) to the variance swap rate on every business day through relation (2.16) holding all other model parameters fixed. This ensures that pricing errors arising from incorrectly estimated implied volatility levels are not mixed with pricing errors that arise when a model is not able to reproduce the smile or skew form of implied volatilities. Then, we calculate the root mean squared error between market-implied and model-implied volatilities for each option contract with strike price in one of the moneyness categories (0.70 to 1.20) on every business day. Table 2.4 provides the mean absolute pricing errors in each moneyness category during 2000-2010, and Table 2.5 contains the results for the two subsamples 2000-2008 (non-crisis period) and 2008-2009 (crisis period). We find that the GB, JD, and SV model provide poor pricing performance for the moneyness categories 0.7-0.8 (out-of-the-money put options). The large pricing errors of the SV can be explained by the fact that volatility risk alone is not able to generate enough excess kurtosis to capture market-implied volatility smiles. Moreover, the poor option pricing performance of the JD model can be traced back to the estimated jump intensity. The jump component implies frequent price jumps of smaller magnitude. This leads to an underestimation of tail risk and to an overestimation of at-the-money implied volatilities. In contrast, the SVJ model has pricing errors that are substantially smaller for out-of-the money put option contracts, since rare and large jumps are able to generate enough excess kurtosis to capture pronounced market-implied volatility smiles. In addition, we compare the pricing errors during 01/2000-09/2008 (non-crisis period) and 09/2008-09/2009 (crisis period). We find that pricing errors are slightly larger during the Financial Crisis, but the results are qualitatively similar in both subsamples.

moneyness	01/2000-12/2010			
	GB	JD	SV	SVJ
0.70 – 0.75	0.1044	0.0818	0.0819	0.0559
0.75 – 0.80	0.0759	0.0523	0.0541	0.0384
0.80 – 0.85	0.0529	0.0370	0.0372	0.0321
0.85 – 0.90	0.0345	0.0319	0.0263	0.0255
0.90 – 0.95	0.0216	0.0303	0.0203	0.0212
0.95 – 1.00	0.0206	0.0282	0.0197	0.0224
1.00 – 1.05	0.0246	0.0295	0.0213	0.0247
1.05 – 1.10	0.0281	0.0358	0.0249	0.0277
1.10 – 1.15	0.0399	0.0408	0.0361	0.0360
1.15 – 1.20	0.0520	0.0510	0.0480	0.0479
average pricing error	0.0454	0.0419	0.0370	0.0332

**Table 2.4: option pricing errors during 2000-2010**

This table reports average absolute errors between the model-implied and market-implied volatilities of option contracts on WTI front-month crude oil futures prices between 2000 and 2010. The values correspond to annual decimals.

moneyness	01/2000-9/2008				09/2008-09/2009			
	GB	JD	SV	SVJ	GB	JD	SV	SVJ
0.70 – 0.75	0.0910	0.0601	0.0642	0.0514	0.1044	0.0903	0.0826	0.0674
0.75 – 0.80	0.0626	0.0409	0.0415	0.0327	0.0848	0.0726	0.0693	0.0576
0.80 – 0.85	0.0450	0.0322	0.0298	0.0292	0.0636	0.0564	0.0542	0.0492
0.85 – 0.90	0.0310	0.0309	0.0232	0.0237	0.0438	0.0411	0.0393	0.0393
0.90 – 0.95	0.0204	0.0310	0.0193	0.0199	0.0322	0.0322	0.0315	0.0326
0.95 – 1.00	0.0195	0.0228	0.0187	0.0218	0.0332	0.0323	0.0324	0.0316
1.00 – 1.05	0.0218	0.0294	0.0189	0.0242	0.0428	0.0402	0.0397	0.0375
1.05 – 1.10	0.0223	0.0332	0.0206	0.0268	0.0494	0.0467	0.0445	0.0418
1.10 – 1.15	0.0291	0.0339	0.0308	0.0330	0.0536	0.0518	0.0477	0.0465
1.15 – 1.20	0.0496	0.0504	0.0556	0.0531	0.0556	0.0544	0.0493	0.0492
average pricing error	0.0392	0.0370	0.0323	0.0311	0.0563	0.0518	0.0491	0.0452

**Table 2.5: option pricing errors during 01/2000-09/2008 and 09/2008-09/2009**

The left table reports average absolute errors between the model-implied and market-implied volatilities of option contracts on WTI front-month crude oil futures prices before the Financial Crisis. The right figure provides the result during the Financial Crisis. The values correspond to annual decimals.

## Hedging Performance

Lastly, we analyze hedging errors of option contracts for our model specifications. The hedging performance provides important information whether the underlying stochastic process can capture the co-movement of futures and option prices. This is especially important for physical traders who manage large hedge portfolios and have to quantify the risk inherent in their financial and physical asset portfolios.

In our empirical test, we compare the in-sample hedging performance of the different model specifications for all option contracts that are traded on succeeding days from January 1, 2000 until December 31, 2010. The hedge portfolio is constructed as follows: we first recalibrate the variance parameter (GB and JD models) or the latent variance state (SV and SVJ models) to each target option contract on every trading day. This ensures that target option contracts are correctly valued in the underlying modeling approach, which minimizes hedging errors that arise due to a poor option pricing performance. Then, we calculate the hedge positions in the futures and option contracts according to the delta and delta-vega hedging strategy (see Subsection 2.2.3). In the delta-vega hedging strategy, we use an option contract from the nearest moneyness category as additional hedging instrument against variance risk. We also test option contracts from other moneyness categories (e.g., at-the-money option contracts) as hedging instruments and obtain the same qualitative results, where hedging errors are generally lower if an option contract with a similar strike price is used as hedging instrument.

Now, we choose such hedge portfolios for each moneyness category. Then, we calculate the hedging errors given by the daily returns of the hedge portfolio minus the interest rate effect for each business day. Table 2.6 shows the absolute hedging errors for the different hedging strategies between 01/2000-12/2010 and 09/2008-09/2009. The hedging errors have similar means and standard deviations for the delta hedging strategy under all model specifications and are larger during the crisis period. In contrast, the hedging errors of the delta-vega hedging strategy, which actively manage variance risk, have significantly lower means and standard deviations for the SV and SVJ models. This confirms our estimation result of weakly correlated futures price and volatility innovations (“unspanned stochastic volatility”). In addition, we analyze the distribution of the hedging errors under all model specifications for the delta and delta-vega hedging strategy. In the delta hedging strategy, we find large hedging errors of -3.6 USD to 1.2 USD between two trading days in all model specifications. These hedging errors can potentially arise due to discrete hedging or price jumps. In order to test the impact of discrete hedging on hedging errors, we conduct a simulation study. We simulate ten thousand futures price returns based on the estimated model parameters and calculate the hedging errors given that the hedge portfolio is rebalanced on a daily basis. In the continuous price dynamics (GB and SV

models), maximal simulated (positive and negative) hedging errors are -0.6 USD and 0.1 USD for the GB model and -0.7 USD and 0.6 USD for the SV model. This shows that discrete hedging alone cannot be the reason for the large empirical hedging errors that are found in the crude oil market. If we consider unhedgeable jump risk in the JD and SVJ model, we obtain hedging errors that cover a wider range: -1.8 USD to 0.1 USD for the JD model and -2.8 USD to 0.5 USD for the SVJ model. The difference between both jump models can be explained by a more extreme jump distribution in the SVJ model. In the delta-vega hedging strategy, empirical hedging errors are significantly reduced. The largest positive and negative real hedging errors for our data set are -0.5 USD and 0.2 USD based on the SV and SVJ model.<sup>38</sup> Once again, the pure diffusive stochastic volatility model underestimates the tails risk as simulated hedging errors only range from only -0.1 USD to 0.1 USD. The SVJ model leads to values of -0.5 USD to 0.5 USD. This means that the lower bound is perfectly fitted, while the upper bound is overestimated in the SVJ model. In summary, ignoring jump risk in modeling approaches results in a significant underestimation of the downside risk of hedge portfolios.

	01/2000-12/2010		09/2008-09/2009	
	Delta	Delta-Vega	Delta	Delta-Vega
GB	-0.0443 (0.3197) [-1.6084, 0.3689]	- - -	-0.1034 (0.6012) [-3.0424, 0.7634]	- - -
JD	-0.0435 (0.3187) [-1.5661, 0.3649]	- - -	-0.1030 (0.6007) [-3.0430, 0.7629]	- - -
SV	-0.0437 (0.3196) [-1.5739, 0.3614]	-0.0024 (0.0211) [-0.0928, 0.0238]	-0.1032 (0.06012) [-3.0414, 0.7632]	-0.0016 (0.0171) [-0.0904, 0.0169]
SVJ	-0.0443 (0.3195) [-1.5364, 0.3588]	-0.0023 (0.0211) [-0.0928, 0.0237]	-0.1028 (0.6006) [-3.0419, 0.7636]	-0.0018 (0.0186) [-0.0926, 0.0176]

**Table 2.6: hedging errors during 2000-2010 and 09/2008-09/2009**

This table reports absolute hedging errors under each model specification for the delta and delta-vega hedging strategies. The underlying data set consists of WTI front-month crude oil futures and option contracts for the years 2000-2010 (complete time period) and 09/2008-09/2009 (crisis period). The table provides posterior means, standard deviations (in parenthesis), and 1% to 99% credibility intervals (in square brackets) for the GB, JD, SV, and SVJ models.

<sup>38</sup>We do not test pseudo delta-vega strategies that hedge variance risk, even if no variance risk exists in the model (see Bakshi, Cao, and Chen (1997)).

	<i>SV</i>	<i>SV*</i>	<i>SVJ</i>	<i>SVJ*</i>
<i>ES</i> <sub>0.999</sub>	0.3550	0.0910	0.4120	0.5316
<i>ES</i> <sub>0.995</sub>	0.2016	0.0618	0.2444	0.1897

**Table 2.7: expected shortfalls for hedge portfolios**

This table provides expected shortfall values for real absolute hedging errors as well as simulated ones (denoted by the asterisk \*) under the SV and SVJ model specification for the 0.5% and 0.1% quantile.

In order to gauge in more detail the impact of unhedgable risk in our hedging portfolios, we compute expected shortfalls for the delta-vega hedging strategy of simulated and actual absolute hedging errors under both stochastic volatility model specifications (see Table 2.7). It can be seen that the SV model significantly underestimates the tail risk, while the simulation-based expected shortfall under the SVJ model is much closer to the actual market-based risk measure. These results confirm again that jump risk must be taken into account if the risk of hedge portfolios has to be quantified accurately.

In summary, we have tested the role of different risk factors for the crude oil futures and option markets from different perspectives. We find that volatility risk is required to capture clustered large returns during economic crisis and times of war, as well as strongly fluctuating variance swap rates over time. However, pure stochastic volatility models cannot capture pronounced implied volatility smiles and the risk inherent in hedge portfolios. This indicates that a further temporary unhedgeable risk factor is priced in the market. Our empirical results show that jump risk is an adequate candidate for such a temporary risk factor. For our data set, jump risk is able to capture quite well pronounced implied volatility smiles and the risk of hedge portfolios. In addition, we find clear evidence for a jump risk premium that is reflected in larger average variance swap rates compared to average squared log-returns. Finally, no evidence for a positive or negative market price of variance risk is found in our data set.

## Chapter 3

# Stochastic Term Structure Modeling Framework

In the previous chapter, we have analyzed different stochastic modeling approaches for a single futures contract. These models are suited for valuing and managing many financial products but fail when it comes to pricing important real options. The valuation of real options requires one to set up an appropriate stochastic term structure model for the common stochastic price behavior of multiple futures contracts. This is a non-trivial task, since specific features of physical energy trading have to be considered in order to obtain a consistent modeling approach. It is not possible, for instance, to apply the classical cost-of-carry relation between spot and futures prices due to physical storage costs. Furthermore, due to limited delivery rates, spot and futures contracts have delivery periods instead of delivery dates. This requires one to distinguish between theoretical “spot” and “futures” contracts with delivery dates, which are often considered in theoretical modeling approaches, and real “spot” and “futures” contracts with delivery periods, which are observed in the market.<sup>1</sup>

This chapter presents a tractable modeling framework for the entire futures price curve that accounts for the specific characteristics of physical and financial energy markets. In the literature, two main strands to model energy price dynamics can be separated. In the first strand, the spot price dynamics or the price dynamics of theoretical futures contracts is modeled exogenously and real futures prices referring to delivery periods are endogenously derived based on the no-arbitrage principle (see, for example, Schwartz (1997), Schwartz and Smith (2000), and Trolle and Schwartz (2009)). The no-arbitrage relation states that each real futures price is equal to the average theoretical futures price

---

<sup>1</sup>In the following, we skip “theoretical” and “real” in front of spot and futures contracts if no ambiguity exists whether the underlying contract has a delivery period or not.

during its delivery period. It turns out that, through the aggregation step, the endogenous price dynamics of real futures contracts are generally intractable and thus difficult to calibrate to market data. In the second strand, the price dynamics of real futures contracts are modeled exogenously (see, for example, Koekebakker and Ollmar (2005) or Benth and Koekebakker (2008)). This has the advantage that model parameters can be directly estimated from historical return data or can be extracted from option price data. The disadvantage is that it is not possible to derive arbitrage-free price dynamics of futures contracts that are not exogenously modeled within the market modeling framework. For example, day-ahead prices remain unspecified if the market model is calibrated to the price dynamics of futures contracts referring to monthly delivery periods (see, for example, Benth and Koekebakker (2008)). This restricts the application of standard market modeling approaches. For instance, storage operators require a stochastic model for day-ahead and futures price dynamics, since they trade in both (i) day-ahead contracts to optimize their physical operation policies and (ii) standardized futures contracts to efficiently hedge their price risks.

In the following, we introduce a modeling framework that tackles the trade-off between tractability and completeness by inverting classical spot and futures price models. We start with a classical market model for the price dynamics of standardized futures contracts referring to delivery periods. We then apply a smooth interpolation function to endogenously derive arbitrage-free price dynamics of theoretical spot and futures contracts relative to the exogenously given price process of real futures contracts instead of vice versa. This completes our modeling approach and makes it capable of valuing a broad range of important energy derivatives, such as power plants, storage facilities, and take-or-pay contracts. Ultimately, our “inverted” approach results in a multi-factor spot price model that depends on observable futures prices instead of latent factors. This direct link can be used to efficiently hedge price risks in customized contracts using liquid exchange-traded products.

This chapter is organized as follows: we first discuss the trade-off between the tractability and completeness of standard modeling frameworks. We then introduce our modeling approach and demonstrate its convenience in the empirical part.

### 3.1 Trade-off between Tractability and Completeness

In general, stochastic modeling approaches for energy price dynamics are similar to classical fixed income models and can be separated into two main classes: (i) market models and (ii) spot and futures price models.

The starting point of market models is the price dynamics of a finite number of traded fu-

tures contracts with fixed delivery periods. This allows one to simplify estimation methods for model parameters, since model parameters can be directly fitted to observable market data. However, market models also lead to undefined price dynamics for delivery contracts with delivery periods that are not exogenously modeled. This restricts the application of market models to standard valuation purposes and prevents their use for the valuation of important real options.

In contrast to market models, spot and futures prices models start with the price dynamics of theoretical delivery contracts referring to delivery dates. It is then possible to derive arbitrage-free futures price dynamics for arbitrary delivery periods based on the risk-neutral valuation approach. For instance, if real futures contracts refer to uniform deliveries in their delivery periods, the no-arbitrage relation between theoretical and real futures prices is given by

$$f_t(\tau_b, \tau_e) = \int_{\tau_b}^{\tau_e} \widehat{w}(u; \tau_b, \tau_e) f_t(u) du, \quad (3.1)$$

where  $f_t(\tau_b, \tau_e)$  corresponds to the real futures price for the delivery period  $(\tau_b, \tau_e]$  and  $f_t(u)$  denotes the theoretical futures price for the delivery date  $u \in (\tau_b, \tau_e]$ . The function  $\widehat{w}(u; \tau_b, \tau_e)$  depends on the settlement procedure of the futures contract. It is equal to

$$\widehat{w}(u; \tau_b, \tau_e) = \frac{\exp(-ru)}{\int_{\tau_b}^{\tau_e} \exp(-rv) dv}$$

if the settlement takes place uniformly during the delivery period.<sup>2</sup> Next, we consider a widely used spot price model to illustrate the problem of fitting spot and futures price dynamics to market data. Schwartz and Smith (2000) separate the log spot price  $s_t$  into a short-term mean reversion component  $\chi_t$  capturing temporary price impacts (e.g., weather shocks) and a long-term component  $\xi_t$  capturing permanent price impacts (e.g., economic growth). In their model, the risk-neutral spot price dynamics is given by<sup>3</sup>

$$\ln s_t = \chi_t + \xi_t, \quad (3.2)$$

---

<sup>2</sup>In real energy markets, futures contracts are settled at discrete dates. This can be easily incorporated by modifying  $\widehat{w}(u; \tau_b, \tau_e)$  (see Benth, Koekebakker, and Ollmar (2007) or Benth and Koekebakker (2008)).

<sup>3</sup>The superscript  $\mathbb{Q}$  is used to indicate that a model parameter can differ between the physical and the risk-neutral measure.

where

$$d\chi_t = -\kappa_\chi^\mathbb{Q}\chi_t dt + \sigma_\chi dw_{\chi,t}^\mathbb{Q}, \quad (3.3)$$

$$d\xi_t = \mu_\xi^\mathbb{Q} dt + \sigma_\xi dw_{\xi,t}^\mathbb{Q} \quad (3.4)$$

with correlated standard Wiener processes  $d[w_{\chi,t}^\mathbb{Q}, w_{\xi,t}^\mathbb{Q}] = \rho_{\chi\xi} dt$ . It follows then that spot prices are log-normally distributed. This allows one to calculate the current futures price curve in closed-form as follows:

$$\begin{aligned} f_t(u) &= \mathbb{E}_t^\mathbb{Q}[s_u] \\ &= \exp\left\{\mathbb{E}_t^\mathbb{Q}[\ln s_u] + \frac{1}{2}\text{Var}_t^\mathbb{Q}[\ln s_u]\right\}, \quad u \geq t, \end{aligned} \quad (3.5)$$

with

$$\begin{aligned} \mathbb{E}_t^\mathbb{Q}[\ln s_u] &= e^{-\kappa_\chi^\mathbb{Q}(u-t)}\chi_t + \xi_t + \mu_\xi^\mathbb{Q}(u-t), \\ \text{Var}_t^\mathbb{Q}[\ln s_u] &= \frac{\sigma_\chi^2}{2\kappa_\chi^\mathbb{Q}}(1 - e^{-2\kappa_\chi^\mathbb{Q}(u-t)}) + \sigma_\xi^2(u-t) + \frac{2\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{\kappa_\chi^\mathbb{Q}}(1 - e^{-\kappa_\chi^\mathbb{Q}(u-t)}). \end{aligned}$$

The estimation of the model parameters  $(\kappa_\chi^\mathbb{Q}, \sigma_\chi, \mu_\xi^\mathbb{Q}, \sigma_\xi, \rho_{\chi\xi})$  requires one to calculate expected average spot prices during the respective delivery period of the traded futures contract. However, the integral equation (3.1) cannot be solved in closed-form for (3.5). Thus, computationally intensive numerical estimation methods are required. This estimation problem is not specific to the Schwartz/Smith model and generally emerges for all non-additive stochastic processes (see Benth, Kallsen, and Meyer-Brandis (2007)).

There are two main approaches used to simplify the estimation of model parameters for non-additive stochastic processes. First, spot or futures price processes are directly fitted to (inconsistent) proxies for unobservable (theoretical) spot or futures prices. For instance, Gibson and Schwartz (1990), Schwartz (1997), Schwartz and Smith (2000), Casassus and Collin-Dufresne (2005), and Cartea and Williams (2008) ignore delivery periods of crude oil or natural gas futures contracts in their empirical studies. Second, model parameters are estimated based on a two-stage estimation approach. In the first estimation step, theoretical spot and futures prices are derived from real futures prices based on an interpolation function. In the second estimation step, model parameters of the underlying spot and futures price processes are fitted to interpolated theoretical spot and futures prices. This approach generally does not ensure consistent parameter estimates, since two independent interpolation functions are used: the interpolation function applied to extract theoretical futures prices and the endogenous interpolation function (futures price curve) implied by the spot price process (see, for example, Koekebakker and Ollmar (2005)).

For additive stochastic processes (e.g., affine-linear models), we obtain tractable price dynamics for real futures contracts (see Bouwman, Raviv, and van Dijk (2012)). This simplifies estimation methods but strongly restricts the number of potential model specifications. For instance, the model specifications considered in the previous chapter (e.g., the Black or Heston model) or in other studies (see, for example, Koekebakker and Ollmar (2005), Benth and Koekebakker (2008), or Trolle and Schwartz (2009)) are not contained in the class of additive stochastic processes. In addition, (standard) affine-linear stochastic models are not able to capture futures price movements that do not depend on time to maturity (e.g., a stochastic summer-winter spreads in the natural gas market). However, such price movements are very important when it comes to the pricing of real option contracts (see Chapter 4).

## 3.2 The Model Framework

In this section, we introduce the theoretical framework behind our energy market modeling approach. The underlying energy market consists of  $m + 1$  real futures contracts with successive delivery periods  $\{(\tau_i, \tau_{i+1}]\}_{i=0}^m$  that are traded until their first delivery dates  $\{\tau_i\}_{i=0}^m$  and refer to uniform deliveries in their delivery periods.<sup>4</sup> The market price of a futures contract with delivery period  $(\tau_i, \tau_{i+1}]$  is denoted by  $f_t^{(i)} = f_t(\tau_i, \tau_{i+1})$ .

Our modeling approach is developed in two major steps. We start with the price dynamics of traded futures contracts and then describe our smooth interpolation approach to complete our pricing framework.

### 3.2.1 Real Futures Contracts

In general, we can model the common stochastic price behavior of the underlying real futures contracts via arbitrary adapted càdlàg martingale processes with finite variations (see Jeanblanc, Yor, and Chesney (2009)).<sup>5</sup> Especially, standard stochastic price processes (e.g., jump diffusion, stochastic volatility, and regime switching models) can be used to

---

<sup>4</sup>In most energy markets, futures contracts have successive delivery periods and are traded until a few business days before their first physical delivery days. For instance, natural gas futures contracts are traded until three business days before their first physical delivery days at the CME, while crude oil futures contracts are traded until three business days before the 25th of the month preceding the respective delivery month at the CME. In the U.S. natural gas market, delivery has to take place as uniformly as possible on an hourly basis. In contrast, in the crude oil market, pipeline operators decide on the delivery day or period on a pro rata basis.

<sup>5</sup>The term “càdlàg” (“continues à droite, limites à gauche”) means that the process is continuous on the right and has finite left limits.

model futures price dynamics. Further, the price dynamics of futures contracts with non-overlapping delivery periods can be modeled separately, since (significant) physical storage costs distort the classical cost-of-carry relation in energy markets.

### 3.2.2 Theoretical Spot and Futures Contracts

The market model does not specify the futures price dynamics that refer to delivery periods  $(\tau_b, \tau_e]$  that are not contained in  $\{(\tau_i, \tau_{i+1}]\}_{i=0}^m$ , i.e.,  $\tau_b \notin \{\tau_i\}_{i=0}^{m+1}$  or  $\tau_e \notin \{\tau_i\}_{i=0}^{m+1}$ . It is not possible, for example, to derive the day-ahead price dynamics from a market model for futures contracts with monthly delivery periods. This leads to an incomplete modeling approach that restricts its practical application for the valuation of important real options.

In our approach, we obtain arbitrage-free price dynamics for all delivery contracts through an interpolation function. The interpolation function  $f_t(u)$  infers theoretical futures prices from real futures prices based on two no-arbitrage conditions and a so called maximum “smoothness” criterion that avoids strongly oscillating futures price curves.<sup>6</sup>

Next, we specify the interpolation function. The first no-arbitrage condition states that two portfolios with the same physical delivery flows must have the same market value at any point in time.

**Condition 1** (*Static No-Arbitrage Condition*)

*The futures price curve satisfies the static no-arbitrage relation at any point in time  $t$ :*

$$f_t(\tau_i, \tau_{i+1}) = \int_{\tau_i}^{\tau_{i+1}} \hat{w}(u; \tau_i, \tau_{i+1}) f_t(u) du, \quad t \in [\tau_0, \tau_i], \quad (3.6)$$

for  $i = 0, \dots, m$ .

The static no-arbitrage relation is imposed even when theoretical futures contracts are non-traded instruments in order to guarantee an arbitrage-free modeling approach for energy markets with arbitrary physical delivery contracts. Moreover, the no-arbitrage principle requires that endogenous futures price dynamics are martingales under the risk-neutral measure regardless of the market model approach.

---

<sup>6</sup>We use the same notation for the futures price curve and the interpolation function. The reason is that the value of the interpolation function  $f_t(u)$  is equal to the theoretical futures price for the delivery date  $u$  at time  $t$ .

**Condition 2** (*Dynamic No-Arbitrage Condition*)

The futures price dynamics satisfy the martingale property

$$f_t(u) = \mathbb{E}_t^{\mathbb{Q}}[f_l(u)], \quad t \leq l \leq u, \quad (3.7)$$

at any point in time  $t \in [\tau_0, u]$ .

These two conditions are both necessary for an arbitrage-free pricing framework, but they do not ensure that endogenous futures price curves are reasonable. Notably, interpolation functions of higher order, which are required to satisfy the static no-arbitrage relation, tend to be strongly oscillating. To circumvent this problem, we impose a so called “maximum smoothness” criterion on the interpolation function. This condition was first introduced by McCulloch (1971) for yield curves and was applied for energy markets by Benth, Koekebakker, and Ollmar (2007). It minimizes the average second derivative of the interpolation function over the underlying time period.

**Condition 3** (*Maximum Smoothness Condition*)

The futures price curve is twice continuously differentiable and minimizes the squared second derivative

$$\min_g \int_t^{\tau_{m+1}} \left( \frac{\partial^2 g_t(u)}{\partial^2 u} \right)^2 du \quad (3.8)$$

with respect to all interpolation functions  $g_t$  with zero derivative in  $\tau_{m+1}$  that satisfy the two no-arbitrage conditions.<sup>7</sup>

These three conditions uniquely define the interpolation function.

**Lemma 3** (*Futures Price Curve*)

The two no-arbitrage conditions and the maximum smoothness condition yield to the following relation between theoretical and real futures prices:

$$f_t(u) = \sum_{i=0}^m \beta_i(u) f_t(\tau_i, \tau_{i+1}), \quad (3.9)$$

$$f_t(\tau_b, \tau_e) = \sum_{i=0}^m \left( \int_{\tau_b}^{\tau_e} \hat{w}(u; \tau_b, \tau_e) \beta_i(u) du \right) f_t(\tau_i, \tau_{i+1}), \quad (3.10)$$

---

<sup>7</sup>The maximum smoothness criterion can also be applied without the zero derivative condition to avoid oscillating interpolation functions.

where the weighting functions  $\{\beta_i(u)\}_{i=0}^m$  fulfill the following conditions:

$$\begin{aligned}\beta_i(u) &= a_i + b_i u + c_i u^2 + d_i u^3 + e_i u^4, \\ \sum_{i=0}^m \beta_i(u) &\equiv 1, \\ \int_{\tau_i}^{\tau_{i+1}} \widehat{w}(u; \tau_i, \tau_{i+1}) \beta_j(u) du &= \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases}, \quad \text{for } i, j = 0, \dots, m.\end{aligned}$$

The spline parameters of the weighting functions  $\{(a_i, b_i, c_i, d_i, e_i)\}_{i=0}^m$  are uniquely determined by the usual spline conditions, the static no-arbitrage condition, and the maximum smoothness criterion (see Appendix A.3).

**Proof:** Benth, Koekebakker, and Ollmar (2007) show that splines of order four are needed to satisfy the maximum smoothness criterion subject to the static no-arbitrage condition. The linear relation (3.9) between theoretical and real futures prices exists due to a linear relation between the spline parameters  $\{(a_i, b_i, c_i, d_i, e_i)\}_{i=0}^m$  and real futures prices (see Appendix A.3). Note that the dynamic no-arbitrage condition for arbitrary martingale processes of real futures contracts is only satisfied for linear relations between theoretical and real futures prices.

Furthermore, the maximum smoothness criterion implies that parallel shifts in real futures prices yield to parallel shifts in the theoretical futures price curve. Thus, weighting functions add up to one. The fact that the integral of the weighting function over a delivery period has to be either zero or one follows directly from the static no-arbitrage condition. ■

Based on Lemma 3, we can derive price dynamics for delivery contracts with arbitrary delivery dates or periods by applying Itô's lemma.

**Lemma 4** (*Spot and Futures Price Dynamics*)

The risk-neutral price dynamics of theoretical and real spot and futures contracts are given by

(i) *spot price dynamics*

$$ds_t = df_t(t) = \sum_{i=0}^m \left( \frac{\partial \beta_i}{\partial t}(t) \right) f_t(\tau_i, \tau_{i+1}) dt + \sum_{i=0}^m \beta_i(t) df_t(\tau_i, \tau_{i+1}), \quad (3.11)$$

where the (theoretical) spot price refers to a rolling immediate delivery date,

(ii) (theoretical) futures price dynamics

$$df_t(u) = \sum_{i=0}^m \beta_i(u) df_t(\tau_i, \tau_{i+1}), \quad (3.12)$$

where the (theoretical) futures price refers to a fixed delivery date  $u$ ,

(iii) (real) futures price dynamics

$$df_t(\tau_b, \tau_e) = \sum_{i=0}^m \left( \int_{\tau_b}^{\tau_e} \hat{w}(u; \tau_b, \tau_e) \beta_i(u) du \right) df_t(\tau_i, \tau_{i+1}), \quad \tau_b < \tau_e, \quad (3.13)$$

where the (real) futures contract refers to an arbitrarily fixed delivery period  $(\tau_b, \tau_e]$ .

