ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF ECONOMICS AND COMMERCIAL LAW
GÖTEBORG UNIVERSITY
154

PRICING OF SOME PATH-DEPENDENT OPTIONS ON EQUITIES AND COMMODITIES

Mats Kjaer

ISBN 91-85169-13-7
ISBN 978-91-85169-13-9
ISSN 1651-4289 print
ISSN 1651-4297 online

Abstract

This thesis brings together three papers about the pricing of European and Bermudan path-dependent options, and one paper about the stochastic modelling of a futures price curve.

Paper one proposes a fast numerical method to compute the price of so called cliquet options with global floor, when the underlying asset follows the Bachelier-Samuelson model. These options often constitute the option part of many capital guaranteed products, and are slow to price with existing Monte Carlo and PDE methods.

Paper two deals with the pricing of swing options, when the logarithm of the underlying asset follows an Ornstein-Uhlenbeck process driven by a jump diffusion. Swing options are Bermudan or American options with multiple exercise rights, and are common on the energy markets.

Paper three investigates the valuation of a natural gas storage facility, when gas trading is permitted on the spot- and futures markets simultaneously. The main idea is to interpret the storage as a swing option and then apply option pricing methods.

Paper four proposes, estimates and evaluates two classes of parsimonious models of the correlation matrix of natural gas futures returns. The individual futures prices follow a Bachelier-Samuelson model with time-dependent volatility.

Preface

About this thesis

This thesis consists of a brief introduction and four appended papers. It is a continuation of my Licentiate of Engineering Degree in Industrial Mathematics from Chalmers University of Technology and is submitted as a partial fulfilment for the degree of Doctor of Philosophy (PhD) in Economics. It corresponds to two years of full time research work and in addition, two years of PhD-level courses are required.

Acknowledgements

I would like to thank:

- My supervisor Professor Lennart Hjalmarsson, for facilitating my transfer to the Department of Economics, finding funding, and proof reading of parts of the manuscript.
- My Texan host and co-author Professor Ehud I. Ronn, for inviting me to the University of Texas at Austin, finding me a position as a Research Fellow, suggesting research topics and providing useful comments on the manuscripts.
- The director of the Centre for Finance, Evert Carlsson, for facilitating my transfer to the Department of Economics.
- The following institutions for providing financial support: The Jan Wallander and Tom Hedelius Foundation, The University of Texas at Austin, The Knut & Alice Wallenberg Foundation, The Dr. Marcus Wallenberg Foundation.
- My current and former colleagues at the Departments of Economics and Mathematics, who have been great company and formed an intellectually stimulating environment.
- My mother Ulrika, father Klaus and younger brother Erik, for always supporting me in my studies.

This thesis is dedicated in memory of my grandfather Hans Kjær (1902–1996), who was the first among my closer relatives to attend Higher Education. In 1927, he obtained his MSc.-degree in Mathematics and Physics from Copenhagen University, Denmark.

Mats Kjaer Göteborg, April 26, 2006

Contents

This thesis consists of a brief introduction and the following four appended papers:

Paper 1 is a revised and shorted version of my Licentiate of Engineering thesis *Pricing of cliquet options with global floor and cap.* It was partly completed in cooperation with the financial software company Front Capital Systems AB, Stockholm, Sweden. Some of the results have been implemented in the C programming language and are available in the Front Arena Trading System.

Introduction

> "At that time, the notion of partial differential equations was very, very strange on Wall Street."
>
> Robert C. Merton, *Derivative Strategies*, March 1998, p. 32.

In 1973, Fischer Black and Myron Scholes published their famous paper "The pricing of options and corporate liabilities" (Black and Scholes [7]). Based on the principle that on a rational market, there are no certain profits, they derived formulas for the theoretical price of European put and call options. Ever since, the development of more complex option types and the use of mathematical tools for their pricing and risk management have exploded.

The papers of this thesis continue this development in in three directions, namely stochastic modelling of asset prices, option pricing and numerical methods for option pricing. The asset classes considered are equities, spot commodities, spot electricity and commodity futures. We will mainly use diffusion models, but a jump-diffusion model is employed in one of the papers.

In this introductory chapter we briefly present the models and ideas used in the appended papers. It is not self-contained, but the theory used is well known and presented more rigorously in several books. Stochastic calculus for Wiener processes is found in Karatzas and Shreves [21] or Øksendahl [28]. For jump diffusions, we refer to Cont and Tankov [8] for an overview and to Protter [31] or Sato [34] for rigorous treatments. In option pricing, Hull [18] and Wilmott [39] are introductory texts that are nevertheless useful when implementing option pricing models in practice. Bingham and Kiesel [5], Karatzas and Shreves [22] and Korn and Korn [23] are more advanced treatments for models driven by Wiener processes. Cont and Tankov [8] cover option pricing in Lévy process market models. Commodity and energy price modelling is the topic of Geman [14] and Ronn [33].

This introduction is organised as follows: Section 1.1 discusses the pricing of path-dependent equity options in the Black-Scholes (B-S) model. Sections 1.2 and 1.3 are about spot price models of consumption commodities and electricity respectively. Futures curve models is the topic of Section 1.4 and in Section 1.5 we illustrate how optimal storage management is equivalent to the pricing of a swing option. We conclude this introduction in Section 1.6 by discussing the estimation and calibration of the models of Sections 1.1 to 1.4.

Throughout this introduction we let $(\Omega, \mathcal{F}, P)$ be a complete probability space equipped with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions as defined in Protter [31].

1.1 Path dependent options in the B-S model

In Paper 1, we discuss the pricing of a class of path-dependent equity options. Here we use the standard Black-Scholes market model with one stock and one risk-free asset. The risk-less rate r and volatility σ are positive and deterministic constants. Since this model is complete, we take P to be the equivalent martingale measure, and expectations with respect to this measure are denoted $\mathbb{E}$. Under P the stock price S_t and bond price B_t satisfy the SDEs

$$\begin{cases} dS_t/S_t &= rdt + \sigma dW_t \\ dB_t/B_t &= rdt, \end{cases}$$

with W_t being a $P-$Wiener process defined on $(\Omega, \mathcal{F}, P)$. Here we take $\mathcal{F}_t$ to be the canonical filtration of the process W_t.

If $T \geq 0$ is a constant point in time, we define a contingent $T-$claim as an $\mathcal{F}_T-$measurable payout $Y \geq 0$ at $t = T$. Now let $0 \leq T_0 < T_1 < \ldots < T_N \leq T$, $N \in \mathbb{N}$ be a set of *monitoring dates.* A *discretely path-dependent European option* has a payoff of the form $Y = H(S_{T_0}, \ldots, S_{T_N})$. General derivatives pricing theory for the Black-Scholes model (see Karatzas and Shreve [22] or Korn and Korn [23]) states that the arbitrage free price V_t at time $0 \leq t \leq T$ of this claim is given by

$$V_t = e^{-r(T-t)}\mathbb{E}[Y|\mathcal{F}_t]. \tag{1.1}$$

This class of path-dependent options includes discretely monitored Asian and lookbacks options (see Hull [18], Chapter 19 for a description of these products). For $T_n \leq t < T_{n+1}$, the Markov property of S_t yields that $V_t = V(t, s_1, s_2, \ldots, s_n, s)$, where $s_k = S_{T_k}$ etc. This means that the number of state variables could be large, but in some situations this number could be reduced significantly. One example on when this is the case is if there exist measurable functions g_k and h such that

$$Y = h\left(\sum_{k=0}^{N-1} g_k(S_{T_k}, S_{T_{k+1}})\right).$$

If we introduce the state variables $\bar{s} = S_{T_n}$ and

$$z = \sum_{k=0}^{n-1} g_k(S_{T_k}, S_{T_{k+1}}),$$

it follows that $V_t = V(t, s, \bar{s}, z)$ for $T_n \leq t < T_{n+1}$. Generalising the arguments used in Andreasen [1] to derive a PDE for discretely monitored Asian options, it follows that $V_t = V(t, s, \bar{s}, z)$ satisfies the PDE (1.2).

$$\begin{cases} \frac{\partial V}{\partial t} + \frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{\partial s^2} + rs\frac{\partial V}{\partial s} - rV = 0,\ T_{n-1} \leq t < T_n, \\ V(T_N, s, \bar{s}, z) = h(z), \\ V(T_n^-, s, \bar{s}, z) = V(T_n, s, s, z + g_n(\bar{s}, s)),\ 1 \leq n \leq N. \end{cases} \tag{1.2}$$

The type of discretely path dependent option considered in Paper 1 is called a *cliquet option with global floor*. Its payoff depends on the stock price returns over the life of the option as follows. The return R_n of S_t over the period $[T_{n-1}, T_n)$ is defined as

$$R_n = \frac{S_{T_n}}{S_{T_{n-1}}} - 1.$$

Truncated returns, $\overline{R}_n = \max(\min(R_n, C), F)$ are returns truncated at some floor and cap levels F and C respectively with $F < C$. A cliquet option with global floor has a payoff Y at time T of

$$Y = B \times \max(\sum_{n=1}^{N} \overline{R}_n, F_g),$$

where the *global floor* F_g is the minimum total return and B is a notional amount which is set to one for the remainder of this thesis. For F_g to be of interest, it must satisfy $NF < F_g$. Consequently we have that $h(z) = \max(z, F_g)$ and

$$g_n(\bar{s}, s) = \max\left(\min\left(\frac{s}{\bar{s}} - 1, C\right), F\right).$$

For this class of options, we may actually reduce the number of state variables by one if we replace s and $\bar{s}$ by $y = s/\bar{s}$.

Since the number of monitoring dates N could easily be $N = 12-48$, computing (1.1) by Monte Carlo simulation or solving (1.2) by finite differences could both be a very time consuming process, even after a dimensionality reduction. The main contribution of Paper 1 is a Fourier-transform based numerical integration method for the computation of (1.1). It relies heavily on the fact that the returns R_n are independent random variables. Compared to existing quasi Monte Carlo and PDE methods, it turns out to be much faster for a given degree of computational accuracy.

1.2 Spot commodity option models

In this section we give a brief overview of some models for spot prices of spot commodities. For a more extensive treatment, we refer to Geman [14] and Ronn [33].

There are many differences between consumption commodities and stocks. These include costly storage, time consuming transportation, and payments that often take place at delivery. Moreover, they are primarily inputs to some production process rather than investments. Below we list some empirical facts about commodity spot prices. Since these may be hard to observe, the most nearby futures price is often used as a proxy.

- **Mean reversion**: In Bessembinder et. al. [4], a strong rate of mean reversion in spot prices for crude oil. Over their period of observation, 44% of a price shock is reversed in eight months on average.

- **Stochastic convenience yield**: Gibson and Schwartz [36] as well as Henker and Milonas [17] report that the implied convenience yield for crude oil may be of a significant magnitude and stochastic in nature.

- **Distributional characteristics**: Many models assume normally distributed log-returns. Henker and Milonas [17] have computed estimates of the first four moments from the empirical distribution and conclude that the log-returns are not normally distributed. Possible explanations include stochastic volatility and/or jumps.

- **Volatility skew**: When computing the implied volatility for different strike prices of vanilla options, Beaglehole and Chebanier [2] report the presence of volatility skew.

Most commodity spot price models are one- or multi-factor *affine models*. They have closed form expressions for futures prices and Fourier transforms of vanilla option prices. Duffie, Filipovic and Schachermayer [12] treat the theoretical properties of affine processes, while their parameter estimation is covered in Singleton [38].

Below we will give some examples of some models. In an early model, Brennan and Schwartz model S_t as a stock in the B-S model paying a continuous dividend yield equalling the net convenience yield y. This model is complete, so the dynamics under the measures P (real world) and Q (risk neutral) are given by

$$\begin{aligned} dS_t/S_t &= \mu dt + \sigma dW_t && (P), \\ dS_t/S_t &= (r-y)dt + \sigma d\hat{W}_t && (Q), \end{aligned} \tag{1.3}$$

respectively, where $\hat{W}_t$ is a $Q-$Wiener process. In Gibson and Schwartz [36], this model is improved by making the convenience yield stochastic and introducing a *market price of convenience yield risk*.

Neither the Brennan-Schwartz nor the Gibson-Schwartz models feature mean reversion. In Schwartz [35], the spot price is considered to be a non-traded asset with $P-$dynamics

$$\begin{aligned} dX_t &= -\alpha X_t dt + \sigma dW_t, \\ S_t &= \exp(f(t) + X_t), \end{aligned} \tag{1.4}$$

where $\alpha > 0$, $\sigma > 0$ and $f(t)$ is a deterministic seasonal trend. Since S_t is not traded, this model is not complete, and an equivalent martingale measure Q is selected by prescribing a *market price of spot price risk* of the form $\lambda(t, X_t) = \lambda_0 + \lambda_1 X_t$ with $\lambda_0 \in \mathbb{R}$, $\lambda_1 > -\alpha/\sigma$, and applying the Girsanov Theorem (see Karatzas and Shreve [21] or Korn and Korn [23] for its formulation and proof). Defining the risk neutral model parameters $\hat{\alpha} = \alpha + \sigma\lambda_1$, $\hat{\beta} = -\sigma\lambda_0$ and $\hat{f}(t) = f(t) + \frac{\hat{\beta}}{\hat{\alpha}}(1 - e^{-\hat{\alpha}t})$ finally yields the $Q-$dynamics of S_t as

$$\begin{cases} dX_t &= -\hat{\alpha}X_t dt + \sigma d\hat{W}_t, \\ S_t &= \exp(\hat{f}(t) + X_t). \end{cases} \tag{1.5}$$

The choice of a linear market price of spot price risk is done for analytical tractability in order to keep the affine structure.

This model could also be expanded by making $f(t)$ and/or the interest rate r stochastic, but these improvements can still not reproduce volatility smiles. Beaglehole and Chebanier [2] solve this by leaving the affine class of models and let the $Q-$dynamics of S_t be given by

$$\begin{cases} dX_t &= \hat{\alpha}(\mu(t) - X_t)dt + \sigma(t, X_t)d\hat{W}_t, \\ S_t &= \exp(X_t). \end{cases} \tag{1.6}$$

In the models presented in this section, European option prices may be computed by evaluating risk neutral expectations. Alternatively, the Feynman-Kac representation formula (see Korn and Korn [23]) states that the price is often given as the unique solution to a parabolic PDE similar to the Black-Scholes PDE. Sometimes this equation has to be solved numerically by some finite difference or finite element method. Wilmott [39] or Eriksson et. al. [13] may be helpful for the implementation of these methods.

1.3 Spot electricity option models

Electric power differs from the consumption commodities of Section 1.2 in that at this moment of writing, it is not possible to store large amounts of electric energy in a feasible manner once it has been produced. Inelastic demand and absence of stocks to smooth supply shocks result in a price dynamics characterised by seasonality, mean reversion and sudden spikes. Accordingly, the concept of convenience yield does not make sense for electricity.

A consequence of these special characteristics of electricity is that it is always modelled as a non-traded asset. Lucia and Schwartz [25] take the crude oil price models from Schwartz [35] and use them to price electricity futures. These models feature seasonality and mean reversion, but have continuous price trajectories. Deng [11] replace the Wiener process of Model (1.4) with a jump diffusion in order to allow discontinuous trajectories. Following Cont and Tankov [8], the class of equivalent martingale measures becomes much larger due to the introduction of jumps. Apart from adding a market price of risk as in (1.5), one may change the jump intensity and distribution, as long as the latter is absolutely continuous with the jump distribution under P. It is common to choose Q such that the logarithm of the price process remains a jump diffusion under Q. If this is the case, the Feynman-Kac representation formula yields that European option prices are often given as the unique solution to a parabolic partial integro-differential equation (PIDE). This equation is similar to the PIDE satisfied by equity option prices in the Merton-model proposed in Merton [27]. The connection between jump-diffusions and PIDEs is studied in detail in Bensoussan and Lions [3] and Cont and Tankov [8]. The latter also discuss how to solve PIDEs numerically by finite differences or finite elements.

Jumps are not spikes however, and several attempts have been made to model this without sacrificing the Markov property. Geman and Roncoroni [15] propose a one-factor model with spikes and use a time-series approach to its estimation under P. They do not discuss option pricing and how to switch from P to Q. Andreasen and Dahlgren [10] specify an affine two-factor model with mean reversion and spikes directly under Q. Consequently they rely fully on futures and option prices for model calibration.

1.4 Commodity futures price models

The futures price $F(t,T)$ of a commodity is the price agreed upon at time $t \geq 0$ for delivery and payment of a pre-specified amount of the underlying commodity at time $T \geq t$ at a pre-specified location (*physical settlement*). That the delivery actually takes place is very rare, and usually the position is liquidated against cash (*financial settlement*). For electricity, it is common that financial settlement against the spot price at maturity is the only way of settlement.

The most well known commodity futures exchanges in the world include Nymex (New York: U.S. crude oil, U.S. natural gas), CBOT (Chicago: U.S. Agricultural products) and LME (London: metals). They offer standardised futures contracts and futures options as well as clearing and settlement services.

In order to price options on futures contracts, or to compute the value at risk (VaR) of a futures portfolio, it may be more convenient to model $F(t,T)$ directly. This was done for a single contract in Black [6], where $F(t,T)$ is assumed to follow the risk neutral dynamics

$$\frac{dF(t,T)}{F(t,T)} = \sigma dW_t. \tag{1.7}$$

Ever since, this *Black-model* is the reference model for the pricing of vanilla options on individual futures contracts.

In order to price options, whose payoff depends on many futures prices with maturities $T_1 < T_2 <, \ldots, T_M$, one must be able to specify the joint dynamics of the discrete futures curve $F(t,T_1), \ldots, F(t,T_M)$ for $t \leq T_1$. Inspired by the development of HJM models (Heath, Jarrow and Morton [16]) in the interest rate area, Cortazar and Schwartz [9] propose the following risk neutral dynamics for $F(t, T_m)$.

$$\frac{dF(t, T_m)}{F(t, T_m)} = \sum_{n=1}^{N} \sigma_{mn}(t) dW_t^n. \tag{1.8}$$

Here $(W_t^1, \ldots, W_N^t)$ are independent Wiener processes, $\sigma_{mn}(t) > 0$ are deterministic so called volatility functions, and we say that the model has N *factors*. If we add the risk less bond B_t and $N \leq M$, the model is complete. Schwartz [35] show that the spot price model (1.5) corresponds to $N = 1$ and $\sigma_{m1}(t) = \sigma e^{-\hat{\alpha}(T_m - t)}$. Similar relations also exist for higher factor models, see Schwartz and Smith [37].

Alternatively, we may prescribe the futures price dynamics as

$$\frac{dF(t, T_m)}{F(t, T_m)} = h_m(t) dW_t^m, \tag{1.9}$$

for $1 \leq m \leq M$. Here $h_m(t) > 0$ is deterministic and $\mathrm{Cov}(W_t^i, W_t^j) = \rho_{ij} t$, $1 \leq i, j \leq M$. The constants ρ_{ij} are elements of a $M \times M$ correlation matrix $\mathbf{C}$ where the rank of $\mathbf{C}$ corresponds to the number of factors. For a discrete futures curve of $M = 12 - 24$ contracts, empirical evidence for crude oil (Schwartz [35]) and copper (Cortazar and Schwartz [9]) suggest that $N = 3$ factors explain up to 99% of the variance. Setting $h_m(t) = \sigma_m$ retrieves the Black model (1.7), but it is more common to let $h_m(t)$ be an increasing increase as the time to maturity $T_m - t$ decrease. The solution of (1.9) is given by

$$F(t, T_m) = F(0, T_m) \exp\left(-\frac{1}{2} \int_0^t h_m^2(u) du + \int_0^t h_m(u) dW_u^m \right),$$

so $F(t, T_m)$ is log-normal and there exist analytical formulas for European vanilla futures options.

1.5 Swing options and storage valuation

One type of derivative that is common on the electric power and natural gas markets, is the *swing option*. It allows flexibility in delivery with respect to both the timing and amount of energy delivered, and is thus a generalisation of American and Bermudan options. Rebonato [32] describes a similar product in the interest rate area called a *chooser flexi-cap*. Pricing of swing options is the topic of Dahlgren [10], Ibáñz [19] and Jaillet, Ronn and Tompaidis [20]. Manoliu [26] and Parsons [29] show how a swing option may be viewed as a real option on a physical storage (of natural gas for example). All these papers have in common that they use some form of dynamic programming to compute swing option prices or storage values. In this section we will illustrate this by giving a very simple example of optimal storage management.

Consider a storage tank containing for example oil or natural gas. It has a capacity of $\bar{Z}$ and is equipped with one outlet. At times $t = 0, 1, 2$ the operator selects the amount of z_t to be withdrawn from the storage during the periods $[0, 1)$, $[1, 2)$ and $[2, 3)$ respectively. Furthermore, we assume that $z_t \in \{0, \bar{Z}/3, 2\bar{Z}/3\}$, and let Z_t be the amount withdrawn up to and including time t. Physical constraints impose the conditions

$$
\begin{aligned}
Z_0 &\in \{0\} \equiv \mathcal{A}_0 \\
Z_1 &\in \{0, \bar{Z}/3, 2\bar{Z}/3\} \equiv \mathcal{A}_1 \\
Z_2 &\in \{0, \bar{Z}/3, 2\bar{Z}/3, \bar{Z}\} \equiv \mathcal{A}_2 \\
Z_3 &\in \{0, \bar{Z}/3, 2\bar{Z}/3, \bar{Z}\} \equiv \mathcal{A}_3.
\end{aligned}
$$

Let S_t be the spot price for the commodity stored and for simplicity we assume that the risk-neutral dynamics of S_t is given by (1.5). Since we assume this to be a small storage, the operator is a price taker, so a time t decision to sell z_t results in a payoff $Y_t = z_t S_t$. The aim of the operator is to maximise the storage value.

Clearly, the storage value V_t at time t satisfies $V_t = V(t, s, Z)$ with $s = S_t$ and $Z = Z_t$, so at $t = 2$ the value is given by the payoff from selling as much as possible

$$V(2, s, Z) = \max_{Z+z_2 \in \mathcal{A}_3} z_2 S_2.$$

At times $1 < t \leq 2$, standard derivatives pricing theory and the Markov property of S_t yield

$$V(t, s, Z) = e^{-r(2-t)} \mathbb{E}[V(2, S_2, Z) | S_t = s].$$

Conditioned on the tank level Z and the decision z_1 at time $t = 1$, the operator receives $Y_1 = z_1 S_1$ *plus* a new storage containing $\bar{Z} - Z - z_1$ and having one decision point left. Maximising over z_1 yields

$$V(1, s, Z) = \max_{Z+z_1 \in \mathcal{A}_2} \left\{ z_1 S_1 + e^{-r} \mathbb{E}[V(2, S_2, Z + z_1) | S_1 = s] \right\},$$

which we recognise as a Bellman-equation. By repetition of this step, we may compute the storage value at $t = 0$ as well.

A swing option is a financial derivative, that mimics the cash flow from the storage described above, but without any delivery of the underlying commodity.

Paper 2 is about the pricing of swing options on electricity and extends the papers cited above in three directions. First, we will allow discontinuous spot price trajectories. Second, the amount of electricity to be delivered is chosen from a closed interval, rather than from a discrete set. Third, at each exercise date, the swing option holder has to fix a vector of amounts for delivery during multiple periods, rather than a scalar amount for delivery during a single period.

Paper 3 is about the valuation of a natural gas storage facility. Unlike the previous papers on the topic by Manoliu [26] and Parsons [29], the storage operator is permitted to trade natural gas on the spot and futures markets simultaneously. This approach enables partial hedging of the storage operations using relatively liquid futures contracts.

1.6 Model estimation and calibration

By Bingham and Kiesel [5], Hull [18] and Rebonato [32], there are basically two main methodologies to select model parameters.

In the *martingale modelling* or *implied* approach, the parameters are chosen such that some distance between model and market prices at some point in time of some set of traded derivatives is minimised. This procedure gives the parameters under the chosen risk neutral measure Q directly, so there is no need to specify Q via P. By Duffie, Filipovic and Schachermayer [12], models that are affine under Q have essentially analytic expressions for futures prices and Fourier transforms of option prices, which facilitates the calibration.

Alternatively, we first estimate the model parameters under P from a time-series of historical data. Second, a parametric subclass of equivalent martingale measures is selected (by introducing some functional forms of the appropriate market prices of risk for example). Third, the parameters that are affected by this change of measure are refitted by the implied method described above. The workhorse of time-series estimation is the Maximum Likelihood (ML) method. Singleton [38] discusses ML-estimation of models that are affine under P, as well as other methods like the method of moments and quasi Maximum Likelihood.

Parameters that are invariant under the change of measure may be chosen by either approach and their number is determined by the choice of Q. In Model (1.4) for example, we can have $\hat{\alpha} = \alpha$ by setting $\lambda_1 = 0$. Thus we can choose to estimate as many parameters as possible from historical data, hopefully optimising the *dynamic fit* of the model, but not retrieving the market prices of futures and futures options. Alternatively, most parameters are implied from traded derivatives, giving a good fit to the current market prices (*static fit*), at the risk of a poorer dynamic fit.

For interest rate models, Rebonato [32] gives the following pieces on advise, which should apply to equities and commodities as well:

- Derivatives that will be used for hedging must be priced exactly by the model.
- The set of derivatives used to imply parameters must be liquidly traded.

In the futures price model (1.8) both the functions $h_m(t)$ and the correlation matrix $\mathbf{C}$ are invariant under the change from P to Q. For commodities, there are usually enough of liquid vanilla options on each single futures contract to imply each $h_m(t)$. In order to imply $\mathbf{C}$ however, we would need actively traded *swaptions* or *calendar spread options*. This is usually not the case, so $\mathbf{C}$ has to be estimated from historical returns. If there are M futures contracts, this means that $M(M-1)/2$ parameters have to be estimated, and $M = 18 - 24$ is not unusual in applications. Pourmahdi [30] report that the variance of the estimator may be reduced significantly if some parsimonious structure is imposed on $\mathbf{C}$. Given such a model of $\mathbf{C}$, the ML-method described in Jöreskog or the Generalised Least Squares (GLS) method proposed in Lee [24] may be used to estimate these parameters.

For many commodities, most of the factors that govern demand and supply, i.e. weather patterns, number of end users, extraction infrastructure and storage capacity, do not change too much from year to year. Moreover they put constraints on the relative price movements on adjacent contracts. Consequently it may actually be possible to describe the correlation structure using a modest number of parameters.

In Paper 4, we propose some parsimonious models of the correlation matrix for natural gas futures returns. To the best of our knowledge, no papers have previously been published on the correlation structure for a commodity, such as natural gas, with a seasonal demand or supply. We believe that our results are particulary helpful for economic agents who seek a parsimonious method of calculating Value-at-Risk (VAR) of a portfolio or compute the price of an exotic option with many underlying maturities.

Bibliography

[1] J. Andreasen. The pricing of discretely sampled asian options: A change of numeraire approach. *Journal of Computational Finance*, 2(1):5–23, Fall 1998.

[2] D. Beaglehole and A. Chebanier. Mean reverting smiles. *Risk Magazine*, April, 2002.

[3] A. Bensoussan and J. Lions. *Impulse Control and Quasi-Variational Inequalities*. Gaulthier-Villars, Paris, 1984.

[4] H. Bessembinder, J. Coughenour, P. Seguin, and M. Smoller. Mean reversion in equilibrium asset prices: evidence from the futures term structure. *Journal of Finance*, 50(1):361–375, March 1995.

[5] R. Bingham, N.H. Kiesel. *Risk neutral valuation: Pricing and hedging of financial derivatives*. Springer, New York, 1998.

[6] F. Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3:167–179, March 1976.

[7] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–659, 1973.

[8] R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall, London, 2003.

[9] A. Cortazar and E. Schwartz. The valuation of commodity-contingent claims. *The Journal of Derivatives*, 4(1):27–39, Summer 1994.

[10] M. Dahlgren. A continuous time model to price commodity based swing options. *Review of Derivatives Research*, 8(1):27–47, 2005.

[11] S. Deng. Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes. Power Working Paper 073, University Of California Energy Institute, 2000.

[12] D. Duffie, D. Filipovic, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13(3):984–1053, 2003.

[13] K. Eriksson, D. Esteep, P. Hansbo, and J. C. *Computational Differential Equations*. Studentlitteratur, Lund, 1996.

[14] H. Geman. *Commodities and commodity derivatives-Modelling and pricing for agriculturals, metals and energy*. John Wiley & Sons, Chishester, 1 edition, 1005.

[15] H. Geman and A. Roncoroni. Understanding the fine structure of electricity prices. *The Journal of Business*, 79(3), 2006.

[16] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuation. *Econometrica*, 60(1):77–105, 1992.

[17] T. Henker and N. Milonas. Price spread and convenience yield behaviour in the international oil market. *Applied Financial Economics*, 11:23–36, 2001.

[18] J. Hull. *Options, futures and other derivatives*. McGraw-Hill, New York, 6 edition, 1999.

[19] A. Ibáñz. Valuation by simulation of contingent claims with multiple early exercise opportunities. *Mathematical Finance*, 14(2):223–248, 2004.

[20] P. Jaillet, E. Ronn, and S. Tompaidis. Valuation of commodity-based swing options. *Management Science*, 50(7):909–921, 2004.

[21] I. Karatzas and S. Shreves. *Brownian Motion and Stochastic Analysis*. Graduate Texts in Mathematics 113. Springer Verlag, New York, 1989.

[22] I. Karatzas and S. Shreves. *Methods of Mathematical Finance*. Applications of Mathematics 39. Springer Verlag, New York, 1998.

[23] E. Korn and R. Korn. *Option Pricing and Portfolio Optimisation-Modern Methods of Financial Mathematics*. Graduate Studies in Mathematics volume 31. American Mathematical Society, 2000.