In (3.11), the spot contract refers to a rolling delivery date. This implies that physical storage capacities are required to trade in the spot contract. Therefore, the spot price process does not have to satisfy the martingale property. Instead, its drift component is equal to the current slope of the futures price curve. In contrast, theoretical and real futures prices refer to fixed delivery dates or periods. Thus, both stochastic processes satisfy the martingale property under the risk-neutral measure.

The key difference to alternative spot price models is that observable instead of theoretical futures price dynamics are modeled exogenously. This means that model prices for traded futures contracts do not have to first be endogenously derived before model parameters can be estimated on market data. Note that the drift component of our spot price process linearly depends on observable futures prices instead of latent factors. As a result, no-arbitrage conditions can be met without restricting the price dynamics of traded futures contracts. This allows us to obtain a spot price process that results in tractable futures price dynamics for real delivery contracts in contrast to standard spot price dynamics. The relations between our “energy market model” and standard normal and log-normal spot and futures price processes are illustrated in Appendix A.4.

### 3.3 Implementation and Empirical Results

In what follows, we give general recommendations for the implementation of our modeling approach for the U.S. crude oil and natural gas markets. In principle, our energy market model is implemented in three steps. First, we have to choose the exogenously given benchmark contracts within the market model component. Second, we have to specify and estimate an appropriate stochastic process for the futures price dynamics based on available market data. Third, we have to complete our pricing framework through the

smooth interpolation function. These three implementation steps will now be illustrated using the U.S. crude oil and natural gas markets as examples.

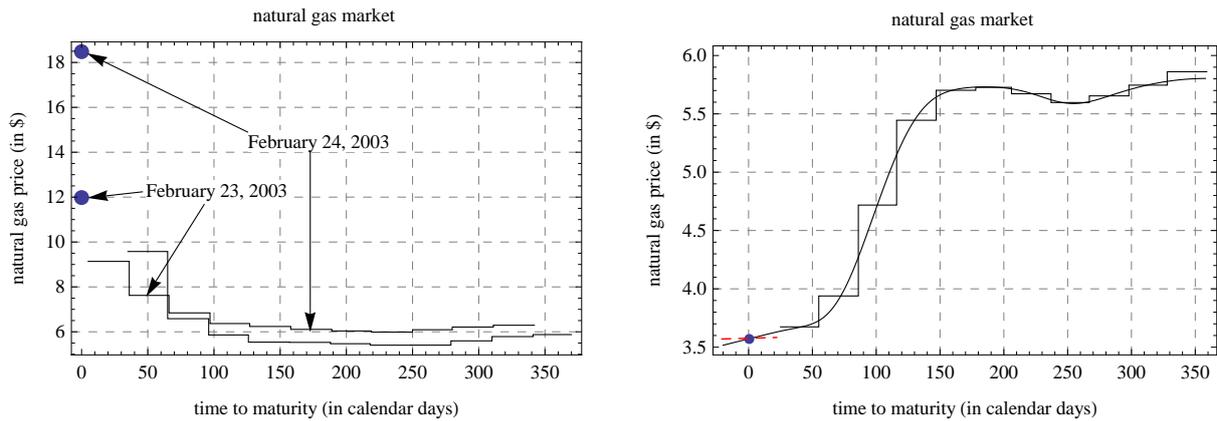
### 3.3.1 Choice of Market Contracts

The first implementation step of our energy market model is to select the futures contracts used within the market modeling approach. Ideally, the market model could be based on futures contracts with non-overlapping short delivery periods that span the whole time period that is relevant for the application of the modeling approach. This would provide a good picture of temporary imbalances between supply and demand in the market, reflect market expectations about future prices, and avoid interpolation and extrapolation errors. Indeed, most exchange-traded futures contracts refer to non-overlapping delivery periods in energy markets so that available market prices can be directly incorporated into the market model.

In the following, we consider the role of futures contracts traded at the CME for crude oil and natural gas markets.<sup>8</sup>

#### Crude Oil

In the U.S. crude oil market, exchange-traded futures contracts refer to delivery periods for each calendar month in the following five years and are traded until three business days before the 25th of the month preceding the delivery month. It seems uncritical that no price information for the current calendar month exists, since crude oil demand and supply do not strongly depend on temporary factors, and physical trading is mostly restricted to delivery contracts having delivery periods after the current month due to pipeline scheduling practice.<sup>9</sup> Thus, the short-end of the futures price curve should behave quite similarly to the front-month futures price, with the result that exchange-traded futures contracts provide a good picture of the crude oil market. Nevertheless, in order to reduce the impact of extrapolation errors, each futures price is carried forward after its expiry date by the underlying market model until its last delivery date.



**Figure 3.1: day-ahead prices, futures prices, and futures price curve**

The left graph shows the natural gas day-ahead price (point) and the futures price curve (solid) on February 23, 2003 and February 24, 2003. The right graph shows the natural gas day-ahead price (point), the natural gas synthetic futures price (red dotted line), and the interpolated natural gas futures price curve (solid line) on August 7, 2009. The synthetic futures price is defined by the no-arbitrage condition between the day-ahead price and the futures price curve, the usual spline conditions, and the maximum smoothness criterion. The natural gas prices refer to physical deliveries at Henry Hub in Louisiana.

## Natural Gas

In the U.S., exchange-traded natural gas futures contracts are traded until three business days before their first delivery dates and refer to delivery periods for each calendar month in the current plus the next twelve years at Henry Hub in Louisiana. However, in contrast to the crude oil market, natural gas demand strongly depends on unpredictable temporary influencing factors. For instance, temporary demand shocks often have a strong impact on market prices of short-dated delivery contracts but little impact on exchange-traded futures contracts referring to physical delivery periods in or after the next calendar month. This means that futures market information provides an incomplete picture of the very short-end of the futures price curve.<sup>10</sup> To see this, consider, as an example, day-ahead and futures prices on February 23, 2003 and February 24, 2003 (see Figure 3.1 (left)). The picture shows that the day-ahead price jumps about 5.50 USD per mmBtu<sup>11</sup>, while the futures price curve only moves slightly. In this case, the different price behavior of

<sup>8</sup>The CME Group is the largest and most liquid futures exchange in the U.S. energy market.

<sup>9</sup>In detail, three business days before the 25th of each month, no information about the current and the next delivery month exists.

<sup>10</sup>For instance, extreme temperatures or delivery problems can cause demand or supply shocks in natural gas or power markets.

<sup>11</sup>The abbreviation mmBtu stands for million British thermal units.

day-ahead and futures contracts can be explained by a strong increase in heating demand due to extreme temperatures on that particular day and an expected warming trend for the following days (source U.S. Energy Information Administration (EIA)).<sup>12</sup> To cope with this problem, we construct a synthetic futures price for the current calendar month based on the observable day-ahead price on each trading day. In detail, we determine the synthetic futures price for the current calendar month so that the interpolation function satisfies the usual spline, no-arbitrage, and maximum smoothness conditions and that the extrapolated day-ahead price is equal to the observable market price of the day-ahead contract. In Figure 3.1 (right), the day-ahead price, the synthetic futures price, the market prices of traded futures contracts, and the smooth interpolation function are shown for one trading day.<sup>13</sup>

### 3.3.2 Specification and Estimation of the Market Models

In the second implementation step of our energy market model, we choose the market model component for the underlying futures price dynamics. The market model should capture the common price behavior of multiple futures contracts as well as the specific distributional properties of return data (e.g., volatility and jump risk). The relevance of specific futures price movements not only depends on their statistical explanatory power but also on the underlying application of the energy market model. Therefore, we can only give several basic remarks how to specify a suitable market model and illustrate the general proceeding for the crude oil and natural gas market, where a concrete specification for a practical application is discussed in the next chapter. To specify the modeling approach, we must decide on the number of underlying risk factors and their stochastic behavior. Here, a principal component analysis can give some indication of the number of risk factors needed to capture the common stochastic behavior of multiple futures prices. Furthermore, statistical tests and market information contained in derivative instruments (e.g., implied volatility smiles) provide useful information for finding suitable return distributions.

Basically, the storage costs associated with holding the physical energy commodity determine, to a great extent, the complexity of energy price dynamics. It can be said that the higher the storage costs, the higher the number of risk factors needed to adequately model the common behavior of the futures price curve. For instance, spreads between

---

<sup>12</sup>See [http://www.eia.gov/naturalgas/weekly/archive/2003/02\\_27/ngupdate.asp](http://www.eia.gov/naturalgas/weekly/archive/2003/02_27/ngupdate.asp) for details.

<sup>13</sup>In most energy markets, multiple contracts with delivery periods in the current month can be used to replace standardized futures contracts after their expiry dates. For instance, day-ahead, two-day-ahead, and balance-of-the-month contracts are traded in European natural gas markets. However, we only have access to day-ahead prices for the U.S. natural gas market.

futures prices referring to different delivery periods are much more volatile in natural gas than in crude oil markets.

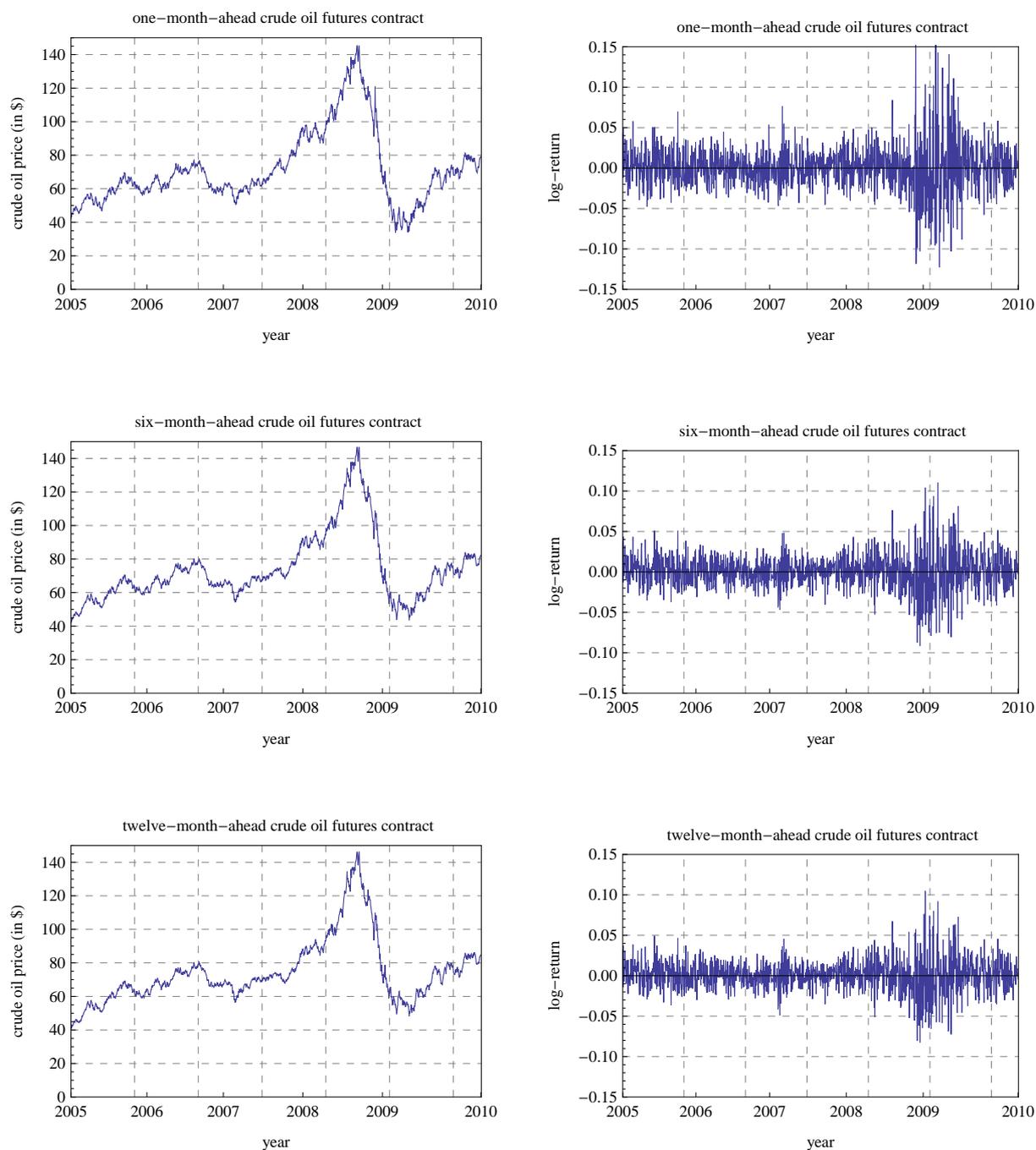
In the following, we explicitly specify and estimate a market model framework for the crude oil and natural gas markets based on historical return data. In both cases, we consider futures contracts with monthly delivery periods within the next year and ignore very long-dated futures contracts due to their low trading volumes and their minor role for most practical applications.

## Crude Oil

The underlying modeling approach should capture the common stochastic behavior of multiple futures prices as well as the distributional properties of each individual time series. In the previous chapter, we extensively analyzed the distributional properties of a single futures contract but not the common stochastic behavior of multiple futures prices, which we will now turn to. In Figure 3.2, it can be seen that one-month-ahead, six-month-ahead, and twelve-month-ahead crude oil futures prices behave very similarly. Thus, a high explanatory power for the common futures price dynamics should be achievable with few state variables. This is statistically confirmed by a principal component analysis, which reveals that 97% of the daily variation in log futures price returns are parallel shifts and that 2% of the daily variation in log futures price returns are twists<sup>14</sup>. The factor loadings further show that twists have a much larger impact on short-dated than on long-dated futures prices (see Figure 3.3). This means that twists mainly arise due to temporary price effects that are particularly important for the valuation of real options (e.g., storage facilities). For this reason, it is often useful to incorporate a twist component in a modeling approach even though twists only explain 2% of the daily variation in log futures price returns. Next, we briefly consider the distributional properties of historical log-return data. In Figure 3.2, we can see that the different time series exhibit moderate price movements with only a very few large returns until the beginning of the Financial Crisis. During the Financial Crisis, crude oil prices collapsed and their volatility levels strongly increased compared to pre-crisis levels (see Section 2.4). For instance, log-returns indicate a 100% higher volatility level between September 2008 and September 2009 compared to the pre-crisis level (January 2005 to September 2008).

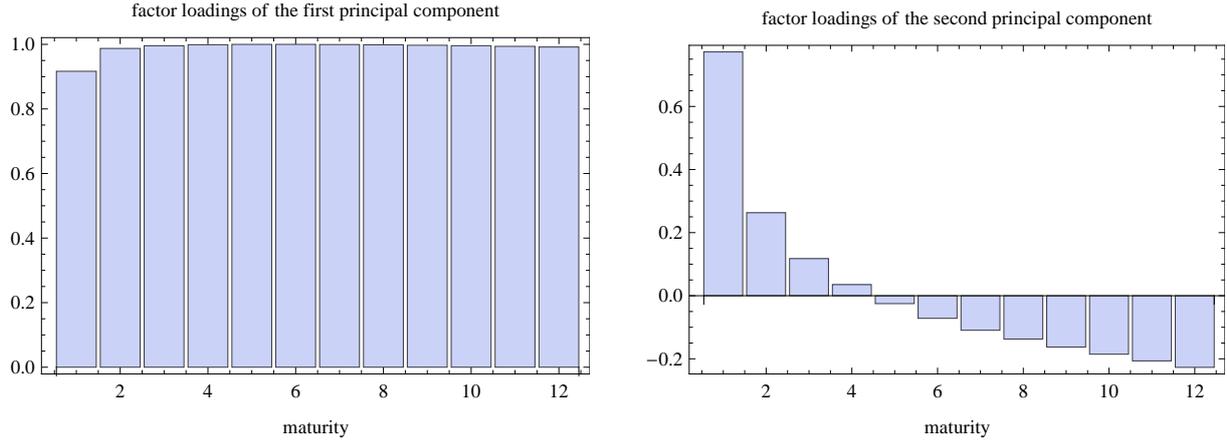
---

<sup>14</sup>This means that the short-end and the long-end of the futures price curve move in opposite directions.



**Figure 3.2: time series of crude oil futures prices and log-returns**

These graphs show one-month-ahead, six-month-ahead, and twelve-month-ahead WTI crude oil futures prices and their log-returns from January 1, 2005 to December 31, 2009.



**Figure 3.3: factor loadings**

These graphs show the factor loading of the first two principal components for daily log-return of WTI crude oil futures prices from January 1, 2005 to December 31, 2009.

Now, the above-mentioned time series properties are captured through a two-factor model with a crisis and non-crisis volatility regime given by

$$df_t^{(i)} = \sum_{j=1}^2 \sigma_t^{(i,j)}(x_t) f_t^{(i)} dw_{f,t}^{(j,\mathbb{Q})}, \quad i = 0, \dots, m, \quad (3.14)$$

where  $w_{f,t}^{(1,\mathbb{Q})}$  and  $w_{f,t}^{(2,\mathbb{Q})}$  are uncorrelated Wiener processes. The regime process  $x_t$  follows a two-state time-homogenous Markov chain with transition rates  $p_{1,2}$  and  $p_{2,1}$ , respectively. In both regimes, we use a constant and an exponential decaying volatility function

$$\sigma_t^{(i,1)}(x_t) = \begin{cases} \sigma_1^{(lg)}, & \text{if } x_t = 1 \\ \sigma_2^{(lg)}, & \text{if } x_t = 2 \end{cases}, \quad \sigma_1^{(lg)}, \sigma_2^{(lg)} > 0, \quad (3.15)$$

$$\sigma_t^{(i,2)}(x_t) = \begin{cases} \sigma_1^{(sh)} \frac{e^{-\kappa_1(\tau_i-t)} - e^{-\kappa_1(\tau_{i+1}-t)}}{\kappa_1(\tau_{i+1}-\tau_i)}, & \text{if } x_t = 1 \\ \sigma_2^{(sh)} \frac{e^{-\kappa_2(\tau_i-t)} - e^{-\kappa_2(\tau_{i+1}-t)}}{\kappa_2(\tau_{i+1}-\tau_i)}, & \text{if } x_t = 2 \end{cases}, \quad \sigma_1^{(sh)}, \sigma_2^{(sh)}, \kappa_1, \kappa_2 > 0, \quad (3.16)$$

where  $\sigma_{rg}^{(lg)}$  (volatility parameter of permanent price impacts) and  $\sigma_{rg}^{(sh)}$  (volatility parameter of temporary price impacts) can differ in both regimes  $rg \in \{1, 2\}$ . Here, we only consider two possible volatility states, instead of a continuous stochastic volatility process, as in the previous chapter. This reduces the dimension of the underlying stochastic process, which is necessary to obtain stable numerical results for many practical applications.

The two volatility components capture parallel shifts and twists in the futures price curve. The first risk factor has the same impact on all log futures prices  $\sigma_t^{(i,1)}(x_t) = \sigma_{rg}^{(lg)}$  in both regimes  $rg \in \{1, 2\}$ . The second risk factor has a stronger impact on short-dated futures prices than on long-dated futures prices.<sup>15</sup> The volatility parameter  $\sigma_{rg}^{(sh)}$  reflects the volatility difference between short- and long-dated futures contracts, while  $\kappa_{rg}$  determines the slope of the volatility function in both regimes  $rg \in \{1, 2\}$ . These two risk factors make it possible to capture shifts and twists in the futures price curve.

In the next step, we briefly describe our estimation approach and our estimation results for the crude oil market model. Note that model parameters can be estimated based on standard methods due tractable price dynamics for observable futures prices. In contrast, standard spot price models result in highly complex endogenous futures prices dynamics that can hardly be fitted to market data. It is generally useful to incorporate all available market information in an estimation algorithm in order to obtain robust estimation results (see Chapter 2). However, for illustration purposes, we only consider daily log-return data of futures contracts referring to the following twelve delivery months from January 1, 2005 to December 31, 2009 in our estimation approach. This requires one to specify the price dynamics under the physical measure. There are two market prices of diffusion risk in each regime  $\eta_f^{(j)}(x_t)\sigma_t^{(i,j)}(x_t)$  for  $j \in \{1, 2\}$  and  $x_t \in \{1, 2\}$  and two market prices of volatility jump risk. The market prices of diffusion risk are explicitly estimated, while both volatility jump risk premia are set equal to zero.<sup>16</sup> It follows then that the physical futures price dynamics is given by

$$df_t^{(i)} = \sum_{j=1}^2 \eta_f^{(j)}(x_t) (\sigma_t^{(i,j)}(x_t))^2 f_t^{(i)} dt + \sum_{j=1}^2 \sigma_t^{(i,j)}(x_t) f_t^{(i)} dw_{f,t}^{(j,\mathbb{P})}. \quad (3.17)$$

The model parameters and state variables are estimated based on the MCMC algorithm, where returns at rolling dates of the front-month futures contract are omitted.<sup>17</sup> In Table 3.1 and Figure 3.4, we give the posterior means, standard deviations, and 99% credibility intervals for all model parameters and plot the filtered latent regime process.

---

<sup>15</sup>The second risk factor captures that market prices of long-dated futures contracts are less volatile than market prices of short-dated futures contracts (Samuelson effect) in the crude oil market. The specific parametric form is used to account for delivery periods of the underlying futures contracts (see Benth and Koekebakker (2008)).

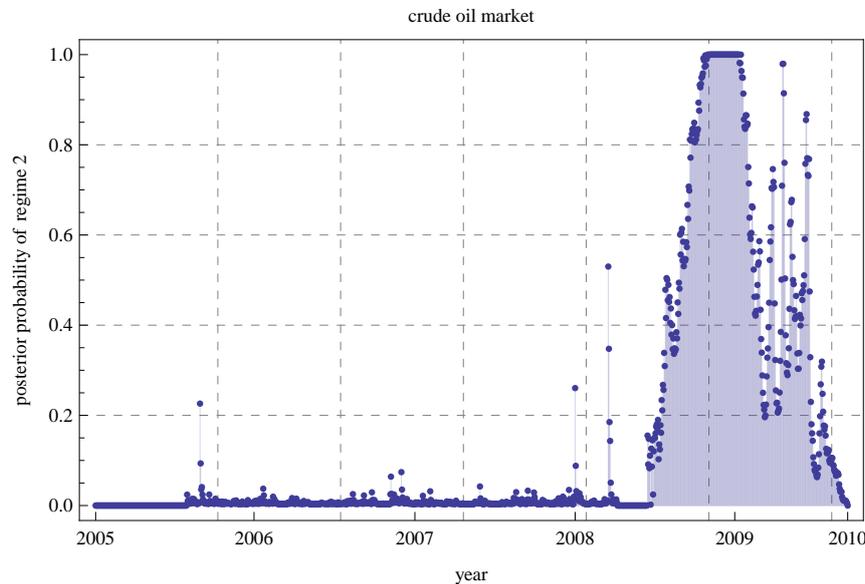
<sup>16</sup>The reason is that volatility jump risk premia cannot be estimated based on return data only. If volatility jump risk premia had been estimated on option price data, we would expect that the probability to switch from the non-crisis regime to the crisis regime would be larger (and/or the probability to switch from the crisis regime to the non-crisis regime would be smaller) under the risk-neutral measure than under the physical measure. The reason is that a negative variance risk premium is found in the crude oil market (see Trolle and Schwartz (2010)).

<sup>17</sup>In Appendix A.5, the MCMC algorithm is specified in detail.

	$\eta_f^{(1)}$	$\eta_f^{(2)}$	$p_{1,2}$	$p_{2,1}$
regime 1	2.9014	0.0416	1.3920	-
	(1.8442)	(1.6683)	(0.7261)	-
	[-1.4927, 7.1443]	[-3.8149, 3.94456]	[0.3123, 2.7654]	-
regime 2	-0.2828	4.79086	-	3.0154
	(3.2217)	(12.0171)	-	(1.4492)
	[-7.1693, 8.7193]	[-13.9418, 25.9551]	-	[0.6012, 7.8809]
	$\kappa$	$\sigma^{(sh)}$	$\sigma^{(lg)}$	-
regime 1	1.764	0.1574	0.2642	-
	(0.504)	(0.0042)	(0.0064)	-
	[0.3783, 5.292]	[0.1306, 0.1613]	[0.2418, 0.2871]	-
regime 2	2.2681	0.5597	0.5523	-
	(0.756)	(0.0437)	(0.03794)	-
	[0.252, 8.064]	[0.4681, 0.6481]	[0.4738, 0.6427]	-

**Table 3.1: model parameter estimates for the crude oil market model**

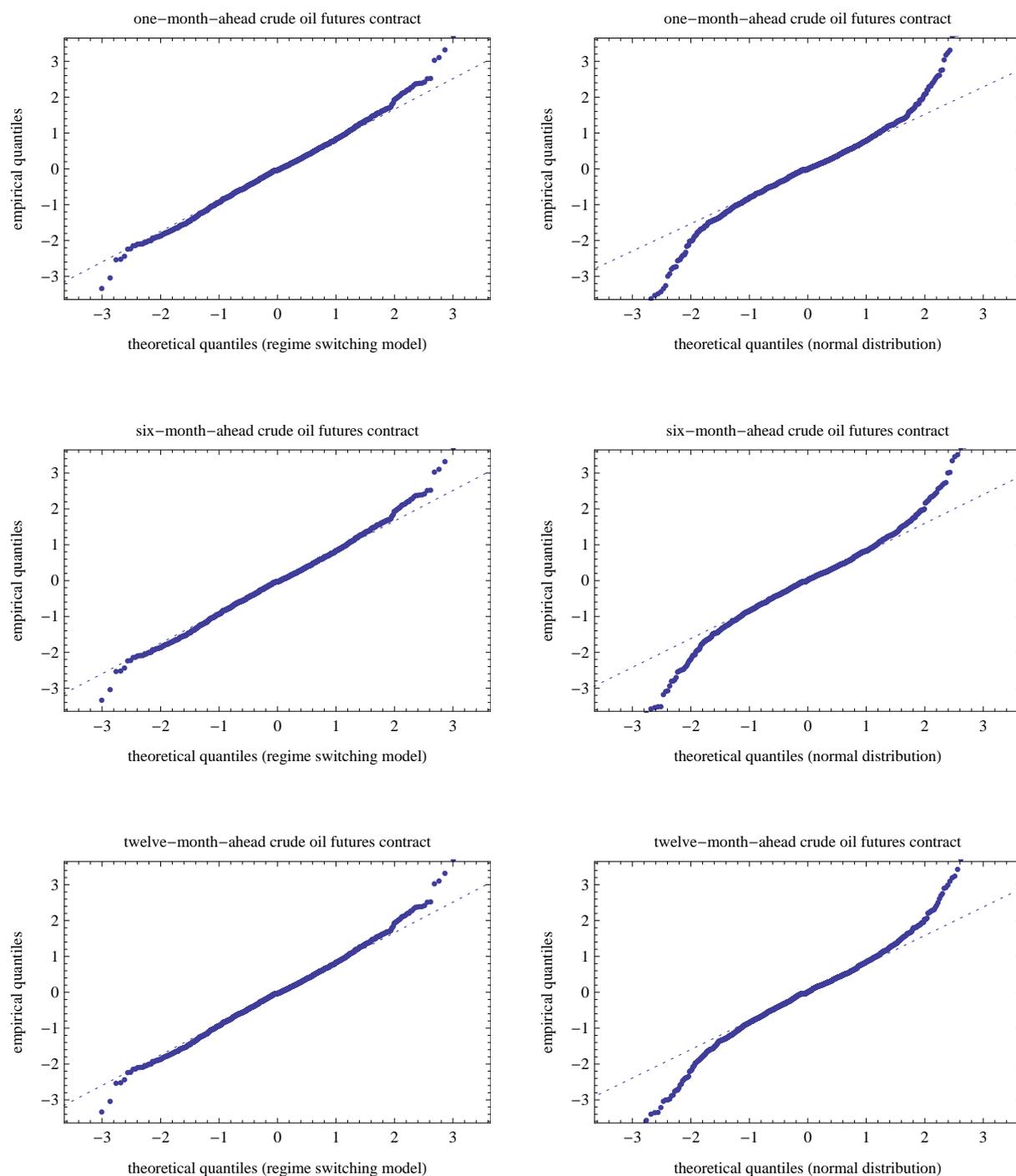
This table reports the means, standard deviations (in parenthesis), and 99% credibility intervals (in square brackets) of the posterior distributions for the crude oil market model (3.17). The model parameters are estimated based on one-month-ahead to twelve-month-ahead WTI crude oil futures contracts from January 1, 2005 to December 31, 2009. The model parameters correspond to annual decimals.



**Figure 3.4: filtered regime process**

This figure shows the estimated posterior probabilities that the crude oil futures price process (3.17) is in the second regime from January 1, 2005 to December 31, 2009.

Indeed, the second regime only captures the Financial Crisis with an approximately 2 times higher short-term and long-term volatility level compared to their pre-crisis levels. The diffusion risk premia are relatively small and statistically insignificant. This is in line with our estimation results from the previous chapter, which show that front-month crude oil futures prices exhibit no significant excess return for different time periods. Moreover, the mean reversion rate of the volatility function  $\kappa_{rg}$  is slightly larger in the financial crisis compared to the pre-crisis period. The reason is that, during the financial crisis, the volatility of one-month-ahead futures prices increased more strongly compared to six-month-ahead futures prices. We use QQ-plots to test the distributional properties of the two-factor regime switching model. The QQ-plots show that crude oil log-returns exhibit a clearly non-normal behavior that can be well captured by the underlying regime switching model (see Figure 3.5).