[24] S. Lee. Constrained estimation in covariance structures. *Biometrika*, 66(3):539–545, 1979.

[25] J. Lucia and E. Schwartz. Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research*, 5(1):5–50, 2002.

[26] M. Manoliu. Storage options valuation using multilevel trees and calendar spreads. *International Journal of Theoretical and Applied Finance*, 7(4):425–464, 2004.

[27] R. Merton. Option pricing when the underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

[28] B. Øksendahl. *Stochastic differential equations - An introduction with applications*. Springer, New York, 5 edition, 1998.

[29] C. Parsons. Valuaing commodity storage contracts: A two-factor tree approach. Preprint WTM Energy Software LLC, 2005.

[30] M. Pourahmadi. Maximum likelihhod estimation of generalised linear models for multivarate normal covariance matrix. *Biometrika*, 87(2):425–435, 2000.

[31] P. Protter. *Stochastic integration and differential equations - A new approach*. Springer, New York, 1990.

[32] R. Rebonato. *Interest rate option models*. John Wiley & Sons, Chichester, 2 edition, 1998.

[33] E. Ronn. *Real options and energy management-Using Options Methodology to Enhance Capital Budgeting Decisions*. RiskBooks, London, 1 edition, 2002.

[34] K. Sato. *Lévy processes and infinitely divisible distributions*. Cambridge University Press, Cambridge, 1999.

[35] E. Schwartz. The stochastic behavior of commodity prices: implications for valuation and hedging. *Journal of Finance*, 52(3):923–972, July 1997.

[36] E. Schwartz and R. Gibson. Stochastic convenience yield and the pricing of oil contingent claims. *Jornal of Finance*, 45(3):959–976, July 1990.

[37] E. Schwartz and J. Smith. Short-term variations and long term dynamics in commodity prices. *Management Science*, 46(7):893–911, 2000.

[38] K. Singleton. Estimation of affine asset pricing models. *Journal of Econometrics*, 102:111–141, 2001.

[39] P. Wilmott. *Derivatives, The Theory and Practice of Financial Engineering*. John Wiley & Sons, Chichester, university edition, 1998.

Paper I

FAST PRICING OF CLIQUET OPTIONS WITH GLOBAL FLOOR

MATS KJAER

Göteborg University

Abstract. We investigate the pricing of cliquet options with global floor, when the underlying asset follows the Bachelier-Samuelson model. These options have a payoff structure, which is a function of the sum of truncated periodic stock returns over the life span of the option.

Fourier integral formulas for the price and Greeks are derived, and a fast and robust numerical integration scheme for the evaluation of these formulas is proposed. This algorithm seems much faster than quasi Monte Carlo and finite difference techniques for a given level of computational accuracy.

1. Introduction

This Millennium started with a recession and rapidly falling stock markets. Investors who had relied on annual returns on investments exceeding 20% suddenly became aware of the risk inherent in owning shares and many turned their attention to safer investments like bonds and ordinary bank accounts. As an attempt to capitalise on this fear of losses, a variety of equity linked products with capital guarantees were introduced on the market. Among the most successful is the so called *cliquet option with global floor*, which is usually packaged with a bond and sold to retail investors under names *like equity linked bond with capital guarantee* or *equity index bond*.

Today, Quasi Monte Carlo and finite difference methods are the most common methods to compute the price and greeks of these options. Since the payoff is rather complex, these methods are relatively time consuming. In this paper, we propose a Fourier integral method, which seems to be faster than these existing methods for a given level of accuracy. Moreover, the method allows us to compute the Greeks directly, avoiding the finite difference approximations of the partial derivatives often employed in the context of Monte Carlo or finite difference methods.

Smaller banks wanting to offer these structured products to their retail clients may lack the scale to support a separate exotic options desk. A fast computational method could allow these banks to hedge their cliquet options with global floor together with their vanilla options, without suffering from risky delays due to slow computations.

For simplicity, we use the standard Bachelier-Samuelson market model, but in separate notes, we show how the method may be used in connection with some more advanced market models.

Date: April 25, 2006.

Key words and phrases. Exotic equity derivatives, Cliquet options with global floor, Bachelier-Samuelson model, Fourier transforms, splines, numerical integration.

JEL classification: C63, G13.

This paper is organised as follows: Section 2 introduces the type of options considered in this paper and Section 3 fixes notation and introduces the market model. The pricing formula is derived in Section 4 and Section 5 discusses some additional payoffs not considered in Section 2 that may be priced with this methodology. A numerical integration scheme is proposed in Sections 6 to 7. The Monte Carlo and PDE methods used as benchmarks are discussed briefly in Section 8 and pricing examples and results from benchmark tests are presented in Section 9. Finally, Section 10 containing conclusions and suggestions for future research concludes the paper.

2. Cliquet options with global floor

Let T be a future point in time, and divide the interval $[0, T]$ into N subintervals called *reset periods* of equal length $\Delta T_n = T_n - T_{n-1}$, where $\{T_n\}_{n=0}^N, T_0 = 0, T_N = T$ are called the *reset days*. The return of an asset with price process S_t over a reset period $[T_{n-1}, T_n)$ is then defined as

$$R_n = \frac{S_{T_n}}{S_{T_{n-1}}} - 1. \tag{2.1}$$

Truncated returns, $\overline{R}_n = \max(\min(R_n, C), F)$ are returns truncated at some floor and cap levels F and C respectively with $F < C$. Absence of floor and/or cap corresponds to $F = -1$ and $C = +\infty$.

A cliquet option with global floor has a payoff Y at time T of

$$Y = B \times \max(\sum_{n=1}^{N} \overline{R}_n, F_g) \tag{2.2}$$

where the F_g is called the *global floor* and B a notional amount which is set to one for the remainder of this paper. More payoffs that may be priced with the methods in this paper are presented in Section 5.

3. Market model

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a complete filtered probability measure generated by the Wiener process W_t. For simplicity take P to be the risk neutral measure, and expectations under this measure are written $\mathbb{E}$. Under this risk neutral measure, the stock price S_t and bond price B_t are assumed to follow the Bachelier-Samuelson dynamics

$$\begin{cases} dS_t/S_t &= rdt + \sigma dW_t, \\ dB_t/B_t &= rdt, \end{cases} \tag{3.1}$$

for $\sigma > 0$ and $r > 0$. This implies that under P, the returns are independent and of the form

$$R_n \sim e^{a+bX_n} - 1, \tag{3.2}$$

where $X_n \sim N(0, 1)$ and

$$\begin{cases} a &= (r - \frac{\sigma^2}{2})(T_n - T_{n-1}) \\ b &= \sigma_n\sqrt{T_n - T_{n-1}}. \end{cases} \tag{3.3}$$

Assuming that the current reset period is m, i.e. $T_{m-1} \leq t < T_m$, this changes to $R_m \sim (s/\bar{s})e^{a_m + b_m X_m} - 1$ where $s = S_t$, $\bar{s} = S_{T_{m-1}}$, $X_m \sim N(0,1)$ and

$$\begin{cases} a_m &= (r - \frac{\sigma^2}{2})(T_m - t) \\ b_m &= \sigma\sqrt{T_m - t}. \end{cases} \tag{3.4}$$

Below we write R_m^t instead of R_m to indicate that s and $\bar{s}$ are known at time t and define $\overline{R}_m^t = \max(\min(R_m^t, C), F)$.

Note 3.1. We may replace the constants σ and r in (3.1) by the time dependent but deterministic non-negative functions $r(t)$ and $\sigma(t)$. Continuous or discrete dividend yields can also be introduced. If the discrete dividend yields in reset period n are denoted $\alpha_l \in (0,1)$, $1 \leq l \leq L$, the coefficients a and b in (3.3) have to be replaced by

$$\begin{aligned} a_n &= \sum_{l=1}^{L} \log(1 - \alpha_l) + \int_{T_n}^{T_{n+1}} \left(r(t) - \frac{1}{2}\sigma^2(t) \right) dt, \\ b_n &= \sqrt{\int_{T_n}^{T_{n+1}} \frac{1}{2}\sigma^2(t) dt}, \end{aligned}$$

and similar for a_m and b_m in (3.4). Here, reset periods of different lengths are allowed. In this generalised model, the random variables $1 + R_n = \exp(a_n + b_n X_n)$ are still log-normal and independent, but not identically distributed.

Note 3.2. We could also let $S_t = \exp(rt + L_t)$, where L_t is a Levy-process under the chosen risk neutral measure P. However, as noted in Cont and Tankov [5], the monthly or quarterly returns mostly used in our context, appear to be much more normally distributed than daily returns. Thus the benefit of using a Levy-process model would be limited for this type of options.

Note 3.3. Dividends are usually paid in discrete amounts, but using a fixed dividend model, like the one by presented in Heath and Jarrow [9], would leave us without an explicit expression for the density of R_n. To avoid this, fixed dividends must be approximated with discrete or continuous dividend yields.

4. Pricing formulas

In this section we derive integral formulas for the price V_t and greeks of a cliquet option with global floor.

Assuming that $t \in [T_{m-1}, T_m)$, we define the *performance up to date*

$$z = \sum_{n=1}^{m-1} \overline{R}_n, \tag{4.1}$$

and the auxiliary variable $A = (N - m + 1)C - F_g + z$. The characteristic function of a random variable X is written $\varphi_X(\xi) = E[e^{i\xi X}]$.

The form of the price formula can be divided into the following three cases, two of which have trivial solutions, whereas the third one requires a more thorough analysis.

(1) $A \leq 0$: Performance has been so poor that the payoff will be F_g regardless of future share price development.
(2) $MF + z \geq F_g$: Performance has been so good that the payoff will be higher than F_g regardless of future share price development. This results in the analytical formulas for the price and the greeks given in Proposition 4.1.
(3) $A > 0$. General case. The formula in Proposition 4.2 is valid. Case II is included in this case but we prefer to treat it separately due to the existence of the analytical formulas for the price and the greeks in Proposition 4.1.

In Case II, we have a portfolio of forward start performance options, a derivative that pays the holder $Y = \max(\frac{S_T}{S_{T_0}} - K, 0)$ at some time $T > T_0$. Hence it is straightforward to compute formulas for the price V_t and we give it without proof in Proposition 4.1. Here $c(t, s, K, T, \sigma, r)$ denotes the formula for the price at time t of a European call option with strike K and maturity K in the Black-Scholes model (3.1) with parameters r and σ. A derivation may be found in Hull [10].

Proposition 4.1. *If $MF + z \geq F_g$, the price V_t of a cliquet option with global floor is given by*

$$\begin{aligned} V_t &= e^{-r(T-t)}\Big\{z + MF + (N-m)\Big[c(0,1,1+F,\Delta T,\sigma,r) - c(0,1,1+C,\Delta T,\sigma,r)\Big] \\ &\quad + e^{r(T_m - t)}\Big[c(0, s/\bar{s}, 1+F, t-T_m, \sigma, r) - c(0, s/\bar{s}, 1+C, t-T_m, \sigma, r)\Big]\Big\}. \end{aligned}$$

The Greeks are found by taking partial derivatives of V_t in Proposition 4.1.

Before stating and proving the formulas for the price and Greeks in the general case, we introduce the random variables $\tilde{R}_n = C - \overline{R}_n$ and $\tilde{R}^t_m = C - \overline{R}^t_m$, which are non-negative. Furthermore, Φ is the distribution function of a $N(0,1)$ random variable, $\phi = \Phi'$ and the constants a_m, a, b and b_m are given in (3.3) and (3.4).

Proposition 4.2. *If $A > 0$, the price V_t of a cliquet option with global floor is given by*

$$V_t = e^{-r(T-t)}\left\{F_g + A^2 \int_{-\infty}^{\infty} \mathrm{sinc}^2(\frac{\xi A}{2}) \times \varphi_{\tilde{R}^t_m}(\xi) \times \left(\varphi_{\tilde{R}_n}(\xi)\right)^{N-m} \frac{d\xi}{2\pi}\right\} \quad (4.2)$$

where

$$\varphi_{\tilde{R}^t_m}(\xi) = e^{i\xi(C-F)} - i\xi \int_0^{C-F} \Phi\left(\frac{a_m - \log(1+C-x)}{b_m}\right) e^{i\xi x} dx,$$

and

$$\varphi_{\tilde{R}_n}(\xi) = e^{i\xi(C-F)} - i\xi \int_0^{C-F} \Phi\left(\frac{a - \log(1+C-x)}{b}\right) e^{i\xi x} dx.$$

Proof. By general derivatives pricing theory, see for example Bingham and Kiesel [2], the price is given by

$$\begin{aligned} V_t &= e^{-r(T-t)} E[\max \sum_{n=1}^{N} \overline{R}_n, F_g) | \mathcal{F}_t] && (4.3) \\ &= e^{-r(T-t)} E[(F_g + \max(z - F_g + ((\overline{R}^t_m + \sum_{n=m+1}^{N} \overline{R}_n), 0)] && (4.4) \end{aligned}$$

since S_t is a Markov process. Using the relations $\overline{R}_n = C - \tilde{R}_n$ and $\overline{R}^t_m = C - \tilde{R}^t_m$ yields

$$\begin{aligned} V_t &= e^{-r(T-t)}\left\{F_g + E[\max(MC - F_g + z - (\tilde{R}^t_m + \sum_{n=m+1}^{N} \tilde{R}_n), 0)]\right\} \\ &= e^{-r(T-t)}\left\{F_g + E[\max(A - (\tilde{R}^t_m + \sum_{n=m+1}^{N} \tilde{R}_n), 0)]\right\}. \end{aligned}$$

By Fourier analysis, see Folland [8] for details, we have

$$\Lambda_A(x) \equiv \max(A - |x|, 0) = A^2 \int_{-\infty}^{\infty} \operatorname{sinc}^2(\frac{\xi A}{2}) e^{i\xi x} \frac{d\xi}{2\pi}.$$

Using this result with $x = \tilde{R}^t_m + \sum_{n=m+1}^{N} \tilde{R}_n$, which is non-negative by construction, gives

$$\begin{aligned} V_t &= e^{-r(T-t)}\left\{F_g + E[A^2 \int_{-\infty}^{\infty} \operatorname{sinc}^2(\frac{\xi A}{2}) e^{i\xi(\tilde{R}^t_m + \sum_{n=m+1}^{N} \tilde{R}_n)} \frac{d\xi}{2\pi}]\right\} \\ &= e^{-r(T-t)}\left\{F_g + A^2 \int_{-\infty}^{\infty} \operatorname{sinc}^2(\frac{\xi A}{2}) E[e^{i\xi(\tilde{R}^t_m + \sum_{n=m+1}^{N} \tilde{R}_n)}] \frac{d\xi}{2\pi}\right\}, \end{aligned}$$

by the Fubini theorem. Independence of returns and identical distribution of $\{\tilde{R}_n\}_{n=m+1}^{N}$ implies that

$$V_t = e^{-r(T-t)}\left\{F_g + A^2 \int_{-\infty}^{\infty} \operatorname{sinc}^2(\frac{\xi A}{2}) E[e^{i\xi \tilde{R}^t_m}](E[e^{i\xi \tilde{R}_n}])^{N-m} \frac{d\xi}{2\pi}\right\}.$$

To arrive at the formula in Proposition 4.2 it remains to compute $E[e^{i\xi \tilde{R}^t_m}]$ and $E[e^{i\xi \tilde{R}_n}]$. But

$$E[e^{i\xi \tilde{R}_n}] = e^{i\xi(C-F)} \cdot P(R_n \le F) + \int_0^{C-F} e^{i\xi x} dP(C - R_n \le x) + 1 \cdot P(R_n > C),$$

so

$$E[e^{i\xi \tilde{R}_n}] = e^{i\xi(C-F)} - i\xi \int_0^{C-F} \Phi\left(\frac{a - \log(1 + C - x)}{b}\right) e^{i\xi x} dx,$$

by (3.2) and partial integration. $E[e^{i\xi \tilde{R}^t_m}]$ is computed analogously. □

The method uses the independence of returns to transform the $(N - m + 1)$-dimensional integral of (4.3) into the set of one dimensional integrals of Proposition 4.2, which may be faster to compute if $(N - m + 1)$ is large enough.

Since differentiation is allowed inside the integral (4.2), expressions similar to (4.2) may be obtained for the greeks in case 3.

Note 4.3. Similar Fourier integral formulas could be derived for the extended market model discussed in Note 2. Since the returns are not identically distributed, all $N - m$ different characteristic functions $\varphi_{\tilde{R}_n}(\xi)$ would have to be evaluated, increasing the computational burden. The approach also works for the Levy-process model of Note 3.

Note 4.4. An alternative approach would be to note that due to the independence of returns, the the density function of $\sum_{n=1}^{N} \overline{R}_n$ is given by the inverse Fourier transform of $(\varphi_{\overline{R}_n})^N$. Knowing this density, the option price $V_t = e^{-r(T-t)}\mathbb{E}[\max(\sum_{n=1}^{N} \overline{R}_n, F_g)|\mathcal{F}_t]$ could be computed by numerical integration. However, $\varphi_{\overline{R}_n}$ are not known explicitly and by (6.1) do not go to zero as $|\xi| \to \infty$, making numerical inversion difficult.

5. Extension to other payoff functions

The methodology used to derive the price formula in Proposition 4.2 can be used to price other related derivatives.

If $C = \infty$, the formulas in Proposition 4.2 is not valid. However, by inserting a large virtual cap C, they can be used to obtain arbitrarily good approximations and an upper bound of the truncation error is given in Proposition 5.1 below. Here $\hat{R}_n = \max(R_n, F)$ is a truncated return with $C = \infty$.

Proposition 5.1. *Let V_t^C and V_t^∞ be the price of floored cliquet options with local caps $C < \infty$ and $C = \infty$ respectively. Then with $\Delta T = T_{n+1} - T_n$*

$$\begin{aligned} V_t^\infty - V_t^C \;\leq\; & 2e^{-r(T-t)}\Big\{e^{r(T_m-t)}c(0, s/\bar{s}, 1+C, T_m - t, \sigma, r) \\ & +(N-m)e^{r\Delta T}c(0, 1, 1+C, \Delta T, \sigma, r)\Big\}. \end{aligned}$$

Proof. Following the first steps of the derivation of Proposition 4.2, the truncation error can be written as

$$\begin{aligned} (V_t^\infty - V_t^C)/e^{-r(T-t)} &= E[\max(\sum_{n=1}^{N} \hat{R}_n, F_g)|\mathcal{F}_t] - E[\max(\sum_{n=1}^{N} \overline{R}_n, F_g)|\mathcal{F}_t] \\ &= E[\max(\hat{R}_m^t + \sum_{n=m+1}^{N} \hat{R}_n, F_g - z) \\ &\quad - \max(\overline{R}_m^t + \sum_{n=m+1}^{N} \overline{R}_n, F_g - z)]. \end{aligned}$$

If $x \geq y$ we have

$$\max(x,a) - \max(y,a) = \begin{cases} 0, & x \leq a, \\ x - a, & x \geq a \geq y, \\ x - y, & y \geq a. \end{cases}$$

Using this with $X = \hat{R}_m^t + \sum_{n=m+1}^{N} \hat{R}_n$, $Y = \overline{R}_m^t + \sum_{n=m+1}^{n} \overline{R}_n$ and $a = F_g - z$ yields

$$\begin{aligned} (V_t^\infty - V_t^C)/e^{-r(T-t)} &= E[X - a; X \geq a \geq Y] + E[X - Y; Y \geq a] \\ &\leq E[X - Y; X \geq a \geq Y] + E[X - Y; Y \geq a] \\ &\leq 2E[X - Y] \\ &= 2E[(\hat{R}_m^t - \overline{R}_m^t) + \sum_{n=m+1}^{N} \hat{R}_n - \sum_{n=m+1}^{N} \overline{R}_n]. \end{aligned}$$

where the inequality follows from the fact that $a \geq Y$ on $X \geq a \geq Y$ and the integrands are non-negative. Computing the expectation and identifying the Black-Scholes call option price formula completes the proof. □

It is also possible to derive a formula similar to that of Proposition 4.2 if a global cap C_g is added, in which case the holder of the derivative receives

$$Y = \min(\max(\sum_{n=1}^{N} \overline{R}_n, F_g), C_g)$$

at time T. To see this, note that

$$\begin{aligned} Y &= \min(\max(z + \overline{R}_m^t + \sum_{n=m+1}^{N} \overline{R}_n, F_g), C_g) \\ &= F_g + \Lambda_A(\tilde{R}_m^t + \sum_{n=m+1}^{N} \tilde{R}_n) - \Lambda_{A-C_g+F_g}(\tilde{R}_m^t + \sum_{n=m+1}^{N} \tilde{R}_n) \end{aligned}$$

and proceed as in the proof of Proposition 4.2. Algebraic manipulations of this type allow us to price other cliquet-style derivatives that appear on the market, for example the *cliquet with global floor and coupon credit* K and the *reversed cliquet*, which pay the holder $Y = \max(\sum_{n=1}^{N} \overline{R}_n - K, F_g)$ and $Y = \max(C_g + \sum_{n=1}^{N} R_n^-, F_g)$ respectively.

6. Numerical computation of the characteristic functions

To compute the pricing formula given in Proposition 4.2, we must compute the characteristic functions

$$E[e^{i\xi\tilde{R}_n}] = e^{i\xi(C-F)} - i\xi \int_0^{C-F} \Phi\left(\frac{a_n - \log(1 + C - x)}{b_n}\right) e^{i\xi x} dx \qquad (6.1)$$

for each ξ. Due to the rapid oscillation of the integrand inside (6.1) for large ξ, this would be computationally very heavy if done directly by numerical integration. The monotonicity and high degree of smoothness of $\Phi(\frac{a_n - \log(1+C-x)}{b_n})$ suggests that interpolation with complete cubic splines over the interval $[0, C - F]$ may be a good idea. Initially this interval is divided into N_p equally long subintervals $[x_n, x_{n+1}]$, $n = 0, \ldots, N_p$ and a cubic polynomial $p_3^{(n)}(x) = c_3^{(n)}x^3 + c_2^{(n)}x^2 + c_1^{(n)}x + c_0^{(n)}$ is assigned to each interval. The coefficients are then chosen such that they interpolate the function at the *spline knots* x_n, $n = 0, \ldots, N_p + 1$ and have continuous first and second derivatives. In addition, we require that the derivative of the spline and the function to be interpolated coincide at the endpoints 0 and $C - F$. For more details about complete cubic spline construction, see De Boor [6] pp. 53-55.

To summarise, the cubic spline approximation $\hat{\Phi}$ of Φ can be written as

$$\hat{\Phi}\left(\frac{a - \log(1 + C - x)}{b}\right) = \sum_{n=0}^{N_p-1} \chi_{[x_n x_{n+1}]}(x) p_3^{(n)}(x),$$

with χ being the indicator function. Replacing Φ by $\hat{\Phi}$ in (6.1) and evaluating the integrals yield an approximation $\hat{\varphi}_{\tilde{R}_n}(\xi)$ of $\varphi_{\tilde{R}_n}(\xi)$ as

$$\begin{aligned}\hat{\varphi}_{\tilde{R}_n}(\xi) &= \sum_{n=0}^{N_p-1}\left\{c_3^{(n)}\left[\frac{e^{i\xi x}}{(i\xi)^4}((i\xi x)^3-3(i\xi x)^2+6i\xi x-6)\right]_{x_n}^{x_{n+1}}\right.\\ &\quad +c_2^{(n)}\left[\frac{e^{i\xi x}}{(i\xi)^3}((i\xi x)^2-2i\xi x+2)\right]_{x_n}^{x_{n+1}}\\ &\quad \left.+c_1^{(n)}\left[\frac{e^{i\xi x}}{(i\xi)^2}(i\xi x-1)\right]_{x_n}^{x_{n+1}}+c_0^{(n)}\left[\frac{e^{i\xi x}}{i\xi}\right]_{x_n}^{x_{n+1}}\right\}.\end{aligned}$$

Despite its horrible appearance, the formula is very fast to evaluate on a computer. To compute the distribution function of a normal random variable at the spline knots, a fractional approximation proposed in Hull [10] is used, which promises five to six correct decimals with little computational effort.

The next proposition states that $\hat{\varphi}$ converges to φ uniformly. We start by stating a lemma, which proof can be found in De Boor [6] on pp. 68-69.

Lemma 6.1. *If $f(x) \in C^{(4)}$, $h = x_n - x_{n-1}$ and $p_3(x)$ is the cubic spline approximation of f on $[a, b]$, then*

$$|f'(x) - p_3'(x)| \leq \frac{h^3}{24}\sup_{x\in[a,b]}\left|\frac{d^4 f}{dx^4}\right|.$$

Proposition 6.2. *Let $\hat{\varphi}_{\tilde{R}_n}(\xi)$ be the approximation of $\varphi_{\tilde{R}_n}(\xi)$ and N_p the number of spline intervals of length $h = (C-F)/N_p$. Then $\hat{\varphi} \to \varphi$ uniformly in ξ when $h \to 0$. More specifically we have that*

$$|\hat{\varphi}_{\tilde{R}_n}(\xi) - \varphi_{\tilde{R}_n}(\xi)| \leq \frac{h^3}{24}(C-F)\sup_{x\in[0,C-F]}\left|\frac{d^4}{dx^4}\Phi\left(\frac{a-\log(1+C-x)}{b}\right)\right|.$$

Proof. Let $E(x) = \hat{\Phi}\left(\frac{a-\log(1+C-x)}{b}\right) - \Phi\left(\frac{a-\log(1+C-x)}{b}\right)$. Then by Lemma 6.1

$$\begin{aligned}|\hat{\varphi}_{\tilde{R}_n}(\xi) - \varphi_{\tilde{R}_n}(\xi)| &= \left|-i\xi\left[E(x)\frac{e^{i\xi x}}{i\xi}\right]_0^{C-F} + i\xi\int_0^{C-F}E'(x)\frac{e^{i\xi x}}{i\xi}dx\right|\\ &= \left|\int_0^{C-F}E'(x)e^{i\xi x}dx\right|\\ &\leq \frac{h^3}{24}(C-F)\sup_{x\in[0,C-F]}\left|\frac{d^4}{dx^4}\Phi\left(\frac{a_n-\log(1+C-x)}{b_n}\right)\right|\end{aligned}$$

by partial integration. Here we have also used the fact that $x = 0$ and $x = C - F$ are points of interpolation with zero error. □

7. A numerical integration scheme

In this section we develop a numerical integration scheme for computation of the pricing formula in Proposition 4.2, which uses the method for computing the characteristic functions proposed in Section 6.

First, the real part of the integrand is even, the imaginary part is odd and the domain of integration is symmetric we have that

$$\begin{aligned} V_t &= e^{-r(T-t)}\Big\{F_g + A^2\int_{-\infty}^{\infty} \mathrm{sinc}^2(\frac{\xi A}{2}) \times \varphi_{\tilde{R}^t_m}(\xi) \times \big(\varphi_{\tilde{R}_n}(\xi)\big)^{N-m}\frac{d\xi}{2\pi}\Big\} \\ &= e^{-r(T-t)}\Big\{F_g + A^2\int_{0}^{\infty} \mathrm{sinc}^2(\frac{\xi A}{2}) \times \mathrm{Re}\,\big\{\varphi_{\tilde{R}^t_m}(\xi) \times \big(\varphi_{\tilde{R}_n}(\xi)\big)^{N-m}\big\}\frac{d\xi}{\pi}\Big\} \quad (7.1)\end{aligned}$$

Only having to integrate the real part over half of the domain, reduces the number of computations by 75%. Since differentiating with respect to a parameter and taking real parts commute, this type of reduction extends to the computation of the greeks as well.

In order to compute the price integral numerically, an artificial upper limit of integration $\bar{\xi}$ is needed. Characteristic functions have a modulus less or equal to one which together with the fact that $\mathrm{sinc}^2(A\xi/2) \geq 0$ gives the following estimate of the truncation error $e(\bar{\xi})$.

$$\begin{aligned} |e(\xi_{max})| &= e^{-r(T-t)}A^2\Big|\int_{\bar{\xi}}^{\infty} \mathrm{sinc}^2(\frac{\xi A}{2}) \times \mathrm{Re}\,\big\{\varphi_{\tilde{R}^t_m}(\xi) \times \varphi_{\tilde{R}_n}(\xi)\big\}\frac{d\xi}{\pi}\Big| \\ &\leq e^{-r(T-t)}\frac{2A}{\pi}\int_{\bar{\xi}A/2}^{\infty} \mathrm{sinc}^2(x)dx. \quad (7.2)\end{aligned}$$

The integral (7.2) is computed numerically for different values of $A\bar{\xi}/2$ and presented in the table below.

$A\xi/2$	10	20	50	100	200	400
$\int_{\bar{\xi}A/2}^{\infty}\mathrm{sinc}^2(x)dx$	0.0521	0.0254	0.0099	0.0040	0.0022	0.0010

Table 1: Truncation errors of the integral (7.2).

Denoting the integrand of (7.1) by ψ yields

$$V_t = e^{-r(T-t)}\left\{F_g + \frac{A^2}{\pi}\int_0^{\infty}\psi(\xi)d\xi\right\}.$$

This integral is then truncated at $\bar{\xi}$, which is set using Table 1 above and approximated with the well known trapezoid rule of numerical quadrature.

$$\int_0^{\bar{\xi}}\psi(\xi)d\xi \approx \sum_{n=0}^{N-1}\left(\frac{\psi(\xi_n)+\psi(\xi_{n+1})}{2}\right)(\xi_{n+1}-\xi_n).$$

Instead of placing the nodes ξ_n uniformly, we try to select them such that the magnitude of the quadrature error contribution e_n from each interval $[\xi_n\ \xi_{n+1}]$ is bounded by some tolerance level ϵ. Starting at $\xi_0 = 0$, this is done iteratively as follows.

According to Eriksson et. al. [7], e_n is bounded by

$$|e_n| \leq \frac{(\xi_{n+1}-\xi_n)^3}{12}\sup_{\xi\in[\xi_n,\xi_{n+1}]}|\psi''(\xi)|.$$

If we require $|e_n| < \epsilon$, a rule for the step length $\Delta\xi_n$ can be obtained as

$$\Delta\xi_n = \left(\frac{12\epsilon}{\sup\limits_{\xi\in[\xi_n,\xi_{n+1}]} |\psi''(\xi)|} \right)^{1/3}.$$

The second derivative $\psi''(\xi)$ is approximated with

$$\psi''(\xi) \approx \frac{\psi(\xi + d\xi) - 2\psi(\xi) + \psi(\xi - d\xi)}{(d\xi)^2}$$

where $d\xi$ is some small number. We also replace $\sup_{\xi\in[\xi_n\xi_{n+1}]} |\psi''(\xi)|$ by $|\psi''(\xi_n)|$, which is justified if the second derivative does not change too much over the interval $[\xi_n, \xi_{n+1}]$.

8. Reference methods

In Section 9, the Fourier method proposed in Sections 4 to 7 will be compared with the following existing pricing methods.

(1) Monte Carlo (MC) simulation using pseudo random numbers.
(2) Quasi Monte Carlo (QMC) using a Fauré sequence.
(3) Partial Differential Equation (PDE) approach using an explicit finite difference (FD) scheme.

An overview of the usage of Monte Carlo and quasi Monte Carlo methods in Finance can be found in Boyle et. al. [3] and Boyle et. al. [4] respectively.

The PDE approach may need some explanations. Similar to the case of discrete Asian options, which is covered in Andreasen [1], it can be proved that the PDE in Proposition 8.1 below holds for the cliquet option with global floor.

Proposition 8.1. *The price $V_t = V(t, s, \bar{s}, z)$ satisfies the partial differential equation*

$$\begin{cases} \frac{\partial V}{\partial t} + \frac{\sigma^2 s^2}{2}\frac{\partial^2 V}{\partial s^2} + rs\frac{\partial V}{\partial s} - rV = 0, \quad T_{n-1} \le t < T_n \\ V(T_N, s, \bar{s}, z) = \max(z, F_g) \\ V(T_n^-, s, \bar{s}, z) = V(T_n, s, s, z + \max(\min(s/\bar{s} - 1, C), F)), \quad 1 \le n \le N. \end{cases}$$

By letting $x = \log(s/\bar{s})$, a PDE with the space dimensions x and z can be derived. This equation is then solved by an explicit finite difference scheme.

Note 8.2. Both the MC, QMC and PDE methods are directly extendable to the extended market model discussed in Note 2, without further computational effort.

9. Numerical results

In order to rank the methods, we compare their accuracy for a given time of computation for two sample derivatives, specified in Table 2 below. They have both existed on the Swedish market.

Option	T	N	F_g	F	C
Cliquet 1	3 years	12	0	-0.05	0.05
Cliquet 2	3 years	36	0	-0.02	0.02

Table 2: Characteristics of two sample cliquet options.

For each option, we compute the price, theta and delta at $t = 0$ and insert these into the Black-Scholes PDE in Proposition 8.1 to obtain the gamma for free. For the Fourier method, the greeks are obtained from evaluation of the integral formulas obtained by differentiating inside the integral of the pricing formula of Proposition 4.2. Finite difference approximations are used to estimate the greeks in the reference methods.

We start by giving some pricing examples for different volatilities when $r = 0.05$.

Option	σ	V	Δ	Θ	Γ
Cliquet 1	0.10	0.0952	0.4451	-0.01008	-1.484
Cliquet 1	0.30	0.0566	0.1154	-0.00167	-0.102
Cliquet 1	0.50	0.0426	0.0567	0.00385	-0.0366
Cliquet 2	0.10	0.0717	0.3339	-0.00602	-1.419
Cliquet 2	0.30	0.0401	0.0804	0.00138	-0.0755
Cliquet 2	0.50	0.0300	0.0398	0.00276	-0.0258

Table 3: Prices and Greeks for at $t = 0$ for Cliquet 1 and 2. The interest rate is $r = 0.05$ per year

All methods are implemented in the C programming language and compiled to a DLL file that is called from a test routine written in Python. Computations are made on a Dell Inspirion 8200 laptop with a 1.6 GHz Pentium® m4 processor and a 256 MB RAM.

For the Monte-Carlo and quasi Monte Carlo methods, the standard error estimated from 100 samples has been used to measure accuracy. In order to obtain this estimate for the quasi Monte Carlo method, a rotation modulo one randomisation is applied to the original Fauré sequence. This method is described in Tuffin [11], where it is used in connection with Fauré sequences for the first time.

The implementation of the Fourier method used in these tests allows the usage of the extended model described in Note 2 of Section 4. A consequence of this is that all the N characteristic functions have to be evaluated, while an implementation where only two characteristic functions have to be evaluated, would be much faster. As remarked in Note 7, the computational effort for the reference methods is not affected significantly by this extension.

The results shown in Figures 1 to 2 refer to the time needed to compute price, the delta, the theta.

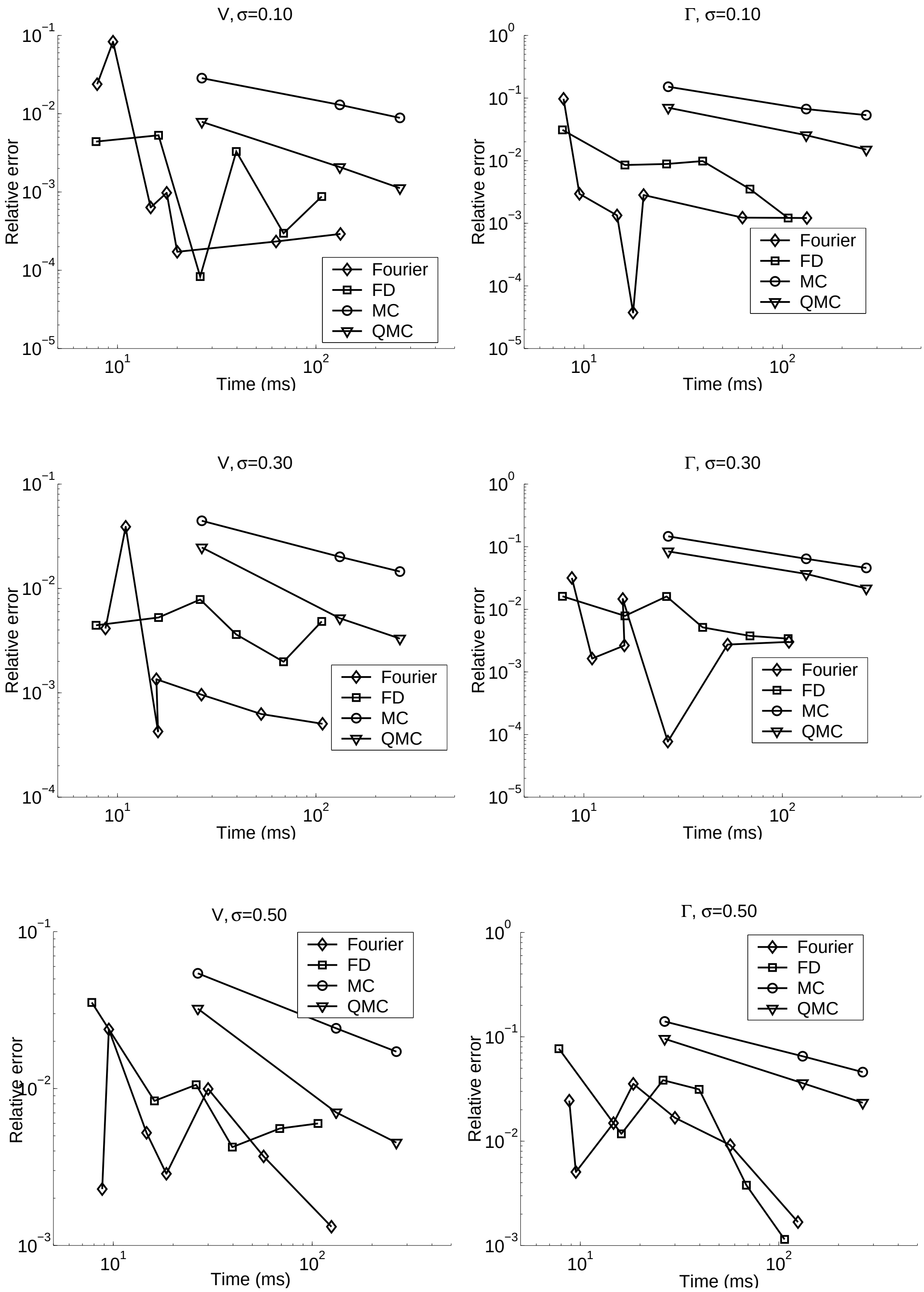

Figure 1: Cliquet 1: Price and gamma for the three methods for different volatilities. Flat interest rate $r = 0.05$.

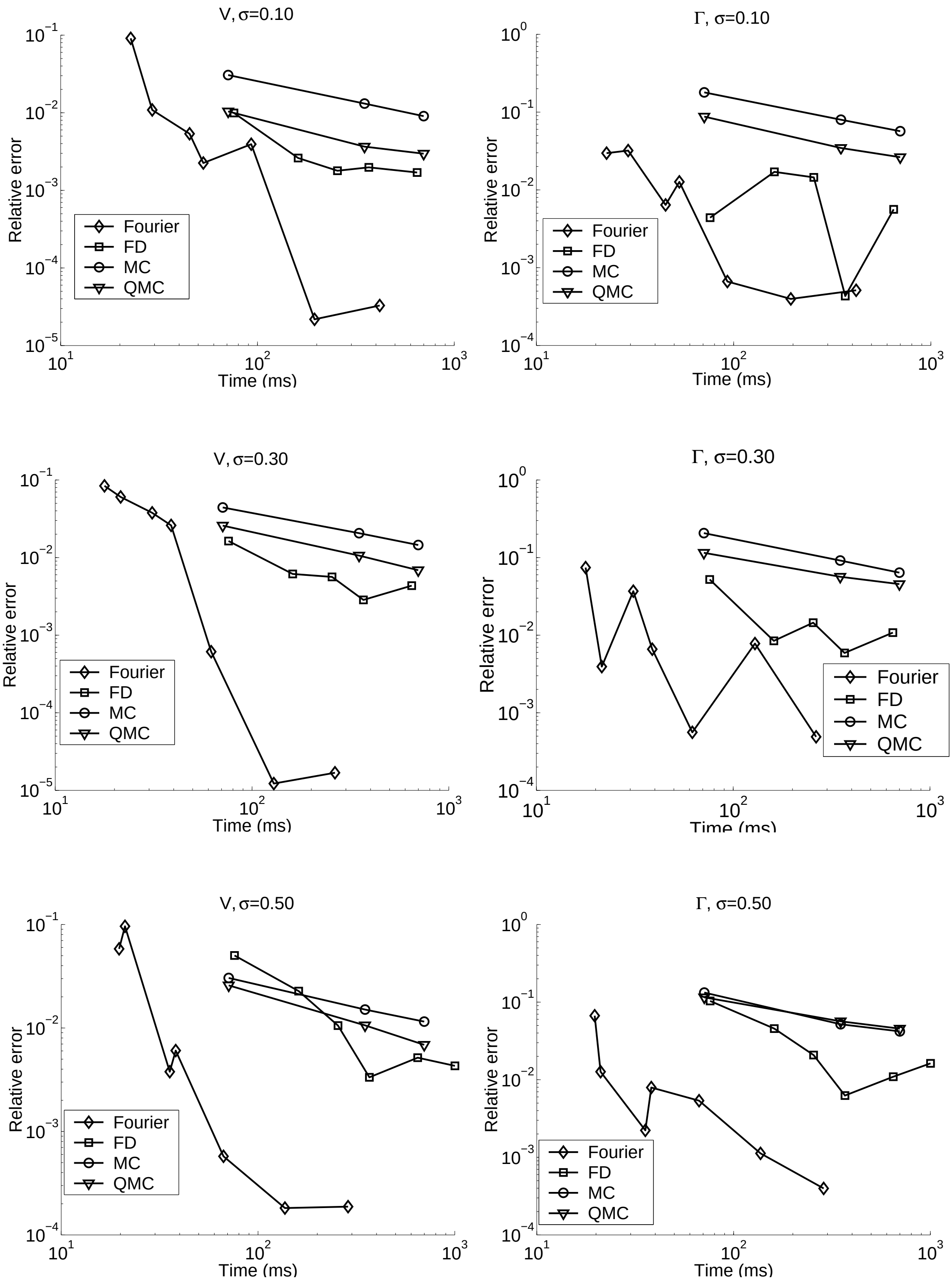

Figure 2: Cliquet 2: Price and gamma for the three methods for different volatilities. Flat interest rate $r = 0.05$.

10. Concluding remarks

Based on the results in Figures 1 to 2, the Fourier integral method outperforms the Monte Carlo and quasi Monte Carlo methods in these two test cases. Compared to the finite difference method it seems particularly fast for case when $N = 36$. This holds even in the presence of dividends, reset periods of unequal length and time dependent interest rate and volatility. The efficiency is achieved by converting the the computation of a multi dimensional integral into the set of one dimensional integrals.

References

[1] J. Andreasen. The pricing of discretely sampled asian options: A change of numeraire approach. *Journal of Computational Finance*, 2(1), Fall 1998.

[2] R. Bingham, N.H. Kiesel. *Risk neutral valuation: Pricing and hedging of financial derivatives.* Springer, New York, 1998.

[3] P. Boyle, M. Broadie, and P. Glasserman. Monte-carlo methods for security pricing. *The Journal of Economic Dynamics & Control*, 21(8):1267–1321, 1997.

[4] P. Boyle, C. Joy, and K. Tan. Quasi monte-carlo methods in numerical finance. *Management Science*, 42:926–938, 1996.

[5] R. Cont and P. Tankov. *Financial Modelling with Jump Processes.* Chapman & Hall, London, 2003.

[6] C. De Boor. *A practical guide to splines.* Springer, New York, 1978.

[7] K. Eriksson, D. Esteep, P. Hansbo, and J. C. *Computational Differential Equations.* Studentlitteratur, Lund, 1996.

[8] G. Folland. *Fourier analysis and its applications.* Brooks-Cole Publishing Company, Pacific Grove, 1992.

[9] R. Heath and P. Jarrow. Ex-dividend stock price behaviour and arbitrage opportunities. *The Journal of Business*, 61:95–108, 1988.

[10] J. Hull. *Options, futures and other derivatives.* McGraw-Hill, New York, 6 edition, 1999.

[11] B. Tuffin. On the use of low-discrepancy sequences in monte-carlo methods. Technical Report no. 1060, I.R.I.S.A, Rennes, France, 1996.

Department of Economics, School of Business, Economics and Law, Göteborg University, Box 600, S-405 30 Göteborg, Sweden

E-mail address: mats.kjaer@economics.gu.se

Paper II

PRICING OF SWING OPTIONS IN A MEAN REVERTING MODEL WITH JUMPS

MATS KJAER

Göteborg University

Abstract. We investigate the pricing of swing options in a model where the logarithm of the spot price is the sum of a deterministic seasonal trend and an Ornstein-Uhlenbeck process driven by a jump diffusion.

First we calibrate the model to Nord Pool electricity market data. Second, the existence of an optimal exercise strategy is proved, and we present a numerical algorithm for computation of the swing option prices. It involves dynamic programming and the solution of multiple parabolic partial integro-differential equations by finite differences.

Numerical results show that adding jumps to a diffusion may result in 2-35% higher swing option prices, depending on the moneyness and timing flexibility of the option.

1. Introduction

One type of derivative that is common on the electric power and natural gas markets, is the *swing option*. It allows flexibility in delivery with respect to both the timing and amount of energy delivered. For many years, it was available in the over the counter (OTC) markets, before its complexity was fully understood. This paper is about the pricing of swing options, with examples taken from the Nord Pool electricity market. However, the proposed pricing framework is applicable on other commodity markets as well.

To the best of our knowledge, the first paper on this topic is the 1995 paper by Thompson [19]. He proposes a lattice based algorithm to price take-or-pay contracts, which is a simple type of swing option. Generalising this approach, Jaillet, Ronn and Tompaidis [12] propose a multi-level lattice method to price swing options on natural gas. They use a trinomial tree discretisation of a continuous time model, where the logarithm of the spot price is a one-factor mean reverting process driven by a Wiener process. This is one of the models used by Lucia and Schwartz [14] to price electricity forwards. Recently, comparably efficient Monte Carlo methods for American options have been developed. These are applied by Ibáñz [11] in the context of swing options, also assuming the model by Lucia and Schwartz [14]. A different approach is taken by Dahlgren [5]. He works in a general one-factor diffusion setting, and shows that under some technical conditions on the drift and

Date: April 26, 2006.

Key words and phrases. Energy derivatives, Swing options, Jump diffusions, parabolic PIDEs, Finite differences.

JEL classification: C13, C61, C63, C65, G13.

This work was supported by the Jan Wallander and Tom Hedelius Foundation.

volatility, the pricing problem can be transformed into solving a set of variational inequalities.

This paper aims at extending the papers cited above in three directions. First, we will allow discontinuous spot price trajectories. Second, the amount of electricity to be delivered is chosen from a closed interval, rather than from a discrete set. Third, at each exercise date, the swing option holder has to fix a vector of amounts for multiple deliveries rather than a scalar amount for a single delivery.

More specifically, we use a model by Deng [7], where the logarithm of the spot price is the sum of a seasonality term and an Ornstein-Uhlenbeck process driven by a jump diffusion. To start with, we note that in the chosen spot price model, the price of a simple European derivative is given by the unique solution to a parabolic partial integro-differential equation (PIDE). Next, we prove the existence of an optimal exercise strategy, and as in the papers cited above, the proof is based on dynamic programming. By combining these two results, we obtain a numerical method for the pricing of swing options. To solve the resulting PIDEs numerically, we modify an operator splitting finite difference method proposed in Cont and Tankov [4].

For model estimation, we use historical Nord Pool spot price data and employ a combination of the least squares method presented in Lucia and Schwartz [14], and the Fourier transform based maximum likelihood approach by Singleton [18]. Traded forwards are used to estimate the market price of risk.

This paper is organised as follows: In Section 2, we briefly discuss how Nord Pool works, followed by some general mathematical assumptions and notation in Section 3. The spot price model is introduced in Section 4, where derivatives pricing is also discussed. Section 5 discusses model estimation and Section 6 introduces swing options. Pricing of these options and a proof of the existence of an optimal exercise strategy is covered in Section 7. Implementation of the numerical algorithms is the topic of Section 8, and parameter estimation and swing option pricing results are presented in Section 9. Finally, Section 10 concludes the paper and gives suggestions for future research.

2. Nord Pool

The spot market on Nord Pool is an auction based day ahead market, where suppliers and consumers from the entire Nordic region place bids for each individual hour during the next day. Bids on eight hour blocks are also allowed. By looking at where supply meets demand, a so-called *system price* is calculated for the entire region. This is not a trivial process since the introduction of block bidding.

The average system price over 24 hours is called the *base load price*. Most derivatives are written on the base load price, from now on referred to as the *spot price*. In particular, forwards for delivery of 1 MWh at a constant load during one day, one week, one month and longer periods are available. These start trading six days, seven weeks and five months prior to the first delivery date, and are financially settled with daily net payments during delivery, making them swaps rather than forwards. Trading is mainly concentrated to those contracts with longer delivery periods. In addition, put and call options on these forwards are offered, but their liquidity is in general very low. Consequently, potential model calibration procedures do not have access to reliable forward curves or implied volatility surfaces. More information about Nord Pool is available at `http://www.nordpool.com`.

3. General assumptions and notation

Before introducing the spot price model, we make some general assumptions of a mathematical nature and fix some notation.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space equipped with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions as defined in Protter [17]. We refer to P as the historical or real world probability measure, and expectations with respect to this measure are denoted $\mathbb{E}_P$. To begin with, the spot price of electricity $S_t, t \geq 0$, is assumed to be a positive semi-martingale belonging to this filtered probability space, but its dynamics will be specified further in Section 4. In addition, it is assumed that the no-arbitrage conditions for the fundamental theorem of asset pricing (see Delbaen and Schachermayer [6] for details) are fulfilled. Hence there exists at least one risk neutral probability measure Q. Risk neutral expectations are simply denoted $\mathbb{E}$.

We also introduce the riskless bank account B_t with dynamics

$$B_t = B_0 e^{rt},$$

for some fixed $r > 0$.

Having fixed a risk neutral measure Q, general derivatives pricing theory (see for example Bingham and Kiesel [3]) shows that the arbitrage free price $V(t)$ of a derivative paying $Y \in L^1(\Omega, \mathcal{F}_T, Q)$, Y bounded from below by a constant, at time $T > t$, is given by

$$V(t) = e^{-r(T-t)}\mathbb{E}[Y|\mathcal{F}_t]. \tag{3.1}$$

Forward prices $F(t,T)$ at time t for delivery at time $T \geq t$ are defined by

$$F(t,T) = \mathbb{E}[S_T|\mathcal{F}_t],$$

and *forward contract prices* $G(t,T)$ by

$$G(t,T) = e^{-r(T-t)}(F(t,T) - K),$$

where K is the strike price. In the case when S_t is a Markov process under Q, we will sometimes write $G(t,T,s)$ and $F(t,T,s)$ to emphasise the dependence on the current spot price $s = S_t$.

4. Spot price model and derivatives pricing

Electric power differs from most other commodities in that at this moment of writing, it is not possible to store large amounts of electric energy in a feasible manner. This means that financial derivatives written on spot electricity cannot be hedged by non-producers. Inelastic demand and absence of stocks to smooth supply shocks result in a price dynamics characterised by seasonality, mean reversion and sudden spikes as seen in Figure 1 below. The model introduced in this section tries to catch some of these features, while still being possible to calibrate in the absence of a liquid vanilla options market.

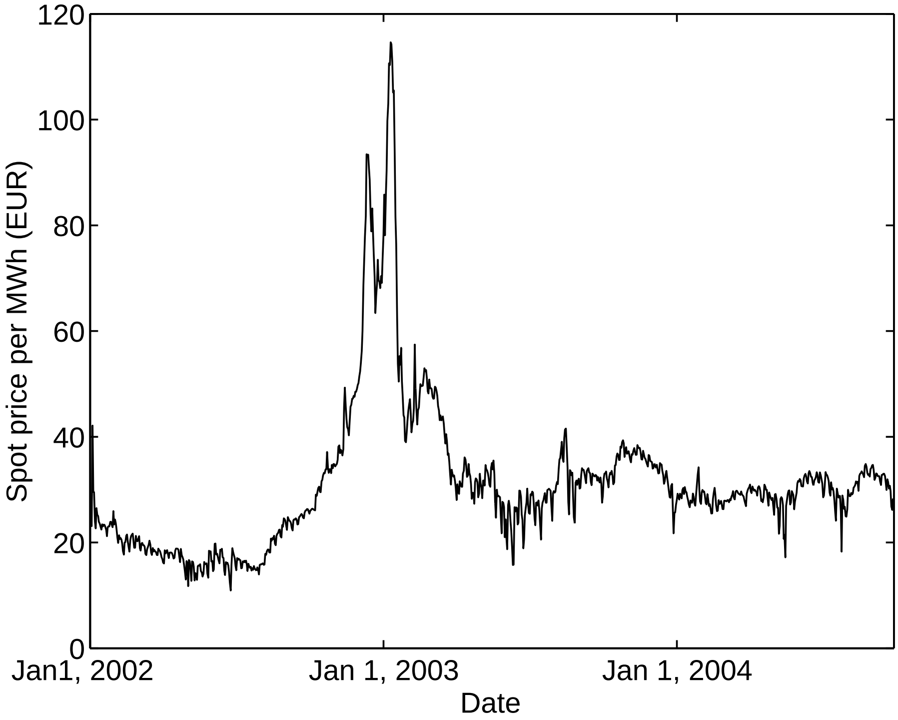

Figure 1: Nord Pool spot price (daily average of system price) from 1 January 2002 to 30 September 30 2004

As a model for S_t, we will use a slightly modified version of the one presented by Deng [7], from now on called the *Deng-model.* Here the spot price follows the P-dynamics

$$\begin{cases} S_t &= \exp(f(t) + X_t) \\ dX_t &= -\alpha X_{t^-} dt + dL_t, \end{cases} \tag{4.1}$$

where $\alpha > 0$ is fixed, $f(t)$ is a deterministic seasonal trend, and L_t is a *compensated* jump diffusion belonging to the filtered probability space defined in Section 3. More specifically, let $\{W_t\}_{t\geq 0}$ be a $P-$Wiener process, $\sigma > 0$ be fixed, and $\{U_t^J\}_{t\geq 0}$ a compound Poisson process independent of $\{W_t\}_{t\geq 0}$, with jump size density f_J under P. Then

$$L_t = \sigma W_t + U_t^J - \lambda_J \mathbb{E}_P[J]t.$$

In addition, we assume that f_J satisfies

$$\int_{\mathbb{R}} e^{2y} f_J(y)dy < \infty, \tag{4.2}$$

which according to Cont and Tankov [4] is sufficient for S_t to have finite second moments under P for all $t \geq 0$.

Setting $\lambda_J = 0$ in (4.1) retrieves the model by Lucia and Schwartz [14], from now on referred to as the *LS-model.* It will be used as a reference model to study the impact on swing option prices from the introduction of jumps.

The solution X_t^{x,t_0} to the SDE (4.1) started at x at time $t_0 < t$ is given by

$$X_t^{x,t_0} = xe^{-\alpha(t-t_0)} - \frac{\lambda_J \mathbb{E}_P[J]}{\alpha}(1 - e^{-\alpha(t-t_0)}) + \sigma \int_{t_0}^t e^{-\alpha(t-s)} dW_s + \sum_{i=N_{t_0}}^{N_t} e^{-\alpha(t-t_i)} J_i, \tag{4.3}$$

where N_t is the number of jumps of U_t^J in $[0,t)$, J_i the jump sizes and t_i the jump times. The density of X_t^{x,t_0} is generally not known explicitly, but the transform analysis by Duffie, Filipovic and Schachermayer [8] shows that the conditional characteristic function under P of $X_{t+\tau}$, $\tau > 0$, given $X_t = x$ is

$$\begin{aligned} \varphi_{X_{t+\tau}^{x,t}}(u) &= \exp\left(iuxe^{-\alpha\tau} - \frac{\sigma^2 u^2}{4\alpha}(1 - e^{-2\alpha\tau})\right) \\ &\quad \times \exp\left(\lambda_J \int_0^\tau \mathbb{E}_P\left[\exp(iuJe^{-\alpha s}) - (1 + iuJe^{-\alpha s})\right]\right), \quad u \in \mathbb{R}. \end{aligned} \tag{4.4}$$

In most cases the integral inside (4.4) cannot be evaluated in closed form, but has to be computed numerically. One exception is the two-sided exponential jump size distribution, with density

$$f_J(x|\lambda_1, \mu_1, \lambda_2, \mu_2) = \begin{cases} \frac{\lambda_1}{\lambda_1+\lambda_2} \frac{e^{-x/\mu_1}}{\mu_1}, & x \geq 0 \\ \frac{\lambda_2}{\lambda_1+\lambda_2} \frac{e^{x/\mu_2}}{\mu_2}, & x < 0. \end{cases} \tag{4.5}$$

Here all parameters are positive and $\lambda_J = \lambda_1 + \lambda_2$. Moreover, (4.4) simplifies to

$$\begin{aligned} \varphi_{X_{t+\tau}^{x,t}}(u) &= \exp\left(iu\{xe^{-\alpha\tau} - \tfrac{\lambda_1\mu_1 - \lambda_2\mu_2}{\alpha}(1 - e^{-\alpha\tau})\} - \tfrac{\sigma^2 u^2}{4\alpha}(1 - e^{-2\alpha\tau})\right) \\ &\quad \times \left(\tfrac{1 - iu\mu_1 e^{-\alpha\tau}}{1 - iu\mu_1}\right)^{\frac{\lambda_1}{\alpha}} \times \left(\tfrac{1 + iu\mu_2 ie^{-\alpha\tau}}{1 + iu\mu_2}\right)^{\frac{\lambda_2}{\alpha}}, \quad u \in \mathbb{R}. \end{aligned} \tag{4.6}$$

Equation (4.6) could be interpreted as U_t^J being the difference between two compound Poisson processes with intensities λ_1 and λ_2, and exponentially distributed jump sizes with parameters μ_1 and μ_2 modelling up and down jumps. In this case, the moment condition (4.2) translates into $\mu_1, \mu_2 < 1/2$. This is the jump distribution that we are going to use in this paper.

Even in the LS-model, which is driven by one Wiener process only, the inability to store spot electricity means that the market is incomplete and Q is not unique. By the Girsanov Theorem (see Karatzas and Shreve [13] or Protter [17]), one family of risk neutral measures is characterised by a *market price of spot price risk* function of the form $\lambda(t, S_t)$. For the sake of analytical tractability, Lucia and Schwartz [14] choose $\lambda \in \mathbb{R}$ constant.

The introduction of jumps makes this class of measures much larger. According to Theorem 9.6 in Cont and Tankov [4], we may change both the jump intensity and jump distribution to any distribution absolutely continuous with f_J. We follow the route taken in Merton [15] for equity derivatives, and do not price jump risk. This may be a dubious assumption for electricity, but in the absence of liquid vanilla options, we feel that pricing the jump risk separately would be difficult. Pricing diffusive risk as in the LS-model gives the $Q-$dynamics for S_t as

$$\begin{cases} S_t &= \exp(f(t) + X_t) \\ dX_t &= (-\sigma\lambda - \alpha X_{t^-})dt + d\tilde{L}_t, \end{cases} \tag{4.7}$$

with $\tilde{L}_t = \sigma\tilde{W}_t + U_t^J - \lambda_J\mathbb{E}_P[J]t$, and $\tilde{W}_t = W_t - \lambda t$ being a $Q-$Wiener process. Another consequence of this choice of Q is that the jump intensity and distribution are invariant under this change of measure, and hence the condition for S_t having second moments under Q is still given by (4.2).