**Figure 3.5: quantile-quantile-plots**

These graphs show the QQ-plots for the one-month-ahead, six-month-ahead, and twelve-month-ahead WTI crude oil futures contracts. In the left panel, we use the mean values of the posterior distributions for all state variables and model parameters in order to determine the respective model distribution. In the right panel, we simply use a normal distribution that is fitted to historical log-return data.

## Natural Gas

In the natural gas market, we again start our empirical analysis with the common price dynamics of the underlying futures contracts and then consider the distributional properties of each individual time series in greater detail. In Figure 3.6, it can be seen that futures prices follow a common trend, but temporary price impacts play a much larger role than for the crude oil market. The principal component analysis confirms our first impression and shows that parallel shifts are the most important risk factor (78%)<sup>18</sup> and twists are the second most important risk factor (6%) of the natural gas futures price curve (see Figure 3.7). Furthermore, we consider the time series properties of historical natural gas prices. It turns out that natural gas prices vary strongly over time and exhibit some extreme price jumps (see Figure 3.6).

In the next step, we incorporate parallel shifts, twists, and price jumps in our market modeling approach. We use a two-factor model with a single jump component for the risk-neutral futures price dynamics

$$df_t^{(i)} = -\lambda_z \bar{\mu}_z^{(i)} f_{t-}^{(i)} dt + \sum_{j=1}^2 \sigma_t^{(i,j)} f_{t-}^{(i)} dw_{f,t}^{(j,\mathbb{Q})} + (e^{\gamma_t^{(i)} z_t} - 1) f_{t-}^{(i)} dn_{f,t}, \quad i = 0, \dots, m, \quad (3.18)$$

where  $w_{f,t}^{(1,\mathbb{Q})}$  and  $w_{f,t}^{(2,\mathbb{Q})}$  are again uncorrelated Wiener processes, and  $n_{f,t}$  is a Poisson process with constant intensity  $\lambda_z$  and random (percentage) jump sizes  $z_t$  that are normally distributed with mean  $\mu_z$  and standard deviation  $\sigma_z$ . The jump compensator  $\bar{\mu}_z^{(i)} = (e^{\gamma_t^{(i)} \mu_z + 0.5(\gamma_t^{(i)})^2 \sigma_z^2} - 1)$  ensures that each futures price process satisfies the martingale property.<sup>19</sup> The volatility functions are parameterized as follows:

$$\sigma_t^{(i,1)} = \sigma^{(lg)} \quad \text{and} \quad \sigma_t^{(i,2)} = \sigma^{(sh)} \frac{e^{-\kappa^{(sh)}(\tau_i - t)} - e^{-\kappa^{(sh)}(\tau_{i+1} - t)}}{\kappa^{(sh)}(\tau_{i+1} - \tau_i)}, \quad \sigma^{(lg)}, \sigma^{(sh)}, \kappa^{(sh)} > 0,$$

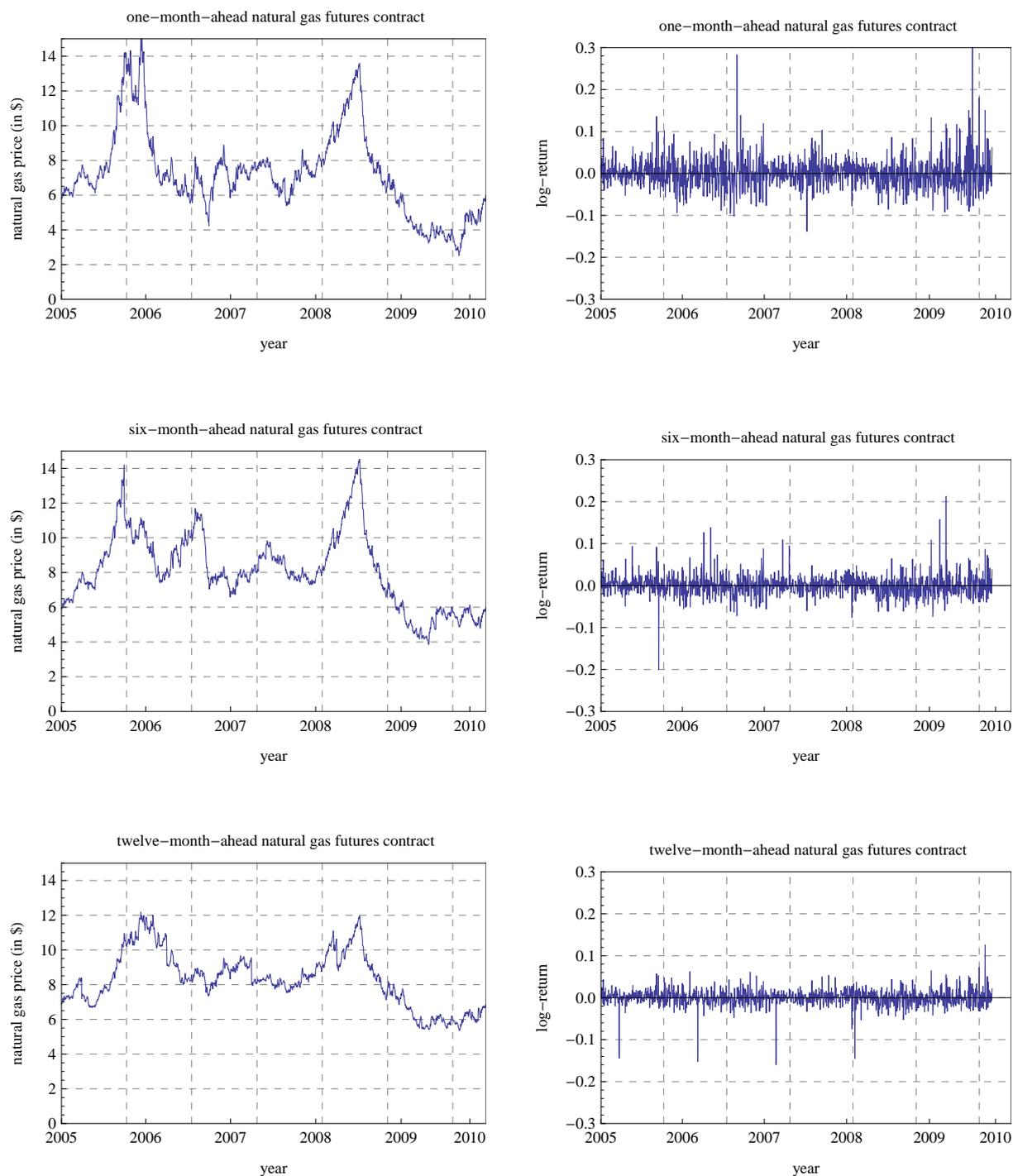
where  $\sigma_1^{(i,1)}$  and  $\sigma_2^{(i,2)}$  again capture shifts and twists in the futures price curve, respectively. The impact of price jumps on different futures contracts is modeled through the weighting function  $\gamma_t^{(i)}$  that is parameterized as follows:

$$\gamma_t^{(i)} = \frac{e^{-\kappa^{(jp)}(\tau_i - t)} - e^{-\kappa^{(jp)}(\tau_{i+1} - t)}}{\kappa^{(jp)}(\tau_{i+1} - \tau_i)}.$$

---

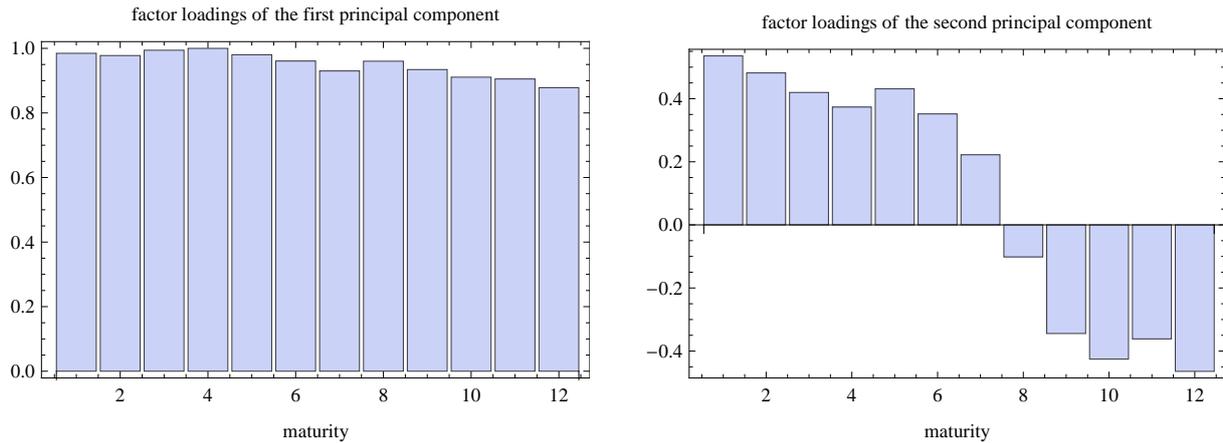
<sup>18</sup>This means that 78% of the daily variation in log futures price returns are parallel shifts.

<sup>19</sup>We assume that the volatility process is constant over time. The reason is that we do not find strongly clustered large returns in historical natural gas return data. The parametric form of the jump compensator follows from normally distributed (percentage) jump sizes.



**Figure 3.6: time series of natural gas futures prices and log-returns**

These graphs show one-month-ahead, six-month-ahead, and twelve-month-ahead natural gas futures prices and their log-returns from January 1, 2005 to December 31, 2009. The natural gas prices refer to a physical delivery at Henry Hub in Louisiana.



**Figure 3.7: factor loadings**

These graphs show the factor loading of the first two principal components for daily log-return of natural gas futures prices from January 1, 2005 to December 31, 2009.

The parametric form of  $\gamma_t^{(i)}$  makes it possible to capture persistent and temporary price jumps: a large  $\kappa^{(jp)}$  causes price jumps to only have an impact on short-dated futures contracts, while a small  $\kappa^{(jp)}$  means that price jumps in short-dated futures contracts are also reflected in long-dated futures contracts.

In the next step, we briefly present our estimation approach for the proposed two-factor jump diffusion model. For illustration purposes, we only take day-ahead prices and futures prices referring to physical deliveries in the following twelve calendar months into account and ignore option market data.<sup>20</sup> The day-ahead prices provide useful market information about temporary price impacts and are incorporated through synthetic futures prices in the market model component.<sup>21</sup> In order to estimate the natural gas market model based on historical return data, it necessary to specify the market prices of diffusion and jump risk. For simplicity, we only estimate the two market prices of diffusion risk  $(\eta_f^{(1)}\sigma_t^{(i,1)}, \eta_f^{(2)}\sigma_t^{(i,2)})$  and set all market prices of jump risk equal to zero.<sup>22</sup> It follows then that the futures price dynamics under the physical measure is given by

$$df_t^{(i)} = \left( -\lambda_z \bar{\mu}_z^{(i)} + \sum_{j=1}^2 \eta_f^{(j)} (\sigma_t^{(i,j)})^2 \right) f_{t-}^{(i)} dt + \sum_{j=1}^2 \sigma_t^{(i,j)} f_{t-}^{(i)} dw_{f,t}^{(j,\mathbb{P})} + (e^{\gamma_t^{(i)} z_t} - 1) f_{t-}^{(i)} dn_{f,t}.$$

The MCMC estimation results for all model parameters and state variables can be found in Table 3.2.

<sup>20</sup>In the previous chapter, we have presented an estimation approach that can be easily applied to incorporate option price in an estimation approach for our jump diffusion market model.

<sup>21</sup>The day-ahead prices are obtained from [http://www.eia.gov/dnav/ng/ng\\_pri\\_fut\\_s1\\_d.htm](http://www.eia.gov/dnav/ng/ng_pri_fut_s1_d.htm).

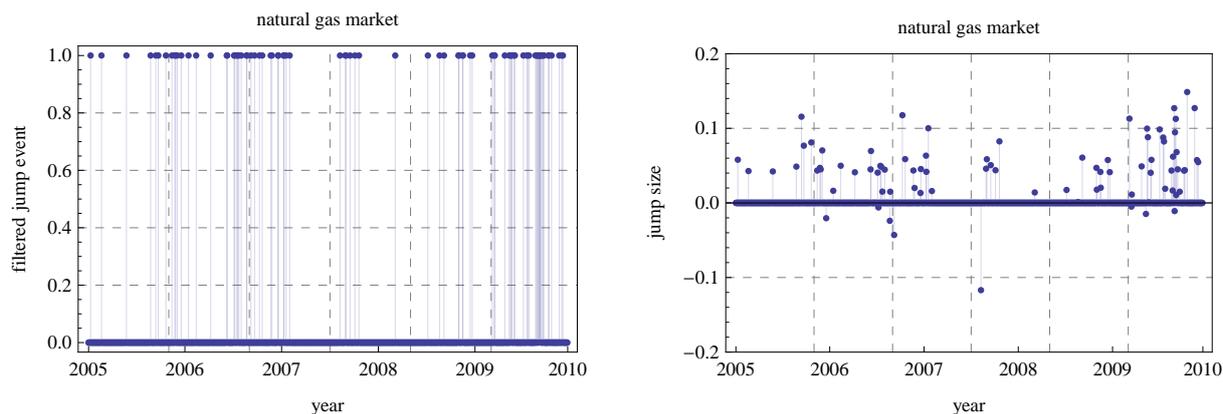
<sup>22</sup>This implies that  $\lambda_z \bar{\mu}_z^{(i)} = \lambda_z^{\mathbb{P}} \bar{\mu}_z^{(i,\mathbb{P})} = \lambda_z^{\mathbb{Q}} \bar{\mu}_z^{(i,\mathbb{Q})}$  for  $i = 0, \dots, m$ .

	$\kappa^{(sh)}$	$\sigma^{(sh)}$	$\sigma^{(lg)}$
diffusion process	1.6651 (0.4793) [0.2767, 4.7934]	0.3821 (0.0092) [0.3654, 0.4234]	0.3196 (0.006) [0.304, 0.334]
risk premia & jump intensity	$\eta_f^{(1)}$ -0.6930 (1.1282) [-3.1104, 1.8169]	$\eta_f^{(2)}$ -1.0133 (0.7818) [-2.7595, 0.4344]	$\lambda_z$ 18.2677 (2.5725) [9.4726, 24.8213]
jump process	$\mu_z$ 0.0363 (0.0148) [0.0077, 0.0768]	$\sigma_z$ 0.1157 (0.0473) [0.0482, 0.2889]	$\kappa^{(jp)}$ 11.9345 (6.1457) [0.5276, 29.4573]

**Table 3.2: model parameter estimates for the natural gas market model**

This table reports means, standard deviations (in parenthesis), and 99% credibility intervals (in square brackets) of the posterior distributions for the natural gas market model. The model parameters are estimated based on log-returns of natural gas futures contracts (synthetic and one-month-ahead to twelve-month-ahead futures contracts) from January 1, 2005 to December 31, 2009. The model parameters correspond to annual decimals.

We find no significant diffusion risk premia, a moderate daily jump intensity of 7.25 percent ( $\lambda_z/252$ ), and a slightly positive mean jump size of 3.63 percent. In Figure 3.8, we plot filtered jump times and jump sizes. It can be seen that jump events are only slightly clustered in the Financial Crisis. In addition, we estimate a large mean reversion parameter for the jump component. Thus, price jumps only have a significant impact on short-dated futures contracts and hardly any impact on long-dated futures contracts.

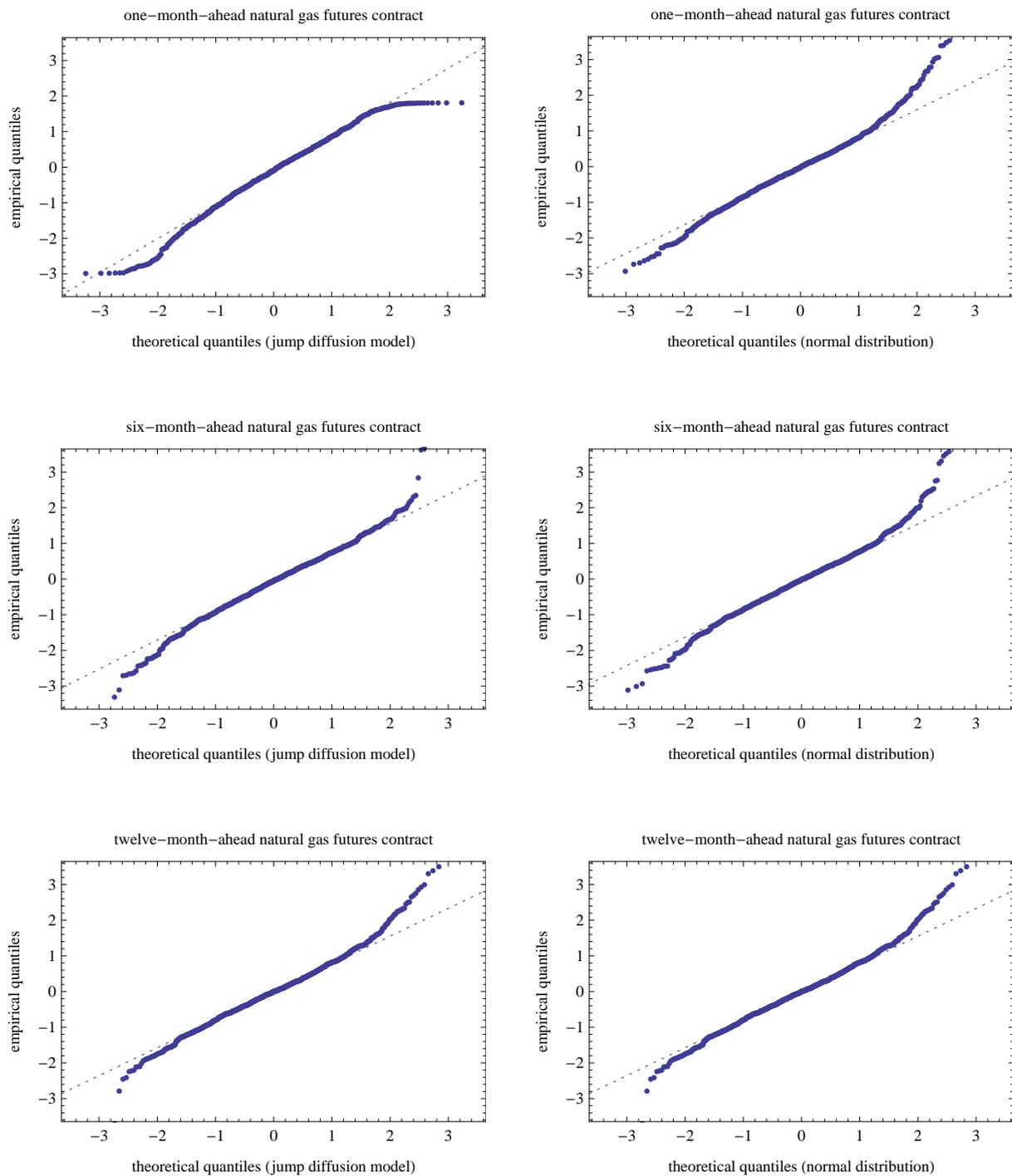


**Figure 3.8: filtered jump times and jump sizes**

These graphs show filtered jump events (left) and filtered mean jump sizes (right) in the natural gas market model (3.18) from January 1, 2005 to December 31, 2009.

As a result, log-returns of long-dated futures contracts are roughly normally distributed in the estimated market model, whereas a clearly non-normal behavior can be observed in the QQ-plots (see Figure 3.9). The QQ-plots show that even if the two-factor jump diffusion model is not really able to capture the distributional properties of the underlying log-return data, the jump component improves the distributional properties significantly compared to a multi-factor Black model.

The reason for the relatively poor distributional properties is that futures prices referring to different maturities jump at different points in times. This can be seen by considering the jump component for different maturities in greater detail. In contrast to our estimation approach, we now identify jump events through a simple two standard deviation filtering approach in two ways. First, a jump event is identified if at least one futures price return is two times larger than its empirical standard deviation. Second, a jump event is filtered out if all futures price returns are two times larger than their empirical standard deviations. In the first filtering approach, we obtain four times more jump events than in the second one. Thus, a high-dimensional return distribution with multiple jump components would be needed to capture the common stochastic behavior of the entire futures price curve adequately. However, a more parsimonious approach, with few stochastic factors, turns out to be more convenient for many pricing applications.



**Figure 3.9: quantile-quantile-plots**

These graphs show the QQ-plots for one-month-ahead, six-month-ahead, and twelve-month-ahead natural gas futures contracts. In the left panel, we use the mean values of the posterior distributions for all state variables and model parameters in order to determine the respective model distribution. In the right panel, we simply use a normal distribution that is fitted to historical log-return data.

### 3.3.3 Completing the Model through the Futures Price Curve

In the last step, we complete our modeling framework through an appropriate interpolation function. The interpolation function is uniquely defined by the usual spline, no-arbitrage, and maximum smoothness conditions and allows us to derive arbitrage-free futures price dynamics based on the respective market model. Importantly, the price dynamics of theoretical and non-standardized futures contracts cannot be simply obtained by taking an appropriate limit or inserting respective delivery periods in the market model. Instead, we have shown that the future price dynamics results from applying Itô's lemma to the interpolation function (3.9) (see Subsection 3.2.2). This leads to spot and futures price dynamics for arbitrary delivery dates or periods given by (3.11), (3.12), and (3.13) subject to the underlying market model specification.

In what follows, we explicitly describe the completion step for the crude oil and natural gas market models.

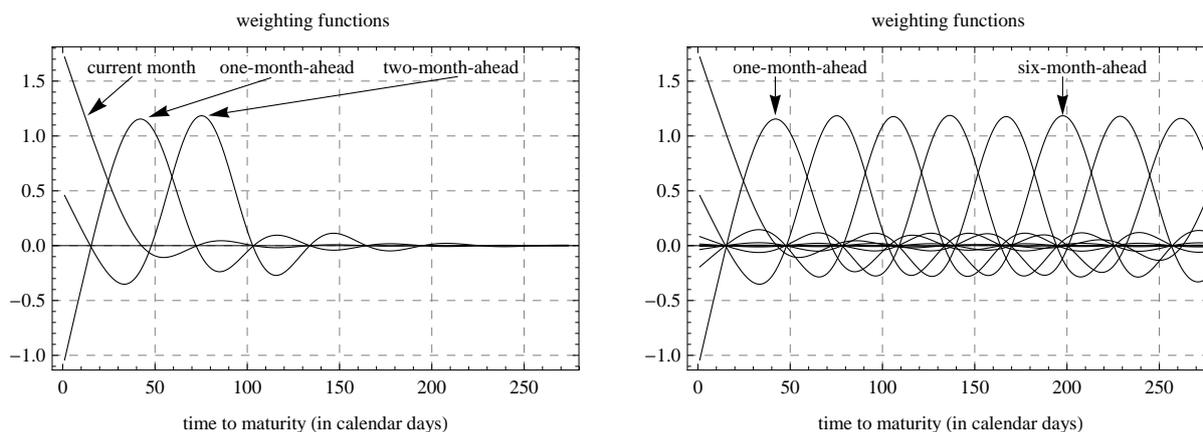
#### Crude Oil

The crude oil market model leads to the following risk-neutral futures price dynamics

$$df_t(u) = \sum_{j=1}^2 \left( \sum_{i=0}^m \beta_i(u) \sigma_t^{(i,j)}(x_t) f_t^{(i)} \right) dw_{f,t}^{(j,\mathbb{Q})}, \quad (3.19)$$

where  $x_t$  follows a two-stage regime switching process and the volatility functions  $\sigma_t^{(i,j)}(\cdot)$  are given in (3.15) and (3.16) for  $j = \{1, 2\}$  and  $i = 0, \dots, m$ . Next, we consider the volatility component in both states in greater detail. The linear relation between theoretical and real futures prices results in a volatility component that linearly depends on real futures prices but is not proportional to the current theoretical futures price. Thus, theoretical futures prices are not log-normally distributed in both regimes. Nevertheless, the volatility component is not decoupled from theoretical futures prices, since theoretical futures prices and volatility levels only differ in their weighting scheme of real futures prices.

In Figure 3.10, we show the weighting function for the following eight futures contracts. The weighting function  $\beta_i(u)$  describes the impact of a change in the  $i$ -th real futures price on the theoretical futures price  $f_t(u)$  for  $i = 0, \dots, m$ . This means that the weighting functions provide the delta hedging positions in exchange-traded futures contracts to hedge price risk of theoretical spot and futures contracts.



**Figure 3.10: weighting functions**

The graphs show the sensitivities of the theoretical futures price curve to changes in the respective real futures price. The left graph shows the sensitivities for the current, one-month-ahead, and two-month-ahead futures contracts. The right graph plots the sensitivities for the following eight futures contracts.

## Natural Gas

In the natural gas market, the futures price dynamics is obtained by inserting the two-factor jump diffusion model in (3.12). It follows then that

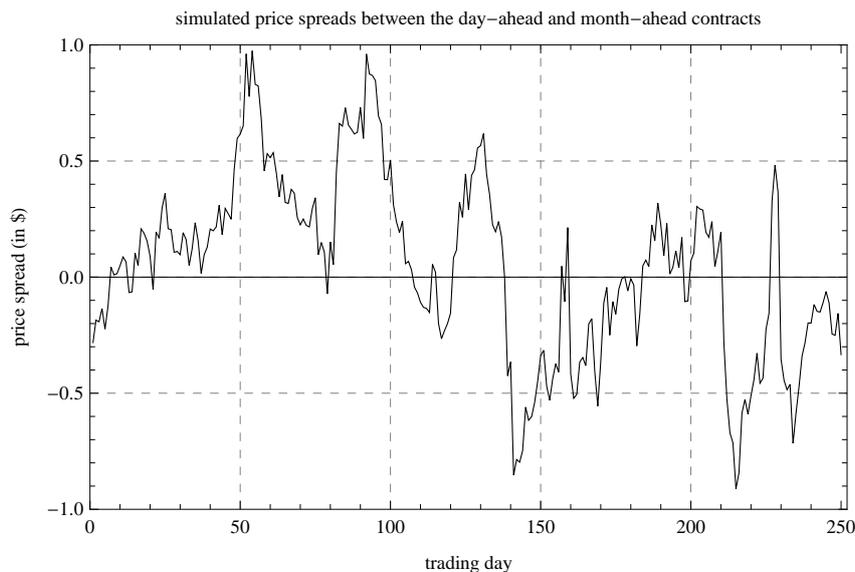
$$df_t(u) = \tilde{\mu}_t(u)dt + \sum_{j=1}^2 \tilde{\sigma}_t^{(j)}(u)dw_{f,t}^{(j,\mathbb{Q})} + \tilde{\gamma}_t(u)dn_{f,t}, \quad (3.20)$$

where

$$\begin{aligned} \tilde{\mu}_t(u) &= -\lambda_z \sum_{i=0}^m \beta_i(u) \bar{\mu}_z^{(i)} f_{t-}^{(i)}, \\ \tilde{\sigma}_t^{(j)}(u) &= \sum_{i=0}^m \beta_i(u) \sigma_t^{(i,j)} f_{t-}^{(i)}, \\ \tilde{\gamma}_t(u) &= \sum_{i=0}^m \beta_i(u) (e^{\gamma_t^{(i)} z_t} - 1) f_{t-}^{(i)}. \end{aligned}$$

Thus, the futures price process follows a jump diffusion model, where jumps in theoretical futures prices are weighted average price jumps  $\sum_{i=0}^m \beta_i(u) (e^{\gamma_t^{(i)} z_t} - 1) f_{t-}^{(i)}$  of real futures contracts. In addition, the same relation between the volatility components of theoretical futures contracts and real futures prices holds true as for the crude oil market model.

To get a feeling for the common stochastic behavior of exogenous and endogenous price dynamics, we plot the spread between the day-ahead and front-month futures price over time within our setting (see Figure 3.11). It can be seen that day-ahead prices fluctuate



**Figure 3.11: simulated price spreads**

This figure shows one representative trajectory for the spread between natural gas day-ahead and front-month futures prices based on the estimated energy market model (3.20).

around stochastic front-month futures prices. This behavior is similar to the mean reversion behavior in standard spot price models and plays a central role in the pricing and hedging of flexible physical assets in natural gas markets.

In summary, the analysis of both market examples shows that the proposed energy market modeling framework is a convenient approach to obtain a consistent spot price process that relies on observable market prices only. It allows to easily incorporate specific characteristics of the underlying energy market, makes use of analytic expressions when estimating parameters associated with the market model component, and through our completion step delivers a complete futures price curve. In contrast, alternative models typically rely on entirely latent factors, which can lead to barely interpretable factor dynamics. Furthermore, the direct link between theoretical spot and futures prices on the one hand and real futures prices on the other hand can be used to efficiently hedge price risk. In the next chapter, we show the convenience of the energy market model for a practical application.