Forward prices are also given by the transform analysis in Duffie, Filipovic, and Schachermayer [8] as

$$\begin{aligned} F(t,T) &= \exp\left(f(T) + (\log S_t - f(t))e^{-\alpha(T-t)}\right) \\ &\quad \times \exp\left(-\frac{\sigma\lambda}{\alpha}(1-e^{-\alpha(T-t)}) + \frac{\sigma^2}{4\alpha}(1-e^{-2\alpha(T-t)})\right) \\ &\quad \times \exp\left(\lambda_J\int_0^{T-t}\mathbb{E}[\exp(Je^{-\alpha s}) - (1+Je^{-\alpha s})]ds\right) \qquad (4.8) \\ &\equiv F_{season}\times F_{diffusion}\times F_{jump}, \end{aligned}$$

showing that forward prices are products of three factors originating from the seasonality trend, diffusion and jumps respectively.

In the double exponential model (4.5), (4.8) simplifies to

$$\begin{aligned} F(t,T) &= \exp\left(f(T) + (\log S_t - f(t))e^{-\alpha(T-t)}\right) \\ &\quad \times \exp\left(-\frac{\sigma\lambda}{\alpha}(1-e^{-\alpha(T-t)}) + \frac{\sigma^2}{4\alpha}(1-e^{-2\alpha(T-t)})\right) \qquad (4.9) \\ &\quad \times \left(\frac{1-\mu_1 e^{-\alpha(T-t)}}{1-\mu_1}\right)^{\frac{\lambda_1}{\alpha}} \times \left(\frac{1+\mu_2 e^{-\alpha(T-t)}}{1+\mu_2}\right)^{\frac{\lambda_2}{\alpha}} \\ &\quad \times \exp\left(-\frac{(\lambda_1\mu_1 - \lambda_2\mu_2)}{\alpha}(1-e^{-\alpha(T-t)})\right). \end{aligned}$$

By (3.1) and the Markov property of X_t, the price at time $t \leq T$ of a simple European derivative with payoff $Y = H(S_T)$ satisfies $V(t) = V(t,x)$, with

$$V(t,x) = e^{-r(T-t)}\mathbb{E}[H(\exp(f(T) + X_T^{t,x})]. \qquad (4.10)$$

We will conclude this section by giving a Feynman-Kac formula, which states that $V(t,x)$ in (4.10) is the unique solution to a parabolic partial integro-differential equation (PIDE). The exact conditions on the payoff H under which this result holds as well as a proof are given in Appendix A.

Let u be a $C^{1,2}$ function[1] with a bounded first $x-$derivative. Then u is in the domain of the operators $\mathcal{D}_x$ and $\mathcal{I}_x$ defined as

$$\begin{aligned} \mathcal{D}_x u &= \frac{\sigma^2}{2}\frac{\partial^2 u}{\partial x^2} + (-\sigma\lambda - \alpha x)\frac{\partial u}{\partial x}, \qquad (4.11) \\ \mathcal{I}_x u &= \lambda_J\int_{\mathbb{R}}\left\{u(t,x+y) - u(t,x) - y\frac{\partial u}{\partial x}\right\} f_J(y)\,dy. \qquad (4.12) \end{aligned}$$

Moreover, u is in the domain of the infitesimal generator of X_t, which is given by $\mathcal{L}_x = \mathcal{D}_x + \mathcal{I}_x$.

[1]This means that u has one continuous $t-$derivative and two continuous $x-$derivatives.

Under some technical conditions on the payoff function H, the option price (4.10) is the unique solution to the PIDE

$$\begin{cases} \frac{\partial V}{\partial t} + \mathcal{L}_x V - rV = 0 \\ V(T,x) = H(e^{f(T)+x}). \end{cases} \tag{4.13}$$

In the absence of jumps, $\mathcal{L}_x = \mathcal{D}_x$, and the PIDE (4.13) reduces to a PDE.

5. Model estimation

In this section we discuss estimation of the model parameters for the Deng and LS models. We use the paradigm of Lucia and Schwartz [14], where the model is first estimated under P using historical spot price data, followed by an estimation of the market price of risk λ from traded forward contracts.

We postulate a periodic seasonal trend of the form

$$f(t|\Theta) = A_0 + \sum_{n=1}^{N} A_n \cos(2\pi f_n t + B_n), \tag{5.1}$$

with the frequencies f_n corresponding to cycles with lengths of one year, three months, one month, one week and three days, capturing the four seasons and the work week. The parameter vector $\Theta = (A_0, A_n, B_n)_{n=1}^N$ is unknown and is to be estimated from data.

The stationary covariance function of the logarithm of the spot price $Y_t = f(t|\Theta) + X_t$ is given by

$$\mathrm{Cov}(Y_{t+\tau}, Y_t) = e^{-\alpha\tau}\mathrm{Var}(Y_t),$$

so α is estimated from sample covariances and variances.

Having found α, Θ is estimated by the non-linear least square method proposed by Lucia and Schwartz [14]. An explicit Euler discretisation of the SDE (4.1) with $y_t = \log S_t$ denoting the sampled process, and time step Δt =1 day, yields

$$y_t = (1-\alpha)y_{t-1} + f(t|\Theta) - (1-\alpha)f(t-1|\Theta) + \epsilon_t, \;\; t = 2, \ldots, T.$$

Here T is the sample size and $\{\epsilon_t\}_{t=2}^T$ a sequence of i.i.d. random variables with mean zero and finite variance. Finally, a least squares estimator can be obtained by minimising

$$F(\Theta) = \frac{1}{T-1}\sum_{t=2}^{T} \left| y_t - \left[f(t|\Theta) - (1-\alpha)f(t-1|\Theta) + (1-\alpha)y_{t-1} \right] \right|^2$$

with respect to Θ.

Next we subtract the estimated seasonal function $f(t|\Theta)$ from $y_t = \log S_t$, which leaves us with discrete observations x_t of the process X_t. These are used to estimate the remaining parameters with the maximum likelihood method. Here we note that the stochastic variables $z_t \equiv x_{t+1} - x_t e^{-\alpha}$ are i.i.d. in both the LS and Deng models. From (4.8), we see that in the LS model, these variables are normally distributed with mean 0 and variance $\frac{\sigma^2}{2\alpha}(1-e^{-2\alpha})$. A standard maximum likelihood estimation is then used to find σ.

In the Deng model, inverse Fourier transformation of the conditional characteristic function (4.4) yields

$$\begin{aligned} f_{(x_{t+1}|x_t)}(x) &= \int_{\mathbb{R}} e^{iux_te^{-\alpha}-iu\frac{\lambda_1\mu_1-\lambda_2\mu_2}{\alpha}(1-e^{-\alpha})-\frac{\sigma^2u^2}{4\alpha}(1-e^{-2\alpha})+\int_0^1\{\varphi_J(ue^{-\alpha s})-1\}ds}e^{-iux}\frac{du}{2\pi} \\ &= \int_{\mathbb{R}} e^{-iu\frac{\lambda_1\mu_1-\lambda_2\mu_2}{a}(1-e^{-\alpha})-\frac{\sigma^2u^2}{4\alpha}(1-e^{-2\alpha})+\int_0^1\{\varphi_J(ue^{-\alpha s})-1\}ds}e^{-iu(x-x_te^{-\alpha})}\frac{du}{2\pi} \\ &\equiv \int_{\mathbb{R}} \varphi_Z(u)e^{-iuz_t}\frac{du}{2\pi}, \end{aligned}$$

where Z is a random variable with the same law as z_t, and characteristic function $\varphi_Z(u)$ under P.

We may now form the log likelihood function as

$$L = -\sum_{t=1}^{T-1}\log\left(\int_{\mathbb{R}}\varphi_Z(u)e^{-iuz_t}\frac{du}{2\pi}\right), \tag{5.2}$$

with α fixed. During the search for the optimal parameters, the Fourier integral (5.2) is approximated by a discrete Fourier transform, which is then computed by the FFT-algorithm. The resulting discretely sampled density is then linearly interpolated when evaluating the likelihood function. More details about approximation of continuous Fourier transforms by discrete ones, and the FFT algorithm can be found in Folland [9]. A similar method is suggested in Singleton [18], but there a Gauss-Hermite quadrature is used instead of FFT to compute the integrals in (5.2).

It remains to estimate the market price of risk λ. Let $\hat{F}(t,T_k)$, $1 \le k \le K$ be the market prices of the traded forward prices for delivery of 1 MWh during one day starting at time T_k. Contracts with longer delivery times may be regarded as a portfolio of these one day delivery forwards. The market price of risk, estimated at time t, is then chosen such that

$$\sum_{k=1}^{K}|\hat{F}(t,T_k) - F(t,T_k)|^2 \tag{5.3}$$

is minimised. Here $F(t,T_k)$ are the model implied forward prices given by (4.8).

6. Description of swing options

Before defining the payoff of a swing option formally, we give an example of a swing option payoff.

Example 1. The contract runs for one year. Every Friday, the holder decides on which days the following week he wants to buy 1 MWh of electricity at 30 EUR/MWh. At least 50 MWh and at most 100 MWh have to be bought in total during the year. The contract is financially settled with net payments every day.

Example 1 is generalised in Definition 6.1 below.

Definition 6.1. *Swing option.* A swing option of class $\mathcal{P}$ is a financial contract with the following payoff characteristics.

(1) *Maturity date:* The contract runs over the period $[0,T]$.
(2) *Strike:* The fixed price K EUR/MWh.

(3) *Swing action times:* The times when the holder is allowed to make decisions are denoted by $\{T_n\}_{n=1}^N$, where $0 \leq T_1 < T_2 < \ldots < T_N < T$.
(4) *Swing action:* At each swing action date T_n, the holder decides on the amount of energy B_n^d MWh to be bought at the fixed price K EUR/MWh over each of the D periods $(T_n^d, T_n^{d+1}]$, $1 \leq d \leq D$. Here $T_n = T_n^1 < T_n^2 < \ldots < T_n^D = T_{n+1}$, and $T_N^{D+1} = T$.
(5) *Allowed amounts per period:* We assume that $B_n^d \in \mathcal{O} \subseteq [0, \infty)$, where $\mathcal{O}$ is either a closed interval $\mathcal{O} = [\underline{B}, \overline{B}]$ or a discrete set. This means that the holder of the swing contract is not allowed to short energy.
(6) *Allowed amount in total:* The holder may buy at least $\underline{M}$ MWh and at most $\overline{M}$ MWh in total. To be of interest, $ND\underline{B} < \underline{M} \leq \overline{M} < ND\overline{B}$.
(7) *Settlement:* All swing contracts are financially settled. To reduce the consequences of a default of the counterpart, net payments occur at times T_n^d, $1 \leq d \leq D$.

In Example 1, $D = 7$, T_n^d is the beginning of day d, and $\mathcal{O} = \{0, 1\}$. Dahlgren [5], Jaillet, Ronn and Tompaidis [12] and Ibanz [11] all focus on the case when $D = 1$. By setting $\mathcal{O} = \{0, 1\}$, $\underline{M} = 0$, $\overline{M} = 1$ and $D = 1$, we see that Bermudan call options belong to class $\mathcal{P}$.

There are endless ways of generalising the swing option class $\mathcal{P}$. For example, one could allow under- or overdrafts at a penalty fee. Making the strike price K depend on calendar time would also be natural and feasible. Finally, one could have contracts where the holder is allowed either to sell or to buy energy, thereby creating a virtual electric energy storage facility.

7. Pricing of swing options

In this section, we show how to price swing options of class $\mathcal{P}$, and derive upper and lower bounds for these prices. The pricing methodology could easily be modified to cope with the other types of swing options discussed at the end of Section 6.

Without loss of generality, we assume that $\underline{B} = 0$ for the remainder of this paper. Any swing option with $\underline{B} > 0$ may be decomposed into a portfolio of $\underline{B}$ of each of the forward contracts $G(T_n, T_n^d)$, $1 \leq n \leq N$, $1 \leq d \leq D$, plus a swing option equal to the original one, but with $\mathcal{O}$ replaced by $[0, \overline{B} - \underline{B}]$, $\underline{M}$ by $\underline{M} - ND\underline{B}$, and $\overline{M}$ by $\overline{M} - ND\underline{B}$.

Initially, we determine the actual payoff induced by each swing action. A time T_n decision to buy B_n^d MWh during the period $(T_n^d, T_n^{d+1}]$ at K EUR/MWh is equivalent to receiving B_n^d forward contracts $G(T_n, T_n^d)$ which deliver 1 MWh each over this period at the fixed price K EUR/MWh. Consequently, a swing action is described by a $D-$vector $\{B_n^1, \ldots, B_n^D\} \in \mathcal{O}^D$. Below we show how to derive the payoff in the case when $\mathcal{O}$ is a closed interval, but discrete state spaces are dealt with similarly.

By maximising the value of the forward contracts $G(T_n, T_n^d)$ received, it is possible to condense this vector into one number $\Delta_n = \sum_{d=1}^D B_n^d$, representing the amount of energy bought due to this swing action. The state space, to which Δ_n belongs, is denoted $\mathcal{S}$, and when $\mathcal{O} = [0, \overline{B}]$, we have that $\mathcal{S} = [0, D\overline{B}]$. Since the forward contracts received are normally not available on the market, they must be priced theoretically by (4.9). More specifically, let $\bar{G}_d(T_n, s)$, $1 \leq d \leq D$, denote the time T_n prices of the D forward contracts with delivery during $(T_n, T_{n+1}]$, sorted by their theoretical value. Here we employ the convention that $d = 1$ corresponds to the most expensive contract. Fixing $\Delta_n \in [(k-1)\overline{B}, k\overline{B})$, where $1 \leq k \leq D$, and buying as much of the most expensive contracts as possible result in a maximised payoff g of

$$g(T_n, s, \Delta_n) = \sum_{d=1}^{k-1} \overline{B}\bar{G}_k(T_n, s) + [\Delta_n - (k-1)\overline{B}]\bar{G}_k(T_n, s). \tag{7.1}$$

By this construction, $g(\cdot, \cdot, \Delta)$ is continuous, piecewise linear and concave, implying that it attains its maximum and minimum on $\mathcal{S}$. This claim holds trivially for finite state spaces. Finally, note that the ordering of the forward contracts according to their theoretical values may change with s, and that if $D = 1$, $g(T_n, s, \Delta_n) = \Delta_n(s - K)$.

Under these assumptions, a swing action means choosing Δ_n without violating the contract constraints, and thereby receiving the amount $g(T_n, s, \Delta_n)$. This is stated formally in Definition 7.1 below.

Definition 7.1. The set of *admissible swing action strategies* $\mathcal{A}$ consists of all sequences $\{\Delta_n\}_{n=1}^{N}$ such that

(1) $\Delta_n \in \mathcal{S}$,
(2) Δ_n is $\in \mathcal{F}_{T_n}-$measurable,
(3) $\sum_{n=1}^{N} \Delta_n \in [\underline{M}, \overline{M}]$.

This definition reflects the fact that the decision must depend on known information only. To keep track of whether the constraint (3) above has been broken or not, we define $Z_t = \sum_{n=1}^{j} \Delta_n$ to be the amount of energy bought up to time $t \in (T_j, T_{j+1}]$. Since S_t and Z_t are both Markov processes, the swing option value $V(t)$ satisfies $V(t) = V(t, s, z)$, where $s = S_t$ and $z = Z_t$.

Before proceeding to to the main pricing theorem, we give upper and lower bounds on the swing option price $V(t, s, z)$ in Proposition 7.2 below. Here $c(t, s, T_n, T_n^d)$ denotes the price at time $t \leq T_n$ of an option that pays

$$Y = e^{-r(T_n^d - T_n)} \max(F(T_n, T_n^d) - K, 0)$$

at time T_n. Due to the affine structure of the model, these option prices may be computed by evaluating Fourier transforms (see Duffie, Filipovic and Schachermayer [8] for details).

In analogy with the definition of $\bar{G}_d(T_n, s)$, we introduce $\vec{G}_k(T_n, s)$, where $1 \leq k \leq D(N-n)$, as all the forward contracts maturing in $[T_n, T_N]$, sorted by their theoretical value. Finally, $\lfloor x \rfloor$ is the integer part of $x \in \mathbb{R}$.

Proposition 7.2. *Let $T_{j-1} \leq t < T_j$, $2 \leq j \leq N$. Then the swing option price $V(t, s, z)$ satisfies*

$$V(t, s, z) \leq \overline{B} \sum_{n=j}^{N} \sum_{d=1}^{D} c(t, s, T_n, T_n^d).$$

$$V(t, s, z) \geq \overline{B} \sup_{M \in [\underline{M}-z, \overline{M}-z]} \left\{ \sum_{k=1}^{\lfloor M/\bar{B} \rfloor} \vec{G}_k(T_n, s) + (M/\bar{B} - \lfloor M/\bar{B} \rfloor)\vec{G}_{\lfloor M/\bar{B} \rfloor + 1}(T_n, s) \right\}$$

Proof. Upper bound: Setting $\underline{M} = 0$ and $\overline{M} = N\overline{B}D$ clearly increases the flexibility and hence the value of the option. In this case, the optimal strategy is trivial: At each swing action date T_n, pick $\overline{B}$ of the forward contracts maturing in $[T_n, T_{n+1})$ with positive value and nothing of the others. This strategy is clearly equivalent to a portfolio of call options.

Lower bound: This is the value of a strategy which involves fixing the entire remaining swing action strategy $\{\Delta_n\}_{n=j}^N$ at time t, which is clearly suboptimal. First we fix $M \in [\underline{M} - z, \overline{M} - z]$, which is the total number of MWh to be bought in the remaining life of the swing option, $[T_j, T_N]$. Next, allocate these M MWh in a feasible manner among the remaining futures, such that the most expensive ones are chosen. Finally we take the supremum over M. □

We now state and prove the main theorem of this section.

Theorem 7.3. *Consider a swing option of class $\mathcal{P}$, and let $1 \le j \le N$. Then the value $V(t, s, z)$ of this swing option is given by*

$$V(t,s,z) = \begin{cases} \sup\limits_{\{\Delta_n\}_{n=j}^N \in \mathcal{A}} \sum\limits_{n=j}^{N} e^{-r(T_n - T_j)} \mathbb{E}\left[g(T_n, S_{T_n}, \Delta_n) | \mathcal{F}_{T_j} \right], & t = T_j, \\ e^{-r(T_j - t)} \mathbb{E}[V(T_j, S_{T_j}, z) | \mathcal{F}_t], & T_{j-1} < t < T_j,\ j > 1, \\ & t < T_1,\ j = 1. \end{cases}$$

Moreover, there exists at least one optimal swing action plan $\{\Delta_n^*\}_{n=j}^N \in \mathcal{A}$ such that the supremum is attained.

Proof. The claim is trivial in the case of a finite state space $\mathcal{O}$, since the number of swing action strategies is finite due to the Bermudan structure of the option. Below we assume that $\mathcal{O} = [0, \overline{B}]$, and prove the claim by backwards induction.

Step 1: The claim holds for $j = N$. At the last swing action date $t = T_N$, we choose the allowed swing action Δ_N that maximizes the payoff, or

$$V(T_N, s, z) = \sup_{\Delta_N \in \mathcal{A}} g(T_N, s, \Delta_N). \tag{7.2}$$

By the continuity of $g(\cdot, \cdot, \Delta)$, the supremum (7.2) is attained by some Δ_N^* on the compact set $\mathcal{S}$. Note that in general, $\Delta_N^* = \Delta_N^*(s, z)$. We also have that $V(T_N, s, z)$ is continuous in z. To prove this claim, observe that the definitions of g and $\bar{G}_k$ imply that

$$|V(T_N, s, z_2) - V(T_N, s, z_1)| \le |z_2 - z_1| \bar{G}_1(T_N, s),$$

so the continuity follows by letting $|z_2 - z_1| \to 0$.

For other times $t \in (T_{N-1}, T_N)$ in this period, we have a European derivative with payoff $V(T_N, s, z)$ at time T_N. Consequently,

$$V(t, s, z) = e^{-r(T_N - t)} \mathbb{E}[V(T_N, S_{T_N}, z) | \mathcal{F}_t], \tag{7.3}$$

by (3.1), proving the claim for $n = N$.

Step 2: Induction assumption. Let j be arbitrary such that $1 < j < N$, and assume the following:

(1) $V(T_j, s, z)$ is given by

$$V(T_j, s, z) = \sup_{\{\Delta_n\}_{n=j}^N \in \mathcal{A}} \sum_{n=j}^{N} e^{-r(T_n - T_j)} \mathbb{E}\left[g(T_n, S_{T_n}, \Delta_n)|\mathcal{F}_{T_j}\right]. \tag{7.4}$$

(2) $V(T_j, s, z)$ is continuous in z.
(3) Given s and z, there exists an optimal swing option strategy $\{\Delta_n^*\}_{n=j}^N$ such that the supremum (7.4) is attained.

Step 3: The claim holds for $j-1$, given the induction assumption in Step 2.

When choosing Δ_{j-1}, we receive a payoff of $g(T_{j-1}, s, \Delta_{j-1})$ *plus* a new swing option with Δ_{j-1} fewer exercise rights, and a first swing action date at time T_j. Combining this with the first assumption of Step 2 yields

$$\begin{aligned} V(T_{j-1}, s, z) &= \sup_{\Delta_{j-1} \in \mathcal{A}} \{g(T_{j-1}, s, \Delta_{j-1}) \\ &\quad + e^{-r(T_j - T_{j-1})} \mathbb{E}[V(T_j, S_{T_j}, z + \Delta_{j-1})|\mathcal{F}_{T_{j-1}}]\} \quad (7.5) \\ &= \sup_{\{\Delta_n\}_{n=j-1}^N \in \mathcal{A}} \sum_{n=j-1}^{N} e^{-r(T_n - T_j)} \mathbb{E}\left[g(T_n, S_{T_n}, \Delta_n)|\mathcal{F}_{T_{j-1}}\right]. \quad (7.6) \end{aligned}$$

It remains to prove the continuity of $V(T_{j-1}, s, z)$ and the existence of a Δ_{j-1}^*, such that the supremum (7.5) is attained. The continuity of $g(T_{j-1}, s, z)$ implies that these two claims follow if

$$\psi(z) \equiv \mathbb{E}[V(T_j, S_{T_j}, z)|\mathcal{F}_{T_{j-1}}]$$

is continuous. However

$$|\psi(z_2) - \psi(z_1)| \leq \mathbb{E}[|V(T_j, S_{T_j}, z_2) - V(T_j, S_{T_j}, z_1)||\mathcal{F}_{T_{j-1}}], \tag{7.7}$$

where the expression inside the expectation (7.7) goes to zero by the induction assumption in Step 2. The integrand may be bound by twice the upper bound given in Proposition 7.2 with $z = 0$, which is integrable. Hence the continuity of ψ follows by the Dominated Convergence Theorem.

For other times t in this period, we have a European derivative with payoff $V(T_{j-1}, s, z)$ at time T_{j-1}.

The theorem now follows by induction. □

Note that the optimal swing action strategy $\{\Delta_n^*\}_{n=1}^N$ may not be unique, since the function $g(\cdot, \cdot, \Delta)$ does not have to be monotone.

In order to use the recursive algorithm implicit in the proof of Theorem 7.3, we need to compute risk neutral conditional expectations of the type (7.3). A simple but time consuming approach by Ibànž [11], is to use Monte Carlo simulation. However, by Section 4, it can also be computed by solving the PIDE (4.13) with end condition $H(s) = V(T_n, s, z)$ for $(t, s) \in [T_{n-1}, T_n) \times [0, \infty)$ with z fixed.

8. Implementation

In this section we discuss the dynamic programming algorithm induced by the proof of Theorem 7.3 and the finite difference method used to solve the PIDE (4.13).

In practice, the pricing problem can only be solved for discrete values of (t, s, z). If the state space $\mathcal{O}$ is finite, then the $z-$variable is already discrete. Otherwise, since we know that $z \in [\underline{M}, \overline{M}]$, this interval is sampled uniformly at N_z points $z_m = m\Delta z$, where $\Delta z = (\overline{M} - \underline{M})/(N_z - 1)$ and $m = 0, \ldots, N_z - 1$.

As remarked in Section 4, we use $x = X_t$ rather than $s = S_t$ as state variable for the reason of numerical efficiency. We start by truncating $x = X_t$ such that $\underline{x} \leq x \leq \overline{x}$, and this interval is sampled uniformly at N_x points x_l similarly to the $z-$variable. In the same fashion, each time interval $[T_{n-1}, T_n]$ between two swing action dates is sampled at N_t points t_k. The spot price s_l at time t_k is then given by $s_l = \exp(f(t_k) + x_l)$, so the x-grid is uniform, but the s-grid is not. Finally, the dynamic programming algorithm implicit in the proof of Theorem 7.3 is implemented as three nested for-loops over the indices n, l, m, where checks are performed in each iteration to see wether a swing action is admissible or not.

To solve the PIDE (4.13) numerically, we use finite differences. When solving the PDE arising in the LS-model, we use a standard Crank-Nicholson scheme, which practical implementation is described in Wilmott [20]. Omitting the index m, and writing $v_k^l = V(t_k, x_l, z_m)$ and $\mathbf{v}_k = (v_k^1, \ldots, v_k^{N_x})^T$, the resulting linear system of equations may be written as

$$\mathbf{A}_1 \mathbf{v}_k = \mathbf{A}_2 \mathbf{v}_{k+1}, \tag{8.1}$$

where the elements of the matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ do not depend on time, thus enabling pre inversion of $\mathbf{A}_1$. Here we have assumed that the solution is linear in s at both boundaries, implying $\frac{\partial^2 V}{\partial s^2} = 0$, or $\frac{\partial^2 V}{\partial x^2} - \frac{\partial V}{\partial x} = 0$. Important properties of this method are unconditional stability and an error proportional to $(\Delta x)^2$ and $(\Delta t)^2$.

For the PIDE (4.13) an operator splitting method similar to the one proposed in Chapter 12:4 of Cont and Tankov [4] is used. The diffusion part

$$\frac{\partial V}{\partial t} + \mathcal{D}_x V - rV$$

is discretised by the Crank-Nicholson method as described above for the LS-model, whereas the integral operator $\mathcal{I}_x$ given in (4.12) is discretised using an explicit scheme. More specifically, if we define

$$\nu_n = \int_{\Delta x(n-\frac{1}{2})}^{\Delta x(n+\frac{1}{2})} f_J(y) dy,$$

then $\mathcal{I}_x$ may be approximated as

$$\begin{aligned} \mathcal{I}_x V(t_{k+1}, x_l, z) &= \lambda_J \int_{\mathbb{R}} \left[V(t_{k+1}, x_l + y, z) - V(t_{k+1}, x_l, z) - y \frac{\partial V}{\partial x} \right] f_J(y)\, dy \\ &\approx \lambda_J \sum_{n=-N}^{N} \left[v_{k+1}^{l+n} - v_{k+1}^{l} - \frac{n}{2}(v_{k+1}^{l+1} - v_{k+1}^{l-1}) \right] \nu_n. \end{aligned}$$

The density f_J is thus truncated at $\pm N\Delta x$, where we set N such that $\sum_{n=-N}^{N} \nu_n$ is close to one. Due to the non-local nature of the integral operator, it may happen that $x_{k+1}^{n+l} \notin [\underline{x}, \overline{x}]$ for some n, implying that extrapolation is necessary outside this interval. To be consistent with the assumed boundary condition $\frac{\partial^2 V}{\partial x^2} - \frac{\partial V}{\partial x} = 0$, we set

$$v_{k+1}^{l+n} = \begin{cases} ae^{x_{l+n}} + b, & x_{l+n} > \overline{x}, \\ ce^{x_{l+n}} + d, & x_{l+n} < \underline{x}, \end{cases}$$

where the unknown coefficients a, b, c and d, are selected to get continuous first derivatives at $\underline{x}$ and $\overline{x}$ respectively. Below we let $\mathbf{v}^e_{k+1}$ denote $\mathbf{v}_{k+1}$ extrapolated by this procedure.

To summarise, the introduction of jumps transforms the linear system (8.1) into

$$\mathbf{A}_1\mathbf{v}_k = \mathbf{A}_2\mathbf{v}_{k+1} + \mathbf{B}\mathbf{v}^e_{k+1}, \tag{8.2}$$

with $\mathbf{A}_1$ and $\mathbf{A}_2$ being the same matrices as in (8.1). The matrix $\mathbf{B}$ is a dense matrix such that

$$\mathbf{B}(l,:)\mathbf{v}^e_{k+1} \approx \mathcal{I}_x V(t_{k+1}, x_l, z),$$

where $\mathbf{B}(l,:)$ denotes the l'th row of $\mathbf{B}$.

9. Results

In this section we present the estimated model parameters and prices of some sample swing options.

First the seasonal function and jump diffusion parameters are estimated using the procedure described in Section 5. The data is daily Nord Pool spot prices over the period 1 January 2002 to 30 September 2004. To estimate the market price of risk, the 15 March 2005 closing prices for forwards with delivery during weeks 10 to 15 are used. The estimated parameters are presented in Tables 1 to 2. Model implied forward curves for both models are displayed in Figure 2 together with the seasonal part of the forward curve, F_{season}, defined in (4.8).

f_n	$f_0 = \infty$	f_1	f_2	f_3	f_4	f_5
A_n	3.3873	0.2930	0.0411	-0.0251	-0.0310	-0.0133
B_n	n.a.	0.4063	1.2807	0.9712	0.6386	1.7629

Table 1: The parameters A_0 and (A_n, B_n), $1 \le n \le 5$, of the seasonal function $f(t|\Theta)$ defined in (5.1) estimated by the procedure of Section 5. The frequencies are $f_0 = \infty$, $f_1 = 1/365$, $f_2 = 4/365$, $f_3 = 12/365$, $f_4 = 52/365$ and $f_5 = 104/365$ cycles per day. The frequency f_0 corresponds to the constant A_0.

Parameter	α	σ	λ_1	μ_1	λ_2	μ_2	λ (MPR)
Lucia-Schwartz	0.0211	0.0711	n.a.	n.a.	n.a.	n.a.	0.0095
Deng	0.0211	0.0370	0.1432	0.0897	0.2355	0.0556	0.0199

Table 2: The parameters α, σ, λ_1, μ_1, λ_2, μ_2 of the Deng and LS models togehter with the market price of spot price risk λ (MPR) estimated by the procedure of Section 5. The time is measured in days.

In Lucia and Schwartz [14], the model calibration is also made with daily Nord Pool data. Although the time period and seasonal function are different, they report $\alpha = 0.0140$ and $\sigma = 0.086$, which is similar to the values reported in Table 2.

Not very surprising, the introduction of jumps results in a smaller diffusive volatility σ, implying that variance has been transferred from the diffusion to the jumps. On average, there are one up (Type 1) jump and two down (Type 2) jumps every

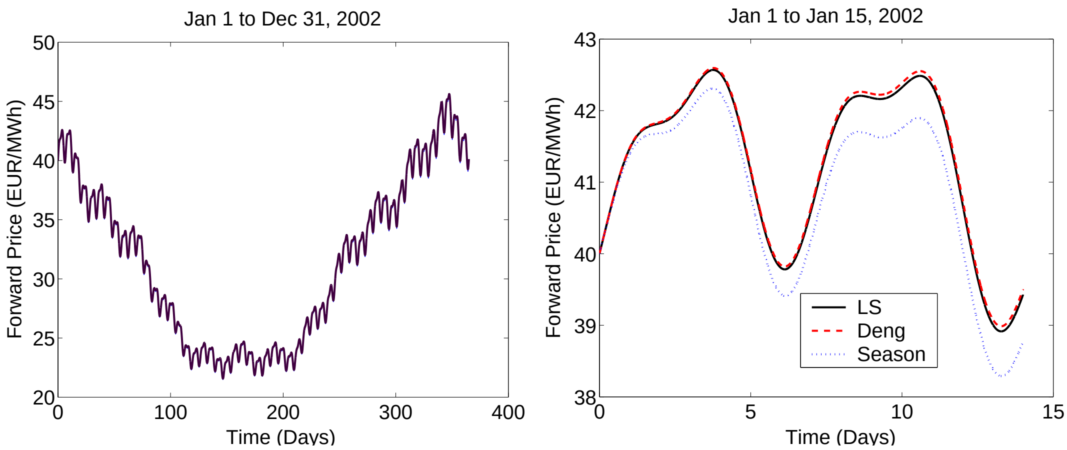

Figure 2: Seasonal part of the forward curve F_{season} given in (4.8) and the model implied forward curves for the LS and Deng models given in (4.9). The spot price on Jan 1 is $s = 40$ EUR/MWh. Model parameters from Tables 1 and 2.

week since $1/\lambda_1 \approx 7$ and $1/\lambda_2 \approx 3.5$. Moreover, the up jumps have a mean of $\mu_1 = 0.0897$, which is almost three standard deviations of the diffusive shock over one day, 0.0370. Consequently, jumps are large and frequent and μ_1 and μ_2 satisfy the moment condition (4.2).

Another effect of the introduction of jumps is that the market price of risk rises. This effect is due to the transfer of variance from the priced volatility, to the unpriced jumps. To maintain the risk premium, the market price of spot price risk per unit of volatility has to increase.

In Figure 2 we see that the model implied forward curves and lie practically on top of each other, and that the risk premium due to uncertainty is about 0.5 EUR. This suggests that the seasonal function is the factor with the largest impact on forward prices.

We now turn our attention to the pricing of swing options. Here we use a Matlab-implementation of the numerical method presented in Section 8. The aim is to investigate the impact on swing option prices in both models, for different strike prices K, number of decision dates N, and total allowed amount $\overline{M}$. All swing options run for one year ($T = 1$) and are priced at $t = 0$. The model parameter values all come from Tables 1 and 2, and the continuously compounded interest rate is $r = 3\%$ per annum.

N	2	4	12	52	364
Deng	842	887	1024	1197	1264
LS	841	884	1010	1167	1228
Difference	0.12%	0.34%	1.39%	2.57%	2.93%

Table 3: Dependence on the number of swing action dates N. Current spot price: $s = 30$ EUR/MWh. Option parameters: $\underline{M} = 0$, $\overline{M} = 100$, $\mathcal{O} = \{0, 1\}$ and $K = 30$ EUR/MWh. Model parameters from Tables 1 and 2.

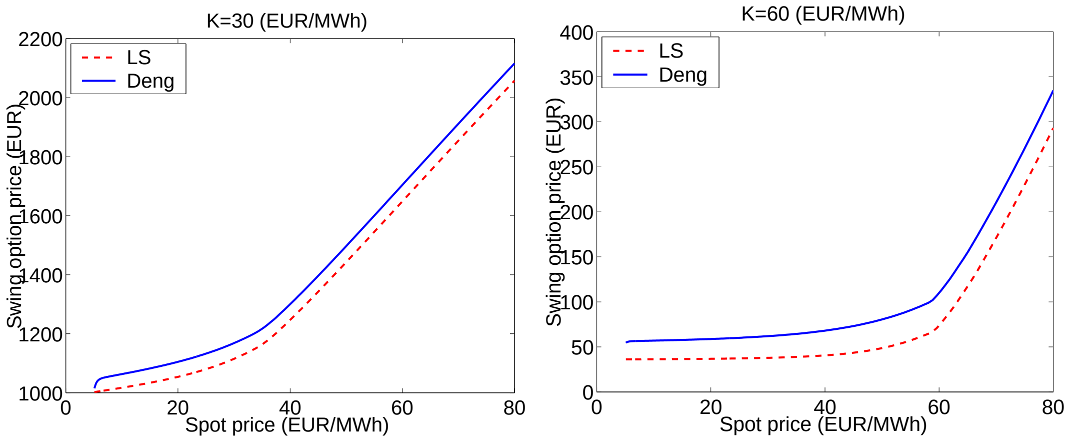

Figure 3: Dependence on the current spot price for $K = 30$ EUR/MWh (left) and $K = 60$ EUR/MWh (right). Option parameters: $\underline{M} = 0$, $\overline{M} = 100$, $N = 52$ and $\mathcal{O} = \{0, 1\}$. Model parameters from Tables 1 and 2.

$\overline{M}$	10	50	100	200	364
Deng	197	774	1197	1518	1559
LS	185	746	1167	1494	1536
Difference	6.49%	3.75%	2.57%	1.61%	1.50%

Table 4: Dependence on $\overline{M}$ for $\underline{M} = 0$. Current spot price: $s = 30$ EUR/MWh. Option parameters: $N = 52$, $\mathcal{O} = \{0, 1\}$ and $K = 30$ EUR/MWh. Model parameters from Tables 1 and 2.

$\underline{M} = \overline{M}$	10	50	100	200	364
Deng	193	748	1131	1281	176
LS	181	721	1104	1262	171
Difference	6.63%	3.74%	2.45%	1.51%	2.92%

Table 5: Dependence on $\overline{M}$ for $\underline{M} = \overline{M}$. Current spot price: $s = 30$ EUR/MWh. Option parameters: $N = 52$, $\mathcal{O} = \{0, 1\}$ and $K = 30$ EUR/MWh. Model parameters from Tables 1 and 2.

Figure 3 and Tables 3 to 5 suggest that the introduction of jumps increases the price of the swing option by 2-6% when $K = 30$, and by up to 35% when $K = 60$. An explanation for this could be the that the fatter tails of the jump diffusion have a big effect on the price only when the swing option is out of the money. This effect seems to be larger when the number of exercise rights $\overline{M}$ is small and the number swing action dates N is large for a given $\overline{B}$. This probably depends on the resulting higher timing flexibility, which in combination with the mean reversion allows the holder "to pick the peaks".

From Table 3, it is evident that swing options with more flexibility in terms of more swing action dates N, and a non-binding lower limit $\underline{M}$, are more expensive.

Tables 4 and 5 also show that the option value increases with $\overline{M}$ if $\underline{M} = 0$, but starts to decrease from $\overline{M}$ around 170 if $\underline{M} = \overline{M}$. Having to buy at an expensive price clearly decreases the value of the option.

10. Conclusion and final remarks

First we introduced the Lucia-Schwartz and Deng models, and suggested a time-series approach to their calibration. Estimation results show that jumps are frequent and of significant magnitude. Moreover, model implied forward prices are dominated by the seasonality part, which suggests that the seasonal function must be modelled with great care.

Second, we proved the existence of an optimal swing action strategy and proposed an algorithm for its computation. The results show that the introduction of jumps could make swing options 2-7% more expensive if in the money. For out of the money options, the difference could be up to 35%. Apart from moneyness, the difference in price between the LS and Deng models seems to depend on the timing flexibility of the swing actions.

The proof of Theorem 7.3 is generic in the sense that it is not restricted by the market model used in this paper. Changing the market model basically means changing the way risk neutral conditional expectations are computed. With this observation in mind, an interesting extension of this paper would be the pricing of electricity swing options in more complex spot price models. This could mean adding spikes, like in Andreasen and Dahlgren [1], making volatility stochastic, as has been done for equities by Heston [10], or having a stochastic seasonality trend as suggested by Lucia and Schwartz [14]. The main obstacle for the use of these models on Nord Pool, is the current lack of liquidity on its forward and options markets, which makes calibration difficult. However, for swing options on electricity markets with more liquid derivatives, such an extension would be feasible.

When the dimensionality increases, PDE/PIDE methods become more time consuming. Instead, Monte Carlo simulation becomes the method of choice, even for American contracts, and this is the route taken by Ibáñz [11]. Hence, further development of Monte Carlo methods for American options would be another natural extension of this paper.

References

[1] J. Andreasen and M. Dahlgren. At the flick of a switch. Working paper, Nordea Markets, Copenhagen, 2005.

[2] A. Bensoussan and J. Lions. *Impulse Control and Quasi-Variational Inequalities.* Gaulthier-Villars, Paris, 1984.

[3] R. Bingham, N.H. Kiesel. *Risk neutral valuation: Pricing and hedging of financial derivatives.* Springer, New York, 1998.

[4] R. Cont and P. Tankov. *Financial Modelling with Jump Processes.* Chapman & Hall, London, 2003.

[5] M. Dahlgren. A continuous time model to price commodity based swing options. *Review of Derivatives Research*, 8(1):27–47, 2005.

[6] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Math. Ann.*, (300):463–520, 1994.

[7] S. Deng. Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes. Power Working Paper 073, University Of California Energy Institute, 2000.
[8] D. Duffie, D. Filipovic, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13(3):984–1053, 2003.
[9] G. Folland. *Fourier analysis and its applications*. Brooks-Cole Publishing Company, Pacific Grove, 1992.
[10] S. Heston. A closed form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.
[11] A. Ibáñz. Valuation by simulation of contingent claims with multiple early exercise opportunities. *Mathematical Finance*, 14(2):223–248, 2004.
[12] P. Jaillet, E. Ronn, and S. Tompaidis. Valuation of commodity-based swing options. *Management Science*, 50(7):909–921, 2004.
[13] I. Karatzas and S. Shreves. *Brownian Motion and Stochastic Analysis*. Graduate Texts in Mathematics 113. Springer Verlag, New York, 1989.
[14] J. Lucia and E. Schwartz. Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research*, 5:5–50, 2002.
[15] R. Merton. Option pricing when the underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.
[16] H. Pham. Optimal stopping of controlled jump diffusion processes: A viscosity solution approach. *Journal of Mathematical Systems, Estimation, and Control*, 8(1):1–27, 1998.
[17] P. Protter. *Stochastic integration and differential equations - A new approach*. Springer, New York, 1990.
[18] K. Singleton. Estimation of affine asset pricing models. *Journal of Econometrics*, 102:111–141, 2001.
[19] A. Thompson. 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(2):271–293, 1995.
[20] P. Wilmott. *Derivatives, The Theory and Practice of Financial Engineering*. John Wiley & Sons, Chichester, university edition, 1998.

Appendix A. Validity of the Feynman-Kac formula

The reasons why we cannot use the Feynman-Kac results from Bensoussan and Lions [2] or Cont and Tankov [4] directly are the unbounded coefficient in front of the first order derivative in $\mathcal{L}_x$ in combination with the insufficient regularity of many financially important payoff functions H.

In order to solve this problem, we change state variable from X_t to $Z_t \equiv \log F(t,T)$, with $F(t,T)$ given in (4.8). An application of Itô's Lemma for jump diffusions (see Protter [17]) gives the Q-dynamics for Z_t as

$$\begin{aligned} dZ_t &= \left(-\frac{\sigma^2}{2}e^{-2\alpha(T-t)} + \lambda_J \mathbb{E}\left[1 - \exp(Je^{-\alpha(T-t)})\right]\right) dt \\ &\quad + \sigma e^{-\alpha(T-t)} d\tilde{W}_t + e^{-\alpha(T-t)} dU_t^J. \end{aligned} \tag{A.1}$$

If we define

$$\begin{aligned} \sigma(t) &= \sigma e^{-\alpha(T-t)}, \\ \gamma(t) &= -\frac{\sigma^2(t)}{2} - \lambda_J \int_{\mathbb{R}} \left[\exp(ye^{-\alpha(T-t)}) - 1 - ye^{-\alpha(T-t)}\right] f_J(y) dy, \end{aligned}$$

and let v be a $C^{1,2}$ function with bounded first $z-$derivative, the infitesimal generator $\mathcal{L}_z$ of Z_t is given by

$$\begin{aligned}\mathcal{L}_z v &= \frac{\sigma^2(t)}{2}\frac{\partial^2 v}{\partial z^2} + \gamma(t)\frac{\partial v}{\partial z} \\ &+\lambda_J \int_{\mathbb{R}} \left[v(t, z + ye^{-\alpha(T-t)}) - v(t,z) - ye^{-\alpha(T-t)}\frac{\partial v}{\partial z}\right] f_J(y)dy.\end{aligned}$$

Note that the coefficients of $\mathcal{L}_z$ only depend on t and are bounded due to the moment condition (4.2).

Since $S_T = F(T,T)$, the option price $V(t,x)$ in (4.10) may be expressed in terms of z as

$$u(t,z) = e^{-r(T-t)}\mathbb{E}[H(\exp(Z_T^{t,z}))] \tag{A.2}$$

$$\equiv e^{-r(T-t)}\mathbb{E}[h(Z_T^{t,z}))]. \tag{A.3}$$

If h is Lipschitz, Proposition 5.3 in Pham [16] yields that the PIDE

$$\begin{cases} \frac{\partial u}{\partial t} + \mathcal{L}_z u - ru = 0, \quad (t,z) \in [0,T) \times \mathbb{R}, \\ u(T,z) = h(z), \quad z \in \mathbb{R}, \end{cases} \tag{A.4}$$

has a unique solution[2] $u \in C^{1,2}([0,T) \times \mathbb{R}) \cap C^0([\mathbb{R})$ given by (A.3).

Next we show that the above result also holds for prices of futures contracts, which have unbounded payoffs given by $h(z) = e^z - K$. Inserting this payoff into (A.3) gives $u(t,z) = e^{-r(T-t)}(e^z - K)$. This function is $C^{1,2}$ and trivially solves (A.4). Furthermore, Theorem 3.3.1 in Lions and Bensoussan [2] ensures the uniqueness of this solution. By the linearity of $\mathcal{L}_z$, (A.3) is the unique solution to (A.4) for payoffs h which are linear combinations of futures contract prices and derivatives with Lipschitz payoffs. This includes call options by the put-call parity.

Finally we wish to switch back to the original state variable x. By (4.8) it is related to z through the relation

$$\begin{aligned} z &= f(T) + xe^{-\alpha(T-t)} - \frac{\sigma\lambda}{\alpha}(1 - e^{-\alpha(T-t)}) + \frac{\sigma^2}{4\alpha}(1 - e^{-2\alpha(T-t)}) \\ &+\lambda_J \int_t^T \mathbb{E}[\exp(Je^{-\alpha(T-s)}) - (1 + Je^{-\alpha(T-s)})]ds \\ &\equiv g(t,x). \end{aligned}$$

Since $u \in C^{1,2}$, we can apply the chain rule to $V(t,x) = u(t, g(t,x))$, which retrieves the PIDE (4.13). It also follows that this PIDE also has a unique solution provided that h is linear combination of futures contracts and derivatives with Lipschitz payoffs.

Department of Economics, Göteborg University, P.O Box 600, S-405 30 Göteborg, Sweden
Email address: mats.kjaer@economics.gu.se

[2] $C^0(\mathbb{R})$ is the space of continuous functions on $\mathbb{R}$.

Paper III

VALUATION OF A NATURAL GAS STORAGE FACILITY

MATS KJAER AND EHUD RONN

Göteborg University & The University of Texas at Austin

Abstract. We investigate the valuation of a natural gas storage facility, where trading is permitted on both the futures and spot markets simultaneously.

The risk neutral futures curve dynamics is specified directly by an HJM-model, whereas the intra-month log-spot price is given by an Ornstein-Uhlenbeck process reverting towards the log-price of the most recently settled futures contract.

We show how to value the storage by dynamic programming and propose a numerical method for the computation of these values. It is also shown that a pure "spot-market" strategy creates the highest storage value. However, numerical examples demonstrate that it is often possible to come close to this value by a strategy involving a large proportion of injections/extractions from futures contracts. Benefits of using futures contracts include the possibility of at-least-partially hedging storage operations with the relatively liquid NYMEX natural gas futures contract.

1. Introduction

Ever since the deregulation of the U. S. natural gas market, this commodity has been traded on local spot markets connected by a nationwide pipeline system. In addition, natural gas futures contracts for delivery at Henry Hub, LA., are traded at the New York Mercantile Exchange (NYMEX). Typically, natural gas is more expensive in the winter than in the summer, and a storage facility could be one way of profiting from these price differences. This paper deals with the valuation of such a storage facility.

Earlier work on this topic by Manoliu [9] uses the swing-option pricing methodology proposed by Jaillet, Ronn and Tompaidis [7] for the valuation of a storage facility. Here transactions are made on a spot market, where the risk-neutral dynamics of the spot-price logarithm is the sum of a deterministic trend and a one-factor Ornstein-Uhlenbeck process. Parsons [10] generalises this approach to a two-factor model, where the mean reversion level is stochastic. Schwartz [11] refers to these models as Model 1 and 2.

In this paper, the storage owner may inject and extract natural gas on the spot and futures markets simultaneously. We show that a pure spot-market strategy similar to the ones proposed in Manoliu [9] and Parsons [10] is the most profitable. However, we also demonstrate that it is possible to come quite close to this value

Date: April 25, 2006.

Key words and phrases. Storage Valuation, Commodity derivatives, Swing Options.

JEL classification: C61, G13.

Mats Kjaer's visit to UT Austin was supported by: The University of Texas at Austin, Knut & Alice Wallensbergs Stiftelse, Wallanderstiftelsen and Tekn. Dr. Marcus Wallenbergs Stiftelse för Internationellt Industriellt Företagande.

using a trading strategy where most of the natural gas injected or extracted comes from NYMEX natural gas futures. This approach enables partial hedging using these relatively liquid contracts.[1]

To proceed, spot and futures market models need to be specified. Based on the work of Heath, Jarrow and Morton [4] in the interest-rate area, Cortazar and Schwartz [2] introduce models for the joint dynamics of commodity futures prices. We will use this framework with a particular functional specification of the volatility to model NYMEX futures prices. Since these contracts have delivery periods starting at the beginning of each calendar month, the spot price process is not entirely determined by the futures prices. Within each calendar month, we let the spot price follow the dynamics of Model 1 in Schwartz [11].

It turns out that an important ingredient of the storage valuation is the pricing of swing options. These options give their holder multiple rights to choose the amount of the commodity to be delivered at multiple occasions, given some constraints. In order to price these options, Jaillet, Ronn and Tompaidis [7] suggest a multi-level tree algorithm, Dahlgren [3] solves multiple variational inequalities, and Ibáñz [6] uses Monte Carlo simulation. In this paper, the multi-level tree method by Jaillet, Ronn and Tompaidis [7], the so-called "trinomial forest," will be utilised.

This paper is organised as follows. Section 2 describes the physical properties of the storage and Section 3 the operational protocols for the storage facility. The futures- and spot-price models are specified in Section 4, and their calibration discussed in Section 5. These models are used for the valuation of the storage, which is the topic of Sections 6 to 8. Details about the numerical implementation are given in Section 9 and some numerical valuation examples in Sections 10 to 11. Section 12 contains conclusions and final comments.

2. Physical properties of the storage facility

The storage facility is modeled analogous to a water tank with one inlet and one outlet, and its properties are specified in Assumption 1 and illustrated in Figure 1. Of course, these assumptions are a simplification of reality, and in Appendix A, we provide an example of the properties of an existing storage facility. We employ the market practice and measure quantities of natural gas in British thermal units (Btu), a measure of energy content equivalent to the SI-unit Joule (J).

Assumption 1. *Physical properties of the storage.*

(1) The storage facility has a fixed capacity of $\overline{Q}$ (Btu).
(2) The storage facility has one injection inlet and one extraction outlet. For any $\overline{q} \in [\overline{Q}, \infty)$, the injection and extraction rates at time $t \geq 0$, q_t^i (Btu/month) and q_t^e (Btu/month) satisfy $q_t^i \in [0, \overline{q}]$ and $q_t^e \in [-\overline{q}, 0]$ respectively. Here $\overline{q}$ is fixed and does not depend on calendar time, outside temperature or storage content.
(3) The injection and extraction rates q_t^e and q_t^i may be changed at most $M \geq \overline{q}/\overline{Q}$ discrete, equidistant times per month. These times are denoted t_m, $m \in \mathbb{N}$ and consequently the period $[t_m, t_{m+1})$ is of duration $1/M$ months.

[1] Hedging intra-month transactions would require access to liquid balance-of-month contracts.

(4) There are no fixed or variable costs of extraction or injection, except for the cost of the natural gas itself.

Parts 1 and 2 of Assumption 1 imply that an empty storage may be filled at the uniform rate $q_t^i = \overline{Q}$ (Btu/month) and a full one emptied at $q_t^i = -\overline{Q}$ (Btu/month). Furthermore, the controls q_t^i and q_t^e are right continuous step functions of time, which are constant over the intervals $[t_m, t_{m+1})$. Below we write $q_m^e \equiv q_{t_m}^e$, $q_m^i \equiv q_{t_m}^i$ and with this notation the amount of energy Q_t stored at time $t \in [t_m, t_{m+1})$ is given by

$$Q_t = Q_0 + \frac{1}{M}\sum_{k=0}^{m-1}(q_k^i - q_k^e) + (q_m^i - q_m^e)(t - t_m). \tag{2.1}$$

For simplicity, we write $Q_m \equiv Q_{t_m}$, and the physical constraints of the storage imply that $Q_m \in [0, \overline{Q}]$ must hold for all m.

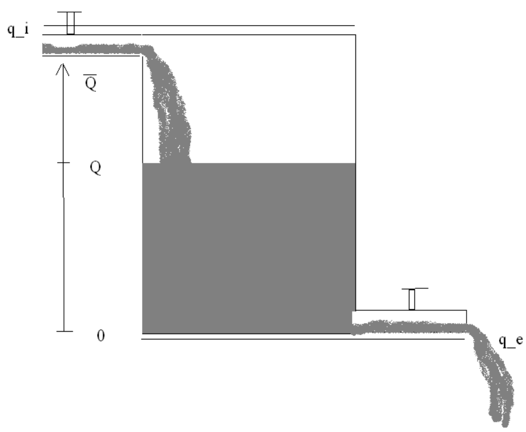

Figure 1: Important characteristics of the storage tank

3. Operation of the storage facility

Given the properties of the storage facility specified in Assumption 1, we postulate that the storage owner buys the storage in the summer and sells it in the winter in order to profit from the summer-winter spread. This behaviour is formalised in Assumption 2.

Assumption 2. *Summer-winter spread strategy.*

(1) The storage is filled during exactly one of the summer months June, July or August, and this is the *injection month.*
(2) The storage is emptied during exactly one of the winter months December, January or February and this is the *extraction month.*

Note that the terms "fill" ("empty") only mean that the storage starts empty (full) at the beginning of the injection (extraction) month and ends up full (empty), but that these processes do not have to be monotonic. For the remainder of this paper, and without loss of generality, we assume that futures contracts are perfectly divisible and that one futures contract delivers $\overline{Q}$ (Btu) at an even rate during one calendar month. In Sections 6 to 8, we prove that the storage valuation problem is indeed scalable.

We also define three *modes of operation,* which the owner may use to implement the summer-winter spread strategy of Assumption 2.

Definition 3.1. *Benchmark mode of operation.*

(1) At June 1, the cheapest of the June, July and August futures contracts (discounted to June 1 prices) is acquired and the most expensive of the December, January and February futures contracts (discounted to June 1 prices) is sold short.

Valuation of a storage run in this way is trivial. The next two modes of operation are more complex and will hopefully result in higher storage values. For the remainder of this paper, acquisition of a negative number of futures is equivalent to short selling.

Definition 3.2. *Futures market mode of operation.*

(1) A particular summer injection month is selected by the acquisition of one futures contract with delivery that month. This decision may be taken at any time prior to the maturity of the futures contract.
(2) A particular winter extraction month is selected by shorting one futures contract with delivery that month analogously with the selection of the summer injection month.

Under the benchmark and futures mode of operation, the storage is injected/extracted by the delivery from the chosen futures contract. This is physically feasible since $\overline{q} \geq \overline{Q}$ by Assumption 1. However, if $\overline{q} > \overline{Q}$, these strategies may not use the full flexibility of the storage. One way of achieving this would be to inject/extract from the spot and futures markets simultaneously, as as described in Definition 3.3.

Definition 3.3. *Extended mode of operation.*

(1) A particular summer injection month is selected by the acquisition of
$$\beta_s \in \left[-\frac{\overline{q}-\overline{Q}}{\overline{Q}}, \frac{\overline{q}}{\overline{Q}}\right]$$
futures contracts with delivery that month. This also implies that the *net* amount bought on the spot market during that month must equal $(1-\beta_s)\overline{Q}$ (Btu). The decision may be taken at any time prior to the maturity of the futures contract.
(2) A particular winter injection month is selected by the acquisition of
$$\beta_w \in \left[-\frac{\overline{q}}{\overline{Q}}, \frac{\overline{q}-\overline{Q}}{\overline{Q}}\right]$$
futures contracts with delivery that month. This also implies that the *net* amount bought on the spot market that month must equal $(-1-\beta_w)\overline{Q}$

(BTU). The decision may be taken at any time prior to the maturity of the futures contract.

(3) Spot market transactions are made in accordance with Assumptions 1 to 2. In particular, this means that the decision about how much to sell or buy on the spot market during the period $[t_m, t_{m+1})$ is taken at time t_m.

When $\beta_s = 0$ and $\beta_w = 0$, we only have spot market transactions, which is the situation described in Manoliu [9] and Parsons [10]. The situation $\beta_s = 1$, $\beta_w = -1$, and abstaining from spot market transactions, is equivalent to the Futures mode of operation.

Before being able to value a storage operated in either of these modes, we need specify the futures- and spot-price dynamics.

4. Spot and futures price models

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, P)$ be a filtered probability space, where $\mathcal{F}_t$ is the $P-$completion of the natural filtration $\sigma(Z_u, 0 \le u \le t)$ of the I-dimensional Wiener process Z_t. For simplicity we choose P to be the risk-neutral probability measure, and the expectation operator with respect to this measure is denoted $\mathbb{E}$. We introduce the risk-free bank account B_t with dynamics

$$B_t = B_0 e^{rt}, \tag{4.1}$$

for some fixed $r > 0$. In addition to this riskfree account, there is a futures market specified in Assumption 3 defined on the probability space introduced above.

Assumption 3. *Futures market.*

(1) There are I futures contracts for delivery during the calendar months 1 to I. Calendar month i begins at time T_i.

(2) The futures price F_t^i, $t \le T_i$, for delivery during $[T_i, T_{i+1})$ follows the $P-$dynamics

$$\frac{dF_t^i}{F_t^i} = \sigma_i e^{-\alpha_i(T_i-t)} dW_t^i, \quad 0 \le t \le T_i. \tag{4.2}$$

Here $\sigma_i > 0$, $\alpha_i > 0$, and W_t^i is a one-dimensional Wiener process.

(3) $\mathrm{Cov}(W_t^i, W_t^j) = \rho_{ij} t$, $1 \le i, j \le I$. The ρ_{ij}'s are elements of an $I \times I$ correlation matrix $\mathbf{C}$, which satisfies the conditions required by such a matrix.

In this model, the price process of each futures contract has its specific volatility σ_i and decay rate α_i. Note that if we set $\alpha_i = 0$, the Black model introduced in Black [1] is retrieved.