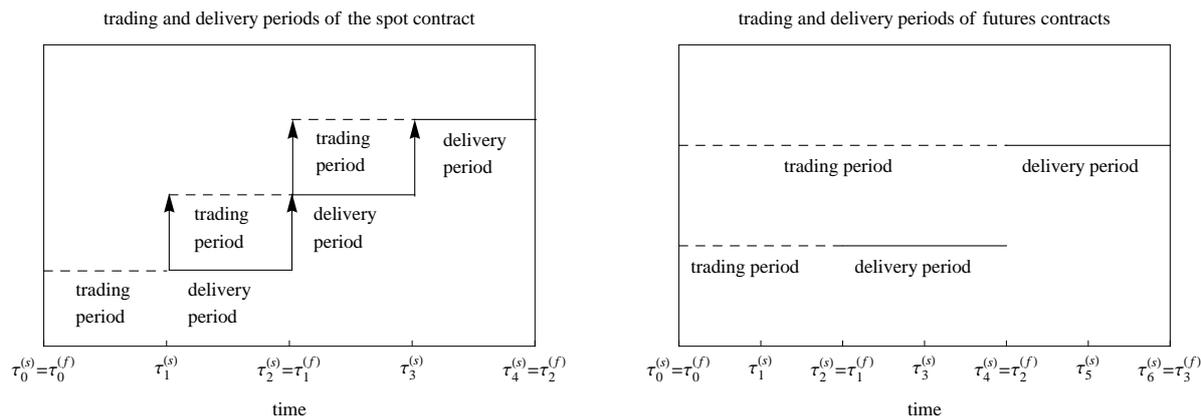
# Chapter 4

## The Valuation of Storage Contracts

In energy markets, physical storage contracts are actively traded in order to allocate storage capacities efficiently among multiple market participants. The traded storage contracts give their holders the right to inject, store, and withdraw the underlying energy commodity subject to maximal injection and withdrawal rates and limited total storage capacities.

In this chapter, we consider the storage valuation problem from a theoretical and practical perspective. We first simplify the underlying path-dependent stochastic optimization problem to make it manageable for numerical valuation algorithms. Namely, we show that (i) the dimension of the initial optimization problem can be reduced to the number of physically non-redundant delivery contracts and that (ii) optimal trading times and volumes can both be restricted to a finite number. This allows us to formulate a low-dimensional discrete stochastic optimization problem without any approximation error for many energy markets.

Then, we pick up our energy market modeling approach from the previous chapter and implement it for the valuation of a concrete natural gas storage contract. This shows how to apply our modeling approach for a concrete practical application and allows us to investigate the value of dynamic storage strategies and the relation between specific price movements and the storage value.



**Figure 4.1: trading and delivery periods of spot and futures contracts**

These graphs illustrate trading and delivery periods of spot and futures contracts. The dashed lines correspond to trading periods and the solid lines to physical delivery periods.

## 4.1 Storage Contracts

In this section, we briefly introduce the underlying energy market environment and the underlying storage contract design. The energy market consists of (i) a single spot contract with a delivery period that is rolled over from  $(\tau_i^{(s)}, \tau_{i+1}^{(s)})$  to  $(\tau_{i+1}^{(s)}, \tau_{i+2}^{(s)})$  at  $\tau_i^{(s)}$  for  $i = 0, \dots, m^{(s)}$  (see Figure 4.1 (left)) and (ii) multiple futures contracts with fixed redundant delivery periods  $\{(\tau_i^{(f)}, \tau_{i+1}^{(f)})\}_{i=0}^{m^{(f)}}$ , i.e.,  $\tau_j^{(f)} \in \{\tau_i^{(s)}\}_{i=0}^{m^{(s)}+1}$  for  $j = 0, \dots, m^{(f)}$  (see Figure 4.1 (right)). The spot contract and all futures contracts can be traded without transaction costs at a spot price of  $s_t$  and futures prices of  $f_t^{(i)} = f_t(\tau_i^{(f)}, \tau_{i+1}^{(f)})$ . In contrast to the previous chapter, we use the superscripts to distinguish between delivery periods of spot and futures contracts.

The storage contract offers its holder the right to inject, withdraw, and store the underlying commodity during the time period  $[\tau_0^{(s)}, \tau_{m^{(s)}+1}^{(s)}]$  subject to local and global technical constraints and injection and withdrawal costs. The technical constraints  $\psi$  consist of (i) constant maximal withdrawal and injection rates  $(\underline{q}, \bar{q})$ , (ii) a maximal storage capacity  $\bar{sv}$ , and (iii) a prearranged final volume in storage  $sv_f$  at  $\tau_{m^{(s)}+1}$ .<sup>1</sup> The injection and withdrawal costs are separated into constant variable operating costs  $d_{inj} \geq 0$  and  $d_{with} \geq 0$  as well as constant injection and withdrawal loss rates  $b_{inj} \geq 0$  and  $b_{with} \geq 0$ .<sup>2</sup> For simplicity, we assume that injection or withdrawal costs for a physical delivery in  $(\tau_i^{(s)}, \tau_{i+1}^{(s)})$  occur at  $\tau_i^{(s)}$  for  $i = 0, \dots, m^{(s)}$ . For instance, if a storage operator decides to deliver or

<sup>1</sup>In general, prearranged penalty payments regulate final payments if the final volume in storage is not equal to  $sv_f$ . In the following, we simply assume that penalty payments ensure that the final volume in storage is equal to  $sv_f$  at the end of the contract period without concretely specifying them.

<sup>2</sup>For instance, injection and withdrawal loss rates are between 0-3% and 0-2% for depleted reservoirs, respectively (see Wu, Wang, and Qin (2011)).

purchase  $q$  units of the underlying energy commodity over the delivery period  $(\tau_i^{(s)}, \tau_{i+1}^{(s)})$ , the following injection and withdrawal costs  $c(q, s)$  occur at  $\tau_i^{(s)}$ :

$$c(q, s) = \begin{cases} -b_{inj}sq - d_{inj}q, & q \geq 0 \text{ (injection)} \\ b_{with}sq + d_{with}q, & q < 0 \text{ (withdrawal)} \end{cases}, \quad i = 0, \dots, m^{(s)},$$

where  $s$  is equal to the spot price at  $\tau_i^{(s)}$  ( $s = s_{\tau_i^{(s)}}$ ). For example, Centrica, the owner of British Gas, offers physical storage contracts via so called “standard bundled units” (SBUs).<sup>3</sup> Each SBU is largely characterized by the following operational constraints: maximal withdrawal and injection rates of  $\underline{q} = -1$  kWh/day and  $\bar{q} = 0.35$  kWh/day and a total storage capacity of  $\bar{v} = 66.6$  kWh. The injection and withdrawal costs are given by variable operating costs of  $d_{inj} = 0.021$  pence/kWh and  $d_{with} = 0.007$  pence/kWh and injection and withdrawal loss rates equal to zero ( $b_{inj} = b_{with} = 0$ ).<sup>4</sup> The contract period starts on April 1st and ends on March 31st of the following year. The initial volume in storage is zero and the storage should be returned with the same volume at the end of the contract period. If the final volume in storage is not zero, the storage volume is auctioned and the owner gets the highest bid less the auction costs.

## 4.2 Static vs. Dynamic Storage Strategies

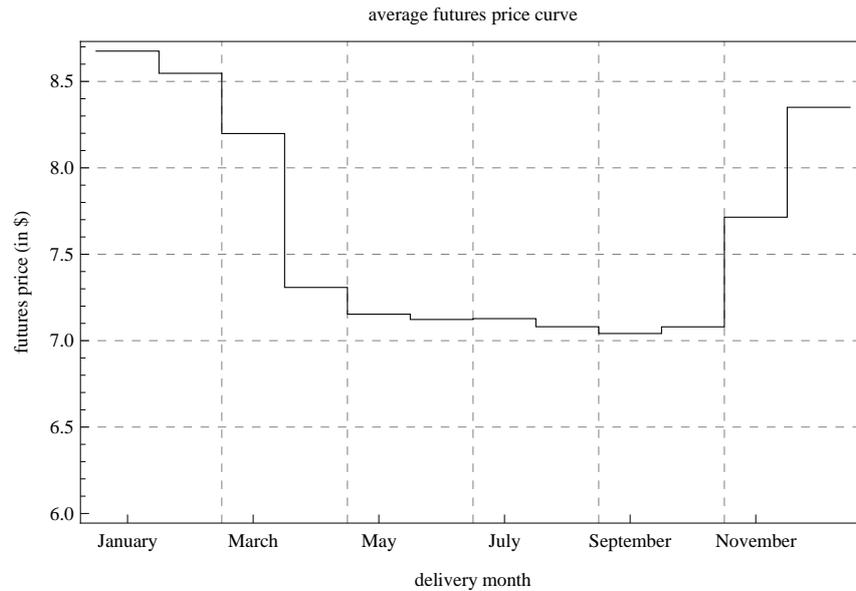
In this section, we consider the main price drivers impacting market-based storage values. For simplicity, we abstract from interest rates effects and injection and withdrawal costs due to their minor impact on optimal trading strategies of storage contracts.

In principle, a physical storage contract allows its holder to store the underlying good subject to specific technical constraints. The storage option is only valuable in markets with a limited total storage capacity (e.g., energy markets), while it is worthless in markets with an unrestricted total storage capacity (e.g., stock markets). In markets with a limited total storage capacity, physical storage options can be exploited through static and dynamic trading strategies. In a static trading strategy, the holder of the storage contract tries to profit from price spreads between futures contracts with different delivery periods at the beginning of the contract period. For instance, static trading strategies in the natural gas market try to profit from the seasonal behavior of natural gas prices. Namely, natural gas futures prices referring to summer months are lower, on average, than those referring to winter months (see Figure 4.2). The summer-winter spread in

---

<sup>3</sup>See <http://www.centrica.com> for details.

<sup>4</sup>In Europe, natural gas prices often refer to megawatt-hours (MWh), where 1 USD per mmBtu corresponds to 0.293071 USD per MWh.



**Figure 4.2: average futures price curve**

This figure shows the average futures price for each calendar month between January, 2005 and December, 2010. The average futures price curve is calculated based on one-month-ahead to twelve-month-ahead natural gas futures contracts. The natural gas futures prices refer to physical deliveries at Henry Hub in Louisiana.

futures contracts can be exploited through a static trading strategy without any price risk. The storage operator can, for instance, fill his storage facility from April to October and deplete it during the winter months November to March subject to the technical constraints. This allows him to earn the summer-winter spread that is, on average, about 1.5 USD per storage unit in the U.S. natural gas market.<sup>5</sup> This “intrinsic” storage value can easily be calculated and provides a trivial lower bound for the fair market value of a storage contract.

In a dynamic trading strategy, the holder of a storage contract adjusts his trading position in spot and futures contracts over time subject to the common stochastic behavior of spot and futures prices. Trivially, storage contracts are worthless, independent of the stochastic spot price process, if the classical cost-of-carry relation between spot and futures prices holds true, as it does in stock markets. This shows that the storage value does not simply depend on the overall price uncertainty in the market, but rather on the stochastic behavior of specific futures price movements. For that reason, it is useful to consider the impact of specific futures price movements on the storage value to get a better understanding of the value of embedded storage options. First, we assume that the futures price curve can vary, but that price spreads between two arbitrary points on the futures price curve never change their signs over time. In this case, optimal injection and

<sup>5</sup>The general optimization problem can be found in Lai, Margot, and Secomandi (2010).

withdrawal decisions do not change over time and are equal to the optimal static trading strategy at the beginning of the contract period. This implies that the value of a storage contract is equal to the intrinsic storage value and that storage contracts have a linear exposure to each point of the futures price curve. In particular, storage contracts have a linear exposure to parallel shifts, which obviously do not change the relative position of futures prices. This implies that a parallel shift in the futures price curve of  $\delta$  increases the storage value of  $\delta$  times the current volume in storage and that the optimal price ranges for injection and withdrawal decisions are shifted by  $\delta$ .

This leads to the following question: why should a dynamic trading strategy be implemented? The only plausible reason to apply a dynamic trading strategy is that futures prices are not well-ordered over time. Indeed, real futures prices change their relative position over time. The main reason for this is the existence of temporary price impacts, which have a strong positive or negative impact on the short-end of the futures price curve and hardly any impact on the long-end of the futures price curve.<sup>6</sup> This “mean reversion” behavior of energy prices is documented in various empirical studies (see Schwartz (1997), Schwartz and Smith (2000), and Cartea and Williams (2008)) and implies that a dynamic storage strategy can be valuable.

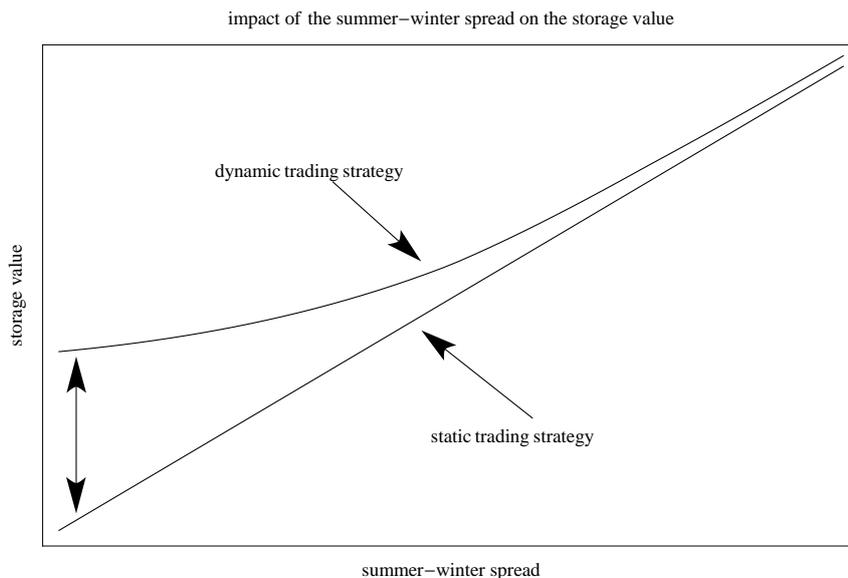
In (multi-factor) stochastic models that allow that futures prices change their relative positions over time, storage values, once again, depend linearly on parallel shifts provided that parallel shifts have no impact on other risk factors of the underlying price dynamics (e.g., affine-linear model with independent risk factors). In contrast, a convex relation between the storage value and the summer-winter spread should exist in the presence of temporary price impacts. Thus, a stochastic summer-winter spread would increase the market value of storage contracts. The reason for our assertion is as follows: if the summer-winter spread narrows, it is more attractive to trade in the spot market to profit from temporary price shocks. This partially compensates for the loss due to a lower intrinsic storage value. Otherwise, if the summer-winter spread widens, the storage operator profits from a larger intrinsic storage value, but short-term trading becomes less profitable. This suggests a convex relation between the summer-winter spread and the storage value (see Figure 4.3).

### 4.3 Stochastic Optimization Problem

In the next step, we formally introduce the storage valuation problem based on two simplistic assumptions. We assume that (i) trading is restricted to the single spot contract

---

<sup>6</sup>In contrast, the long-end of the futures price curve is generally well-ordered.



**Figure 4.3: impact of the summer-winter spread on the storage value**

This figure shows the expected convex relation between the summer-winter spread and the storage value.

and that (ii) trading times are restricted to the rolling dates  $\{\tau_i^{(s)}\}_{i=0}^m$  of the spot contract. Later, we show that both assumptions have no impact on optimal storage strategies. In the following, we skip the superscripts, since all delivery periods refer to the spot contract. In a rational market, the storage value is equal to the maximal expected risk-neutral trading profit of all admissible trading strategies

$$sw_{\tau_0}(s_{\tau_0}, sv_{\tau_0}; \psi) = \sup_{\{q_{\tau_i}\}_{i=0}^m \in ad(sv_{\tau_0}; \psi)} \mathbb{E}_{\tau_0}^{\mathbb{Q}} \left[ \sum_{i=0}^m e^{-r(\tau_i - \tau_0)} (-q_{\tau_i} s_{\tau_i} + c(q_{\tau_i}, s_{\tau_i})) \right], \quad (4.1)$$

where

$$ad(sv_{\tau_0}; \psi) = \{ \{q_{\tau_j}\}_{j=0}^m \mid q_{\tau_j} \in [\max\{-sv_{\tau_j}, \underline{q}\}, \min\{\overline{sv} - sv_{\tau_j}, \overline{q}\}], sv_{\tau_{m+1}} = sv_f \}.$$

In (4.1),  $sv_{\tau_0}$  is the initial volume in storage,  $r$  is the constant risk-free interest rate, and  $ad(sv_{\tau_0}; \psi)$  is the set of all admissible trading strategies subject to the initial volume in storage  $sv_{\tau_0}$  and the local and global constraints  $\psi = (\underline{q}, \overline{q}, \overline{sv}, sv_f)$ . The volume in storage changes over time subject to the (endogenous) optimal trading strategy  $\{q_{\tau_i}^*\}_{i=0}^m$  as follows:

$$sv_{\tau_i} = sv_{\tau_0} + \sum_{j=0}^{i-1} q_{\tau_j}^*, \quad \text{for } i = 0, \dots, m+1.$$

This means that  $sv_{\tau_i}$  is equal to the initial volume in storage plus the aggregated delivery volume until  $\tau_i$ , where  $q_{\tau_j}^* < 0$  means that the underlying energy commodity is sold at the market (withdrawal decision), and  $q_{\tau_j}^* > 0$  means that the underlying energy commodity is bought at the market (injection decision) at  $\tau_j$  for  $j = 0, \dots, i - 1$  and  $i = 0, \dots, m$ .

## 4.4 Theoretical Results

In the previous section, we have formulated the stochastic optimization problem (4.1) under two simplistic assumptions: (i) trading is restricted to the spot contract and (ii) trading times are restricted to the rolling dates  $\{\tau_i\}_{i=0}^m$  of the spot contract. Next, we argue that neither of these assumptions has an impact on the storage value, even when embedded storage options can be continuously exploited in spot and various futures contracts instead of being restricted to discrete trading decisions in the single spot contract.

First, we show that trading can be restricted to non-redundant physical delivery contracts having delivery periods that cannot be duplicated by other traded delivery contracts. The reason for this is the simple no-arbitrage argument that two trading strategies with the same physical delivery flows must have the same expected market value at any point in time. In our market environment, the spot contract is the only non-redundant physical delivery contract, whereas all futures contracts have redundant physical delivery periods. This implies that trading can be restricted to the spot contract without influencing the storage value.<sup>7</sup> It is important to keep in mind that the no-arbitrage argument can only be applied in markets without transaction costs. Otherwise, it may be optimal to trade in futures contracts in order to reduce transaction costs.

Second, it is easy to show that optimal trading times can be restricted to the rolling dates of the spot contract  $\{\tau_i\}_{i=0}^m$  without reducing the expected storage value under  $\mathbb{Q}$ . The reason is that (i) interest rates effects have no impact on early exercise strategies, since payoffs are linked to physical delivery flows instead of trading times, and that (ii) active trading within the trading period of the spot contract is purely speculative and does not exploit any physical storage options. This means, in summary, that the optimization problem (4.1) really expresses the fair storage value in our market environment.

In the next step, we analyze the impact of local and global trading constraints  $\psi$  on the storage value (see Kaminski, Feng, and Pang (2008)).

---

<sup>7</sup>This does not mean that a physical duplication strategy leads to the same payoff, but rather to the same expected payoff under the pricing measure. In particular, trading strategies in spot markets can be much riskier than trading strategies in futures markets.

**Proposition 4** (*Scale-Invariance and Concavity*)

The market value of a storage contract is

(i) *scale-invariant of degree one:*

$$sw_{\tau_i}(s_{\tau_i}, \alpha sv_{\tau_i}; \alpha\psi) = \alpha sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \psi) \quad \text{for } \alpha > 0, i = 0, \dots, m,$$

(ii) *concave in each constraint  $\psi_j$ :*

$$sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \psi) \geq \lambda sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \widehat{\psi}_j) + (1 - \lambda) sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \widetilde{\psi}_j),$$

where

$$\begin{aligned} \psi &= (\psi_1, \dots, \psi_{j-1}, \lambda \widehat{\psi}_j + (1 - \lambda) \widetilde{\psi}_j, \psi_{j+1}, \dots, \psi_4), \\ \widehat{\psi}_j &= (\psi_1, \dots, \psi_{j-1}, \widehat{\psi}_j, \psi_{j+1}, \dots, \psi_4), \\ \widetilde{\psi}_j &= (\psi_1, \dots, \psi_{j-1}, \widetilde{\psi}_j, \psi_{j+1}, \dots, \psi_4) \end{aligned}$$

for  $0 \leq \lambda \leq 1$  and  $j = 1, \dots, 4$ .

**Sketch of the proof:** In Appendix A.6, we provide the theoretical proofs of both relations. The idea behind both proofs is as follows: if an admissible trading strategy  $q = \{q_{\tau_j}\}_{j=i}^m$  for a storage contract with current volume in storage  $sv_{\tau_i}$  and local and global constraints  $\psi$  results in a payoff  $y = \{-q_{\tau_j} s_{\tau_j} + c(q_{\tau_j}, s_{\tau_j})\}_{j=i}^m$ , then  $\alpha q$  is an admissible trading strategy for a storage contract with current volume in storage  $\alpha sv_{\tau_i}$  and local and global constraints  $\alpha\psi$  and leads to a payoff  $\alpha y$ . Hence, it follows that

$$\alpha sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \psi) \leq sw_{\tau_i}(s_{\tau_i}, \alpha sv_{\tau_i}; \alpha\psi), \quad \alpha > 0.$$

It follows then that

$$\begin{aligned} sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \psi) &= sw_{\tau_i}(s_{\tau_i}, \frac{1}{\alpha} \alpha sv_{\tau_i}; \frac{1}{\alpha} \alpha\psi) \geq \frac{1}{\alpha} sw_{\tau_i}(s_{\tau_i}, \alpha sv_{\tau_i}; \alpha\psi) \\ \Rightarrow \alpha sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \psi) &\geq sw_{\tau_i}(s_{\tau_i}, \alpha sv_{\tau_i}; \alpha\psi) \\ \Rightarrow \alpha sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \psi) &= sw_{\tau_i}(s_{\tau_i}, \alpha sv_{\tau_i}; \alpha\psi), \quad i = 0, \dots, m. \end{aligned}$$

The concavity of the storage value can be shown with similar arguments. ■

In the last step, we consider optimal trading volumes at rolling dates of the spot contract. In principle, the operator of a storage contract can choose his optimal trading position in the spot contract from a continuum of admissible trading volumes  $q_{\tau_i} \in [\max\{-sv_{\tau_i}, \underline{q}\}, \min\{\bar{sv} - sv_{\tau_i}, \bar{q}\}]$  for  $i = 0, \dots, m$ . However, it is not possible to test all trading volumes for optimality in numerical valuation algorithms. Therefore, trading volumes must be restricted to a finite number before the numerical calculation. In the following, we show that trading volumes can be restricted to either the maximal injection rate, maximal withdrawal rate, or zero (strong bang-bang condition) or to multipliers of the largest common factor of the local and global constraints  $\psi$  (weak bang-bang condition) without affecting the fair storage value.

**Definition 1** (*Strong Bang-Bang Condition*)

The trading strategy  $\{q_{\tau_i}\}_{i=0}^m$  satisfies the strong bang-bang condition if each  $q_{\tau_i}$  is equal to

(i) the maximal withdrawal rate

$$q_{\tau_i} = \underline{q},$$

(ii) or the maximal injection rate

$$q_{\tau_i} = \bar{q},$$

(iii) or zero

$$q_{\tau_i} = 0$$

for  $i = 0, \dots, m$ .

It turns out that optimal trading strategies of storage contracts satisfy the strong bang-bang condition only if the spot contract refers to an immediate delivery date instead of a delivery period. In such a theoretical market environment, traders can continuously adjust their storage volume to market prices. The optimal trading strategy is then simply given by: (i) selling at the maximal withdrawal rate if the first derivative of the storage value with respect to the volume in storage is smaller than the current spot price minus the withdrawal costs, (ii) buying at the maximal injection rate if the first derivative of the storage value with respect to the volume in storage is larger than the current spot price plus the injection costs, or (iii), if neither of these conditions is met, doing nothing (see Thompson, Davison, and Rasmussen (2009)).<sup>8</sup>

---

<sup>8</sup>Of course, we require that the storage value is differentiable with respect to the volume in storage.

In real market environments, physical delivery contracts have delivery periods. Thus, physical delivery rates cannot be adjusted continuously. Here, it seems reasonable, at first glance, that optimal trading positions should be equal to aggregated delivery volumes of the optimal continuous trading strategy over the underlying delivery periods of real delivery contracts.<sup>9</sup> This means that all trading volumes between  $\underline{q}$  and  $\bar{q}$  can be optimal from an ex-ante perspective. Interestingly, our initial intuition is wrong and trading volumes can also be restricted to a finite number in real market environments.

**Definition 2** (*Weak Bang-Bang Condition*)

The trading strategy  $\{q_{\tau_i}\}_{i=0}^m$  satisfies the weak bang-bang condition when all trading volumes are multipliers of the largest common factor of the local and global constraints  $lcf(\psi)$ , i.e.,  $q_{\tau_i} = j \cdot lcf(\psi)$  for  $j \in \mathbb{Z}$  and  $i = 0, \dots, m$ .

In the next step, we show that the optimal trading strategy of a storage contract satisfies the weak bang-bang condition if the initial volume in storage is a multiplier of  $lcf(\psi)$ .<sup>10</sup>

**Proposition 5** (*Weak Bang-Bang Property*)

In our market environment, the optimal trading strategy of a storage contract satisfies the weak “bang-bang” condition if the initial volume in storage is a multiplier of  $lcf(\psi)$ .

**Proof:** We use the dynamic programming principle to prove that the optimal trading strategy satisfies the weak bang-bang condition based on backward induction. The dynamic programming principle states that optimal trading volumes are chosen as to maximize the current payoff plus the continuation value of the storage contract (objective function). We show that the continuation value function

$$cv_{\tau_i}(s_{\tau_i}, sv_{\tau_{i+1}}) = e^{-r(\tau_{i+1}-\tau_i)} \mathbb{E}_{\tau_i}^{\mathbb{Q}} [sw_{\tau_{i+1}}(s_{\tau_{i+1}}, sv_{\tau_{i+1}}; \psi)], \quad i = 0, \dots, m,$$

is piecewise linear with kink points on the weak bang-bang grid  $wbb(\psi)$  that contains all storage volumes that can be reached by a weak bang-bang strategy. Thus, all potential maxima of the objective function are also on the weak bang-bang grid and the optimal trading strategy satisfies the weak bang-bang condition.

---

<sup>9</sup>This means that we consider the optimal trading strategy in the theoretical spot contract over the delivery period of a real delivery contract. Then, the aggregated delivery volume is equal to the expected time to sell at the maximum rate times the maximal withdrawal rate minus the expected time to buy at the maximum rate times the maximal injection rate.

<sup>10</sup>Secomandi (2010) shows a similar result for a discrete spot price process.

In detail, we prove the following two theoretical properties of the optimal trading strategy and the continuation value function:

(i) **trading strategy:**

optimal trading volumes  $\{q_{\tau_i}^*\}_{i=0}^m$  are equal to

$$q_{\tau_i}^* = j \cdot lcf(\psi), \quad j = (lcf(\psi))^{-1} \max\{-sv_{\tau_i}, \underline{q}\}, \dots, (lcf(\psi))^{-1} \min\{\bar{sv} - sv_{\tau_i}, \bar{q}\}$$

for  $i = 0, \dots, m$ ,

(ii) **continuation value function:**

the continuation value function is concave and piecewise-linear in  $sv_{\tau_{i+1}}$  with kink points on the weak bang-bang grid:

$$cv_{\tau_i}(s_{\tau_i}, sv_{\tau_{i+1}} + \delta) - cv_{\tau_i}(s_{\tau_i}, sv_{\tau_{i+1}}) = \beta_{\tau_i}(sv_{\tau_{i+1}})\delta, \quad sv_{\tau_{i+1}} \in wbb(\psi), 0 \leq \delta < lcf(\psi),$$

where  $\beta(sv_{\tau_{i+1}})$  is decreasing in  $sv_{\tau_{i+1}}$  for  $i = 0, \dots, m$ .

The proof is conducted by backward induction based on the dynamic programming principle. It is easy to show that the first condition is satisfied at the last trading date  $\tau_m$  if  $sv_{\tau_m}$  is on the weak bang-bang grid  $wbb(\psi)$ . The reason is that the final volume in storage must be equal to the prearranged volume  $sv_f$ . This leads to an optimal trading volume  $q_{\tau_m}^*$  at  $\tau_m$  given by

$$q_{\tau_m}^*(sv_{\tau_m}) = sv_f - sv_{\tau_m}$$

for all admissible trading strategies  $\{q_{\tau_i}\}_{i=0}^{m-1}$ . Thus, the optimal trading strategy satisfies the weak bang-bang condition if  $sv_{\tau_m} \in wbb(\psi)$ , since  $sv_f \in wbb(\psi)$ .

Next, we calculate the continuation value at  $\tau_{m-1}$  based on the optimal trading strategy at  $\tau_m$ . The continuation value function can be derived from the optimal trading decision at  $\tau_m$ . It is given by

$$cv_{\tau_{m-1}}(s_{\tau_{m-1}}, sv_{\tau_m}) = (sv_{\tau_m} - sv_f)e^{-r(\tau_m - \tau_{m-1})}\beta_{\tau_{m-1}},$$

where

$$\beta_{\tau_{m-1}} = \mathbb{E}_{\tau_{m-1}}^{\mathbb{Q}}[s_{\tau_m}] + \mathbf{1}_{\{sv_{\tau_m} \geq sv_f\}} \mathbb{E}_{\tau_{m-1}}^{\mathbb{Q}}[c(-1, s_{\tau_m})] - \mathbf{1}_{\{sv_{\tau_m} < sv_f\}} \mathbb{E}_{\tau_{m-1}}^{\mathbb{Q}}[c(1, s_{\tau_m})].$$

It is concave, since  $\mathbb{E}_{\tau_{m-1}}^{\mathbb{Q}}[c(-1, s_{\tau_m})] \leq -\mathbb{E}_{\tau_{m-1}}^{\mathbb{Q}}[c(1, s_{\tau_m})]$ , and is piecewise-linear in the volume in storage with a single kink point at  $sv_f$ , since the expected cost function has a kink point at  $sv_{\tau_m} = sv_f$ .