For $t_1 \le t_2 \le T_i$, the SDE (4.2) has the solution

$$F_{t_2}^i = F_{t_1}^i \exp\left\{ -\frac{\sigma_i^2}{4\alpha_i} e^{-2\alpha_i T_i}(e^{2\alpha_i t_2} - e^{2\alpha_i t_1}) + \sigma_i \int_{t_1}^{t_2} e^{-\alpha_i(T_i-u)} dW_u^i \right\}. \tag{4.3}$$

Consequently, if $\min\{T_i, T_j\} \ge t_2$, $\log(F_{t_2}^i/F_{t_1}^i)$ and $\log(F_{t_2}^j/F_{t_1}^j)$ have a bivariate normal distribution with mean vector $\mu(t_1, t_2)$ and covariance matrix $\mathbf{C}(t_1, t_2)$ with elements

$$\mu_i(t_1, t_2) = -\frac{\sigma_i^2}{4\alpha_i} e^{-2\alpha_i T_i}(e^{2\alpha_i t_2} - e^{\alpha_i t_1}), \tag{4.4}$$

$$C_{ij}(t_1, t_2) = \rho_{ij} \frac{\sigma_i \sigma_j}{\alpha_i + \alpha_j} e^{-\alpha_i T_i} e^{-\alpha_j T_j} \left(e^{(\alpha_i+\alpha_j)t_2} - e^{(\alpha_i+\alpha_j)t_1} \right), \tag{4.5}$$

for $i, j \in \{1, \ldots, I\}$. Finally, it can be shown that the correlation between $\log(F^i_{t_2}/F^i_{t_1})$ and $\log(F^j_{t_2}/F^j_{t_1})$ equals ρ_{ij} in the limit when $|t_2 - t_1| \to 0$.

The futures market specified above is sufficient to perform a valuation of a storage facility operated in the futures market mode. In the extended mode, a spot market model must first be specified. By Assumption 2, spot market transactions are made only during the injection and extraction months, implying that we do not need to compare spot prices across months. Thus it suffices to specify the spot price within each month separately.

Assumption 4. *Intra month spot market.*

(1) For $T_i \le t < T_{i+1}$, the spot price S^i_t follows the $P-$dynamics

$$\begin{cases} S^i_t &= \exp\left[f\left(t\right) + X_t\right] \\ dX_t &= -\hat{\alpha}_i X_t \, dt + \hat{\sigma}_i \, dW^i_t, \quad X_{T_i} = 0, \end{cases} \tag{4.6}$$

where $f(t)$ is a deterministic function, $\hat{\alpha}_i > 0$, $\hat{\sigma}_i > 0$ and W^i_t the Wiener process used to drive F^i_t.

(2) For $T_i \le t < T_{i+1}$, the spot price parameters of month i are related to the futures price parameters of this month through the relations

$$\begin{aligned} \hat{\alpha}_i &= \alpha_i, \\ \hat{\sigma}_i &= \sigma_i \frac{\hat{\alpha}_i (T_{i+1} - T_i)}{1 - \exp\left[-\hat{\alpha}_i (T_{i+1} - T_i)\right]}, \end{aligned}$$

which implies $\hat{\sigma}_i > \sigma_i$. A detailed motivation for this assumption is found in Appendix B.

(3) At the beginning of the month, $\mathbb{E}[S^i_t | \mathcal{F}_{T_i}] = e^{r(t_m - T_i)} F^i_{T_i}$ for $T_i \le t < T_{i+1}$, meaning that the discounted initial intra month futures curve is flat.

(4) Let t_m be the times when the storage operator could change the injection and extraction rates, introduced in Part 3 of Assumption 1. A spot market deal clocked at time $t_m = T_i + m/M$ is delivered at an even rate during $[t_m, t_{m+1})$ at the price $S^i_{t_m}$.

Absence of arbitrage implies that spot and futures prices are related through the relation

$$F^i_t = \frac{1}{M} \mathbb{E}\Big[\sum_{m=0}^{M-1} e^{-r(t_m - T_i)} S^i_{t_m} \Big| \mathcal{F}_t \Big], \tag{4.7}$$

which holds for *any* choice of $\hat{\alpha}_i > 0$ and $\hat{\sigma}_i > 0$ by Part 3 of Assumption 4. This is the reason why these parameters have to be specified separately in Part 2 of Assumption 4. Moreover, for $T_{i-1} \le t < T_i$ the seasonal function $f(t)$ is given by

$$f(t) = \log F^i_{T_i} + r(t - T_i) - \frac{\hat{\sigma}^2_i}{4\hat{\alpha}_i} \left[1 - e^{-2\hat{\alpha}_i (t - T_i)}\right], \tag{4.8}$$

so

$$S^i_t = F^i_{T_i} \exp\left(r(t - T_i) - \frac{\hat{\sigma}^2_i}{4\hat{\alpha}_i} \left[1 - e^{-2\hat{\alpha}_i (t - T_i)}\right] + \hat{\sigma}_i \int_{T_i}^{t} e^{-\hat{\alpha}_i (t - u)} dW^i_u \right). \tag{4.9}$$

5. Model calibration

Let indices i, j refer to calendar months with $i = 1$ being January. For the estimation of the futures market parameters, we simplify matters by assuming a structure of the correlation matrix and constant parameters within each season, i.e.

$$\begin{aligned}
&\alpha_i = \alpha_s,\ \sigma_i = \sigma_s, && && i = 6, 7, 8,\\
&\alpha_i = \alpha_w,\ \sigma_i = \sigma_w, && && i = 12, 1, 2,\\
&\rho_{ij} = \exp(-\lambda_s |T_i - T_j|), && \lambda_s > 0, && i, j = 6, 7, 8,\\
&\rho_{ij} = \exp(-\lambda_w |T_i - T_j|), && \lambda_w > 0, && i, j = 12, 1, 2.
\end{aligned}$$

We estimate λ_s and λ_w by fitting the functional form $\exp(-\lambda|T_i - T_j|)$ to a sample correlation matrix in the least squares sense. This functional form will generate a valid correlation matrix since it is the autocorrelation function of a stationary Ornstein-Uhlenbeck process with mean reversion parameter $\lambda > 0$. Summer and winter sample correlation matrices are estimated from historical log-returns. The data used is NYMEX closing prices from the six month period before the first contract of the season matures (Jan 1 to May 27 for the summer, June 1 to Nov 27 for the winter). We compute correlation matrices for the years 2003 and 2004 and use the average matrix as target for the least squares fitting.

Results are shown in Table 1 below.

summer	Jun	Jul	Aug	**winter**	Dec	Jan	Feb
Jun	1.0000	0.9785	0.9574	Dec	1.0000	0.9844	0.9690
Jul	0.9785	1.0000	0.9785	Jan	0.9844	1.0000	0.9844
Aug	0.9574	0.9785	1.0000	Feb	0.9690	0.9844	1.0000

Table 1: Estimated summer and winter correlation matrices. Least squares estimation result: $\lambda_s = 0.2610$ and $\lambda_w = 0.1887$. The unit is per annum.

Having obtained the correlation matrices, we estimate α_s, α_w, σ_s and σ_w from Nymex futures option prices. In the Bloomberg information system, the price of an option on the futures contract for month i is quoted in terms of its Black implied volatility $\bar{\sigma}_i$, which is related to the parameters of our model through the expression

$$\frac{\sigma_i^2}{2\alpha_i}\left[1 - e^{-2\alpha_i(T_i - t)}\right] = \bar{\sigma}_i^2 (T_i - t), \tag{5.1}$$

with T_i being the maturity of the option and $t < T_i$ the date of quotation.

The estimation procedure for (α_s, σ_s) and (α_w, σ_w) may now be described with the following steps:

(1) Take monthly implied volatilities for the eight most recent months for the three contracts.
(2) Estimate (α, σ) from Equation 5.1 above in the least squares sense (two unknowns, twenty four equations).
(3) Fix α, but refit σ using Equation 5.1 with α and the most recent Black implied volatilities. This ensures a best fit to the most recent implied volatilities.

Since the valuation date of our storage is May 31, 2005, Black implied volatilities sampled on the last trading day of each month from September 2004 to April 2005 for the summer months and March 2004 to October 2004 for the winter months. The results are displayed in Table 2 below. Here σ_{LS} refers to the volatility obtained from the least squares estimation in Step 2 of the estimation procedure and σ to the recalculated volatility from Step 3.

Season	α	σ_{LS}	σ
summer	1.151	0.439	0.385
winter	1.468	0.600	0.675

Table 2: The parameters of the futures price model estimated from Black implied volatilities obtained from Bloomberg's. The unit is per annum.

From Table 2 we see that both volatilities and mean reversion factors are higher during the winter than the summer. This could be attributable to the fact that the demand peak of natural gas occurs in the winter.

6. Valuation - futures market mode of operation

In this section we derive the value of a storage run in the futures market mode defined in Definition 3.2.

Futures prices refer to payments occurring at the start of delivery of the contract. In particular, they are $P-$martingales, which together with properties of the max-function imply that there is no reward for early acquisition of a particular contract. Thus a necessary condition for optimality may be formulated verbally as:

> *"Do nothing until the beginning of the summer/winter season. Then, at the beginning of each calendar month, decide whether to buy/short the prompt month contract or waiting."*

This condition together with backwards induction yield the value $V_s(t)$ of the summer injection operations as

$$
\begin{aligned}
V_s(T_8) &= -F^8_{T_8}, && (6.1)\\
V_s(T_7) &= \max\{-F^7_{T_7}, -e^{-r(T_8-T_7)}F^8_{T_7}\}, && (6.2)\\
V_s(T_6) &= \max\{-F^6_{T_6}, e^{-r(T_7-T_6)}\mathbb{E}[V_s(T_7)|\mathcal{F}_{T_6}]\}, && (6.3)\\
&= \max\{-F^6_{T_6}, e^{-r(T_7-T_6)}\mathbb{E}[\max\{-F^7_{T_7}, -e^{-r(T_8-T_7)}F^8_{T_7}\}|\mathcal{F}_{T_6}]\}, && (6.4)\\
V_s(t) &= e^{-r(T_6-t)}\mathbb{E}[V_s(T_6)|\mathcal{F}_t], \quad t \le T_6. && (6.5)
\end{aligned}
$$

This value is negative and is simply the cost of filling the storage in an optimal manner. Similar reasoning gives the value $V_w(t)$ of the winter extraction as

$$\begin{aligned} V_w(T_2) &= F^2_{T_2}, & (6.6)\\ V_w(T_1) &= \max\{F^1_{T_1}, e^{-r(T_2-T_1)}F^2_{T_1}\}, & (6.7)\\ V_w(T_{12}) &= \max\{F^{12}_{T_{12}}, e^{-r(T_1-T_{12})}\mathbb{E}[\max\{F^1_{T_1}, e^{-r(T_2-T_1)}F^2_{T_1}\}|\mathcal{F}_{T_{12}}]\}, & (6.8)\\ V_w(t) &= e^{-r(T_{12}-t)}\mathbb{E}[V_w(T_{12})|\mathcal{F}_t], \quad t \le T_{12}. & (6.9) \end{aligned}$$

The expressions (6.3) and (6.8) can be computed analytically by Lemma 6.1 below. Here $\Phi(x)$ is the distribution function of a normally distributed random variable.

Lemma 6.1. *Let $c_1, c_2 \in \mathbb{R}$ and consider a derivative with payoff*

$$Y = \max\{c_1 F^{i+1}_{T_{i+1}}, c_2 F^{i+2}_{T_{i+1}}\}$$

at time T_{i+1}. If F^{i+1}_t and F^{i+2}_t follow the risk-neutral dynamics specified in Assumption 3, the price $V_{ex}(t)$, $t < T_{i+1}$, of this derivative is given by

$$V_{ex}(t) = \begin{cases} c_1 e^{-r(T_{i+1}-t)} F^{i+1}_t \Phi(D) + c_2 e^{-r(T_{i+1}-t)} F^{i+2}_t \Phi(\Sigma - D), & c_1 > 0, c_2 > 0, \\ c_1 e^{-r(T_{i+1}-t)} F^{i+1}_t \Phi(-D) + c_2 e^{-r(T_{i+1}-t)} F^{i+2}_t \Phi(D - \Sigma), & c_1 < 0, c_2 < 0, \\ e^{-r(T_{i+1}-t)} \max\{c_1 F^{i+1}_t, c_2 F^{i+2}_t\}, & c_1 \ge 0 \ge c_2 \text{ or} \\ & c_1 \le 0 \le c_2, \end{cases}$$

where

$$\begin{aligned} \Sigma^2 &= C_{i+1,i+1}(t, T_{i+1}) + C_{i+1,i+1}(t, T_{i+1}) - 2C_{i+1,i+2}(t, T_{i+1}), \\ D &= \frac{-\log\left(c_2 F^{i+2}_t / c_1 F^{i+1}_t\right) + \Sigma^2/2}{\Sigma}. \end{aligned}$$

Here the covariance matrix elements $C_{ij}(t, T)$ above are given in (4.5).

The proof is found in Appendix C.1. Unfortunately, it does not seem possible to derive analytical expressions for $V_s(t)$, $t < T_6$ and $V_w(t)$, $t < T_{12}$. Instead, the conditional expectations (6.5) and (6.9) are computed by Monte Carlo simulation: We use the explicit expressions (6.3) and (6.8) of the payoff together with the known elements of the mean vector (4.3) and covariance matrix (4.4) of the joint futures price distribution.

Finally, expressions (6.1) to (6.9) yield V_s and V_w homogeneous of degree one with respect to scaling of $\overline{Q}$ as long as $\overline{q} \ge \overline{Q}$. In particular, this means that the assumption that one futures contract fills the storage in one month could be made without loss of generality. Finally, by (4.3), storage values are also homogenous of degree one when the current futures prices $\{F^6_t, F^7_t, F^8_t\}$ and $\{F^{12}_t, F^1_t, F^2_t\}$ are multiplied by some common factor.

7. Valuation of swing options

In order to prepare for the valuation of a storage run in the extended mode, we briefly discuss the pricing of swing options. More details about the pricing of such options may be found in Jaillet, Ronn and Tompaidis [7] or Dahlgren [3].

Let $\Delta_m \in \mathbb{R}$ be the *net* amount bought ($\Delta_m \ge 0$) or sold ($\Delta_m \le 0$) on the spot market during $[t_m, t_{m+1})$. Since it is never optimal to buy and sell on the spot market during the same period, the flow from the β futures contracts together with the sequence $\{\Delta_m\}_{m=0}^{M-1}$ is sufficient to describe the operation of a storage run in the extended mode. The set of physically possible spot market strategies given a particular β is defined in Definition 7.1.

Definition 7.1. The set of *physically possible spot market strategies* $\mathcal{P}(\beta)$ consists of all sequences $\{\Delta_m\}_{m=0}^{M-1}$ such that

1. $\beta \in \left[-\left(\frac{\overline{q}-\overline{Q}}{\overline{Q}}\right), \frac{\overline{q}}{\overline{Q}}\right]$ in summer, $\beta \in \left[-\frac{\overline{q}}{\overline{Q}}, \left(\frac{\overline{q}-\overline{Q}}{\overline{Q}}\right)\right]$ in winter,

2. $\Delta_m \in \left[-\frac{\overline{q}}{M}, \left(\frac{\overline{q}-\beta\overline{Q}}{M}\right)\right]$ for $\beta \geq 0$, $\Delta_m \in \left[-\left(\frac{\overline{q}+\beta\overline{Q}}{M}\right), \frac{\overline{q}}{M}\right]$ for $\beta < 0$,

3. $Q_N = \sum_{m=0}^{N}(\Delta_m + \beta\overline{Q}/M) \in [0, \overline{Q}]$, $N = 1, \ldots, M-1$,

4. $\sum_{m=0}^{M-1} \Delta_m = (1-\beta)\overline{Q}$ in summer, $\sum_{m=0}^{M-1} \Delta_m = (-1-\beta)\overline{Q}$ in winter.

Condition 4 of Definition 7.1 simply says that the storage is empty at the beginning of the summer injection month and full at the end and that we have the opposite situation for the winter extraction month.

A financial contract that allows its holder to choose when and how much of a commodity to buy or deliver, without violating a set of pre-specified rules, is called a *swing option*. The class of swing options encountered in this paper is described formally in Definition 7.2.

Definition 7.2. A swing option of class $\mathcal{S}(\beta)$ for swing month i with M exercise opportunities $t_m = T_i + m/M$, $m \in \{1, \ldots, M\}$, allows the holder, at each t_m to choose $\Delta_m \in \mathcal{P}(\beta)$. The holder thereby immediately receives the amount $Y_m = -\Delta_m S_{t_m}^i$.

By Definition 7.2, there is one type of swing option for the summer and one for the winter for each β. Requiring that Δ_m is chosen at time t_m, leads to the following definition of the set of admissible spot market strategies $\mathcal{A}(\beta)$.

Definition 7.3. The set of admissible spot market strategies $\mathcal{A}(\beta)$ consists of all sequences $\{\Delta_m\}_{m=0}^{M-1} \in \mathcal{P}(\beta)$ such that $\Delta_m \in \mathcal{F}_{t_m}$.

The spot price $s = S_t^i$ and storage level $Q = Q_t$ are both Markov processes, so the value process SW_t^i, $T_i \leq t \leq T_{i+1}$ of a swing option of class $\mathcal{S}(\beta)$ satisfies

$$\mathrm{SW}_t^i = \mathrm{SW}^i(t, s, Q), \tag{7.1}$$

for any fixed and feasible β.

Below, we price a summer swing option is priced; the winter swing option is priced analogously, and all results derived in this section hold for the winter swing option as well.

The core of the valuation is the Bellman equation

$$\begin{aligned} \mathrm{SW}^i(t_M, s, \overline{Q}) &= 0, &(7.2)\\ \mathrm{SW}^i(t_m, s, Q) &= \sup_{\Delta_m \in \mathcal{A}(\beta)} \Big\{ -\Delta_m s + e^{-r(t_{m+1}-t_m)} \mathbb{E}[\mathrm{SW}^i(t_{m+1}, S^i_{t_{m+1}}, Q_{m+1}) | \mathcal{F}_{t_m}] \Big\}, \end{aligned}$$

with $Q_{m+1} = Q + \Delta_m + \beta\overline{Q}/M$ being the storage content at time t_{m+1} given the swing action Δ_m and the flow from the β futures contracts. Also, the storage is full at the end of the injection month. Since the payoff $Y_m = -\Delta_m S_{t_m}$ is continuous in Δ_m, and Δ_m belongs to a closed and bounded set, the supremum is attained by some Δ^*_m, which in general depends on t_m, S_{t_m} and Q.

Repeating (7.2) until $m = 0$ yields

$$\mathrm{SW}^i(t_0, s, 0) = \sup_{\{\Delta_m\}_{m=0}^{M-1} \in \mathcal{A}(\beta)} \sum_{m=0}^{M-1} e^{-r(t_m - t_0)} \mathbb{E}[-\Delta_m S^i_{t_m} | \mathcal{F}_{t_0}], \tag{7.3}$$

where we recall that the storage tank is empty before the beginning of the summer injection month. Prior to the start of the swing month, i.e. if $t \le T_i$, we have

$$\begin{aligned} \mathrm{SW}^i_t &= e^{-r(T_i - t)} \mathbb{E}[\mathrm{SW}^i(t_0, S^i_{t_0}, 0) | \mathcal{F}_t] &(7.4)\\ &= e^{-r(T_i - t)} \mathbb{E}[\mathrm{SW}^i(t_0, F^i_{T_i}, 0) | \mathcal{F}_t] &(7.5)\\ &= \mathrm{SW}^i(t, F^i_t, 0), &(7.6) \end{aligned}$$

by general derivatives pricing theory and the fact that $S^i_{t_0} = F^i_{T_i}$ by (4.9). A number of useful properties of swing option prices are collected in Proposition 7.4 below, which proof is found in Appendix C.2.

Proposition 7.4. *Let S^i_t follow the spot price dynamics (4.9). Then the following properties hold for the price SW^i_t of a class $\mathcal{S}(\beta)$ swing option:*

(1) *The swing option price SW^i_t is homogeneous of degree one with respect to scaling of $\overline{Q}$ and $\overline{q}$ by a common factor.*
(2) *For $t \le T_i$ we have that*

$$SW^i(t, F^i_t, Q_0) = k_i F^i_t,$$

where $k_i \equiv SW^i(t_0, 1, Q)$ is deterministic and only depends on the market and storage parameters $\hat{\alpha}_i$, r, $\hat{\sigma}_i$, ,$|T_{i+1} - T_i|$, β, $\overline{Q}$, $\overline{q}$ and M.
(3) *Let $\hat{\sigma}_i = 0$ and/or $\hat{\alpha}_i = 0$. Then*

$$SW^i(t_0, F^i_{T_i}, Q_0) = \begin{cases} (-1+\beta) F^t_{T_i}, & \text{(summer)},\\ (1+\beta) F^t_{T_i}, & \text{(winter)}. \end{cases}$$

.

The constant $k_i \equiv \mathrm{SW}^i(t_0, 1, Q)$ is computed by the recurrence relation (7.2), and the details of the numerical implementation are discussed in Section 9.

8. Valuation - extended mode of operation

In this mode of operation, the storage owner must decide which month is to be the injection or extraction month, as well as how many futures to buy and sell during this month. As a consequence of this decision, the owner receives β futures contracts *plus* a swing option of class $\mathcal{S}(\beta)$ for that particular month. More specifically, if deciding for month i at time $t \leq T_i$, the owner receives the amount Y_i given by

$$\begin{aligned} Y_i &= -\beta e^{-r(T_i-t)} F_t^i + \mathrm{SW}_t^i(\beta) \\ &= (-\beta e^{-r(T_i-t)} + k_i(\beta)) F_t^i \quad (8.1) \\ &\equiv c_i(\beta) F_t^i \quad (8.2) \end{aligned}$$

where (8.1) follows by Part 2 of Proposition 7.4. Arguments similar to the ones used at the beginning of Section 6 implies that there is no reward for early decisions.

First, assume that β is fixed. Then arguments similar to the ones used in Section 6 to derive (6.1) to (6.9) give the value V_s of the summer injection as

$$\begin{aligned} V_s(T_8) &= c_8(\beta) F_{T_8}^8, \quad (8.3) \\ V_s(T_7) &= \max\left\{ c_7(\beta) F_{T_7}^7, e^{-r(T_8-T_7)} c_8(\beta) F_{T_7}^8 \right\}, \quad (8.4) \\ V_s(T_6) &= \max\left\{ c_6(\beta) F_{T_6}^6, e^{-r(T_7-T_6)} \mathbb{E}[V_s(T_7)|\mathcal{F}_{T_6}] \right\}, \quad (8.5) \\ V_s(t) &= e^{-r(T_6-t)} \mathbb{E}[V_s(T_6)|\mathcal{F}_t], \quad t \leq T_6, \quad (8.6) \end{aligned}$$

with analogous expressions being valid for the winter extraction value $V_w(t)$. Again, from Section 6 we note that (8.5) may be evaluated analytically by Lemma 6.1, and Part 1 of Proposition 7.4 tells us that the valuation problem is scalable in $\overline{Q}$ and $\overline{q}$.

Expressions (8.3) to (8.6) give one storage value for each feasible β. In Proposition 8.1 below, we show that $\beta = 0$ is always an optimal choice for both the summer injection and winter extraction.

Proposition 8.1. *Let β be as in Definition 7.1. Then $V_s(t)$, $t \leq T_8$ and $V_w(t)$, $t \leq T_2$, are maximised for $\beta = 0$.*

Proof. A portfolio of β futures plus an $\mathcal{S}(\beta)$ swing option may be viewed as a swing option, which set of admissible spot market strategies $\hat{\mathcal{A}}(\beta)$ consists of all sequences $\{\widehat{\Delta}_m\}_{m=0}^{M-1}$ such that

1. $\widehat{\Delta}_m \in \left[-\left(\frac{\overline{q} - \beta\overline{Q}}{M} \right), \frac{\overline{q}}{M} \right]$ when $\beta \geq 0$, $\widehat{\Delta}_m \in \left[-\frac{\overline{q}}{M}, \left(\frac{\overline{q} + \beta\overline{Q}}{M} \right) \right]$ when $\beta < 0$,
2. $Q_N = \sum_{m=0}^{N} \hat{\Delta}_m \in [0, \overline{Q}]$, $N = 1, \ldots, M-1$,
3. $\sum_{m=0}^{M-1} \hat{\Delta}_m = \overline{Q}$ in summer, $\sum_{m=0}^{M-1} \hat{\Delta}_m = -\overline{Q}$ in winter).

The admissible spot market strategies $\{\Delta_m\}_{m=0}^{M-1}$ of an $\mathcal{S}(0)$ swing option satisfy Points 2 and 3 above, but with Point 1 replaced by $\Delta_m \in [-\overline{q}/M, \overline{q}/M]$. Thus $\mathcal{A}(0) \supseteq \hat{\mathcal{A}}(\beta)$, with equality if $\beta = 0$, which concludes the proof. □

Proposition 8.1 tells us that choosing to take part of the injection/extraction from a futures contract, reduces the storage flexibility. Hence making all injections/extractions on the spot market maximises the storage value. Note that this result did not use any properties about our spot- and futures price models. Thus it holds for any market model, provided that the no-arbitrage condition 4.7 holds.

What Proposition 8.1 does *not* tell us, is how much value is lost by choosing $\beta \neq 0$. We investigate that numerically in Section 11.

9. Numerical implementation

In this section we start by discussing the implementation of the Monte Carlo valuation computation of expressions (6.5) and (8.6). The swing option valuation is also described in more detail.

Given the current futures price vector $\mathbf{F}_t \equiv (F_t^n, F_t^{n+1}, F_t^{n+2})$, the risk-neutral distribution of $\mathbf{F}_{T_i} = (F_{T_i}^n, F_{T_i}^{i+1}, F_{T_i}^{i+2})$ is joint log-normal with mean vector μ and covariance matrix $\mathbf{C}$ specified in (4.3) to (4.4). We start by letting $\mathbf{C} = \mathbf{UDU}^T$ be the eigenvalue decomposition of the symmetric matrix $\mathbf{C}$, and define $\mathbf{A} = \mathbf{UD}^{1/2}$, which exists since $\mathbf{C}$ is positive semi definite. Then the Monte Carlo algorithm may be described with the following steps.

(1) Generate a 3×1 vector $\mathbf{x}$ of i.i.d. $N(0,1)$ random variables and let $\mathbf{y} = \mu + \mathbf{Ax}$. The vector $\mathbf{y}$ has mean μ and covariance matrix $\mathbf{C}$.
(2) Compute the simulated futures prices as $\mathbf{F}_{T_i} = \mathbf{F}_t \exp(\mathbf{y})$, where the multiplication and exponential function are interpreted componentwise.
(3) Compute a simulated payoff V by inserting the simulated futures prices into the analytical expressions for $V_s(T_6)$ and $V_w(T_{12})$. Compute the total value as the sum and save the value.
(4) Repeat N_{sim} times, compute the sample average and discount.

The homogeneity of the max-function would allow the use one of the three futures contracts as a numeraire, which would reduce the dimensionality from three to two. In our application however, computational speed is not critical, so we refrain from doing this improvement.

We now turn to the implementation of the swing option valuation. Here we only show how a summer swing option of class $\mathcal{S}(0)$ is implemented, with the other types are treated analogously.

First, the state variables (t, S, Q) are discretised. The storage level Q is discretised as $Q_k = k\overline{Q}/K$ with $k \in \{0, \ldots, K\}$. A discrete swing action may then be written $\Delta_m = \delta_m \overline{Q}/K$, $\delta_m \in \{k_1, \ldots, k_2\}$ for some integers $k_1 \leq k_2$. Here we assume that $\overline{Q}$, $\overline{q}$, M and K are chosen such that all extremal swing actions belong to this discrete set of swing actions. The rationale for capturing these bang-bang strategies is discussed briefly in Note 9.1 below.

Following Jaillet, Ronn and Tompaidis [7], $X_t = \log S_t - f(t)$ is used as state variable and this stochastic process is discretised into a trinomial tree. We let the time grid coincide with the opportunities t_m, so the time step Δt satisfies $\Delta t = t_{m+1} - t_m = 1/M$ months. The x-discretisation at time t_m is $x_l = l\Delta x$ for $l \in \{-m, \ldots, +m\}$, and by (4.6) the spot price grid at time t_m becomes $s_l = \exp[f(t_m) + x_l]$, where $f(t)$ is given in Equation (4.8). Hull and White [5] show that $\Delta x = \hat{\sigma}_i\sqrt{3\Delta t}$ optimises accuracy while maintaining stability, and with this choice, the risk neutral transition probabilities from x_l become

$$\begin{cases} p_d^l &= \frac{1}{6} + \frac{\hat{\alpha}_i\sqrt{\Delta t}}{\hat{\sigma}_i\sqrt{3}} x_l \quad \text{Down,} \\ p_s^l &= \frac{2}{3} \quad \text{Same,} \\ p_u^l &= \frac{1}{6} - \frac{\hat{\alpha}_i\sqrt{\Delta t}}{\hat{\sigma}_i\sqrt{3}} x_l \quad \text{Up.} \end{cases} \tag{9.1}$$

The discrete version of the Bellman equation (7.3) now becomes

$$
\begin{aligned}
\mathrm{SW}^i(t_M, s_l, Q_K) &= 0 \\
\mathrm{SW}^i(t_m, s_l, Q_k) &= \max_{\delta_m \in \mathcal{A}} \Big\{ -\delta_m \overline{Q}/K s_l + \frac{1}{1+r\Delta t} \Big[p_d^l \mathrm{SW}^i(t_{m+1}, s_{l-1}, Q_{k+\delta_m}) + \\
& \quad p_s^l \mathrm{SW}^i(t_{m+1}, s_l, Q_{k+\delta_m}) + p_u^l \mathrm{SW}^i(t_{m+1}, s_{l+1}, Q_{k+\delta_m}) \Big] \Big\}, \quad (9.2)
\end{aligned}
$$

Equation (9.2) may be implemented with nested for-loops over the indices m, k, l, where checks are made in each iteration if the spot market transaction $\Delta_m = \delta_m \Delta Q$ is feasible or not.