In the next step, we conduct that the weak bang-bang condition is satisfied at  $\tau_{i-1}$  if the first condition holds true for  $\tau_i$  and the second condition holds true for  $\tau_{i-1}$ . The dynamic programming principle states that the optimal trading strategy at  $\tau_{i-1}$  is given by

$$q_{\tau_{i-1}}^* = \arg \max_{q_{\tau_{i-1}}} \underbrace{\left( -q_{\tau_{i-1}} s_{\tau_{i-1}} + c(q_{\tau_{i-1}}, s_{\tau_{i-1}}) \right)}_{\text{payoff}} + \underbrace{cv_{\tau_{i-1}}(s_{\tau_{i-1}}, sv_{\tau_{i-1}} + q_{\tau_{i-1}})}_{\text{continuation value}}$$

subject to  $q_{\tau_{i-1}} \in [\max\{-sv_{\tau_i}, \underline{q}\}, \min\{\overline{sv} - sv_{\tau_i}, \overline{q}\}]$ . The payoff function and the continuation value function are concave and piecewise-linear with kink points at zero  $q_{\tau_{i-1}} = 0$  and  $\{q_{\tau_{i-1}} + sv_{\tau_{i-1}} \mid q_{\tau_{i-1}} + sv_{\tau_{i-1}} \in wbb(\psi)\}$ , respectively. Therefore, the objective function is concave and piecewise-linear and its potential maxima are at zero, the kink points of the continuation value function, and the boundary points  $\max\{-sv_{\tau_i}, \underline{q}\}$  and  $\min\{\overline{sv} - sv_{\tau_i}, \overline{q}\}$ . Especially, the optimal trading strategy satisfies the weak bang-bang condition if  $sv_{\tau_{i-1}} \in wbb(\psi)$ . In addition, the piecewise linear objective function implies that optimal trading strategies for two storage contracts with storage volumes  $sv_{\tau_{i-1}} + \delta$  and  $sv_{\tau_{i-1}}$ , where  $0 < \delta < lcf(\psi)$ , are equal to  $q_{\tau_{i-1}}^{(\delta)} = q_{\tau_{i-1}}^{(0)} = \underline{q}$  (maximal withdrawal volume),  $q_{\tau_{i-1}}^{(\delta)} = q_{\tau_{i-1}}^{(0)} = \overline{q}$  (maximal injection volume), or  $q_{\tau_{i-1}}^{(\delta)} - q_{\tau_{i-1}}^{(0)} = -\delta$ .

In the last step, we prove that the continuation value function is piecewise-linear with kink points on the weak bang-bang grid at  $\tau_{i-1}$  if (i) and (ii) hold true for  $\tau_i$ :

$$\begin{aligned} \Delta cv_{\tau_{i-1}}(s_{\tau_{i-1}}, sv_{\tau_i}, \delta) &= cv_{\tau_{i-1}}(s_{\tau_{i-1}}, sv_{\tau_i} + \delta) - cv_{\tau_{i-1}}(s_{\tau_{i-1}}, sv_{\tau_i}) \\ &= \beta(sv_{\tau_i})\delta \end{aligned}$$

if  $sv_{\tau_i} \in wbb(\psi)$  and  $0 \leq \delta < lcf(\psi)$ . In the following, we set the interest rate process equal to zero to reduce notational burden. The left-hand side can be rewritten based on the dynamic programming principle as follows:

$$\begin{aligned} \Delta cv_{\tau_{i-1}}(s_{\tau_{i-1}}, sv_{\tau_i}, \delta) &= \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \max_{q_{\tau_i}^{(\delta)}} \left( -q_{\tau_i}^{(\delta)} s_{\tau_i} + c(q_{\tau_i}^{(\delta)}, s_{\tau_i}) + cv_{\tau_i}(s_{\tau_i}, sv_{\tau_i} + \delta + q_{\tau_i}^{(\delta)}) \right) \right] \\ &\quad - \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \max_{q_{\tau_i}^{(0)}} \left( -q_{\tau_i}^{(0)} s_{\tau_i} + c(q_{\tau_i}^{(0)}, s_{\tau_i}) + cv_{\tau_i}(s_{\tau_i}, sv_{\tau_i} + q_{\tau_i}^{(0)}) \right) \right] \end{aligned}$$

subject to

$$\begin{aligned} \max\{-sv_{\tau_i}, \underline{q}\} &\leq q_{\tau_i}^{(0)} \leq \min\{\overline{sv} - sv_{\tau_i}, \overline{q}\} \\ \max\{-(sv_{\tau_i} + \delta), \underline{q}\} &\leq q_{\tau_i}^{(\delta)} \leq \min\{\overline{sv} - (sv_{\tau_i} + \delta), \overline{q}\}. \end{aligned}$$

Now, we can exploit that  $q_{\tau_i}^{(\delta)} = q_{\tau_i}^{(0)} = \underline{q}$ , or  $q_{\tau_i}^{(\delta)} = q_{\tau_i}^{(0)} = \bar{q}$ , or  $q_{\tau_i}^{(\delta)} - q_{\tau_i}^{(0)} = -\delta$ . Thus, we can rewrite the continuation value as follows:

$$\begin{aligned} \Delta cv_{\tau_{i-1}}(s_{\tau_{i-1}}, sv_{\tau_i}, \delta) &= \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{q_{\tau_i}^{(\delta)} = \underline{q}\}} (cv_{\tau_i}(s_{\tau_i}, sv_{\tau_i} + \delta + \underline{q}) - cv_{\tau_i}(s_{\tau_i}, sv_{\tau_i} + \underline{q})) \right] \\ &+ \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{q_{\tau_i}^{(\delta)} = \bar{q}\}} (cv_{\tau_i}(s_{\tau_i}, sv_{\tau_i} + \delta + \bar{q}) - cv_{\tau_i}(s_{\tau_i}, sv_{\tau_i} + \bar{q})) \right] \\ &+ \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{q < q_{\tau_i}^{(\delta)} \leq 0\}} (\delta s_{\tau_i} + c(-\delta, s_{\tau_i})) \right] \\ &+ \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{0 < q_{\tau_i}^{(\delta)} < \bar{q}\}} (\delta s_{\tau_i} - c(\delta, s_{\tau_i})) \right], \end{aligned}$$

where  $\mathbf{1}_{\{q_{\tau_i}^{(\delta)} = x\}}$  is one when the optimal trading strategy at  $\tau_i$  is equal to  $x$ . In the first two cases, the optimal trading strategies for both storage contracts lead to the same payoff at  $\tau_i$ . Thus, the difference between the expected storage value is equal to the expected difference between the continuation values at the next trading date  $\tau_i$ . Then, we can use the induction anchor to show that

$$\begin{aligned} \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{q_{\tau_i}^{(\delta)} = \underline{q}\}} (cv_{\tau_i}(s_{\tau_i}, sv_{\tau_i} + \delta + \underline{q}) - cv_{\tau_i}(s_{\tau_i}, sv_{\tau_i} + \underline{q})) \right] &= \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{q_{\tau_i}^{(\delta)} = \underline{q}\}} \beta_{\tau_i}(sv_{\tau_i} + \underline{q}) \delta \right] \\ &= \underbrace{\mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{q_{\tau_i}^{(\delta)} = \underline{q}\}} \beta_{\tau_i}(sv_{\tau_i} + \underline{q}) \right]}_{\hat{\beta}_1} \delta, \\ \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{q_{\tau_i}^{(\delta)} = \bar{q}\}} (cv_{\tau_i}(s_{\tau_i}, sv_{\tau_i} + \delta + \bar{q}) - cv_{\tau_i}(s_{\tau_i}, sv_{\tau_i} + \bar{q})) \right] &= \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{q_{\tau_i}^{(\delta)} = \bar{q}\}} \beta_{\tau_i}(sv_{\tau_i} + \bar{q}) \delta \right] \\ &= \underbrace{\mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{q_{\tau_i}^{(\delta)} = \bar{q}\}} \beta_{\tau_i}(sv_{\tau_i} + \bar{q}) \right]}_{\hat{\beta}_2} \delta. \end{aligned}$$

In the last two cases, both storage contracts have the same continuation value after the next rolling date, which means that the difference between storage values referring to storage volumes  $sv_{\tau_i} + \delta$  and  $sv_{\tau_i}$  is equal to the difference in the expected payoff at  $\tau_i$ . The difference in expected payoffs is piecewise-linear due to the piecewise linearity of the payoff function

$$\begin{aligned} \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{q < q_{\tau_i}^{(\delta)} \leq 0\}} (\delta s_{\tau_i} + c(-\delta, s_{\tau_i})) \right] &= \underbrace{\mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{q < q_{\tau_i}^{(\delta)} \leq 0\}} (s_{\tau_i} + c(-1, s_{\tau_i})) \right]}_{\hat{\beta}_3} \delta, \\ \mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{0 < q_{\tau_i}^{(\delta)} < \bar{q}\}} (\delta s_{\tau_i} - c(\delta, s_{\tau_i})) \right] &= \underbrace{\mathbb{E}_{\tau_{i-1}}^{\mathbb{Q}} \left[ \mathbf{1}_{\{0 < q_{\tau_i}^{(\delta)} < \bar{q}\}} (s_{\tau_i} - c(1, s_{\tau_i})) \right]}_{\hat{\beta}_4} \delta. \end{aligned}$$

Thus, the continuation value is piecewise-linear

$$\Delta cv_{\tau_{i-1}}(s_{\tau_{i-1}}, sv_{\tau_i}, \delta) = \underbrace{(\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 + \hat{\beta}_4)}_{\beta(sv_{\tau_i})} \delta,$$

where the intercept  $\beta(sv_{\tau_i})$  depends on unknown market information at  $\tau_i$  from an ex-ante perspective. The concavity of the continuation value function at  $\tau_{i-1}$  follows directly from the concavity of the continuation value function at  $\tau_i$ . ■

In summary, the storage value can be expressed by a numerically manageable discrete optimization problem without any approximation error

$$sw_{\tau_0}(s_{\tau_0}, sv_{\tau_0}; \psi) = \sup_{\{q_{\tau_i}\}_{i=0}^m} \mathbb{E}_{\tau_0}^{\mathbb{Q}} \left[ \sum_{i=0}^m e^{-r(\tau_i - \tau_0)} (-q_{\tau_i} s_{\tau_i} + c(q_{\tau_i}, s_{\tau_i})) \right], \quad (4.2)$$

where

$$q_{\tau_i} \in \{q = \underline{q} + j \cdot lcf(\psi) \mid q \in [\max\{-sv_{\tau_i}, \underline{q}\}, \min\{\bar{sv} - sv_{\tau_i}, \bar{q}\}]\} \text{ for } i = 0, \dots, m.$$

Interestingly, our theoretical results are not restricted to standard storage contracts but can also be applied to flexible delivery contracts with take-or-pay clauses (ToP contracts). The ToP contract design is widely used to share price and quantity risk among contracting parties through (i) minimal and maximal delivery rates  $\underline{q}$  and  $\bar{q}$ , (ii) minimal and maximal total cumulative delivery volumes  $\underline{sv}$  and  $\bar{sv}$ , and (iii) a fixed or indexed delivery price  $d_{with} = s^{(ToP)}$ . Thompson (1995) gives a simple example of a ToP contract: the take-or-pay contract gives its holder the right to purchase up to one unit of natural gas in each month for a fixed delivery price of 1.50 USD per unit, but also requires payment of a financial penalty if the total purchase volume is less than 7.2 units over the underlying contract year. The market value of such a take-or-pay contract can be calculated by the following duplication strategy: (i) a long position of 7.2 futures contracts with a uniform delivery of one unit over one year and (ii) a storage contract with an initial volume in storage of  $sv_{\tau_0} = 4.8$ , a maximal injection rate of  $\bar{q} = 0.6$  per month, a maximal withdrawal rate of  $\underline{q} = -0.4$  per month, a total storage capacity of  $\bar{sv} = 7.6$ , and a cost function with  $b_{inj} = d_{with} = b_{with} = 0$  and  $d_{with} = 1.5$ .<sup>11</sup> The only difference to a standard storage contract is that the total cumulative delivery volume (final volume in storage) need not be equal to a specific prearranged value. Instead, it must be in the interval  $[0, 4.8]$ . This can easily be incorporated through modifying the contract payments at the end of the contract period to  $\gamma \mathbf{1}_{\{sv_{\tau_{m+1}} \notin [0, 4.8]\}}$  for a sufficiently large  $\gamma > 0$ . It turns out that our theoretical results can also be applied for piecewise-linear final payments in the volume in storage.<sup>12</sup> Thus, all our theoretical results referring to the general specification hold true for storage and take-or-pay contracts.

---

<sup>11</sup>For simplicity, we assume that the interest rate process is zero.

<sup>12</sup>In our proof, we only require that the payoff function and the continuation value function are piecewise-linear (see induction anchor).

It is important to keep in mind that the operation costs of storage facilities are much lower than the fixed delivery prices of take-or-pay contracts. As a result, take-or-pay contracts are more similar to standard option contracts than storage contracts are.

## 4.5 Numerical Example

Now, we use a simple numerical example to illustrate that the optimal storage strategy satisfies the weak bang-bang condition.

The underlying storage contract is characterized by the following contract terms: a contract period of  $[\tau_0, \tau_3]$ , an initial and final volume in storage of  $sv_{\tau_0} = sv_f = 2$ , maximal withdrawal and injection rates of  $(\underline{q}, \bar{q}) = (-2, 1)$ , a total storage capacity of  $\bar{sv} = 5$ , and a zero cost function  $c(\cdot, \cdot) \equiv 0$ .

The holder of the storage contract can trade in a spot contract with a rolling delivery period  $(\tau_i, \tau_{i+1}]$  for  $i \in \{0, 1, 2\}$ . The current spot and futures prices for the delivery periods  $\{(\tau_i, \tau_{i+1}]\}_{i=0}^2$  are given by

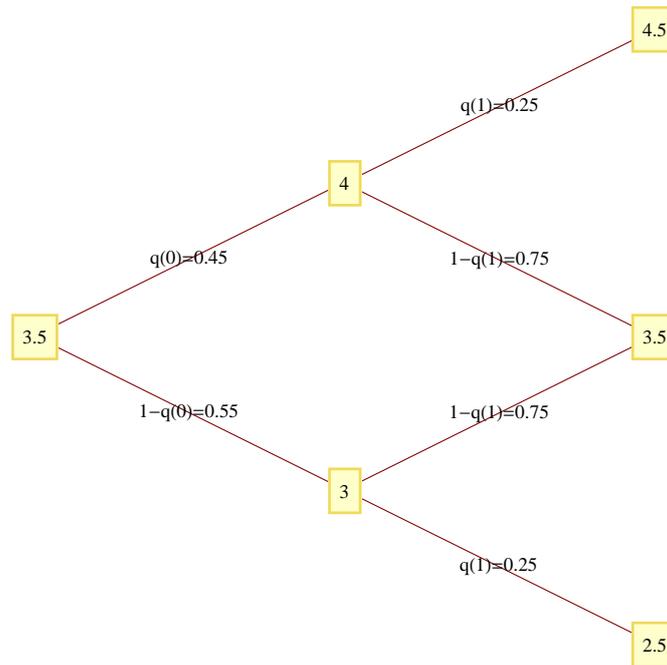
$$s_{\tau_0} = f_{\tau_0}(\tau_0, \tau_1) = 3.5, \quad f_{\tau_0}(\tau_1, \tau_2) = 3.45, \quad \text{and} \quad f_{\tau_0}(\tau_2, \tau_3) = 3.475, \quad (4.3)$$

where the interest rate process is set equal to zero. It is already known that the holder of the storage contract only has to trade in the spot contract at  $\{\tau_i\}_{i=0}^2$  to maximize the expected storage value under the pricing measure. For that reason, we only have to specify a time-discrete spot price process for the valuation of the underlying storage contract. In our example, we use a simple binomial tree to model a “mean reverting” spot price process. The mean reversion behavior is captured by setting the probability of an up move in the down state equal to the probability of a down move in the up state (see Figure 4.4). This allows us to calculate unique risk-neutral probabilities  $q(0) = 0.45$  and  $q(1) = 0.25$  for the given spot price process based on the initial futures price curve (4.3). Next, we recursively calculate the optimal trading strategy for all admissible volumes in storage in order to show that the weak bang-bang property is satisfied. At the last trading date  $\tau_2$ , the optimal trading strategy is simply given by

$$q_{\tau_2}^* = 2 - sv_{\tau_2}, \quad 1 \leq sv_{\tau_2} \leq 4.$$

This allows us to calculate the continuation value functions  $cv_{\tau_1}(4, sv_{\tau_2})$  and  $cv_{\tau_1}(3, sv_{\tau_2})$  for the up and down state at  $\tau_1$  as follows:

$$\begin{aligned} cv_{\tau_1}(4, sv_{\tau_2}) &= 0.25(sv_{\tau_2} - 2)4.5 + 0.75(sv_{\tau_2} - 2)3.5 = 3.75sv_{\tau_2} - 7.5 \quad (\text{up state}), \\ cv_{\tau_1}(3, sv_{\tau_2}) &= 0.75(sv_{\tau_2} - 2)3.5 + 0.25(sv_{\tau_2} - 2)2.5 = 3.25sv_{\tau_2} - 6.5 \quad (\text{down state}). \end{aligned}$$



**Figure 4.4: spot price dynamics**

This figure shows the discrete spot price process and the risk-neutral probabilities for our numerical example.

The dynamic programming principle then leads to the optimal trading strategy  $q_{\tau_1}^*$  at  $\tau_1$ :

$$q_{\tau_1}^* = \arg \max_{\substack{1 \leq sv_{\tau_1} + q_{\tau_1} \leq 4 \\ -2 \leq q_{\tau_1} \leq 1}} (3.75 (sv_{\tau_1} + q_{\tau_1}) - 7.5 - 4q_{\tau_1}) = 1 - sv_{\tau_1} \quad (\text{up state}),$$

$$q_{\tau_1}^* = \arg \max_{\substack{1 \leq sv_{\tau_1} + q_{\tau_1} \leq 4 \\ -2 \leq q_{\tau_1} \leq 1}} (3.25 (sv_{\tau_1} + q_{\tau_1}) - 6.5 - 3q_{\tau_1}) = 1 \quad (\text{down state}).$$

In the up state, the spot price at  $\tau_1$  is larger than the expected spot prices at  $\tau_2$ . Thus, the optimal trading strategy is equal to the maximal withdrawal volume given that the prearranged final volume in storage can be reached. In the down state, the optimal trading strategy is equal to the maximal injection rate, since the spot price at  $\tau_1$  is smaller than the expected spot price at  $\tau_2$  and the maximal injection volume is an admissible trading strategy independent of the current volume in storage. Thus, the optimal trading strategy satisfies the weak bang-bang condition if  $sv_{\tau_1} \in \{0, 1, 2, 3\}$ .

Finally, we calculate the continuation value and the optimal trading strategy at  $\tau_0$ . This

requires to calculate the storage value in both states at  $\tau_1$  subject to the volume in storage:

$$\begin{aligned} 3.75 (sv_{\tau_1} + \underbrace{(1 - sv_{\tau_1})}_{q_{\tau_1}^*}) - 7.5 - 4 \underbrace{(1 - sv_{\tau_1})}_{q_{\tau_1}^*} &= 4 sv_{\tau_1} - 7.75 && \text{(up state),} \\ 3.25 (sv_{\tau_1} + \underbrace{1}_{q_{\tau_1}^*}) - 6.5 - 3 \underbrace{1}_{q_{\tau_1}^*} &= 3.25 sv_{\tau_1} - 6.25 && \text{(down state).} \end{aligned}$$

Then, the continuation value function at  $\tau_0$  is the weighted sum of these (piecewise) linear functions and is thus also (piecewise) linear. It is given by

$$cv_{\tau_0}(3.5, sv_{\tau_1}) = 0.45 (4 sv_{\tau_1} - 7.75) + 0.55 (3.25 sv_{\tau_1} - 6.25) = 3.5875 sv_{\tau_1} - 6.925.$$

In this case, we only have to consider a single side condition, since the initial volume in storage is known. The optimal trading strategy at  $\tau_0$  is equal to one, since

$$q_{\tau_0}^* = \arg \max_{-2 \leq q_{\tau_0} \leq 1} (3.5875 (sv_{\tau_0} + q_{\tau_0}) - 6.925 - 3.5 q_{\tau_0}) = 1.$$

Thus,  $sv_{\tau_1}$  is equal to three and the initial storage value is  $sw_{\tau_0}(3.5, 2; \psi) = 0.3375$ . This shows that the optimal trading strategy satisfies the weak bang-bang property in our discrete numerical example.

## 4.6 Natural Gas Storage Valuation

In this section, we apply our energy market modeling approach to the valuation of a concrete storage contract in the U.S. natural gas market.

The underlying market consists of the day-ahead contract (spot contract) and futures contracts with successive monthly delivery periods.<sup>13</sup> The contract design of the underlying storage contract is chosen similar to the SBUs offered by Centrica. This means that the contract period lasts from April 1st to March 31st of the following year, maximal withdrawal and injection volumes are given by  $\underline{q} = -3$  mmBtu/day and  $\bar{q} = 1$  mmBtu/day, and the total storage capacity is equal to  $\bar{sv} = 180$  mmBtu. The variable injection and withdrawal costs are given by  $d_{inj} = 0.07$  USD/mmBtu and  $d_{with} = 0.02$  USD/mmBtu and no natural gas is lost during the injection or withdrawal process, i.e.,  $b_{inj} = b_{with} = 0$ . The initial volume in storage is zero and the storage must be returned with the same

---

<sup>13</sup>In the U.S. natural gas market, short-dated delivery contracts are traded over-the-counter, while long-dated delivery contracts are exchange-traded products. The day-ahead contract is the most important and transparent short-dated delivery contract. For other delivery contracts, market prices are neither publicly available, nor can they be obtained from the Bloomberg database.

volume at the end of the contract period.

Now, we specify our stochastic modeling approach for the underlying price dynamics based on the fundamental consideration of the storage valuation problem in Section 4.2. This means that we set up a stochastic price process that captures temporary price impacts, a stochastic summer-winter spread, and parallel shifts. It is no problem to incorporate these risk factors in our energy market modeling approach. In contrast, market models and standard spot price models are either not able to model the common price dynamics of day-ahead and futures contracts (market models), cannot be used to model a stochastic summer-winter spread over time (affine-linear spot price models), or are difficult to calibrate to market data (non affine-linear spot price models).

In the energy market model, we assume that a linear relation between the underlying risk factors  $\{x_t^{(j)}\}_{j=1}^3$  and the futures prices of the market model  $\{f_t^{(i)}\}_{i=1}^{12}$  exists.<sup>14</sup> This has the advantage that parallel shifts can be handled analytically. It follows then that the price dynamics of a futures contract with an arbitrary delivery period  $(\tau_b, \tau_e]$  is given by

$$df_t(\tau_b, \tau_e) = \sum_{j=1}^3 \left( \int_{\tau_b}^{\tau_e} \widehat{w}(u; \tau_b, \tau_s) \beta_j(u) du \right) dx_t^{(j)}, \quad (4.4)$$

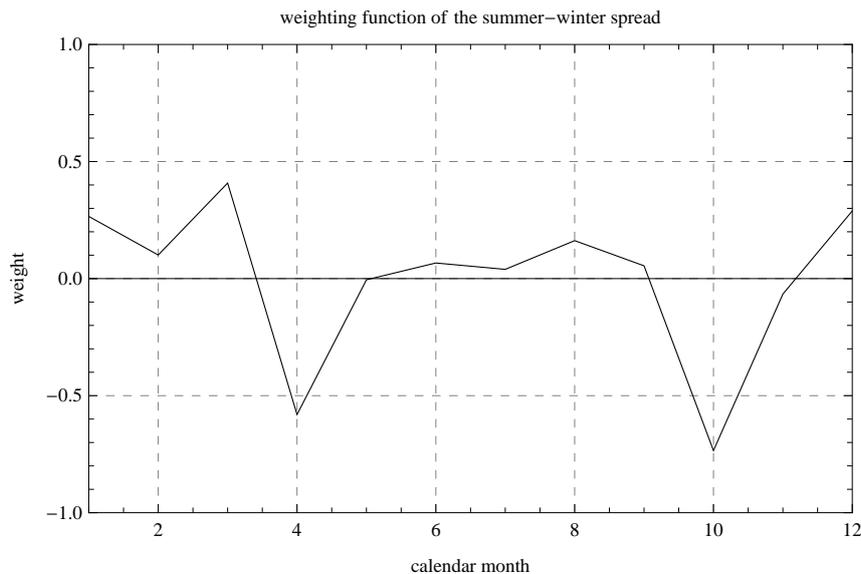
where  $\{\beta_j(u)\}_{j=1}^3$  are maximal smooth splines that satisfy the two no-arbitrage conditions (3.6) and (3.7), as well as the maximum smoothness criterion (3.8). In (4.4), spline parameters are directly link to the underlying risk factors to shorten notation. Now, we specify the risk factors and their impact on the futures price curve for the above market model. It is useful to link the different risk factors  $\{x_t^{(j)}\}_{j=1}^3$  to observable market data in order to directly capture relevant futures price movements for the underlying application and to increase the robustness of estimation results.

In our case, we choose the individual risk factors and their impact on the futures price curve according to our fundamental consideration of the storage valuation problem in Section 4.2. The first risk factor  $x_t^{(1)}$  is set equal to the arithmetic average futures price of the following twelve monthly delivery periods and has the same impact on all futures prices. The second risk factor  $x_t^{(2)}$  is equal to the summer-winter spread between the October and January futures contracts. The impact of the summer-winter spread on the different futures prices of the market model is estimated using the ordinary least square estimator based on return data between February and October.<sup>15</sup> In Figure 4.5, we show the estimated impact of the summer-winter spread on the futures prices of the market model (weighting function). As expected, the weighting function is close to zero for the

---

<sup>14</sup>The  $i$ -th futures price refers to a uniform delivery in the  $i$ -th calendar month for  $i = 1, \dots, 12$ .

<sup>15</sup>We do not use market data between October and January, because the summer-winter spread is not observable during these calendar months.



**Figure 4.5: weighting function of the summer-winter spread**

This figure shows the estimated impact of the summer-winter spread on each futures price.

off-peak season from May to September, negative for the beginning and end of the peak season, and positive in the winter months. The third risk factor  $x_t^{(3)}$  captures temporary price risks that have no impact on the futures prices of the market model. In contrast to the first two risk factors, we have to apply a simple filtering approach, similar to the one in the previous chapter, to extract temporary price impacts from historical market data. In detail, we calculate the price residuals between the model price for the two-day-ahead contract and the observable day-ahead contract at the following day at any business day. Then, we subtract the weighted movements of the first two risk factors from these price residuals to obtain returns of the third risk factor. The impact of the third risk factor on the futures price curve is uniquely determined by the smooth interpolation approach based on the additional market information about the day-ahead price.

Lastly, we have to decide on the stochastic modeling approach of the risk factors based on their underlying time series. In order to maintain a two-dimensional optimization problem, we do not use complex stochastic processes, such as stochastic volatility models, to model return distributions but rather use a one-dimensional hyperbolic distribution for each risk factor

$$x_{t_{i+1}}^{(j)} - x_{t_i}^{(j)} \sim HB(\alpha, \beta, \delta, \mu),$$

where  $x_{t_i}^{(j)}$  is the  $i$ -th observation of the  $j$ -th risk factor for  $j \in \{1, 2, 3\}$  and  $i = 1, \dots, n$ . The density function of the hyperbolic distribution is given by

$$f_{hb}(x) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha\gamma K_1(\gamma\sqrt{\alpha^2 - \beta^2})} e^{\beta(x-\mu) - \alpha\sqrt{\gamma^2 + (x-\mu)^2}}, \quad \mu \in \mathbb{R}, \alpha, \gamma > 0, 0 \leq |\beta| < \alpha,$$

where  $K_1$  denotes the modified Bessel function of the third kind with index 1 (see Eberlein and Keller (1995)).

This uniquely determines our modeling approach. The model parameters are estimated based on U.S. natural gas data from January 1, 2005 to December 31, 2010.<sup>16</sup> We first subtract the empirical mean from each underlying time series and then estimate the model parameters of the hyperbolic distribution using the maximum likelihood method, subject to

$$\mathbb{E}_{t_i}^{\mathbb{Q}}[x_{t_{i+1}}^{(j)} - x_{t_i}^{(j)}] = \mu + \frac{\gamma\beta K_2(\gamma\sqrt{\alpha^2 - \beta^2})}{\sqrt{\alpha^2 - \beta^2} K_1(\gamma\sqrt{\alpha^2 - \beta^2})} = 0.$$

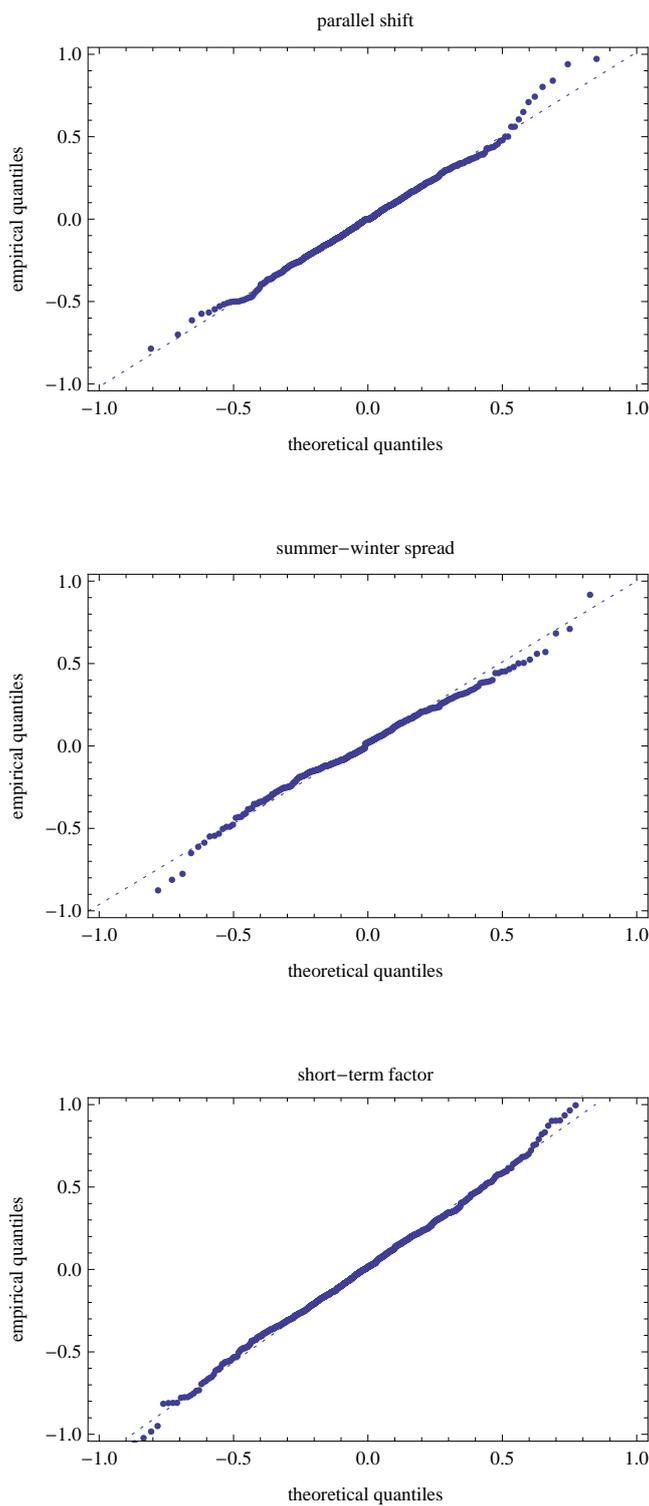
The ancillary condition ensures that the martingale property is satisfied. In Figure 4.6, we test the distributional properties of the underlying return distributions using QQ-plots and find that the return data are well captured by the hyperbolic distribution. The estimated model framework uniquely determines the theoretical storage value via the discrete optimization problem (4.2), where the interest rate process is set equal to zero. In our case, optimal trading times can be restricted to the last trading dates of the day-ahead contract and optimal trading volumes can be restricted to selling one, two, or three units, doing nothing, or buying one unit. This is because the largest common factor of the maximal injection rate  $\bar{q} = 1$ , the maximal withdrawal rate  $\underline{q} = -3$ , the total storage capacity  $\bar{sv} = 180$ , and the final volume in storage  $sv_f = 0$  is equal to one.

There are various algorithms that can be used to solve the optimization problem numerically. Boogert and de Jong (2008) propose a simulation-based algorithm to approximate the continuation value function of a storage contract based on the Longstaff and Schwartz (2001) algorithm.

Bardou, Bouthemy, and Pages (2009) apply a quantization tree approach to numerically determine storage values. In our empirical tests, we use the Longstaff-Schwartz (LS) algorithm, which has provided good approximation results in previous studies (see Boogert and de Jong (2008), Neumann and Zachmann (2009), and Boogert and de Jong (2011)). Based on the initial futures price curve from March 29, 2012 (see Figure 4.7 (left)), we

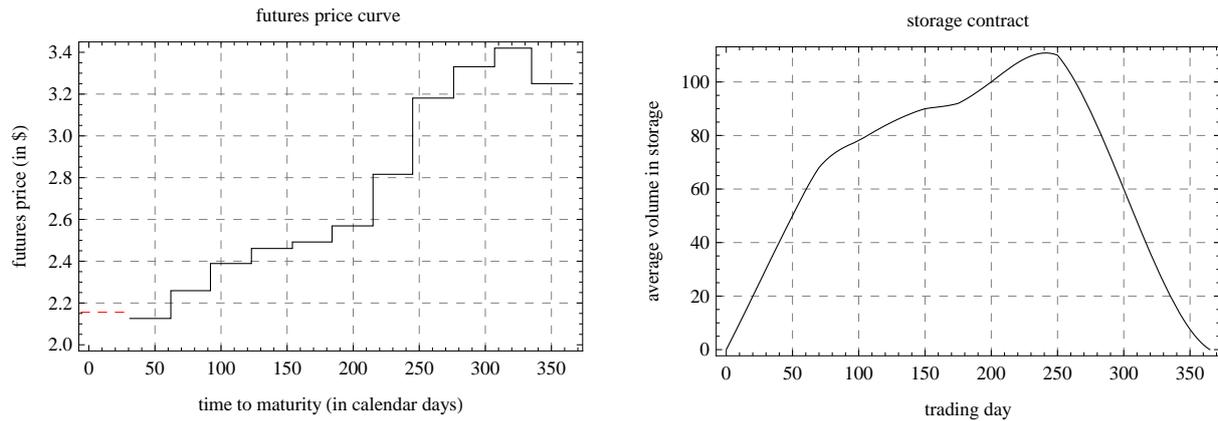
---

<sup>16</sup>It is generally not necessary to estimate the price dynamics of parallel shifts, because parallel shifts can be easily extracted from market data and their pricing impact can be determined analytically. Nevertheless, for the sake of completeness, we calibrate all three model components based on the maximum likelihood estimator.



**Figure 4.6: quantile-quantile-plots**

These graphs show the QQ-plots for parallel shifts (top), the summer-winter spread (middle), and the short-term risk factor (bottom) based on the estimated hyperbolic return distributions. The underlying data set consists of natural gas day-ahead prices and futures prices from January, 2005 to December, 2010. The natural gas prices refer to physical deliveries at Henry Hub in Louisiana.



**Figure 4.7: initial futures price curve and average volume in storage**

These left graph shows the natural gas futures price curve on March 29, 2012. The natural gas futures prices refer to physical deliveries at Henry Hub in Louisiana. The right graph shows the average volume in storage for a contract period of one year.

generate ten thousand price paths through our energy market model. In the simulation run, we ignore parallel shifts that have no impact on the fair storage volume and can subsequently be incorporated in the optimal trading strategy (see Section 4.2).

In the LS algorithm, we recursively calculate the storage values for all admissible volumes in storage based on the simulated price scenarios. We start with the final trading day and determine the storage value for all admissible storage volumes for each simulated price path. Then, we choose a parametric function (e.g., polynomials) and regress the state variables at  $\tau_{m-1}$  on the known storage values at  $\tau_m$  for all admissible volumes in storage in order to approximate the unknown continuation value function. In our case, we have two state variables: the summer-winter spread and the temporary risk factor. Next, we use the dynamic programming principle to calculate the optimal trading strategy based on the current payoff and the approximated continuation value function. This allows us to calculate the storage value for each volume in storage and each price path based on the “optimal” trading strategy at  $\tau_{m-1}$ . This procedure is repeated until the first trading date and provides an approximation method for the optimal trading strategy and the true storage value.

In the LS algorithm, we only have to decide on the parametric function to approximate the continuation value. We have tested several specifications and obtain the best results<sup>17</sup> with polynomials of order three given by

$$cv(z_1, z_2, sv_{\tau_i}) = \sum_{k=0}^3 \sum_{j=0}^{3-i} h_{k,j}(sv_{\tau_i}) z_1^k z_2^j, \quad sv_{\tau_i} \in wbb(-3, 1, 180, 0), h_{i,j} \in \mathbb{R}. \quad (4.5)$$

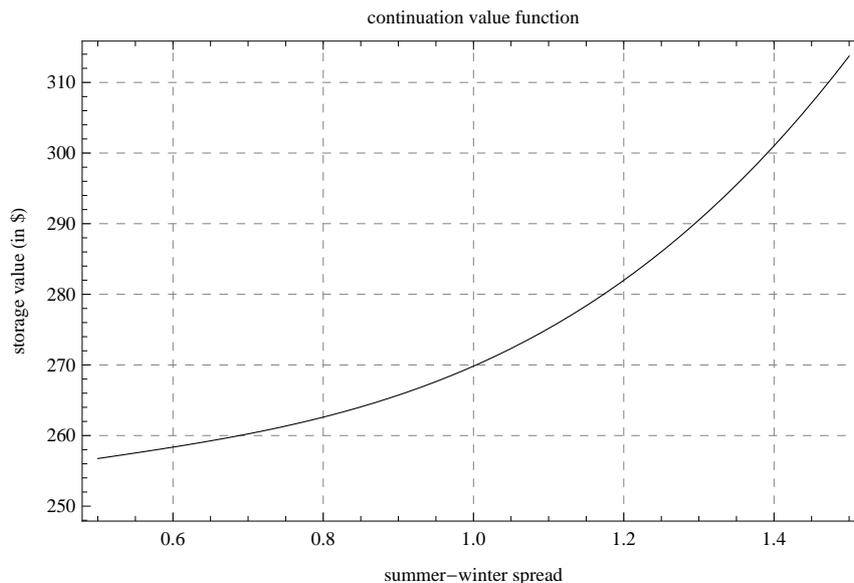
In (4.5), we choose the observable spread between the day-ahead and the front-month futures price as well as the observable summer-winter spread for the two observable state variables  $(z_1, z_2)$  at any trading date. This simplifies the implementation of the derived optimal trading strategy compared to using the unobservable temporary risk factor directly in real applications. Our simulation results show that a dynamic trading strategy in the day-ahead contract increases the expected storage value by 112 percent over a static trading strategy in futures contracts (168 USD (static) to 357 USD (dynamic)). However, the optimal dynamic trading strategy is highly risky, having a standard deviation of 92 USD given that parallel shifts and the stochastic summer-winter spread are actively hedged, while the temporary risk factor remains unhedged.<sup>18</sup> Moreover, we find that the price spread between day-ahead and front-month futures contracts largely determines the optimal storage strategy and that the storage value is, as expected, a convex function in the summer-winter spread (see Figure 4.8). When we look at the average volume in storage (see Figure 4.7 (right)), we find that optimal trading volumes are close to the maximal injection rate of one at the beginning of the contract period. There are two main reasons for this: first, the initial volume in storage is zero. Thus, an injection decision increases the storage value more strongly than for a more filled storage facility. Second, the futures price curve is relatively steep, which means that delaying an injection decision is relatively costly. The injection rate is substantially lower between the 70th and 260th trading days. During this period, storage operators mainly try to profit from temporary imbalances in the markets. In the peak season, the optimal trading strategy is to sell at an average rate of about 2.25. Interestingly, the short-term optimization has such a strong impact on the optimal trading strategy that the expected maximal volume in storage is only 60-65% of the total storage capacity over the underlying contract period.

In summary, we have illustrated some of the advantages of our market modeling approach over standard modeling techniques. It was shown that the storage valuation problem depends on the common stochastic behavior of day-ahead prices and the summer-winter

---

<sup>17</sup>This means that we obtain the highest values for the storage contract when we apply the optimal trading strategy to various simulated price paths.

<sup>18</sup>It is possible to hedge price risks that arise due to parallel shifts and the stochastic summer-winter spread component. In contrast, temporary price risk can only be hedged if another short-dated delivery contract is traded at the market.



**Figure 4.8: continuation value function**

This figure shows the relation between the continuation value function (4.5) and the summer-winter spread. The continuation value function refers to a storage contract with a volume in storage equal to zero at the 75th trading day ( $sv_{\tau_{75}} = 0$ ).

spread. This can be captured by neither standard affine-linear stochastic processes nor by standard market models. In a standard affine-linear spot price process, the futures price curve is shocked dependent on time to maturity. Thus, a stochastic summer-winter spread, which does not depend on time to maturity, cannot be modeled in the affine-linear framework. Furthermore, alternative model specifications would imply complex model prices for futures contracts referring to delivery periods, which are difficult to consistently calibrate to market data. In contrast to affine-linear spot price dynamics, market models are able to model arbitrary futures price movements but not the common stochastic price behavior of the day-ahead and futures contracts. Thus, market models cannot be applied for the valuation of storage contracts. In addition, we have theoretically shown that optimal trading volumes can be restricted to the weak bang-bang grid. This saves a lot of computational time compared to finer grids. However, it turns out that optimal trading volumes are mostly equal to the maximal injection or withdrawal rate. Thus, weak bang-bang strategies only slightly increase the storage value compared to strong bang-bang strategies (less than 2 percent in our example).

# Chapter 5

## Conclusion and Outlook

In this thesis, we examine stochastic modeling approaches for a single futures contract and for the entire futures price curve from different perspectives.

Our first contribution is to provide a novel approach to incorporate option market data through a suitably aggregated option portfolio in an estimation algorithm. The option portfolio is constructed in a way that its market value linearly depends on the latent variance state. This allows one to filter out latent variance states by solving linear equations instead of using non-linear option price formulas. Hereby, it is possible to obtain consistent and more robust estimation results for model parameters, latent state variables, and risk premia without increasing computational time considerably compared to standard estimation methods ignoring option market data. The estimation method does not depend on any specific characteristic of energy markets, and thus can be applied to various other financial instruments.

Our second contribution is a comprehensive empirical analysis of the crude oil futures price dynamics. The empirical results provide a refined view of the role of different risk factors for the pricing and hedging of derivative instruments. We show that a stochastic volatility component is required to capture the distributional properties of historical return data and that a jump component only leads to a slight further improvement. However, the picture changes when we compare the option pricing performance of the different model specifications. Here, we find that volatility risk remains an important factor in explaining strongly fluctuating variance swap rates over time. Though, the shape of observed market-implied volatility smiles can only be reproduced by an additional jump component, while pure stochastic volatility models are not able to capture the tails of implied risk-neutral return distributions. In addition, we shed light on the compensation for taking over different risk factors. We find an insignificant market price of variance risk and a significant aggregated market price of jump risk in the stochastic volatility model

with jumps. In pure stochastic volatility models, we obtain unreliable market prices of variance risk that emphasize the importance of an appropriate model specification for estimating risk premia. As a result, we conclude that the aggregated variance risk premia found in Trolle and Schwartz (2010) should be traced back to a non-zero market price of jump risk and not to a non-zero market price of variance risk as suggested by Doran and Ronn (2008). We also consider the role of volatility and price jump risk from the hedging perspective. Here, we find that both volatility and jump risk are important, although in rather different ways. We show that the risk of hedge portfolios can be considerably reduced if variance risk is actively managed. However, we also find strong evidence for an unhedgeable risk factor that leads to rare but large hedging errors in all our hedge portfolios. As a result, a pure diffusive price process severely underestimates the riskiness of hedge portfolios in the crude oil market. The large hedging errors are mainly caused by price jumps in the futures contract but, in some cases, price jumps in the respective option contract also result in large hedging errors. We suspect that a jump component in the volatility process might be another risk factor that should be considered in order to fully understand and quantify the riskiness of hedge portfolios. We leave this question open for further research.

Our third contribution is a consistent term structure model for energy price dynamics. The modeling approach accounts for the specific features of physical energy trading and remains tractable from an application point of view, whereas existing modeling approaches are either incomplete or intractable. For instance, standard spot price models lead to intractable price dynamics of futures contracts with delivery periods, and market models are not able to capture the common price dynamics of multiple delivery contracts with overlapping delivery periods. We solve the trade-off between tractability and completeness by means of a market model component and a smooth interpolation function. This allows us to, first, specify an arbitrage-free modeling approach based on observable market prices instead of on latent factors and, second, to derive arbitrage-free (theoretical) spot and future price dynamics depending on the same observable market price in a linear albeit time-dependent structure. The interpolation function completes the pricing framework and allows us to value and hedge a broad range of important energy derivatives.

Lastly, we apply our energy market modeling approach for the valuation of a natural gas storage contract. We first simplify the underlying dynamic stochastic optimization problem in several ways. We show that trading can be restricted to physically non-redundant delivery contracts and that trading times and volumes can both be restricted to a finite number. This allows us to formulate a tractable discrete dynamic stochastic optimization problem for numerical calculations. Then, we concretely specify our energy market model for the U.S. natural gas market and assess the market value of a concrete storage contract. Since our energy market modeling approach is not restricted to a specific commodity, it

would be interesting to analyze its advantages over other modeling approaches for different practical applications in other important commodity markets. We leave these issues for future research.

# Appendix A

## Appendix

### A.1 Option Pricing Functions

The functions  $h_t^{(1)}(\cdot)$  and  $h_t^{(2)}(\cdot)$  for the SVJ model at  $t = 0$  are given by

$$\begin{aligned}
 h_0^{(1)}(\tau, f, v, \phi) &= \exp\left\{-2\frac{\kappa_v^{\mathbb{Q}}\theta_v^{\mathbb{Q}}}{\sigma_v^2}\left[\ln\left(1 - \frac{(\xi_v - \kappa_v + (1+i\phi)\rho_{f,v}\sigma_v)(1 - e^{-\xi_v\tau})}{2\xi_v}\right)\right]\right. \\
 &\quad - \frac{\kappa_v^{\mathbb{Q}}\theta_v^{\mathbb{Q}}}{\sigma_v^2}(\xi_v - \kappa_v^{\mathbb{Q}} + (1+i\phi)\rho_{f,v}\sigma_v)\tau \\
 &\quad + \lambda_z(1 + \bar{\mu}_z^{\mathbb{Q}})[(1 + \bar{\mu}_z^{\mathbb{Q}})^{i\phi}e^{0.5i\phi(1+i\phi)(\sigma_z^{\mathbb{Q}})^2} - 1]\tau - \lambda_z i\phi\bar{\mu}_z^{\mathbb{Q}}\tau \\
 &\quad \left. + \frac{i\phi(i\phi + 1)(1 - e^{-\xi_v\tau})}{2\xi_v - (\xi_v - \kappa_v^{\mathbb{Q}} + (1+i\phi)\rho_{f,v}\sigma_v)(1 - e^{-\xi_v\tau})}v + i\phi \ln[f]\right\}, \\
 h_0^{(2)}(\tau, f, v, \phi) &= \exp\left\{\lambda_z[(1 + \bar{\mu}_z^{\mathbb{Q}})^{i\phi}e^{0.5i\phi(i\phi-1)(\sigma_z^{\mathbb{Q}})^2} - 1]\tau - \lambda_z i\phi\bar{\mu}_z^{\mathbb{Q}}\tau\right. \\
 &\quad - 2\frac{\kappa_v^{\mathbb{Q}}\theta_v^{\mathbb{Q}}}{\sigma_v^2}\left[\ln\left(1 - \frac{(\xi_v^* - \kappa_v^{\mathbb{Q}} + i\phi\rho_{f,v}\sigma_v)(1 - e^{-\xi_v^*\tau})}{2\xi_v^*}\right)\right] \\
 &\quad - \frac{\kappa_v^{\mathbb{Q}}\theta_v^{\mathbb{Q}}}{\sigma_v^2}[\xi_v^* - \kappa_v^{\mathbb{Q}} + i\phi\rho_{f,v}\sigma_v]\tau \\
 &\quad \left. + \frac{i\phi(i\phi - 1)(1 - e^{-\xi_v^*\tau})}{2\xi_v^* - (\xi_v^* - \kappa_v^{\mathbb{Q}} + i\phi\rho_{f,v}\sigma_v)(1 - e^{-\xi_v^*\tau})}v + i\phi \ln[f]\right\},
 \end{aligned}$$

where

$$\begin{aligned}
 \xi_v &= \sqrt{(\kappa_v^{\mathbb{Q}} - (1+i\phi)\rho_{f,v}\sigma_v)^2 - i\phi(1+i\phi)\sigma_v^2} \\
 \xi_v^* &= \sqrt{(\kappa_v^{\mathbb{Q}} - i\phi\rho_{f,v}\sigma_v)^2 - i\phi(i\phi - 1)\sigma_v^2}.
 \end{aligned}$$

If we insert  $\lambda_z = 0$  in the above formulas, we obtain  $h_t^{(1)}(\cdot)$  and  $h_t^{(2)}(\cdot)$  for the pure stochastic volatility model. In the GB and JD model,  $h_t^{(1)}(\cdot)$  and  $h_t^{(2)}(\cdot)$  are given by the limes of  $\sigma_v \rightarrow 0$ .

## A.2 MCMC Algorithm

In this section, we present the concrete prior and conditional posterior distributions for our MCMC algorithm (see Section 2.3). We use uninformative prior distributions for all parameters (prior distributions with large standard deviations). The only two exceptions are the jump size variance and jump intensity parameters. Here, we choose an informative prior distributions to identify jumps as rare events that are associated with large futures price returns. Table A.1 shows the parameters of each prior distribution.

First, we provide the update blocks of our estimation approach, where the updating sequence is chosen randomly in the MCMC algorithm. The update blocks for the Gibbs Sampler as well as the components that are updated with the Metropolis-Hastings algorithm are given by

(i) *state variables:*

(a) jump events and jump sizes:

$$\begin{aligned} p(\delta n_{f,t_i} | \mu_f^{\mathbb{P}}, \bar{\eta}_f, \lambda_z, v_{t_i}, y_{t_i}, z_{t_i}) &\sim \text{Ber}, \\ p(z_{t_i} | \mu_f^{\mathbb{P}}, \bar{\eta}_f, \mu_z^{\mathbb{P}}, (\sigma_z^{\mathbb{P}})^2, v_{t_i}, y_{t_i}, \delta n_{f,t_i}) &\sim \mathcal{N}, \end{aligned}$$

where  $\mu_f^{\mathbb{P}} = -\lambda_z \bar{\mu}_z^{\mathbb{Q}}$  denotes the price jump compensator.

(b) variance states:

$$p(v_{t_i} | u_{-v_{t_i}}, vs, y) : \text{Metropolis},$$

where  $u_{-v_{t_i}}$  corresponds to the vector of model parameters and state variables excluding the current variance state  $v_{t_i}$ .

(ii) *model parameters:*

(a) model parameters of the jump process:

$$\begin{aligned} p(\lambda_z | \delta n_f) &\sim \mathcal{B}, \\ p(\mu_z^{\mathbb{P}} | (\sigma_z^{\mathbb{P}})^2, z) &\sim \mathcal{N}, \\ p((\sigma_z^{\mathbb{P}})^2 | \mu_z^{\mathbb{P}}, z) &\sim \mathcal{IG}, \\ p(c_z^{\mathbb{Q}} | \kappa_v^{\mathbb{Q}}, \theta_v^{\mathbb{Q}}, v, vs) &: \text{Metropolis}, \end{aligned}$$

where  $\delta n_f = \{\delta n_{f,t_i}\}_{i=1}^n$ ,  $v = \{v_{t_i}\}_{i=1}^n$ , and  $z = \{z_{t_i}\}_{i=1}^n$ .

(b) correlation parameter:

$$p(\rho_{f,v} | \mu_f^{\mathbb{P}}, \bar{\eta}_f, \kappa_v^{\mathbb{P}}, \theta_v^{\mathbb{P}}, \sigma_v, n_f, v, y, z) : \text{Metropolis.}$$

(c) model parameters of the drift component:

$$\begin{aligned} p(\mu_f^{\mathbb{P}} | \bar{\eta}_f, \kappa_v^{\mathbb{P}}, \theta_v^{\mathbb{P}}, \sigma_v, \rho_{f,v}, n_f, v, y, z) &\sim \mathcal{N}, \\ p(\bar{\eta}_f | \mu_f^{\mathbb{P}}, \kappa_v^{\mathbb{P}}, \theta_v^{\mathbb{P}}, \sigma_v, \rho_{f,v}, n_f, v, y, z) &\sim \mathcal{N}. \end{aligned}$$

(c) model parameters of the variance process:

$$\begin{aligned} p(\kappa_v^{\mathbb{P}}, \theta_v^{\mathbb{P}} | \mu_f^{\mathbb{P}}, \bar{\eta}_f, \sigma_v, \rho_{f,v}, n_f, v, y, z) &\sim \mathcal{N}, \\ p(\sigma_v^2 | \kappa_v^{\mathbb{P}}, \theta_v^{\mathbb{P}}, v) &\sim \mathcal{IG}, \\ p(\kappa_v^{\mathbb{Q}}, \theta_v^{\mathbb{Q}} | \kappa_v^{\mathbb{P}}, \theta_v^{\mathbb{P}}, c_z^{\mathbb{Q}}, v, vs) &: \text{Metropolis.} \end{aligned}$$

model parameter	mean ( $\mu$ )	variance ( $\sigma^2$ )	shape ( $\alpha$ )	scale ( $\beta$ )	p	q	distribution type
$\mu_f^{\mathbb{P}}$	0	1	-	-	-	-	$\mathcal{N}(\mu, \sigma^2)$
$\bar{\eta}_f$	-0.5	1	-	-	-	-	$\mathcal{N}(\mu, \sigma^2)$
$\lambda_z$	-	-	-	-	2	40	$\mathcal{B}(p, q)$
$\mu_z^{\mathbb{P}}$	0	1	-	-	-	-	$\mathcal{N}(\mu, \sigma^2)$
$(\sigma_z^2)^{\mathbb{P}}$	-	-	4	0.03	-	-	$\mathcal{IG}(\alpha, \beta)$
$(\alpha, \beta)$	$\vec{0}$	$\vec{I}$	-	-	-	-	$\mathcal{N}(\mu, \sigma^2)$
$\sigma_v^2$	-	-	4	0.0001	-	-	$\mathcal{IG}(\alpha, \beta)$
$\sigma_f^2$	-	-	4	0.001	-	-	$\mathcal{IG}(\alpha, \beta)$
$\rho_{f,v}$	-	-	-	-	-	-	$\mathcal{U}[-1, 1]$

**Table A.1: prior distributions for the GB, JD, SV, and SVJ models**

This table gives the concrete prior distributions (daily decimals) for all model parameters in the GB, JD, SV, and SVJ models. The model parameters  $\alpha$  and  $\beta$  correspond to the drift component of the variance process after a re-parameterization. In the above table,  $\mathcal{N}$  refers a normal distribution,  $\mathcal{B}$  refers to a beta distribution,  $\mathcal{IG}$  refers to an inverse gamma distribution, and  $\mathcal{U}[-1, 1]$  refers to a uniform distribution on the interval  $[-1, 1]$ . Further,  $\vec{0}$  corresponds to a vector of zeros ( $2 \times 1$ ), and  $\vec{I}$  is the identity matrix ( $2 \times 2$ ).

Second, we describe the individual conditional posterior distributions in greater detail.

- *Drift Parameters:*  $\mu_f^{\mathbb{P}}$  and  $\bar{\eta}_f$

We apply the Gibbs Sampling approach to update the price jump compensator  $\mu_f^{\mathbb{P}} = -\lambda_z \bar{\mu}_z^{\mathbb{Q}}$  based on a conjugated normal prior distribution with mean  $a_0$  and variance  $b_0$ . In the SVJ model, the posterior distribution is given by

$$\mu_f^{\mathbb{P}} \sim \mathcal{N}(a, b)$$

with mean and variance equal to

$$\begin{aligned} a &= b \left( \sum_{i=1}^n \left( \frac{\tilde{\varepsilon}_{f,t_i} - \frac{\rho_{f,v}}{\sigma_v} \tilde{\varepsilon}_{v,t_i}}{(1 - \rho_{f,v}^2) v_{t_i}} \right) + \frac{a_0}{b_0} \right), \\ b &= \left( \sum_{i=1}^n \left( \frac{1}{(1 - \rho_{f,v}^2) v_{t_i}} \right) + \frac{1}{b_0} \right)^{-1}, \end{aligned}$$

where

$$\begin{aligned} \tilde{\varepsilon}_{f,t_i} &= y_{t_i} - \bar{\eta}_f v_{t_i} - z_{t_i} \delta n_{f,t_i}, \\ \tilde{\varepsilon}_{v,t_i} &= v_{t_{i+1}} - \kappa_v^{\mathbb{P}} \theta_v^{\mathbb{P}} - (1 - \kappa_v^{\mathbb{P}}) v_{t_i}. \end{aligned}$$

Similarly, the market price of diffusion risk  $\bar{\eta}_f = \eta_f + 0.5$  is updated based on a conjugated normal prior distribution with mean  $\bar{a}_0$  and variance  $\bar{b}_0$ . This leads to a conditional posterior distribution given by

$$\bar{\eta}_f \sim \mathcal{N}(\bar{a}, \bar{b})$$

with mean and variance equal to

$$\begin{aligned} \bar{a} &= \bar{b} \left( \sum_{i=1}^n \left( \frac{\tilde{\varepsilon}_{f,t_i} - \frac{\rho_{f,v}}{\sigma_v} \tilde{\varepsilon}_{v,t_i}}{1 - \rho_{f,v}^2} \right) + \frac{\bar{a}_0}{\bar{b}_0} \right), \\ \bar{b} &= \left( \sum_{i=1}^n \left( \frac{v_{t_i}}{1 - \rho_{f,v}^2} \right) + \frac{1}{\bar{b}_0} \right)^{-1}, \end{aligned}$$

where

$$\begin{aligned} \tilde{\varepsilon}_{f,t_i} &= y_{t_i} - \mu_f^{\mathbb{P}} - z_{t_i} \delta n_{f,t_i}, \\ \tilde{\varepsilon}_{v,t_i} &= v_{t_{i+1}} - \kappa_v^{\mathbb{P}} \theta_v^{\mathbb{P}} - (1 - \kappa_v^{\mathbb{P}}) v_{t_i}. \end{aligned}$$

In the JD model, the drift component is equal to  $\widehat{\mu}_f^{\mathbb{P}} = -\lambda_z \overline{\mu}_z^{\mathbb{Q}} + \overline{\eta}_f \sigma_f^2$ . It follows then that the posterior distribution is given by

$$\widehat{\mu}_f^{\mathbb{P}} \sim \mathcal{N}(\tilde{a}, \tilde{b})$$

with

$$\begin{aligned} \tilde{a} &= \tilde{b} \left( \sum_{i=1}^n \frac{(y_{t_i} - z_{t_i} \delta n_{f,t_i})}{\sigma_f^2} + \frac{\tilde{a}_0}{\tilde{b}_0} \right), \\ \tilde{b} &= \left( \frac{n}{\sigma_f^2} + \frac{1}{\tilde{b}_0} \right)^{-1} \end{aligned}$$

if the prior distribution of  $\mu_f^{\mathbb{P}}$  is normally distributed with mean  $\tilde{a}_0$  and variance  $\tilde{b}_0$ .