Note 9.1. The reason for including the extremal swing actions is that numerical experiments suggest that most, but not all swing actions are extremal. This indicates that not truncating these swing options could be a good heuristic for reducing the $Q-$discretisation error.

10. Valuation examples - swing options

In this section we investigate how the summer and winter swing options of class $\mathcal{S}(0)$ depend on some model and storage parameters. To compute the prices, a Matlab implementation of the numerical swing option pricing method described in Section 9 is used. The specific storage and model parameters together with the valuation results are shown in Table 3 and in Figure 2 below.

$\hat{\sigma}_i \backslash \hat{\alpha}_i$	0.0	1.0	2.0	3.0
0.0	-10000	-10000	-10000	-10000
0.3	-10000	-9996	-9984	-9976
0.6	-10000	-9983	-9967	-9952
0.9	-10000	-9979	-9951	-9928

$\hat{\sigma}_i \backslash \hat{\alpha}_i$	0.0	1.0	2.0	3.0
0.0	10000	10000	10000	10000
0.3	10000	10008	10016	10024
0.6	10000	10017	10033	10048
0.9	10000	10025	10049	10071

Table 3: The dollar price at the beginning of the summer (left) and winter (right) swing options of class $\mathcal{S}(0)$ for different values of $\hat{\sigma}_i$ and $\hat{\alpha}_i$ (annualised). Market parameters: $F^i_{T_i} = \$10,000$ USD, $r = 0.03$ (per annum). Storage parameters: $\overline{Q} = 10^4$ (mmBtu), $\overline{q} = 2\overline{Q}$ (Btu/month), $M = 30$, and $K = 100$ levels.

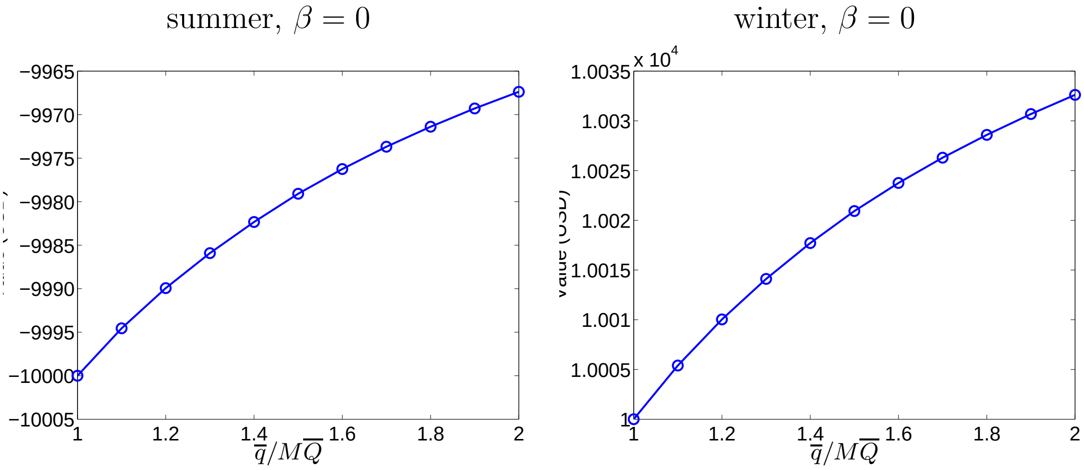

Figure 2: Dependence on $\overline{q}$ for the summer and winter swing options of class $\mathcal{S}(0)$. Market parameters: $F^i_{T_i} = \$10,000$, $r = 0.03$ (per annum), $\hat{\alpha}_i = 2.0$ (per annum) and $\hat{\sigma}_i = 0.6$ (per annum). Storage parameters: $\overline{Q} = 10^4$ (mmBtu), $M = 30$, $K = 100$ levels.

The results in Table 3 show that the swing option prices increase with $\hat{\alpha}_i$ for fixed $\hat{\sigma}_i$ and with $\hat{\sigma}_i$ for fixed $\hat{\alpha}_i$. A large value of $\hat{\alpha}_i$ means that today's spot price will have a greater impact on future spot prices, making prediction easier. Conversely, if $\hat{\alpha}_i = 0$, impact of the current spot price on future spot prices is minimal, and the proof of Proposition 7.4 shows that any feasible strategy is optimal.

A larger value of $\hat{\sigma}_i$ increases the probability of large spot prices, which also creates larger trading gains.

As expected, Figure 1 reveals that the value of all four types of swing options increase with $\overline{q}$, but the rate of this increase decreases.

11. Valuation examples - storage operations

Now we to the valuation of the storage operations. First we investigate how the summer and winter storage values on June 1 depend on model and storage parameters in the different modes of operation defined in Section 3. In light of Proposition 8.1, we use $\beta = 0$ for the extended mode, meaning that all transactions are on the spot market exclusively. However, we will also try $\beta = 1$ for the summer and $\beta = -1$ for the winter in order to see how much storage value is lost by using futures.

The Monte Carlo computations of the June 1 value of the winter operations use the same 3×10^6 pseudo random numbers for all examples. Results are given in Tables 4 to 7.

	Mode	$\alpha_i = 0.0$	$\alpha_i = 1.0$	$\alpha_i = 2.0$	$\alpha_i = 3.0$	$\alpha_i = 4.0$
$\sigma_i = 0.0$	B	-9950	-9950	-9950	-9950	-9950
	F	-9950	-9950	-9950	-9950	-9950
	E($\beta = 0$)	-9950	-9950	-9950	-9950	-9950
	E($\beta = 1$)	-9950	-9950	-9950	-9950	-9950
$\sigma_i = 0.3$	B	-9950	-9950	-9950	-9950	-9950
	F	-9885	-9887	-9881	-9872	-9862
	E($\beta = 0$)	-9985	-9878	-9863	-9845	-9826
	E($\beta = 1$)	-9985	-9878	-9864	-9846	-9827
$\sigma_i = 0.6$	B	-9950	-9950	-9950	-9950	-9950
	F	-9809	-9812	-9800	-9782	-9762
	E($\beta = 0$)	-9809	-9795	-9765	-9729	-9691
	E($\beta = 1$)	-9985	-9895	-9766	-9731	-9693
$\sigma_i = 0.9$	B	-9950	-9950	-9950	-9950	-9950
	F	-9733	-9737	-9719	-9692	-9662
	E($\beta = 0$)	-9773	-9711	-9667	-9613	-9556
	E($\beta = 1$)	-9773	-9812	-9669	-9616	-9580

Table 4: June 1 dollar value of the summer injection for different values of annualised σ_i and α_i. Here B=Benchmark, F=Futures market mode, and E=Extended mode. Parameters: $F^6_{T_6} = F^7_{T_6} = F^8_{T_6} = 10000$ USD, $r = 0.03$ (per annum), $\lambda = 0.3$ (per annum), $M = 30$, $\overline{Q} = 10^4$ (mmBtu), $\overline{q} = 2\overline{Q}$ (Btu/month), and $K = 100$ levels.

	Mode	$\alpha_i = 0.0$	$\alpha_i = 1.0$	$\alpha_i = 2.0$	$\alpha_i = 3.0$	$\alpha_i = 4.0$
$\sigma_i = 0.0$	B	9851	9851	9851	9851	9851
	F	9851	9851	9851	9851	9851
	E($\beta = 0$)	9851	9851	9851	9851	9851
	E($\beta = -1$)	9851	9851	9851	9851	9851
$\sigma_i = 0.3$	B	9851	9851	9851	9851	9851
	F	10152	10090	10077	10078	10083
	E($\beta = 0$)	10152	10100	10095	10105	10119
	E($\beta = -1$)	10152	10100	10095	10004	10118
$\sigma_i = 0.6$	B	9851	9851	9851	9851	9851
	F	10479	10356	10329	10330	10339
	E($\beta = 0$)	10479	10375	10365	10386	10414
	E($\beta = -1$)	10479	10374	10364	10384	10412
$\sigma_i = 0.9$	B	9851	9851	9851	9851	9851
	F	10806	10623	10508	10583	10597
	E($\beta = 0$)	10806	10651	10637	10669	10713
	E($\beta = -1$)	10806	10650	10636	10666	10709

Table 5: June 1 dollar value of the winter extraction for different values of annualised σ_i and α_i. Here B = Benchmark, F = Futures market, E = Extended. Parameters: $F^6_{T_6} = F^7_{T_6} = F^8_{T_6} = \$10000$, $r = 0.03$ (per annum), $\lambda = 0.3$ (per annum), $M = 30$, $\overline{Q} = 10^4$ (mmBtu), $\overline{q} = 2\overline{Q}$ (Btu/month), and $K = 100$ levels.

Parameters	Mode	$\overline{q}/\overline{Q} = 1.0$	$\overline{q}/\overline{Q} = 1.2$	$\overline{q}/\overline{Q} = 1.4$	$\overline{q}/\overline{Q} = 1.6$	$\overline{q}/\overline{Q} = 1.8$
$\sigma_i = 0.0$	E($\beta = 0$)	-9950	-9950	-9950	-9950	-9950
$\alpha_i = 0.0$	E($\beta = 1$)	-9950	-9950	-9950	-9950	-9950
$\sigma_i = 0.3$	E($\beta = 0$)	-9887	-9884	-9882	-9880	-9879
$\alpha_i = 1.0$	E($\beta = 1$)	-9887	-9984	-9882	-9881	-9879
$\sigma_i = 0.3$	E($\beta = 0$)	-9872	-9864	-9857	-9852	-9848
$\alpha_i = 3.0$	E($\beta = 1$)	-9872	-9864	-9858	-9853	-9949
$\sigma_i = 0.9$	E($\beta = 0$)	-9737	-9729	-9723	-9718	-9714
$\alpha_i = 1.0$	E($\beta = 1$)	-9737	-9729	-9723	-9719	-9715
$\sigma_i = 0.9$	E($\beta = 0$)	-9691	-9667	-9649	-9635	-9623
$\alpha_i = 3.0$	E($\beta = 1$)	-9691	-9667	-9650	-9636	-9625

Table 6: June 1 dollar value of the summer injection for different values of $\overline{q}$ and combinations of α_i and σ_i. Here S=Spot market and C=Combined mode of operation. Parameters: $F^6_{T_6} = F^7_{T_6} = F^8_{T_6} = 10000$ USD, $r = 0.03$ (per annum), $\lambda = 0.3$ (per annum), $M = 30$, $\overline{Q} = 10^4$ mmBTU, and $K = 100$ levels.

Parameters	Mode	$\overline{q}/\overline{Q} = 1.0$	$\overline{q}/\overline{Q} = 1.2$	$\overline{q}/\overline{Q} = 1.4$	$\overline{q}/\overline{Q} = 1.6$	$\overline{q}/\overline{Q} = 1.8$
$\sigma_i = 0.0$	E($\beta = 0$)	9851	9851	9851	9851	9851
$\alpha_i = 0.0$	E($\beta = -1$)	9851	9851	9851	9851	9851
$\sigma_i = 0.3$	E($\beta = 0$)	10091	10094	10096	10097	10099
$\alpha_i = 1.0$	E($\beta = -1$)	10091	10094	10096	10097	10098
$\sigma_i = 0.3$	E($\beta = 0$)	10078	10087	10093	10098	10102
$\alpha_i = 3.0$	E($\beta = -1$)	10078	10086	10092	10097	10101
$\sigma_i = 0.9$	E($\beta = 0$)	10623	10632	10639	10644	10648
$\alpha_i = 1.0$	E($\beta = -1$)	10623	10632	10638	10643	10647
$\sigma_i = 0.9$	E($\beta = 0$)	10584	10610	10630	10646	10659
$\alpha_i = 3.0$	E($\beta = -1$)	10584	10609	10629	10644	10656

Table 7: June 1 dollar value of the winter extraction for different values of $\overline{q}$ and combinations of α_i and σ_i. Here S=Spot market and C=Combined mode of operation. Parameters: $F^6_{T_6} = F^7_{T_6} = F^8_{T_6} = 10000$ USD, $r = 0.03$ (per annum), $\lambda = 0.3$ (per annum), $M = 30$, $\overline{Q} = 10^4$ mmBTU, and $K = 100$ levels.

As expected by Proposition 8.1, Tables 4 to 7 show that the extended mode with $\beta = 0$ is the most valuable, but the differences compared to the $\beta_s = 1$, $\beta_w = -1$ strategy are minimal for all parameter values. Not less interesting, the futures mode is competitive roughly when α_i is less than 1.5 or the storage flexibility $\overline{q}/\overline{Q}$ is less than 1.2 to 1.6.

As expected, higher volatility in terms of large σ_i yields higher values. Finally, the marginal value increase caused by increasing $\overline{q}$ increases with increasing values of α_i and σ_i.

We conclude this section by performing a valuation of the storage using the model parameters from Tables 1 to 2 in Section 5.

Mode	V_s	V_w	$V_s + V_w$
Benchmark	-6264	7793	1528
Futures market	-6229	8085	1856
Extended, $\beta_s = 0$, $\beta_w = 0$	-6221	8109	1888
Extended, $\beta = 1$, $\beta_w = -1$	-6222	8108	1887

Table 8: May 27, 2005 dollar values of the summer and winter operations for all modes when the model parameters are taken from Tables 1 to 2. May 27, 2005 futures prices: $F_{T_6}^6 = 6315$, $F_{T_6}^7 = 6280$, $F_{T_6}^8 = 6350$, $F_{T_6}^{12} = 7600$, $F_{T_6}^1 = 7920$, $F_{T_6}^2 = 7950$. Other parameters: $r = 0.03$ (per annum), $M = 30$, $\overline{Q} = 10^4$ (mmBtu), $\overline{q}/\overline{Q} = 2$, $K = 100$ levels.

Here the differences between the three modes are quite small. The most important takeaway from Table 8 however, is the importance of taking the decision of summer injection or winter extraction based on some kind of dynamic optimisation approach. This increases the storage value by 20% compared to the static benchmark approach.

12. Concluding remarks

We define three modes of operation of a natural gas storage facility and prove that a pure spot-market strategy creates the highest storage value.

However, we also demonstrate that it is often possible to come very close to this value by a combined spot-futures strategy, where the net injection/extraction from the spot market over the injection/extraction month is zero. For some market and storage parameters, including the ones we estimated from real market data, the pure futures market mode is also competitive.

One advantage with the combined spot-futures strategy is that NYMEX futures could be used for partial hedging. In the futures mode, using our complete futures model, the hedging could even be perfect.

Our approach is by far the most general one. One could extend the injection/extraction month to the entire summer/winter periods, or starting operations (futures or extended mode) during the now passive Spring and Autumn months. This would lead to a further increase of the storage value.

If the results of this paper were to be used in an actual storage valuation, Assumption 1 would have to be extended to incorporate the features of a real facility outlined in Appendix A. Our valuation framework of Sections 6 to 8 could be extended to hold under these assumptions as well. In this case, it is likely that the futures mode of operation becomes even more competitive, since injection/extraction at an even rate minimises the total flow and hence reduces the losses.

References

[1] F. Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3:167–179, March 1976.

[2] A. Cortazar and E. Schwartz. The valuation of commodity-contingent claims. *The Journal of Derivatives*, 4(1):27–39, Summer 1994.

[3] M. Dahlgren. A continuous time model to price commodity based swing options. *Review of Derivatives Research*, 8(1):27–47, 2005.

[4] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuation. *Econometrica*, 60(1):77–105, 1992.
[5] J. Hull and W. A. Numerical procedures for implementing term structure models i: single factor models. *Journal of Derivatives*, pages 7–16, Fall 1994.
[6] A. Ibáñz. Valuation by simulation of contingent claims with multiple early exercise opportunities. *Mathematical Finance*, 14(2):223–248, 2004.
[7] P. Jaillet, E. Ronn, and S. Tompaidis. Valuation of commodity-based swing options. *Management Science*, 50(7):909–921, 2004.
[8] W. Magrabe. The value of an option to exchange one asset for another. *Journal of Finance*, 33:177–186, 1978.
[9] M. Manoliu. Storage options valuation using multilevel trees and calendar spreads. *International Journal of Theoretical and Applied Finance*, 7(4):425–464, 2004.
[10] C. Parsons. Valuaing commodity storage contracts: A two-factor tree approach. Preprint WTM Energy Software LLC, 2005.
[11] E. Schwartz. The stochastic behavior of commodity prices: implications for valuation and hedging. *Journal of Finance*, 52(3):923–972, July 1997.

Appendix A. A sample natural gas storage facility

The characteristics of this natural gas storage is supplied by Dr. Kevin G. Kindall, Commercial Division, ConocoPhillips.

(1) Capacity: $\overline{Q} = 10^6$ (mmBtu).
(2) Maximal injection rate: The maximal injection rate $\overline{q}_i$ depend on the storage level Q according to the following table:

$Q/\overline{Q}$	$\overline{q}_i$(Btu/month)
0-50%	$0.167\overline{Q}$
50-100%	$0.140\overline{Q}$

(3) Maximal extraction rate: The maximal extraction rate $\overline{q}_e$ depend on the storage level Q according to the following table:

$Q/\overline{Q}$	$\overline{q}_e$(Btu/month)
0-10%	$0.250\overline{Q}$
10-16%	$0.333\overline{Q}$
16-30%	$0.375\overline{Q}$
30-35%	$0.475\overline{Q}$
30-35%	$0.500\overline{Q}$

(4) Injection and extraction rate changes: Usually done on a daily basis.
(5) Costs: Injection cost: 2.18c/mmBTU, extraction cost: 1.95c/mmBTU. Due to leakage, 3.59% of the injected but none of the extracted volume is lost.

The analytical framework of Sections 2, 6, 7 and 8 could be extended to incorporate these features without much difficulty.

Appendix B. Motivation of the assumptions about $\hat{\sigma}_i$ and $\hat{\alpha}_i$

Consider Model 1 in Schwartz [11], in which the risk neutral dynamics of the spot price S_t for $t \geq 0$ is given by

$$\left\{ \begin{array}{rcl} S_t & = & \exp\left[f(t) + X_t\right] \\ dX_t & = & -\hat{\alpha} X_t\, dt + \hat{\sigma}\, dW_t, \quad X_0 = 0, \end{array} \right. \tag{B.1}$$

with $\hat{\alpha} > 0$, $\hat{\sigma} > 0$ and $f(t)$ deterministic for $t \geq 0$. Our spot price S_t^i for month i, introduced in Assumption 4, follows (B.1) for $T_i \leq t < T_{i+1}$, but is not defined for other times. This means that we are not allowed to compare spot prices across months, and as a consequence we cannot find a relationship between the spot- and futures-market parameters. Nevertheless, we think that the similarities between (B.1) and our model are so great that we make the following assumption.

The relationship between the spot- and futures-market parameters for our spot price model is the same as for the model (B.1).

This relationship is derived as follows. Let $t_m = T_i + m(T_{i+1} - T_i)/M$, $t \leq t_m$ and write

$$G_t^{t_m} = \exp\left\{ f(t_m) + [\log S_t - f(t)]\, e^{-\hat{\alpha}(t_m - t)} + \frac{\hat{\sigma}^2}{4\hat{\alpha}}\left[1 - e^{-2\hat{\alpha}(t_m - t)}\right] \right\}.$$

Then by Schwartz [11]

$$F_t^i = \frac{1}{M}\sum_{m=0}^{M-1} e^{-r(t_m - t_0)} G_t^{t_m} \tag{B.2}$$

$$\equiv h(t, S_t), \tag{B.3}$$

for $t \leq T_i$. By (B.2), h is a continuous and strictly increasing function of S_t for each t, which implies that h has a well defined inverse $S_t = h^{-1}(t, F_t^i)$. An application of Itô's Lemma to (B.2) and some rearranging yield

$$\frac{dF_t^i}{F_t^i} = \hat{\sigma} e^{-\hat{\alpha}(T_i - t)} \left(\frac{\sum_{m=0}^{M-1} e^{-(r+\alpha)(t_m - t_0)} G_t^{t_m}}{\sum_{m=0}^{M-1} e^{-r(t_m - t_0)} G_t^{t_m}} \right) dW_t$$

$$\equiv \hat{\sigma} e^{-\hat{\alpha}(T_i - t)} \beta(t, F_t^i) dW_t. \tag{B.4}$$

Here the definition of the function β is enabled by the fact that h is invertible. Moreover, $\beta(t, F_t^i) \in (0, 1)$, but since it is not constant, F_t^i is not log-normal. In the case when the terms $e^{-r(t_m - t_0)} G_t^{t_m}$ are all equal, we obtain

$$\beta = \sum_{m=0}^{M-1} e^{-\alpha(t_m - t_0)} \tag{B.5}$$

$$\approx \frac{1 - \exp(-\hat{\alpha}(T_{i+1} - T_i))}{\hat{\alpha}(T_{i+1} - T_i)}, \tag{B.6}$$

where the sum on the first line is approximated by an integral. Setting $\beta(t, F_t^i) = \beta$ in (B.4) yields an approximate risk neutral dynamics for F_t^i as

$$\frac{dF_t^i}{F_t^i} = \beta \hat{\sigma} e^{-\hat{\alpha}(T_i - t)} dW_t. \tag{B.7}$$

If, on the other hand, we had started with prescribing that the risk neutral dynamics for F_t^i is given by

$$\frac{dF_t^i}{F_t^i} = \sigma e^{-\alpha(T_i - t)} dW_t, \tag{B.8}$$

a comparison between (B.7) and (B.8) gives

$$\begin{aligned}\alpha &= \hat{\alpha}\\ \sigma &= \beta\hat{\sigma}.\end{aligned}$$

In the Black model, when $\hat{\alpha} = 0$, $\beta(t, F_t^i) = 1$, implying that $\hat{\sigma} = \sigma$ in this case.

Appendix C. Proofs

C.1. **Proof of Lemma 6.1.**

Proof. The case when c_1 and c_2 do not have the same sign is trivial.

If c_1 and c_2 are of equal sign, the pricing formulas are derived by a change of numeraire approach similar to the one used by Magrabe [8]. This includes rewriting the payoff Y as

$$\begin{aligned}Y &= \max\{c_1 F_{T_{i+1}}^{i+1}, c_2 F_{T_{i+1}}^{i+2}\}\\ &= F_{T_{i+1}}^{i+1} \max\left\{c_1, c_2 \frac{F_{T_{i+1}}^{i+2}}{F_{T_{i+1}}^{i+1}}\right\}.\end{aligned}$$

Using the futures price F_t^{i+1} as numeraire and letting $\hat{Y} = Y/F_{T_{i+1}}^{i+1}$ and $\hat{F}_t = F_t^{i+2}/F_t^{i+1}$ yield

$$\hat{Y} = \max\left\{c_1, c_2 \hat{F}_{T_{i+1}}\right\}.$$

By (4.5), $\log(\hat{F}_{T_{i+1}}/\hat{F}_t)$ is a random variable with variance

$$\Sigma^2(t, T_{i+1}) = C_{i+1,i+1}(t, T_{i+1}) + C_{i+2,i+2}(t, T_{i+1}) - 2C_{i+1,i+2}(t, T_{i+1}),$$

so if X denotes a standard normal random variable, we have that

$$\hat{F}_T = \hat{F}_t \exp\left(-\frac{1}{2}\Sigma^2(t, T_{i+1}) + \Sigma(t, T_{i+1})X\right)$$

under the equivalent martingale measure $\hat{P}$ implied by the chosen numeraire. Furthermore, under this numeraire, the exchange option value is given by

$$\hat{V}_{ex}(t) = \hat{\mathbb{E}}[\hat{Y}|\mathcal{F}_t],$$

where $\hat{\mathbb{E}}$ is the expectation with respect to $\hat{P}$. Assuming $c_1 \geq 0$ and $c_2 \geq 0$ and proceeding similarly to the derivation of the Black-Scholes call option formula yield

$$\hat{V}_{ex}(t) = c_1\Phi(D) + c_2\hat{F}_t\Phi(\Sigma - D),$$

with

$$D = \frac{-\log\left(\frac{c_2 F_t^{i+2}}{c_1 F_t^{i+1}}\right) + \frac{1}{2}\Sigma^2}{\Sigma}.$$

Finally, switching back to the bank account B_t as numeraire, and recalling that $F_t^{T_{i+1}}$ refers to a payment occurring at time T_{i+1} give us the pricing formula

$$V_{ex}(t) = c_2 e^{-r(T_{i+1}-t)} F_t^{i+1}\Phi(D) + c_2 e^{-r(T_{i+1}-t)} F_t^{i+2}\Phi(\Sigma - D).$$

The case when $c_1 < 0$, $c_2 < 0$ is proved analogously. □

C.2. **Proof of Proposition 7.4.**

Part 1. If we multiply $\overline{Q}$ and $\overline{q}$ by the same factor $\kappa > 0$, it follows by Definition 7.1, that all physically possible swing actions $\Delta_m \in \mathcal{P}(\beta)$ are multiplied by κ as well. Inserting $\kappa\Delta_m$ into the Bellman equation (7.2) proves the claim.

Part 2. By (4.9), S_t is linear in $F^i_{T_i}$ if started at time t_0, so (7.3) may be rewritten as

$$\begin{aligned}\mathrm{SW}^i(t_0, F^i_{T_i}, Q_0) &= F^i_{T_i} \sup_{\{\Delta_m\}_{m=0}^{M-1} \in \mathcal{A}(\beta)} \sum_{m=0}^{M-1} e^{-r(t_m - t_0)} \mathbb{E}[\Delta_m (S^i_{t_m}/F^i_{T_i}) | \mathcal{F}_{t_0}] \\ &= \mathrm{SW}^i(t_0, 1, Q) F^i_{T_i} \\ &\equiv k_i F^i_{T_i},\end{aligned}$$

where k_i only depends on model and storage parameters but not on the futures price. Consequently we have for $t \le T_i$

$$\begin{aligned}\mathrm{SW}^i(t, F^i_t, Q_0) &= \mathbb{E}[\mathrm{SW}^i(t_0, F^i_{T_i}, Q_0) | \mathcal{F}_t] \\ &= k_i F^i_t.\end{aligned}$$

Part 3. By (4.9), $e^{-r(t-t_0)} S^i_t$, $t \ge t_0$ is a $P-$martingale for $\hat{\sigma}_i = 0$ and/or $\hat{\alpha}_i = 0$. Consequently, the discrete time gains process Y_n, $n = 0, \ldots, M-1$ defined as

$$Y_n = \sum_{m=0}^{n} -\Delta_m e^{-r(t_m - t_0)} S^i_{t_m},$$

is a $P-$martingale, implying that every feasible spot market strategy $\{\Delta_m\}_{m=0}^{M-1} \in \mathcal{A}(\beta)$ is optimal. Fixing one such strategy, plugging it into (7.3) and using the no-arbitrage condition (4.7) complete the proof.

Department of Economics, School of Business, Economics and Law, Göteborg University, Box 600, 405 30 Göteborg, Sweden
E-mail address: `mats.kjaer@economics.gu.se`

Department of Finance, McCombs School of Business, 1 University Station, B6600, University of Texas at Austin, Austin, TX. 78712-1179, U. S. A.
E-mail address: `eronn@mail.utexas.edu`

Paper IV

MODELLING THE CORRELATION MATRIX OF NATURAL GAS FUTURES RETURNS

MATS KJAER AND EHUD RONN

Göteborg University & The University of Texas at Austin

Abstract. We specify and estimate two types of parsimonious models of the correlation matrix of natural gas futures returns. The first uses a correlation function to generate the matrix, and the second one employs a factor analytic approach.

There are two important findings. First, there is a distinct drop in correlation between two futures contracts for every April month located between their two maturities. Our suggested models are able to reproduce this effect, which we also explain in micro-economic terms. Second, examples suggest that three to five parameters are sufficient to obtain a good fit to an 18×18 sample correlation matrix.

1. Introduction

Natural gas futures contracts for delivery at Henry Hub, LA., have been traded on the New York Mercantile Exchange (NYMEX) since April 1990. These futures contracts are currently amongst the most actively-traded commodity contracts, attracting hedgers as well as speculators.

Activities such as portfolio selection, Value-at-Risk (VaR) calculation and exotic option pricing often involve estimation of the correlation matrix between returns of futures of different maturities. Typically, the first estimator to consider for this task is the historical sample correlation matrix (SCM). However, having M futures contracts means that $\frac{1}{2}M(M-1)$ parameters need to be estimated. Pourmahdi [9] reports that the SCM is unstable for large M, and that rather large sample sizes are required in order to obtain acceptable confidence intervals.

In this paper we propose and evaluate some parsimonious correlation matrix models for natural gas futures returns. The aim is to reproduce the most important stylised facts of the SCM, using far fewer parameters. Most of the factors that govern demand and supply — such as weather patterns, number of end users, extraction infrastructure and storage capacity — do not change substantially from one year to the next. Thus we believe that *volatilities* do change from year to year, but that correlations may be more stable and possible to describe with some suitable mathematical functions.

Date: April 25, 2006.

Key words and phrases. Correlation matrix modelling, Futures price modelling, Factor analysis, Least squares estimation, Maximum likelihood estimation.

JEL classification: C13, C51.

Mats Kjaer's visit to UT Austin was supported by: The University of Texas at Austin, The Knut & Alice Wallensberg Foundation, and the Dr. Marcus Wallenberg Foundation.

To the best of our knowledge, no papers have previously been published on the correlation structure for a commodity, such as natural gas, with a seasonal demand or supply. We believe that our results are particulary helpful for economic agents who seek a parsimonious method of calculating Value-at-Risk (VAR) of a portfolio. There are two main contributions of this paper. First, we are able to obtain a good fit of an 18×18 sample correlation matrix with models using three to five parameters. Second, we explain the distinct drop in correlation found between two futures contacts for each April month located between their maturities. Most of our proposed models are capable of reproducing this drop.