- *Jump Sizes and Jump Times:  $\delta n_f$  and  $z$*

We use the Gibbs Sampler to draw jump times and jump sizes conditional on the remaining model parameters (see Eraker, Johannes, and Polson (2003)). The posterior distribution of each jump size  $z_{t_i}$  depends on the mean jump size, jump size variance, drift parameters, and the variance state at  $t_i$  for  $i = 1, \dots, n$ . It is given by

$$z_{t_i} \sim \mathcal{N}(c_{t_i}, d_{t_i})$$

with

$$\begin{aligned} c_{t_i} &= d_{t_i} \left( \frac{\delta n_{f,t_i}}{v_{t_i}} (y_{t_i} - (\mu_f^{\mathbb{P}} + \overline{\eta}_f v_{t_i})) + \frac{\mu_z^{\mathbb{P}}}{(\sigma_z^{\mathbb{P}})^2} \right), \\ d_{t_i} &= \left( \frac{\delta n_{f,t_i}}{v_{t_i}} + \frac{1}{(\sigma_z^{\mathbb{P}})^2} \right)^{-1}. \end{aligned}$$

If we condition on the jump size distribution and the overall jump intensity  $\lambda_z$ , we can sample jump events from a Bernoulli distribution as follows:

$$\delta n_{f,t_i} \sim \text{Ber}(\tilde{\lambda}_{t_i})$$

with

$$\begin{aligned}\tilde{\lambda}_{t_i} &= \frac{p(\delta n_{f,t_i} = 1|u_1)}{p(\delta n_{f,t_i} = 1|u_1) + p(\delta n_{f,t_i} = 0|u_1)}, \\ p(\delta n_{f,t_i} = 1|u_1) &\propto \lambda_z \exp\left(\frac{(y_{t_i} - (\mu_f^{\mathbb{P}} + \bar{\eta}_f v_{t_i}) - z_{t_i})^2}{2v_{t_i}}\right), \\ p(\delta n_{f,t_i} = 0|u_1) &\propto (1 - \lambda_z) \exp\left(\frac{(y_{t_i} - (\mu_f^{\mathbb{P}} + \bar{\eta}_f v_{t_i}))^2}{2v_{t_i}}\right),\end{aligned}$$

where  $u_1 = \{\mu_f^{\mathbb{P}}, \bar{\eta}_f, \lambda_z, v_{t_i}, y_{t_i}, z_{t_i}\}$ . We identify jump times and jump sizes based on independent futures price and variance innovations ( $\rho_{f,v} = 0$ ) in order to reduce the impact of noise introduced by the latent variance states on our filtering results.

- *Jump Intensity:  $\lambda_z$*

In our discrete setup, the total number of jumps is beta distributed

$$\lambda_z \sim \mathcal{B}(e, f)$$

with

$$\begin{aligned}e &= e_0 + \sum_{i=1}^n \delta n_{f,t_i}, \\ f &= f_0 + n - \sum_{i=1}^n \delta n_{f,t_i}\end{aligned}$$

if the prior distribution of  $\lambda_z$  is beta distributed with parameters  $(e_0, f_0)$ .

- *Jump Size Parameters:  $\mu_z^{\mathbb{P}}$  and  $(\sigma_z^{\mathbb{P}})^2$*

The log jump sizes are normally distributed so that the mean and the variance parameter  $\mu_z^{\mathbb{P}}$  and  $(\sigma_z^{\mathbb{P}})^2$  can be standardly updated. We assume that the prior distribution of the mean parameter is normally distributed with mean  $g_0 = 0$  and variance  $h_0 = 1$ . It follows then that the posterior distribution is given by

$$\mu_z^{\mathbb{P}} \sim \mathcal{N}(g, h)$$

with

$$\begin{aligned}g &= h \left( \frac{\sum_{i=1}^n z_{t_i}}{(\sigma_z^{\mathbb{P}})^2} + \frac{g_0}{h_0} \right), \\ h &= \left( \frac{n}{(\sigma_z^{\mathbb{P}})^2} + \frac{1}{h_0} \right)^{-1}.\end{aligned}$$

We use again an inverse gamma distributed prior for  $(\sigma_z^{\mathbb{P}})^2$  with parameters  $l_0$  and  $m_0$ . This results in a conditional posterior distribution given by

$$(\sigma_z^{\mathbb{P}})^2 \sim \mathcal{IG}(l, m)$$

with

$$\begin{aligned} l &= l_0 + \frac{n}{2}, \\ m &= m_0 + \sum_{i=1}^n \frac{(z_{t_i} - \mu_z^{\mathbb{P}})^2}{2}. \end{aligned}$$

- *Variance States:  $v$*

The conditional posterior distribution for each variance state  $v_{t_i}$  can be expressed throughout two analytically tractable components (see Section 2.3 for more details)

$$p(v_{t_i} | u_{-v_{t_i}}, v_s, y) \propto p(v_{t_i}, y | u_2) p(\widehat{v}_{s_{t_i, \tau_i}} | u_3),$$

where  $u_2 = \{\mu_f^{\mathbb{P}}, \bar{\eta}_f, \kappa_v^{\mathbb{P}}, \theta_v^{\mathbb{P}}, \sigma_v, \rho_{f,v}, \delta n_f, v_{-v_{t_i}}, z\}$ ,  $u_3 = \{\kappa_v^{\mathbb{Q}}, \theta_v^{\mathbb{Q}}, c_z^{\mathbb{Q}}, v_{t_i}\}$  and  $v_{-v_{t_i}}$  corresponds to the vector of latent variance states excluding  $v_{t_i}$ .

The first component  $p(v_{t_i}, y | u_2)$  captures the dependence between preceding and succeeding variance states and the preceding and current futures price log-returns. It is given by

$$p(v_{t_i}, y | u_2) \propto v_{t_i}^{-1} \exp(-\omega_1) \exp(-(\omega_2 + \omega_3)) \quad (\text{A.1})$$

with

$$\begin{aligned} \omega_1 &= \frac{(y_{t_i} - (\mu_f^{\mathbb{P}} + \bar{\eta}_f v_{t_i}) - z_{t_i} \delta n_{f,t_i})^2}{2v_{t_i}}, \\ \omega_2 &= \frac{(v_{t_i} - (v_{t_{i-1}} + \kappa_v^{\mathbb{P}}(\theta_v^{\mathbb{P}} - v_{t_{i-1}})) - \rho_{f,v} \sigma_v (y_{t_{i-1}} - (\mu_f^{\mathbb{P}} + \bar{\eta}_f v_{t_{i-1}}) - z_{t_{i-1}} \delta n_{f,t_{i-1}}))^2}{2(1 - \rho_{f,v}^2) \sigma_v^2 v_{t_{i-1}}}, \\ \omega_3 &= \frac{(v_{t_{i+1}} - (v_{t_i} + \kappa_v^{\mathbb{P}}(\theta_v^{\mathbb{P}} - v_{t_i})) - \rho_{f,v} \sigma_v (y_{t_i} - (\mu_f^{\mathbb{P}} + \bar{\eta}_f v_{t_i}) - z_{t_i} \delta n_{f,t_i}))^2}{2(1 - \rho_{f,v}^2) \sigma_v^2 v_{t_i}}. \end{aligned}$$

The second component  $p(\widehat{v}_{s_{t_i, \tau_i}} | u_3)$  incorporates the additional option market information through non-adjusted variance swap rates into our filtering approach.<sup>1</sup>

---

<sup>1</sup>If no variance swap rate is available, we set  $p(\widehat{v}_{s_{t_i, \tau_i}} | u_3)$  equal to one.

It is given by

$$p(\widehat{v}_{S_{t_i, \tau_i}} | u_{-v_{t_i}}, v_{t_i}) = \frac{1}{\sqrt{2\pi}\sigma_e} \exp\left(-\frac{(\widehat{v}_{S_{t_i, \tau_i}} - \widehat{v}_{S_{t_i, \tau_i}}^{mod})^2}{2\sigma_e^2}\right), \quad (\text{A.2})$$

where

$$\widehat{v}_{S_{t_i, \tau_i}}^{mod} = \theta_v^{\mathbb{Q}} + \frac{1 - e^{-\kappa_v^{\mathbb{Q}}(\tau_i - t_i)}}{\kappa_v^{\mathbb{Q}}(\tau_i - t_i)}(v_{t_i} - \theta_v^{\mathbb{Q}}) + c_z^{\mathbb{Q}}. \quad (\text{A.3})$$

It is not possible to sample directly from such a non-standard distribution function. Therefore, we use the random walk Metropolis-Hastings algorithm with a proposal density given by

$$\tilde{p}(v_{t_i}^{(g-1)}, v_{t_i}^{(g)}) \sim \mathcal{N}(v_{t_i}^{(g-1)}, (0.2(\sigma_v)^{(g-1)})^2),$$

where  $g$  is the current simulation run.

- *Variance Parameters:*  $\kappa_v^{\mathbb{P}}$ ,  $\theta_v^{\mathbb{P}}$ ,  $\sigma_v$ , and  $\sigma_f^2$

The drift parameters  $\kappa_v^{\mathbb{P}}$  and  $\theta_v^{\mathbb{P}}$  of the volatility process are updated according to Asgharian and Bengtsson (2006). The model parameters are re-parameterized as  $\alpha = \kappa_v^{\mathbb{P}}\theta_v^{\mathbb{P}}$  as well as  $\beta = 1 - \kappa_v^{\mathbb{P}}$ . We use a multivariate normal prior distribution with a zero mean vector  $\vec{0}$  and covariance matrix given by the identity matrix  $\vec{I}$  of dimension  $(2 \times 2)$ . The posterior distribution of  $\alpha$  and  $\beta$  is then given by

$$(\alpha, \beta) \sim \mathcal{N}(u, W)$$

with mean and covariance given by

$$u = W \left( \vec{I}^{-1} \vec{0} + \frac{1}{(1 - \rho_{f,v}^2)\sigma_v^2} Z^T z \right),$$

$$W = \left( \vec{I}^{-1} + \frac{1}{(1 - \rho_{f,v}^2)\sigma_v^2} Z^T Z \right)^{-1},$$

where

$$z = \begin{pmatrix} \frac{v_{t_2} - \rho_{f,v}\sigma_v\tilde{\varepsilon}_{f,t_1}}{\sqrt{v_{t_1}}} \\ \frac{v_{t_3} - \rho_{f,v}\sigma_v\tilde{\varepsilon}_{f,t_2}}{\sqrt{v_{t_2}}} \\ \vdots \\ \frac{v_{t_{n+1}} - \rho_{f,v}\sigma_v\tilde{\varepsilon}_{f,t_{n-1}}}{\sqrt{v_{t_n}}} \end{pmatrix}, \quad Z = \begin{pmatrix} (\sqrt{v_{t_1}})^{-1} & \sqrt{v_{t_1}} \\ (\sqrt{v_{t_2}})^{-1} & \sqrt{v_{t_2}} \\ \vdots & \vdots \\ (\sqrt{v_{t_n}})^{-1} & \sqrt{v_{t_n}} \end{pmatrix},$$

$$\tilde{\varepsilon}_{f,t_i} = y_{t_i} - (\mu_f^{\mathbb{P}} + \bar{\eta}_f v_{t_i}) - z_{t_i} \delta n_{f,t_i}.$$

The parameter  $\sigma_v^2$  is drawn from a conjugated inverse gamma distribution with parameters  $\alpha_0$  and  $\beta_0$ :

$$\sigma_v^2 \sim \mathcal{IG}(\alpha, \beta)$$

with

$$\begin{aligned} \alpha &= \alpha_0 + \frac{n}{2}, \\ \beta &= \beta_0 + \sum_{i=2}^n \frac{(v_{t_i} - \kappa_v^{\mathbb{P}} \theta_v^{\mathbb{P}} - (1 - \kappa_v^{\mathbb{P}}) v_{t_{i-1}})^2}{2v_{t_{i-1}}}, \end{aligned}$$

where we set  $\rho_{f,v} = 0$  as done in Eraker, Johannes, and Polson (2003).

In the JD model, the variance parameter  $\sigma_f^2$  is constant. Its prior distribution is assumed to be inverse gamma distributed with parameters  $\gamma_0$  and  $\chi_0$ . It follows then that the conditional posterior distribution is given by

$$\sigma_f^2 \sim \mathcal{IG}(\gamma, \chi)$$

with

$$\begin{aligned} \gamma &= \frac{n}{2} + \gamma_0, \\ \chi &= \frac{1}{2} \sum_{i=1}^n (y_{t_i} - \mu_f^{\mathbb{P}} - z_{t_i} \delta n_{f,t_i})^2 + \chi_0. \end{aligned}$$

- *Correlation Coefficient:  $\rho_{f,v}$*

The correlation coefficient  $\rho_{f,v}$  depends on the co-movements of the futures price and variance process. We obtain the following posterior distribution for a uniform prior distribution  $\mathcal{U}[-1, 1]$ :

$$\begin{aligned} p(\rho_{f,v} | u_4, n_f, v, y, z) &\propto p(v, y | \rho_{f,v}, u_3, n_f, z) p(\rho_{f,v}) \\ &\propto p(v, y | \rho_{f,v}, u_3, n_f, z) \mathbf{1}_{\{-1 \leq \rho_{f,v} \leq 1\}}. \end{aligned}$$

The function  $p(v, y | \rho_{f,v}, u_3, n_f, z)$  corresponds to the joint likelihood function of variance states and futures price returns. It is given by

$$p(v, y | \rho_{f,v}, u_4) \propto \left( \frac{1}{\sqrt{1 - \rho_{f,v}^2}} \right)^n \exp \left( - \sum_{i=1}^n \frac{\left( \tilde{\varepsilon}_{f,t_i} - \frac{\rho_{f,v}}{\sigma_v} \tilde{\varepsilon}_{v,t_i} \right)^2}{2(1 - \rho_{f,v}^2)v_{t_i}} \right)$$

with

$$\begin{aligned} \tilde{\varepsilon}_{f,t_i} &= y_{t_i} - (\mu_f^{\mathbb{P}} + \bar{\eta}_f v_{t_i}) - z_{t_i} \delta n_{f,t_i}, \\ \tilde{\varepsilon}_{v,t_i} &= v_{t_{i+1}} - \kappa_v^{\mathbb{P}} \theta_v^{\mathbb{P}} - (1 - \kappa_v^{\mathbb{P}}) v_{t_i}. \end{aligned}$$

In the above formulas,  $u_4 = \{\mu_f^{\mathbb{P}}, \bar{\eta}_f, \kappa_v^{\mathbb{P}}, \theta_v^{\mathbb{P}}, \sigma_v, n_f, v, y, z\}$  denotes the vector of relevant model parameters besides  $\rho_{f,v}$ . We sample from the posterior distribution by a random walk Metropolis-Hastings algorithm as in Asgharian and Bengtsson (2006). The proposal density is t-distributed with 6.5 degrees of freedom and standard deviation 0.015.

- *Risk-Neutral Parameters:*  $\kappa_v^{\mathbb{Q}}, \theta_v^{\mathbb{Q}}, \mu_z^{\mathbb{Q}},$  and  $\sigma_z^{\mathbb{Q}}$

In order to use variance swap rates  $vs_{t_i, \tau_i}$  in the filtering approach, we have to estimate the risk neutral parameters  $\kappa_v^{\mathbb{Q}}, \theta_v^{\mathbb{Q}}, \mu_z^{\mathbb{Q}},$  and  $\sigma_z^{\mathbb{Q}}$  simultaneously. We assume that variance swap rates  $vs_{t_i, \tau_i}$  are observed with independent normally distributed error terms with mean 0 and variance  $\sigma_e^2$  and choose an improper prior distribution for  $\kappa_v^{\mathbb{Q}}$ . This allows us to compute the posterior distribution for the risk neutral parameters

$$\begin{aligned} p(\kappa_v^{\mathbb{Q}} | u_5, v, vs) &\propto \prod_{i=1}^n p(vs_{t_i, \tau_i} | u_5, \kappa_v^{\mathbb{Q}}, v_{t_i}) \\ &\propto \exp \left( - \frac{\sum_{i=1}^n \left( \widehat{vs}_{t_i, \tau_i} - \frac{\kappa_v^{\mathbb{P}} \theta_v^{\mathbb{P}}}{\kappa_v^{\mathbb{Q}}} - \frac{1 - e^{-\kappa_v^{\mathbb{Q}}(\tau_i - t_i)}}{\kappa_v^{\mathbb{Q}}(\tau_i - t_i)} \left( v_{t_i} - \frac{\kappa_v^{\mathbb{P}} \theta_v^{\mathbb{P}}}{\kappa_v^{\mathbb{Q}}} \right) - c_z^{\mathbb{Q}} \right)^2}{2\sigma_e^2} \right), \end{aligned} \quad (\text{A.4})$$

where  $\widehat{vs}_{t_i, \tau_i}$  is the non-adjusted variance swap rate and  $u_5 = \{\kappa_v^{\mathbb{P}}, \theta_v^{\mathbb{P}}, c_z^{\mathbb{Q}}\}$ . In (A.4), we have used the fact that the product from the drift parameters of the volatility dynamics has to be equal under both measures  $\kappa_v^{\mathbb{Q}} \theta_v^{\mathbb{Q}} = \kappa_v^{\mathbb{P}} \theta_v^{\mathbb{P}}$ . We use again a random walk Metropolis-Hastings algorithm to sample from such an intractable conditional posterior distribution. The proposal density for the Metropolis-Hastings step is

centered around the parameter value of the last iteration step  $(\kappa_v^{\mathbb{Q}})^{g-1}$ :

$$\tilde{p}((\kappa_v^{\mathbb{Q}})^{(g-1)}, (\kappa_v^{\mathbb{Q}})^{(g)}) \sim \mathcal{N}((\kappa_v^{\mathbb{Q}})^{(g-1)}, (0.2(\kappa_v^{\mathbb{Q}})^{(g-1)})^2)$$

The simulated value of  $\theta_v^{\mathbb{Q}}$  can then be derived from  $\theta_v^{\mathbb{Q}} = \frac{\kappa_v^{\mathbb{P}} \theta_v^{\mathbb{P}}}{\kappa_v^{\mathbb{Q}}}$ . The risk neutral jump parameters  $\mu_z^{\mathbb{Q}}$  and  $\sigma_z^{\mathbb{Q}}$  can only be updated throughout the aggregated variance jump compensator  $c_z^{\mathbb{Q}}$ . The posterior distribution of  $c_z^{\mathbb{Q}}$  is equal to

$$\begin{aligned} p(c_z^{\mathbb{Q}} | \kappa_v^{\mathbb{Q}}, \theta_v^{\mathbb{Q}}, v, vs) &\propto \prod_{i=1}^n p(\widehat{v}_{s_{t_i, \tau_i}} | \kappa_v^{\mathbb{Q}}, \theta_v^{\mathbb{Q}}, c_z^{\mathbb{Q}}, v_{t_i}) \\ &\propto \exp \left( - \frac{\sum_{i=1}^n \left( \widehat{v}_{s_{t_i, \tau_i}} - \theta_v^{\mathbb{Q}} - \frac{1 - e^{-\kappa_v^{\mathbb{Q}}(\tau_i - t_i)}}{\kappa_v^{\mathbb{Q}}(\tau_i - t_i)} (v_{t_i} - \theta_v^{\mathbb{Q}}) - c_z^{\mathbb{Q}} \right)^2}{2\sigma_e^2} \right), \end{aligned}$$

an improper prior distribution is assumed for  $c_z^{\mathbb{Q}}$ . We apply a random walk Metropolis-Hastings algorithm to generate samples for  $c_z^{\mathbb{Q}}$ . The proposal density for the Metropolis-Hastings step is centered around the value of the last iteration step  $(c_z^{\mathbb{Q}})^{(g-1)}$ :

$$\tilde{p}((c_z^{\mathbb{Q}})^{(g-1)}, (c_z^{\mathbb{Q}})^{(g)}) \sim \mathcal{N}((c_z^{\mathbb{Q}})^{(g-1)}, (0.2(c_z^{\mathbb{Q}})^{(g-1)})^2).$$

### A.3 Smooth Futures Price Curve

Benth, Koekebakker, and Ollmar (2007) show that polynomial splines of order four are required in order to satisfy both no-arbitrage conditions and the maximum smoothness criterion

$$f_t(u) = \begin{cases} \tilde{a}_0 + \tilde{b}_0 u + \tilde{c}_0 u^2 + \tilde{d}_0 u^3 + \tilde{e}_0 u^4, & u \in [\tau_0, \tau_1] \\ \tilde{a}_1 + \tilde{b}_1 u + \tilde{c}_1 u^2 + \tilde{d}_1 u^3 + \tilde{e}_1 u^4, & u \in [\tau_1, \tau_2] \\ \vdots & \vdots \\ \tilde{a}_m + \tilde{b}_m u + \tilde{c}_m u^2 + \tilde{d}_m u^3 + \tilde{e}_m u^4, & u \in [\tau_m, \tau_{m+1}] \end{cases}.$$

The spline parameters are uniquely determined by the following linear equation system:

$$\begin{pmatrix} 2H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} y \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}. \quad (\text{A.5})$$

In (A.5), the matrix  $A$  and the vector  $b$  are defined by

(i) the static no-arbitrage condition

$$f_t^{(i)} = \int_{\tau_i}^{\tau_{i+1}} \hat{w}(u; \tau_i, \tau_{i+1}) f_t(u) du, \quad i = 0, \dots, m,$$

(ii) the slope of the interpolation function at the last delivery date  $\tau_{m+1}$

$$\frac{\partial f_t}{\partial u}(\tau_{m+1}) = 0$$

(iii) the usual spline conditions

$$\begin{aligned} 0 &= \Delta_i \tilde{a} \tau_{i+1}^4 + \Delta_i \tilde{b} \tau_{i+1}^3 + \Delta_i \tilde{c} \tau_{i+1}^2 + \Delta_i \tilde{d} \tau_{i+1} + \Delta_i \tilde{e}, \\ 0 &= 4\Delta_i \tilde{a} \tau_{i+1}^3 + 3\Delta_i \tilde{b} \tau_{i+1}^2 + 2\Delta_i \tilde{c} \tau_{i+1} + \Delta_i \tilde{d}, \\ 0 &= 12\Delta_i \tilde{a} \tau_{i+1}^2 + 6\Delta_i \tilde{b} \tau_{i+1} + 2\Delta_i \tilde{c}, \end{aligned}$$

where  $\Delta_i \tilde{a} = \tilde{a}_{i+1} - \tilde{a}_i$ ,  $\Delta_i \tilde{b} = \tilde{b}_{i+1} - \tilde{b}_i$ ,  $\Delta_i \tilde{c} = \tilde{c}_{i+1} - \tilde{c}_i$ ,  $\Delta_i \tilde{d} = \tilde{d}_{i+1} - \tilde{d}_i$ , and  $\Delta_i \tilde{e} = \tilde{e}_{i+1} - \tilde{e}_i$  for  $i = 0, \dots, m-1$ .

The matrix  $H$  is derived from the maximum smoothness condition

$$\min_y y^T H y,$$

where

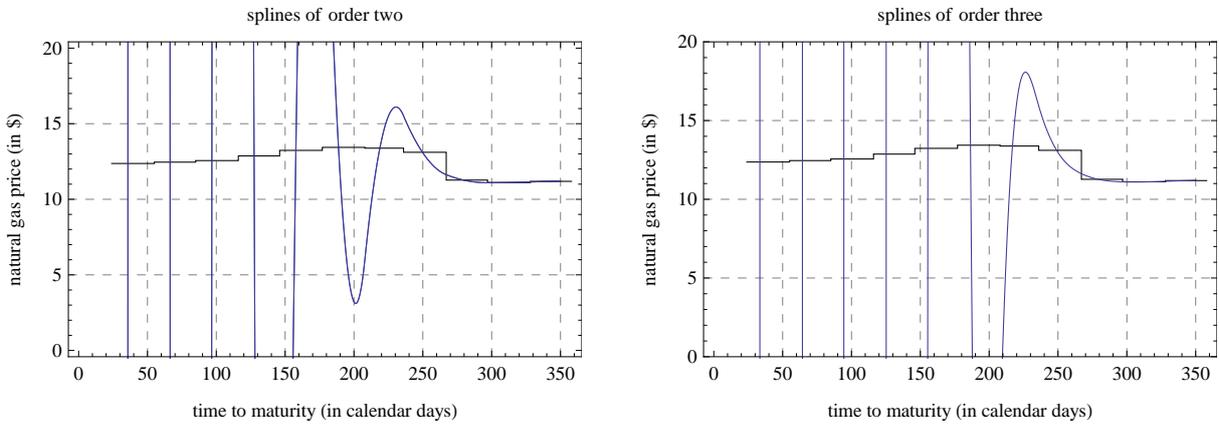
$$H = \begin{pmatrix} h_1 & & 0 \\ & \ddots & \\ 0 & & h_{m+1} \end{pmatrix}, h_j = \begin{pmatrix} \frac{144}{5}\Delta_j^5 & 18\Delta_j^4 & 8\Delta_j^3 & 0 & 0 \\ 18\Delta_j^4 & 12\Delta_j^3 & 6\Delta_j^2 & 0 & 0 \\ 8\Delta_j^3 & 6\Delta_j^2 & 4\Delta_j & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$y = (\tilde{a}_0, \tilde{b}_0, \tilde{c}_0, \tilde{d}_0, \tilde{e}_0, \dots, \tilde{a}_m, \tilde{b}_m, \tilde{c}_m, \tilde{d}_m, \tilde{e}_m), \text{ and } \Delta_j^k = \tau_{j+1}^k - \tau_j^k,$$

for  $k = 1, \dots, 5$  and  $j = 0, \dots, m$ .

The linear equation system (A.5) implies that the futures price curve linearly depends on real futures prices.

The role of the maximum smoothness criterion to avoid extreme inter-or extrapolation values can be exemplarily illustrated for quadratic and cubic splines without a maximum smoothness condition (see Figure A.1).



**Figure A.1: splines without the maximum smoothness criterion**

The graphs show two interpolation functions without the maximum smoothness criterion (quadratic (left) and cubic splines (right)) for real futures price on December 28, 2005.

## A.4 Normal and Log-Normal Market Models

The best way to show the link of our modeling approach to standard models is to consider a one-factor normal and log-normal price dynamics for real futures contracts:

$$df_t^{(i)} = \sigma^{(i)}(f_t^{(i)})^\gamma dw_{f,t}^{\mathbb{Q}} \quad \text{for } i = 0, \dots, m, \quad (\text{A.6})$$

where  $w_{f,t}^{\mathbb{Q}}$  is a standard Wiener process. It follows then that

$$df_t(u) = \sum_{i=0}^m \beta_i(u) df_t^{(i)} = \sum_{i=0}^m \beta_i(u) \sigma^{(i)}(f_t^{(i)})^\gamma dw_{f,t}^{\mathbb{Q}}.$$

First, if real futures price returns are normally distributed, i.e.,  $\gamma = 0$ , theoretical futures prices are also normally distributed

$$df_t(u) = \sum_{i=0}^m \beta_i(u) \sigma^{(i)} dw_{f,t}^{\mathbb{Q}}.$$

Second, if real futures price are log-normally distributed, i.e.,  $\gamma = 1$ , theoretical futures price returns are log-normally distributed if and only if the volatility curve is flat ( $\sigma^{(i)} = \sigma(t)$ ):

$$\begin{aligned} df_t(u) &= \sum_{i=0}^m \beta_i(u) \sigma^{(i)} f_t^{(i)} dw_{f,t}^{\mathbb{Q}} \\ &= \sigma(t) \sum_{i=0}^m \beta_i(u) f_t^{(i)} dw_{f,t}^{\mathbb{Q}} \\ &= \sigma(t) f_t(u) dw_{f,t}^{\mathbb{Q}}. \end{aligned}$$

## A.5 MCMC Algorithm

In this section, we describe our estimation procedure for both market models. Namely, we present the Markov chain Monte Carlo algorithm for a two-factor regime switching model with jumps that contains both market model specifications.

First, we discretize the log futures price dynamics as follows:

$$\begin{aligned}
y_{t_k}^{(i)} &= \log f_{t_{k+1}}^{(i)} - \log f_{t_k}^{(i)} \\
&= -\lambda_z \bar{\mu}_z^{(i)} + \left( \sum_{j=1}^2 (\eta_f^{(j)}(x_{t_k}) - 0.5) (\sigma_{t_k}^{(i,j)}(x_{t_k}))^2 \right) + \sum_{j=1}^2 \sigma_{t_k}^{(i,j)}(x_{t_k}) \varepsilon_{f,t_k}^{(j,\mathbb{P})} + \gamma_{t_k}^{(i)} z_{t_k} \delta n_{f,t_k} \\
&= \tilde{\mu}_{t_k}^{(i)}(x_{t_k}) + \tilde{\sigma}_{t_k}^{(i)}(x_{t_k}) \varepsilon_{f,t_k}^{\mathbb{P}} + \tilde{\gamma}_{t_k}^{(i)} z_{t_k} \delta n_{f,t_k},
\end{aligned}$$

where

$$\begin{aligned}
\tilde{\mu}_{t_k}^{(i)}(x_{t_k}) &= -\lambda_z \bar{\mu}_z^{(i)} + \sum_{j=1}^2 (\eta_f^{(j)}(x_{t_k}) - 0.5) (\sigma_{t_k}^{(i,j)}(x_{t_k}))^2, \\
\tilde{\sigma}_{t_k}^{(i)}(x_{t_k}) &= (\sigma_{t_k}^{(i,1)}(x_{t_k})^2 + \sigma_{t_k}^{(i,2)}(x_{t_k})^2)^{0.5}, \\
\tilde{\gamma}_{t_k}^{(i)} &= \gamma_{t_k}^{(i)}
\end{aligned}$$

for  $k = 1, \dots, n$ . In the above formulas,  $\varepsilon_{f,t_k}^{(j,\mathbb{P})}$  and  $\varepsilon_{f,t_k}^{\mathbb{P}}$  are normally distributed random variables with zero means and standard deviations of one,  $\delta n_{f,t_k}$  is a Bernoulli distributed random variable with parameter  $\lambda_z$  for  $j \in \{1, 2\}$ ,  $i = 1, \dots, 12$ , and  $k = 1, \dots, n$ .<sup>2</sup> The state process  $x_t$  is discretized as

$$x_{t_{k+1}} = x_{t_k} + (2 - x_{t_k}) \delta n_{x,t_k}^{(1,2)} + (1 - x_{t_k}) \delta n_{x,t_k}^{(2,1)},$$

where  $\delta n_{x,t_k}^{(1,2)}$  and  $\delta n_{x,t_k}^{(2,1)}$  are again Bernoulli distributed random variables with parameters  $q_{1,2}$  and  $q_{2,1}$ .

Next, we introduce the prior distributions and the updating steps for our MCMC algorithm.

- *Drift Parameters (Market Price of Diffusion Risk):  $\eta_f$*

The drift components are updated based on log-returns of the one-month-ahead and twelve-month-ahead futures contracts in both regimes. In both cases, we use a conjugated normal prior distribution with mean  $a_0 = 0$  and variance  $b_0 = 1$ . The

---

<sup>2</sup>The time distance between two observations  $\delta t$  is equal to one business day ( $\delta t = 1$ ). The time-discretization results in a discretization error. Johannes, Kumar, and Polson (1999) document that the effect of time-discretization in the Poisson arrivals is minimal.

posterior distribution is then a normal distribution with mean  $a^{(rg)}$  and variance  $b^{(rg)}$ :

$$\tilde{\mu}(rg) \sim \mathcal{N}(a^{(rg)}, b^{(rg)}),$$

where

$$a^{(rg)} = \frac{\sum_{k=1}^{n^{(rg)}} \mathbf{1}_{\{x_{t_k}=rg\}} \frac{y_{t_k}^{(i)} - \tilde{\gamma}_{t_k}^{(i)} z_{t_k}}{\tilde{\sigma}_{t_k}^{(i)}(x_{t_k})}}{1 + n^{(rg)} b_0},$$

$$b^{(rg)} = \frac{1}{1 + n^{(rg)}}.$$

The superscript  $rg$  denotes the state of the regime switching process and  $n^{(rg)}$  is equal to the number of observations in the respective regime  $rg \in \{1, 2\}$ . The drift parameters are then converted to risk premia estimates by solving the linear equation system.

- *Transition Probabilities:*  $q_{1,2}$  and  $q_{2,1}$

We choose conjugated priors that are beta distributed with parameters  $c_0^{(1,2)} = 3.75$ ,  $d_0^{(1,2)} = 71.25$ ,  $c_0^{(2,1)} = 3.75$ , and  $d_0^{(2,1)} = 71.25$  for both transition probabilities. Then, the posterior distributions are beta distributed with parameters given by

$$p_{1,2} \sim \mathcal{B}(c^{(1,2)}, d^{(1,2)}),$$

$$p_{2,2} \sim \mathcal{B}(c^{(2,1)}, d^{(2,1)}),$$

where

$$c^{(1,2)} = c_0^{(1,2)} + n^{(1,2)}, \quad d^{(1,2)} = d_0^{(1,2)} + n^{(1)} - n^{(1,2)}, \quad (\text{A.7})$$

$$c^{(2,1)} = c_0^{(2,1)} + n^{(2,1)}, \quad d^{(2,1)} = d_0^{(2,1)} + n^{(2)} - n^{(2,1)}. \quad (\text{A.8})$$

In (A.7) and (A.8),  $n^{(1,2)}$  ( $n^{(2,1)}$ ) is equal to the number of switches from state one to state two (two to one).

- *Volatility Parameters:*  $\sigma^{(sh)}$ ,  $\sigma^{(lg)}$  and  $\kappa$

The volatility parameters are updated via the Gibbs Sampler based on log-returns referring to different maturities. First, the twelve-month-ahead futures contract is used to estimate the long-term volatility parameters  $\hat{\sigma}^{(lg)} = \{\sigma^{(lg)}, \sigma_1^{(lg)}, \sigma_2^{(lg)}\}$ . Second, the short-term volatility parameters  $\hat{\sigma}^{(sh)} = \{\sigma^{(sh)}, \sigma_1^{(sh)}, \sigma_2^{(sh)}\}$  are estimated via the difference between the log-returns of a short-dated (one-month-ahead to three-month-ahead) futures contract and long-dated (nine-month-ahead to twelve-

		$\sigma^{(lg)}$		$\sigma^{(sh)}$		$\kappa$	
		$e_0$	$f_0$	$e_0$	$f_0$	$e_0$	$f_0$
crude oil	state 1	10	0.00248	10	0.0003866	10	0.00727
	state 2	10	0.0111	10	0.0048572	10	0.00782
natural gas	state 1	10	0.0028871	10	0.0022859	10	0.00038

**Table A.2: prior distributions for the crude oil and natural gas market models**

The table specifies the inverse gamma prior distributions for the short-term and long-term volatility parameters and the slope parameter.

month-ahead) futures contract. Third, the slope parameters  $\hat{\kappa} = \{\kappa, \kappa_1, \kappa_2\}$  are estimated from the difference between the log-returns of a short-dated (one-month-ahead to three-month-ahead) futures contract and a medium-dated (three-month-ahead to nine-month-ahead) futures contract.

In all cases, we first separate the log-return data in both regimes, update the variance of the underlying time series, and then convert the sampled values in the respective model parameters. We use an inverse gamma prior distribution for all variance parameters, where the concrete prior parameters  $e_0^{(rg)}$  and  $f_0^{(rg)}$  can be found in Table A.2. The posterior distributions are inverse gamma distributed with parameters

$$(\tilde{\sigma}_{t_k}^{(i)}(x_{t_k}))^2 = \mathcal{IG}(e^{(rg)}, f^{(rg)}),$$

where

$$\begin{aligned} e^{(rg)} &= e_0^{(rg)} + 0.5n^{(rg)} \\ f^{(rg)} &= f_0^{(rg)} + 0.5 \sum_{k=1}^{n^{(rg)}} \hat{y}_{t_k}^2, \quad rg \in \{1, 2\}, \end{aligned}$$

The adjusted log-returns are given by

- (i) short-term volatility and slope parameters:

$$\hat{y}_{t_k} = (y_{t_k}^{(\tilde{m}t_1)} - (-\lambda_z \bar{\mu}_z^{(\tilde{m}t_1)} + \tilde{\gamma}_{t_k}^{(\tilde{m}t_1)} z_{t_k})) - (y_{t_k}^{(\tilde{m}t_2)} - (-\lambda_z \bar{\mu}_z^{(\tilde{m}t_2)} + \tilde{\gamma}_{t_k}^{(\tilde{m}t_2)} z_{t_k})),$$

- (ii) long-term volatility parameters:

$$y_{t_k}^{(\tilde{m}t_1)} - (-\lambda_z \bar{\mu}_z^{(\tilde{m}t_1)} + \tilde{\gamma}_{t_k}^{(\tilde{m}t_1)} z_{t_k}),$$

where  $\tilde{m}t_1$  and  $\tilde{m}t_2$  are the maturities considered in the update step.<sup>3</sup>

---

<sup>3</sup>The maturity effect only has a minor impact on the estimation results.

- *State Variables:  $x$*

The current regime is filtered out based on the previous and next states of the regime process and the transition probabilities, where the prior probability of each state is 0.5. It follows then that the posterior probability in the  $g$ -th iteration is given by

$$x_{t_k} - 1 \sim \text{Ber}(\tilde{q})$$

$$\begin{aligned} \tilde{q} &= \frac{l^c(x_{t_k} = 2|u_1)}{l^c(x_{t_k} = 1|u_1) + l^c(x_{t_k} = 2|u_1)}, \\ q_{pr}(1) &= (1 - q_{1,2}^{(g)})(2 - x_{t_{k-1}}^{(g)}) + q_{2,1}^{(g)}(x_{t_{k-1}}^{(g)} - 1), \\ q_{nx}(1) &= (1 - q_{1,2}^{(g)})(2 - x_{t_{k+1}}^{(g-1)}) + q_{2,1}^{(g)}(x_{t_{k+1}}^{(g-1)} - 1), \\ q_1 &= \frac{q_{pr}(1)q_{nx}(1)}{q_{pr}(1)q_{nx}(1) + (1 - q_{pr}(1))(1 - q_{nx}(1))}, \\ q_2 &= \frac{(1 - q_{pr}(1))(1 - q_{nx}(1))}{q_{pr}(1)q_{nx}(1) + (1 - q_{pr}(1))(1 - q_{nx}(1))}, \\ l^c(x_{t_k} = 1|u_1) &\propto q_1 \exp\left\{\frac{(y_{t_k}^{(i)} - \tilde{\mu}(1))^2}{2(\tilde{\sigma}_{t_k}^{(i)}(1))^2}\right\}, \\ l^c(x_{t_k} = 2|u_1) &\propto q_2 \exp\left\{\frac{(y_{t_k}^{(i)} - \tilde{\mu}(2))^2}{2(\tilde{\sigma}_{t_k}^{(i)}(2))^2}\right\}, \end{aligned}$$

where  $u_1 = \{y_{t_k}^{(i)}, \tilde{\mu}(1), \tilde{\mu}(2), \tilde{\sigma}_{t_k}^{(i)}(1), \tilde{\sigma}_{t_k}^{(i)}(2)\}$ .

- *Jump Sizes and Jump Times:  $\delta n_f$  and  $z$*

We use the Gibbs Sampler to draw jump times and jump sizes subject to the remaining model parameters (see Eraker, Johannes, and Polson (2003)).

The jump events are filtered out based on a random futures contract. The jump sizes are drawn independently in order to update the mean reversion parameter for the jump process. The posterior jump size distribution depends on the average jump size, the jump size variance, and the variance state at the respective point in time  $t_k$ . It is given by

$$z_{t_k} \sim \mathcal{N}(g_{t_k}, h_{t_k}),$$

where

$$g_{t_k} = \frac{h_{t_k}}{\tilde{\gamma}_{t_k}^{(i)}} \left( \frac{\delta n_{f,t_k}}{(\tilde{\sigma}_{t_k}^{(i)}(x_{t_k}))^2} (y_{t_k}^{(i)} - \tilde{\mu}_{t_k}^{(i)}(x_{t_k})) + \frac{\mu_z}{(\sigma_z)^2} \right),$$

$$h_{t_k} = \frac{1}{(\tilde{\gamma}_{t_k}^{(i)})^2} \left( \frac{\delta n_{f,t_k}}{(\tilde{\sigma}_{t_k}^{(i)}(x_{t_k}))^2} + \frac{1}{(\sigma_z)^2} \right)^{-1}.$$

It can be seen that a higher variance level  $(\tilde{\sigma}_{t_k}^{(i)}(x_{t_k}))^2$  reduces the impact of log-returns relative to the average jump size  $\mu_z$  and increases the variance of the jump size distribution.

If we condition on the jump size distribution as well as the overall jump intensity  $\lambda_z$ , we can sample jump events from a Bernoulli distribution as follows:

$$\delta n_{f,t_k} \sim \text{Ber}(\tilde{\lambda}_{t_k})$$

with

$$\tilde{\lambda}_{t_k} = \frac{p(\delta n_{f,t_k} = 1|u_2)}{p(\delta n_{f,t_k} = 1|u_2) + p(\delta n_{f,t_k} = 0|u_2)},$$

$$p(\delta n_{f,t_k} = 1|u_2) \propto \lambda_z e^{-\frac{(y_{t_k}^{(i)} - \tilde{\mu}_{t_k}^{(i)}(x_{t_k}) - \tilde{\gamma}_{t_k}^{(i)} z_{t_k})^2}{2(\tilde{\sigma}_{t_k}^{(i)}(x_{t_k}))^2}},$$

$$p(\delta n_{f,t_k} = 0|u_2) \propto (1 - \lambda_z) e^{-\frac{(y_{t_k}^{(i)} - \tilde{\mu}_{t_k}^{(i)}(x_{t_k}))^2}{2(\tilde{\sigma}_{t_k}^{(i)}(x_{t_k}))^2}},$$

where  $u_2 = \{x_{t_k}, y_{t_k}^{(i)}, z_{t_k}, \tilde{\mu}_{t_k}^{(i)}(x_{t_k}), \tilde{\gamma}_{t_k}^{(i)}, \lambda_z\}$ .

- *Jump Intensity:*  $\lambda_z$

In our discrete setup, the total number of jumps is binomial distributed if the prior distribution of  $\lambda_z$  is beta distributed with parameters  $l_0 = 2$  and  $m_0 = 40$ :

$$\lambda_z \sim \mathcal{B}(l, m),$$

where

$$l = l_0 + \sum_{k=1}^n \delta n_{f,t_k}$$

$$m = m_0 + n - \sum_{k=1}^n \delta n_{f,t_k}$$

- *Jump Size Parameters:  $\mu_z$ ,  $\sigma_z$ , and  $\gamma$*

The log jump sizes are normally distributed so that the mean and the variance parameters can be standardly updated. We assume that the prior distribution of the mean parameter is normally distributed with mean  $u_0 = 0$  and standard deviation  $w_0 = 1$ , we obtain the following posterior distribution

$$\mu_z \sim \mathcal{N}(u, w)$$

with

$$u = \frac{u_0 \sigma_z^2 + w_0^2 \sum_{k=1}^n z_{t_k}}{\sigma_z^2 + n w_0^2},$$

$$w = \frac{\sigma_z^2 w_0^2}{\sigma_z^2 + n w_0^2}.$$

We use again an inverse gamma distribution as a prior for  $\sigma_z^2$  with parameters  $\alpha_0 = 5$  and  $\beta_0 = 0.03$  and obtain

$$\sigma_z^2 \sim \mathcal{IG}(\alpha, \beta),$$

where

$$\alpha = \alpha_0 + 0.5n$$

$$\beta = \beta_0 + 0.5 \sum_{k=1}^n (z_{t_k} - \mu_z)^2$$

The slope parameter is extracted through the difference between the filtered short-dated jumps and long-dated jumps.

Now, we have specified all individual updating steps. We only have to set the jump intensity to zero for the crude oil market model and the transition probabilities to zero for the natural gas market model in order to apply the MCMC algorithm for our model specifications.

## A.6 Normalization of Storage Contracts

The scale-invariance of degree one follows directly from

$$\begin{aligned}
\alpha sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \psi) &= \alpha \sup_{\{q_{\tau_j}\}_{j=i}^m \in ad(sv_{\tau_i}; \psi)} \mathbb{E}_t^{\mathbb{Q}} \left[ \sum_{j=i}^m e^{-r(\tau_j - \tau_i)} (-q_{\tau_j} s_{\tau_j} + c(q_{\tau_j}, s_{\tau_j})) \right] \\
&= \sup_{\{q_{\tau_j}\}_{j=i}^m \in ad(sv_{\tau_i}; \psi)} \mathbb{E}_t^{\mathbb{Q}} \left[ \sum_{j=i}^m e^{-r(\tau_j - \tau_i)} (-\alpha q_{\tau_j} s_{\tau_j} + c(\alpha q_{\tau_j}, s_{\tau_j})) \right] \\
&\leq \sup_{\{q_{\tau_j}\}_{j=i}^m \in ad(\alpha sv_{\tau_i}; \alpha \psi)} \mathbb{E}_t^{\mathbb{Q}} \left[ \sum_{j=i}^m e^{-r(\tau_j - \tau_i)} (-q_{\tau_j} s_{\tau_j} + c(q_{\tau_j}, s_{\tau_j})) \right] \\
&\leq sw_{\tau_i}(s_{\tau_i}, \alpha sv_{\tau_i}; \alpha \psi)
\end{aligned}$$

for  $\alpha > 0$  and  $i = 0, \dots, m$ . It follows then that

$$\begin{aligned}
sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \psi) &= sw_{\tau_i}(s_{\tau_i}, \frac{1}{\alpha} \alpha sv_{\tau_i}; \frac{1}{\alpha} \alpha \psi) \geq \frac{1}{\alpha} sw_{\tau_i}(s_{\tau_i}, \alpha sv_{\tau_i}; \alpha \psi) \\
\Rightarrow \alpha sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \psi) &\geq sw_{\tau_i}(s_{\tau_i}, \alpha sv_{\tau_i}; \alpha \psi) \\
\Rightarrow \alpha sw_{\tau_i}(s_{\tau_i}, sv_{\tau_i}; \psi) &= sw_{\tau_i}(s_{\tau_i}, \alpha sv_{\tau_i}; \alpha \psi), \quad i = 0, \dots, m.
\end{aligned}$$

The same arguments can be used to show that the storage value is concave in each technical constraint. The initial point are two storage contracts with technical constraints  $\widehat{\psi}_j$  and  $\widetilde{\psi}_j$  and optimal trading strategies  $\widehat{q}^*$  and  $\widetilde{q}^*$ . Then, the trading strategy  $q = \lambda \widehat{q}^* + (1 - \lambda) \widetilde{q}^*$  is an admissible trading strategy for technical constraints  $\psi$  with an expected payoff that is equal to  $\lambda$  times the expected payoff of  $\widehat{q}^*$  plus  $1 - \lambda$  times the expected payoff of  $\widetilde{q}^*$ .

# Bibliography

- Acharya, Viral V., Lars A. Lochstoer, and Tarun Ramadorai, 2011, Limits to Arbitrage and Hedging: Evidence from Commodity Markets, *Working Paper*.
- Asgharian, Hossein, and Christoffer Bengtsson, 2006, Jump Spillover in International Equity Markets, *Journal of Financial Econometrics* 4, 167–203.
- Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical Performance of Alternative Option Pricing Models, *The Journal of Finance* 52, 2003–2049.
- Bakshi, Gurdip, and Nikunj Kapadia, 2003, Delta-Hedged Gains and the Negative Market Volatility Risk Premium, *The Review of Financial Studies* 16, 527–566.
- Bakshi, Gurdip, and Dilip Madan, 2000, Spanning and Derivative-Security Valuation, *Journal of Financial Economics* 55, 205–238.
- Bardou, Olivier, Sandrine Bouthemy, and Gilles Pages, 2009, Optimal Quantization for the Pricing of Swing Options, *Applied Mathematical Finance* 16, 183–217.
- Barone-Adesi, Giovanni, and Robert E. Whaley, 1987, Efficient Analytic Approximation of American Option Values, *The Journal of Finance* 42, 301–320.
- Bates, David S., 2000, Post-'87 Crash Fears in the S&P 500 Futures Option Market, *Journal of Econometrics* 94, 181–238.
- Benth, Fred Espen, Jan Kallsen, and Thilo Meyer-Brandis, 2007, A Non-Gaussian Ornstein-Uhlenbeck Process for Electricity Spot Price Modeling and Derivatives Pricing, *Applied Mathematical Finance* 14, 153–169.
- Benth, Fred Espen, and Steen Koekebakker, 2008, Stochastic Modeling of Financial Electricity Contracts, *Energy Economics* 30, 1116–1157.
- Benth, Fred Espen, Steen Koekebakker, and Fridthjof Ollmar, 2007, Extracting and Applying Smooth Forward Curves from Average-Based Commodity Contracts with Seasonal Variation, *The Journal of Derivatives* 15, 52–66.

- Black, Fischer, 1976, The Pricing of Commodity Contracts, *Journal of Financial Economics* 3, 167–179.
- Boogert, Alexander, and Cyriel de Jong, 2008, Gas Storage Valuation Using a Monte Carlo Method, *The Journal of Derivatives* 15, 81–98.
- Boogert, Alexander, and Cyriel de Jong, 2011, Gas Storage Valuation Using a Multi-Factor Price Process, *Journal of Energy Markets* 4, 29–52.
- Bouwman, Kees E., Eran Raviv, and Dick van Dijk, 2012, An Arithmetic Modeling Framework for the Term Structure of Electricity Prices, *Working Paper*.
- Branger, Nicole, Eva Krautheim, Christian Schlag, and Norman Seeger, 2012, Hedging Under Model Misspecification: All Risk Factors are Equal, but Some are More Equal Than Others ..., *Journal of Futures Markets* 32, 397–430.
- Branger, Nicole, and Christian Schlag, 2008, Can Tests Based on Option Hedging Errors Correctly Identify Volatility Risk Premia, *Journal of Financial and Quantitative Analysis* 43, 1055–1090.
- Breedon, Douglas T., and Robert H. Litzenberger, 1978, Prices of State-Contingent Claims Implicit in Option Prices, *The Journal of Business* 51, 621–651.
- Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, Model Specification and Risk Premia: Evidence from Futures Options, *The Journal of Finance* 62, 1453–1490.
- Brooks, Chris, and Marcel Prokopczuk, 2011, The Dynamics of Commodity Prices, *Working Paper*.
- Carr, Peter, and Liuren Wu, 2009, Variance Risk Premiums, *The Review of Financial Studies* 22, 1311–1341.
- Cartea, Alvaro, and Thomas Williams, 2008, UK Gas Market: The Market Price of Risk and Applications to Multiple Interruptible Supply Contracts, *Energy Economics* 30, 829–846.
- Casassus, Jaime, and Pierre Collin-Dufresne, 2005, Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates, *The Journal of Finance* 60, 2283–2331.
- Chernov, Mikhail, and Eric Ghysels, 2000, A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation, *Journal of Financial Economics* 56, 407–458.

- Cont, Rama, and Peter Tankov, 2002, Calibration of Jump-Diffusion Option-Pricing Models: A Robust Non-Parametric Approach, *Working Paper*.
- DeGroot, Morris H., 1990, *Optimal Statistical Decisions*. (McGraw-Hill).
- Dempster, Michael A. H., Elena Medova, and Ke Tang, 2010, Long and Short Term Jumps in Commodity Futures Prices, *Working Paper*.
- Doran, James S., and Ehud I. Ronn, 2008, Computing the Market Price of Volatility Risk in the Energy Commodity Markets, *Journal of Banking & Finance* 32, 2541–2552.
- Eberlein, Ernst, and Ulrich Keller, 1995, Hyperbolic Distributions in Finance, *Bernoulli* 1, 281–299.
- Eraker, Bjorn, 2004, Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices, *The Journal of Finance* 59, 1367–1403.
- Eraker, Bjorn, Michael Johannes, and Nicholas Polson, 2003, The Impact of Jumps in Volatility and Returns, *The Journal of Finance* 58, 1269–1300.
- Falkowski, Michal, 2011, Financialization of Commodities, *Contemporary Economics* 5, 4–17.
- Geman, Stuart, and Donald Geman, 1984, Stochastic Relaxation, Gibbs Distribution and the Bayesian Restoration of Images, in *Pattern Analysis and Machine Intelligence* (IEEE Transactions, ).
- Geyer, Charles J., 1992, Practical Markov Chain Monte Carlo, *Statistical Science* 7, 473–483.
- Gibson, Rajna, and Eduardo S. Schwartz, 1990, Stochastic Convenience Yield and the Pricing of Oil Contingent Claims, *The Journal of Finance* 45, 959–976.
- Hain, Martin, Marliese Uhrig-Homburg, and Nils Unger, 2012, Risk Factors and Their Associated Risk Premia: An Empirical Analysis of the Crude Oil Market, *Working Paper*.
- Hastings, W. Keith, 1970, Monte Carlo Sampling Methods Using Markov Chains and Their Applications, *Biometrika* 57, 97–109.
- Heston, Steven L., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *The Review of Financial Studies* 6, 327–343.

- Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi, 2004, Bayesian Analysis of Stochastic Volatility Models with Fat-Tails and Correlated Errors, *Journal of Econometrics* 122, 185–212.
- Jeanblanc, Monique, Marc Yor, and Marc Chesney, 2009, *Mathematical Methods for Financial Markets*. (Springer Verlag).
- Johannes, Michael, Rohit Kumar, and Nicholas G. Polson, 1999, State Dependent Jump Models: How Do US Equity Indices Jump?, *Working Paper*.
- Jones, Christopher S., 2003, The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets, *Journal of Econometrics* 116, 181–224.
- Kaminski, Vincent, Youyi Feng, and Zhan Pang, 2008, Value, Trading Strategies and Financial Investment of Natural Gas Storage Assets, *Working Paper*.
- Kang, Sang Baum, and Xuhui Pan, 2011, Do Variance Risk Premia Predict Commodity Futures Returns? Evidence from the Crude Oil Market, *Working Paper*.
- Koekebakker, Steen, and Fridthjof Ollmar, 2005, Forward Curve Dynamics in the Nordic Electricity Market, *Managerial Finance* 31, 73–94.
- Lai, Guoming, Francois Margot, and Nicola Secomandi, 2010, An Approximate Dynamic Programming Approach to Benchmark Practice-Based Heuristics for Natural Gas Storage Valuation, *Operations Research* 58, 564–582.
- Larsson, Karl, and Marcus Nossman, 2011, Jumps and Stochastic Volatility in Oil Prices: Time Series Evidence, *Energy Economics* 33, 504–514.
- Liu, Jun, and Jun Pan, 2003, Dynamic Derivative Strategies, *Journal of Financial Economics* 69, 401–430.
- Liu, Jun S., Wing H. Wong, and Augustine Kong, 1994, Covariance Structure of the Gibbs Sampler with Applications to the Comparisons of Estimators and Augmentation Schemes, *Biometrika* 81, 27–40.
- Longstaff, Francis A., and Eduardo S. Schwartz, 2001, Valuing American Options by Simulation: A Simple Least-Squares Approach, *The Review of Financial Studies* 14, 113–147.
- McCulloch, J. Huston, 1971, Measuring the Term Structure of Interest Rates, *The Journal of Business* 44, 19–31.

- Merton, Robert C., 1973, Theory of Rational Option Pricing, *The Bell Journal of Economics and Management Science* 4, 141–183.
- Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller, 1953, Equation of State Calculations by Fast Computing Machines, *The Journal of Chemical Physics* 21, 1087–1092.
- Neumann, Anne, and Georg Zachmann, 2009, Expected vs. Observed Storage Usage: Limits to Intertemporal Arbitrage, in A. Credi, eds.: *The Economics of Natural Gas Storage: A European Perspective* (Springer Verlag, Berlin ).
- Pan, Jun, 2002, The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study, *Journal of Financial Economics* 63, 3–50.
- Roberts, Gareth O., and Richard L. Tweedie, 2008, Understanding MCMC, *Working Paper*.
- Schwartz, Eduardo S., 1997, The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging, *The Journal of Finance* 52, 923–973.
- Schwartz, Eduardo S., and James E. Smith, 2000, Short-Term Variations and Long-Term Dynamics in Commodity Prices, *Management Science* 46, 893–911.
- Secomandi, Nicola, 2010, Optimal Commodity Trading with a Capacitated Storage Asset, *Management Science* 56, 449–467.
- Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde, 2002, Bayesian Measures of Model Complexity and Fit, *Journal of the Royal Statistical Society (Series B)* 64, 583–639.
- Thompson, Andrew C., 1995, Valuation of Path-Dependent Contingent Claims with Multiple Exercise Decisions over Time: The Case of Take-or-Pay, *Journal of Financial and Quantitative Analysis* 30, 271–293.
- Thompson, Matt, Matt Davison, and Henning Rasmussen, 2009, Natural Gas Storage Valuation and Optimization: A Real Options Application, *Naval Research Logistics* 56, 226–238.
- Trolle, Anders B., and Eduardo S. Schwartz, 2009, Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives, *The Review of Financial Studies* 22, 4423–4461.
- Trolle, Anders B., and Eduardo S. Schwartz, 2010, Variance Risk Premia in Energy Commodities, *The Journal of Derivatives* 17, 15–32.

Tsay, Ruey S., 2005, *Analysis of Financial Time Series*. (Wiley Interscience) 2ed edn.

Uhrig-Homburg, Marliese, and Nils Unger, 2012, An Energy Market Modeling Approach for Valuing Real Options, *Working Paper*.

Wu, Owen Q., Derek D. Wang, and Zhenwei Qin, 2011, Seasonal Energy Storage Operations with Limited Flexibility, *Working Paper*.