For non-seasonal commodities and interest rates, research has focused on so-called *factor-analytic models*. Here the first three factors are commonly known as *level, slope* and *curvature,* and they generate most of the correlation structure. This three-factor phenomenon has been observed for interest rates (Rebonato [10]), copper futures (Cortazar and Schwartz [3]) and crude oil futures (Schwartz [11]). Longstaff et. al. [7] use this observation and historical LIBOR rates to estimate the historical factors, and fit the factor weights such that the mean squared error between model implied and market swaption prices is minimised.

In research fields such as geophysics, meteorology, statistics, engineering, social sciences and medicine, the literature concerning correlation modelling is vast. Apart from the factor-analytic approach described above, several other classes of methods are available. We will employ the *correlation-function modeling approach,* where the correlation matrix elements are specified directly from a correlation function of some stochastic process.

Most research work has been done for stationary[1] correlation functions, and Buell [2] and Stein [12] provide methods for their contruction, as well as libraries of such functions. Dunsmuir and Nott [4] show how to derive non-stationary correlation functions from stationary ones and apply this to the modeling of the power output from an array of wind mills.

We will specify a continuous-time model for the joint dynamics of futures prices, and propose some nested non-stationary correlation function and factor-analytic models. First, the model parameters of the individual futures prices are estimated separately with the Maximum Likelihood (ML) method. Second, the correlation matrix models are estimated with both the Un-weighted Least Squares (ULS) and ML methods. Jöreskog [5] discusses the properties of the ML method in this context. Finally, we compute confidence intervals for the estimated parameters, and perform tests to evaluate the models.

This paper is organised as follows. Section 2 specifies the joint futures price dynamics. The data format and quality is discussed in Section 3 and some stylised facts of the sample correlation matrix are given in Section 4. This is followed by the specification of the correlation function and historical factor models in Section 5, and their estimation and evaluation in Section 6. Estimation and test results are presented in Section 7 and Section 8 contains conclusions and final remarks.[2]

[1]The correlation function $C(m,n)$, with $m,n \in \mathbb{Z}$, is called stationary if $C(m,n) = C(m+k, n+k)$ for $k \in \mathbb{Z}$.

[2]In Appendix A, we provide proofs of the important positive semi-definiteness of our models. We also provide the reader with a toolbox, that permit verification of positive semi-definiteness of models other than the ones used in this paper.

2. Futures price model

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, P)$ be a filtered probability space, where $\mathcal{F}_t$ is the P-completion of the natural filtration $\sigma(Z_u, 0 \le u \le t)$ of the M-dimensional Wiener process Z_t. Here P is the real world probability measure, and expectations with respect to this measure are denoted by $\mathbb{E}$. On this probability space, we introduce M futures contracts for delivery during the calendar months 1 to M. Here calendar month m begins at time T_m and under P, the futures price F_t^m, $0 \le t \le T_m$, for delivery during $(T_m, T_{m+1}]$ follows the SDE

$$\frac{dF_t^m}{F_t^m} = \mu_m dt + \sigma_m e^{-\alpha_m(T_m-t)} dW_t^m, \tag{2.1}$$

where $\mu_m \in \mathbb{R}$, $\sigma_m > 0$, $\alpha_m \in \mathbb{R}$, and W_t^m is a one-dimensional Wiener process. Furthermore, $\mathrm{Cov}(W_t^m, W_t^n) = C(m,n)t$, where $C(m,n)$ is an element of an $M \times M$ correlation matrix $\mathbf{C}$. In other words,

(1) $\mathbf{C}$ is symmetric and positive semi-definite.
(2) $\mathbf{C}(m,m) = 1$ and $\mathbf{C}(m,n) \in [-1,1]$, $1 \le m,n \le M$.

If $\alpha_m = 0$ in (2.1), the model introduced in Black [1] is retrieved. It serves as the benchmark model for pricing vanilla futures options.

Next we define the log-returns of F_t^m as

$$R_{mt} \equiv \log(F_t^m / F_{t-\Delta t}^m),$$

which by (2.1) are normally distributed. If Δt is "small" the mean and variance of R_{mt} are given by

$$\mathbb{E}[R_{mt}] = \left[\mu_m - \frac{\sigma_m^2}{2} e^{-2\alpha_m(T_m-t)}\right]\Delta t, \tag{2.2}$$

$$\mathrm{Var}\,(R_{mt}) = \sigma_m^2 e^{-2\alpha_m(T_m-t)}\Delta t, \tag{2.3}$$

respectively, and it follows that

$$\mathrm{Corr}\,(R_{ms}, R_{mt}) = \delta_{st},$$
$$\mathrm{Corr}\,(R_{mt}, R_{nt}) = \mathbf{C}(m,n),$$

where δ_{st} is the Kronecker delta. Furthermore, we define *standardised returns* x_{mt} as

$$x_{mt} = \frac{R_{mt} - \mathbb{E}[R_{mt}]}{\sqrt{\mathrm{Var}\,(R_{mt})}}. \tag{2.4}$$

If we write $\mathbf{x}_t = (x_{1t}, \ldots, x_{Mt})$, then $\{\mathbf{x}_t\}_{t=0}^{\infty}$ is an i.i.d. sequence of normally distributed random variables with mean zero and covariance matrix $\mathbf{C}$, which is the matrix that we are going to model. [3]

[3]Correlation only captures linear dependencies. More general dependence structures can created using *copula* techniques and Nielsen [8] provides an overview of this topic. The methods proposed in this paper may be viewed as special case of a time dependent Gaussian copula approach.

3. The data

We will use daily closing prices from NYMEX to compute daily log returns, and we employ the convention of labelling contracts by their delivery month. For example, "Jan06" refers to the futures contract with delivery in January 2006. Currently NYMEX offers futures contracts for 72 consecutive months, and in Figure 1 we give an example of a futures curve. Clearly, this is a seasonal commodity, caused by a combination of the relative difficulty of storing natural gas, and the increased demand during the winter months. This suggests that the correlation matrix could be non-stationary.

Inspection of daily trading volumes reveals that the contracts with delivery more than 18 – 24 months away have poor liquidity. Nevertheless, NYMEX also posts closing prices for illiquid contracts, and in Figure 2 below we illustrate the potential danger of using these for estimation purposes. We would expect the Jan09 – Dec09 correlation to be close to unity, since we find it unlikely that there is any information available to the market that would permit discrimination between two contracts maturing four and five years into the future. This is not the case in Figure 2 however, though there are very few trades in the 2009-year contracts during 2004.

With the above observation in mind, we will estimate our correlation matrices from $M = 18$ consecutive contracts. Since we expect the correlation matrix to be non-stationary, the delivery months of the contracts must stay fixed, and we use one year of daily closing prices prior to the maturity of the first contract. The data sets used in this paper are specified in Table 1 below, where T denotes the sample size.

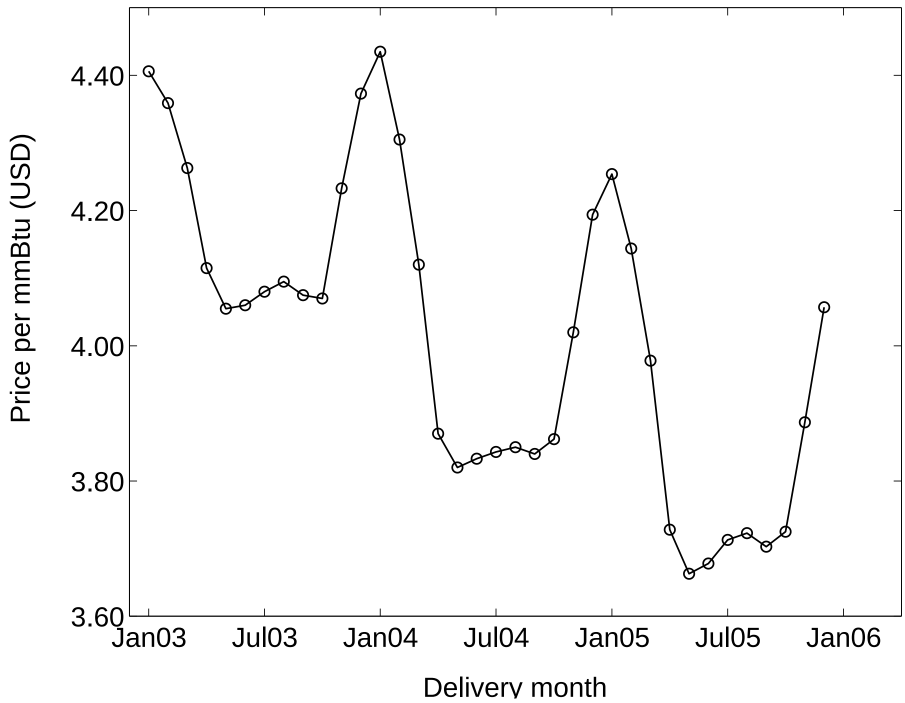

Figure 1: Nymex closing prices in USD/mmBtu for the Jan03 to Dec05 contracts on December 05, 2002.

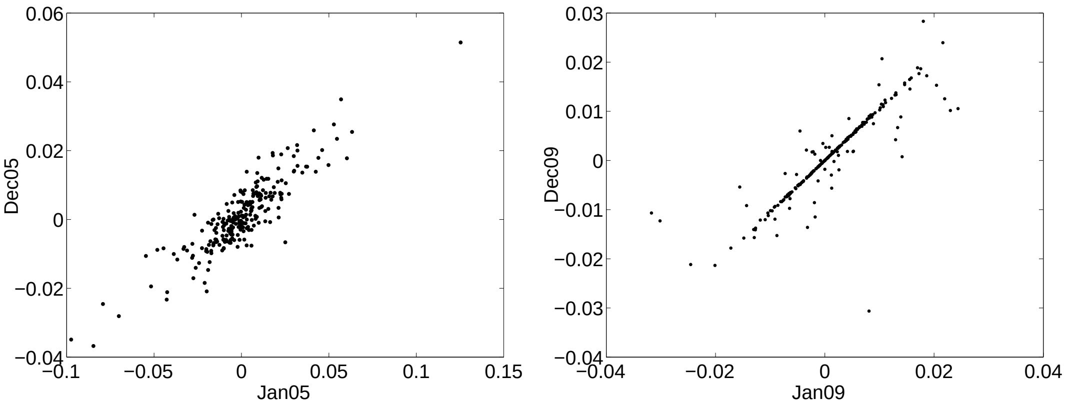

Figure 2: Scatter plots of daily 2004 log-returns (246 samples) for the Jan05-Dec05 (left) and Jan09-Dec09 (right) contract pairs. The sample correlation coefficient for the Jan05-Dec05 contracts is 0.8939 and for the Jan09-Dec09 contracts 0.8730.

Data set	First contract	Last contract	Data sampling period	T
DS1	Jan03	Jun04	Jan 2002 - Dec 2002	246
DS2	Apr03	Sep04	Apr 2002 - Mar 2003	246
DS3	Jul03	Dec04	Jul 2002 - Jun 2003	246
DS4	Oct03	Mar05	Oct 2002 - Sep 2003	246
DS5	Jan04	Jun05	Jan 2003 - Dec 2003	246
DS6	Apr04	Sep05	Apr 2003 - Mar 2004	246
DS7	Jul04	Dec05	Jul 2003 - Jun 2004	246
DS8	Oct04	Mar06	Oct 2003 - Sep 2004	246
DS9	Jan05	Jul05	Jan 2004 - Dec 2004	246

Table 1: The data sets used in this paper. The sample size T equals the number of trading days during the sample period minus one.

4. Stylised facts of sample correlations

The aim of this section is to provide an empirical study of the sample correlation matrix, which will be useful for understanding the intuition behind the parsimonious models introduced in Section 5.

In Figures 3 and 4, we display the sample correlation matrix $\mathbf{S}$ of the standardised returns $\{\mathbf{x}_t\}_{t=1}^T$ computed from the data set DS5. Here we have estimated μ_m, α_m and σ_m in (2.1) from the log-returns $\{R_{mt}\}_{t=1}^T$ by standard maximum likelihood estimation, and computed the standardised returns $\{x_{mt}\}_{t=1}^T$ from (2.4).

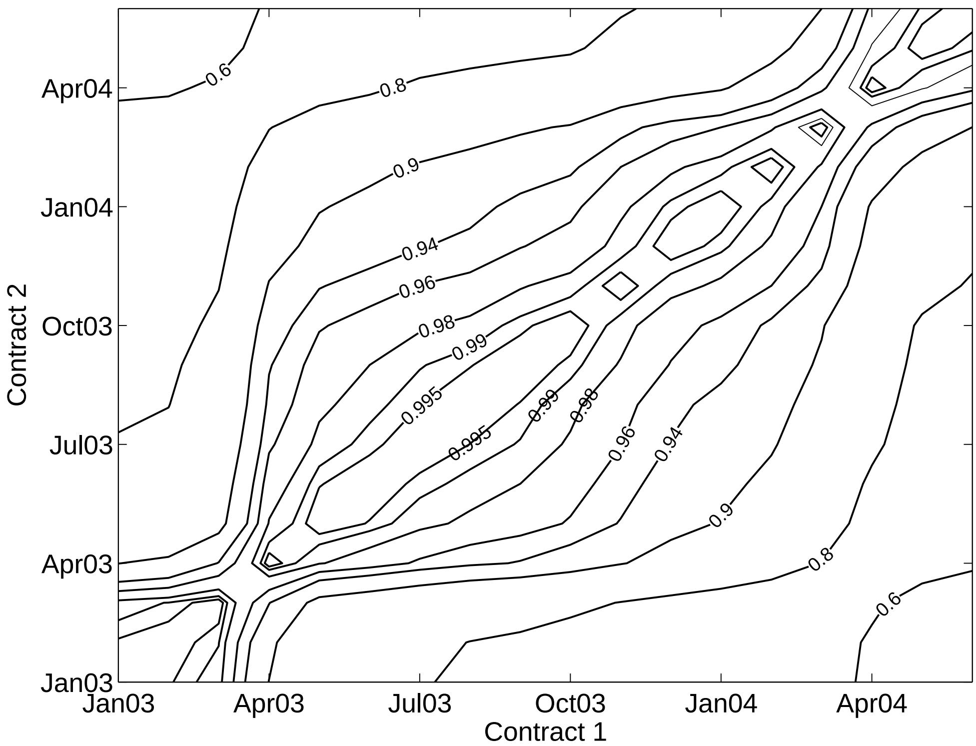

Figure 3: Contour plot of the sample correlaton matrix of standardised returns computed from DS5.

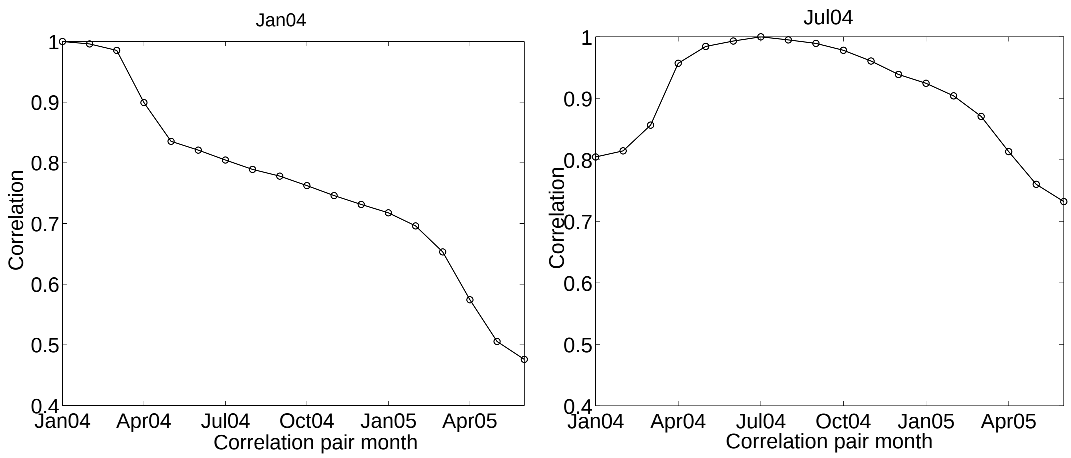

Figure 4: Sample correlation of standardised returns between the Jan04 (left) and the Jul04 (right) contracts and the other contracts of DS5.

Some of the most prominent features observed in this Figures 3 and 4 sample are listed in Stylised Facts 1 (SF1) below. They are also present in the sample correlation matrices of standardised returns computed from the other data sets of Table 1, which we do not display due to limited space.

Stylised Facts 1. *Shorter dated contracts correlations.*

(1) The function $\mathbf{S}(m,n)$ is decreasing in $|m-n|$.
(2) The function $\mathbf{S}(m,n)$ is non-negative.
(3) The correlations $\mathbf{S}(m,m+k)$, with m fixed, has a downward jump every time $m+k$ passes an April month. This means that $\mathbf{S}$ is non-stationary.
(4) The decay between successive Aprils does not occur at a uniform rate.

Although Section 3 emphasised that NYMEX natural gas futures trading is heavily concentrated to the first M =18-24 contracts, there may arise a need to model correlations among longer dated contracts. For example, this could be a part of computing the VaR of an over the counter (OTC) transaction, or when evaluating investments in natural gas extraction infrastructure. In this case financial intuition suggest that that the following Stylised Facts 2 (SF2) should hold.

Stylised Facts 2. *Longer dated contracts correlations.*

(1) $\lim_{k\to\infty}\mathbf{S}(m+k,n+k)=1$ for fixed m,n. Information about two futures contracts with distant maturities is the same.
(2) $\lim_{k\to\infty}\mathbf{S}(m,m+k)=A\in[0,1]$. Correlations between a close by and a distant futures pair need not go to zero.

Next, we perform a factor analysis of the sample correlation matrix of standardised returns of our data sets. The results are displayed in Table 2 and Figures 5 and 6 below.

Factor	1	2	3	4
DS1	93.82%	4.95%	0.69%	0.26%
DS2	92.16%	5.68%	1.04%	0.56%
DS3	89.18%	8.98%	0.80%	0.44%
DS4	88.60%	8.42%	1.60%	0.44%
DS5	86.25%	9.12%	2.61%	1.02%
DS6	88.91%	9.06%	1.02%	0.32%
DS7	91.57%	7.14%	0.46%	0.40%
DS8	91.24%	6.54%	0.99%	0.69%
DS9	94.69%	3.59%	1.00%	0.46%

Table 2: Distribution of variance among the first four factors of the sample correlation matrix of standardised returns. Data sets: DS1 to DS9. Software used: Matlab's `pcacov` function.

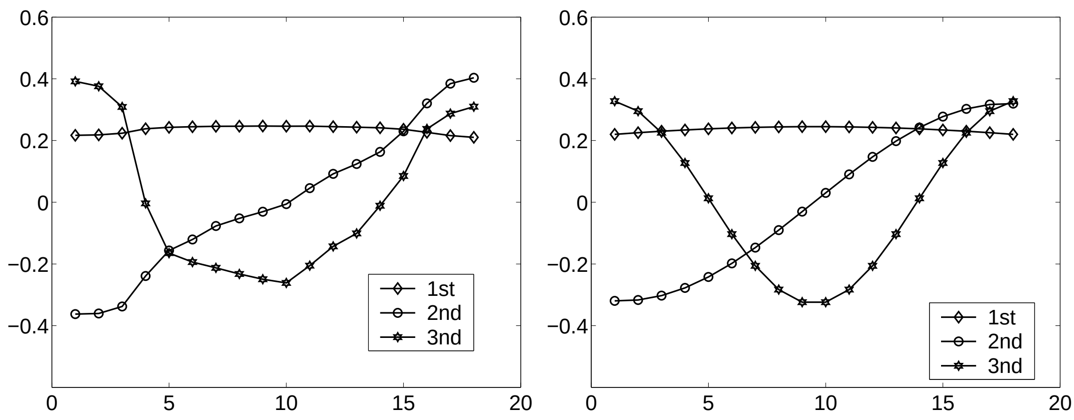

Figure 5: The first three factors for the sample correlation matrix $\mathbf{S}$ (left) and a fitted stationary correlation matrix $\mathbf{C}$ (right) of standardised returns computed from DS5. Here the elements of $\mathbf{C}$ are given by $C(m,n) = \exp(-\bar{\theta}|m-n|)$. The parameter $\bar{\theta}$ is the ULS estimator to be introduced in Section 6. Software used: Matlab's `pcacov` and `fminsearch` functions.

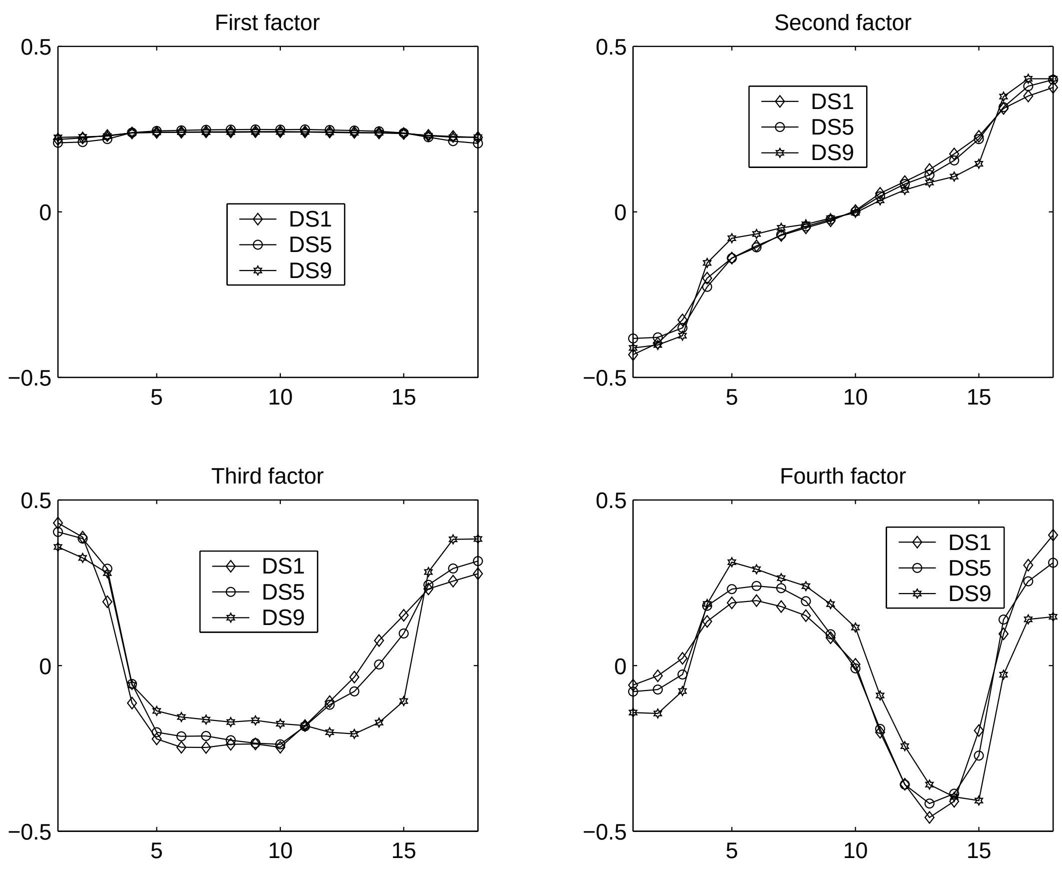

Figure 6: The first four factors of the sample correlation matrix of standardised returns. Data sets: DS1, DS5 and DS9. Software used: Matlab's `pcacov` function.

Figures 5 and 6 and Table 2 suggest that Stylised Facts 3 (SF3) below should hold for the factors.

Stylised Facts 3. *Factor analysis.*

(1) The first three factors explain more than 97% of the total variance, and the distribution of variance among the first two factors is quite stable over time (Table 2).
(2) The first three factors may be interpreted as level, slope and curvature, and they are quite stable over time (Figure 6).
(3) The slope and curvature factors of the SCM appear to be perturbed versions of the corresponding factors of a fitted stationary correlation matrix (Figure 5).

5. Structured models of the correlation matrix

In this section, we introduce the correlation function (CM) and factor analytic (FM) model classes of $\mathbf{C}$. We will write $\mathbf{C}(\theta)$ to emphasise that the elements of this matrix depend on a parameter vector θ.

Starting with the correlation function models, and inspired by Stylised Facts 1, we define the function $N_A(m,n)$ as the number of April months between calendar months m and n. Here we use the convention that $N_A(m,m) = 0$, and that for a fixed m, $N_A(m, m+k)$ increases as $m+k$ goes from April to May. Having done this, we specify three nested correlation functions in Definition 5.1 below.

Definition 5.1. *Correlation function models.* The correlation function models CM1, CM2 and CM3, have correlation matrices $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$ respectively, with elements given by the correlation functions

$$\begin{aligned}
C_1(m,n) &= \exp\left[-\theta_1|m-n|\right], \ \theta_1 \geq 0, \\
C_2(m,n) &= \exp\left[-\theta_1|m-n| - \theta_3 N_A(m,n)\right], \ \theta_1 \geq 0, \ \theta_3 \geq 0, \\
C_3(m,n) &= \exp\left[-\theta_1|m-n| - \theta_2|m-n|^2 - \theta_3 N_A(m,n)\right], \ \theta_1 \geq 0, \ \theta_2 \geq 0 \ \theta_3 \geq 0,
\end{aligned}$$

respectively.

In Proposition A.2 in Appendix A, we prove that the functions of Definition 5.1 are positive semi-definite. By comparing Stylised Facts 1 with Definition 5.1, it follows that CM1 is able to reproduce SF1(1) and SF1(2), CM2 SF1(1) to SF1(3), and CM3 SF1(1) to SF1(4) respectively. Since we only rely on data from $M = 18$ maturities, we will not try to model Stylised Facts 2.

The main advantages of the correlation function models are that they are easy to understand, and that they allow inter- and extrapolation to non-traded maturities provided that one believes that this makes economical sense. Given a particular functional form, the main difficulty is to prove positive semi-definiteness.[4] An alternative approach is to check positive semi-definiteness *ex-post*, and set all negative eigenvalues to zero. Rebonato [10] reports that this procedure alters the matrix elements very little, because the data seems to ensure that only the smallest eigenvalues by magnitude are negative.

[4] In Appendix A, we collect a number of techniques for creating valid correlation functions, which could serve as a toolbox for readers who wish to create their own correlation function models.

We now turn our attention to the factor analytic models. First we will fit a structured *covariance* matrix $\mathbf{\Sigma}(\theta)$ to data, and obtain $\mathbf{C}(\theta)$ from $\mathbf{\Sigma}(\theta)$ by normalisation. The rationale for this procedure is explained at the end of this section.

Assume that we intend to model the correlation matrix of a particular data set with sample correlation matrix $\mathbf{S}$. By Stylised Facts 3 the first three factors change relatively little from year to year. Inspired by this, we first compute a historical sample correlation matrix $\mathbf{S}_h$ from past data. Due to the April-hump perturbation in the second and third factors shown in Figure 6, it is important that $\mathbf{S}_h$ contains correlations between futures contracts maturing the same calendar months as those used to compute $\mathbf{S}$. For the data sets of Table 1, this means that $\mathbf{S}_h$ and $\mathbf{S}$ could be computed from the first and second parts of the data set. Alternatively, $\mathbf{S}_h$ can be computed from DS1 and $\mathbf{S}$ from DS5 and so on, which is the approach taken in this paper.

By the Spectral Theorem, $\mathbf{S}_h$ may be decomposed as

$$\mathbf{S}_h = \mathbf{U}\mathbf{D}\mathbf{U}^T, \tag{5.1}$$

where by convention column one of the orthogonal matrix $\mathbf{U}$ corresponds to the level, column two to the slope and so on. The factor analytic models introduced in Definition 5.2 below is a standard class of models used in many papers, for example Jöreskog [5] and Lee [6].

Definition 5.2. *Factor analytic model of order M_f.* Let $1 \leq M_f \leq M$, and let $\mathbf{V}_h$ be an $M \times M_f$ matrix containing the first M_f columns of the matrix $\mathbf{U}$ given in (5.1). Moreover, let $\mathbf{\Psi}$ be an $M_f \times M_f$ diagonal matrix with elements

$$\mathbf{\Psi}(m, m) = \theta_m^2,\ \theta_m \in \mathbb{R},\ 1 \leq m \leq M_f.$$

Finally if $\epsilon \in \mathbb{R}$ and $\mathbf{I}$ denotes the $M \times M$ identity matrix, then the factor analytic model of order M_f is given by

$$\mathbf{\Sigma} = \mathbf{V}_h \mathbf{\Psi} \mathbf{V}_h^T + \epsilon^2 \mathbf{I}.$$

Below we will write FMi for a historical factor model with $M_f = i$ factors. From Definition 5.2 it follows that the parameter vector $\theta = (\theta_1, \ldots, \theta_{M_f}, \epsilon)$ is to be estimated from the sample correlation matrix $\mathbf{S}$ of standardised returns.

The resulting covariance matrix $\mathbf{\Sigma}$ is positive semi-definite by construction and if $\epsilon^2 > 0$, then it is positive definite and hence non-singular. These properties also hold for the correlation matrix $\mathbf{C}$ obtained by normalising $\mathbf{\Sigma}$.

Requiring $\mathbf{\Sigma}$ to to be a correlation matrix from the beginning would lead to complicated constraints on the parameters, and avoiding this is the rationale for first fitting $\mathbf{\Sigma}$ and then normalising to get $\mathbf{C}$. Note that even though $\mathbf{\Sigma}$ is the best fit to data in some metric, this does not have to hold for $\mathbf{C}$. Since $\mathbf{C}$ and $\mathbf{\Sigma}$ are close however, we think that the impact of the normalisation on the goodness of fit will be small.

Unlike the correlation function models, the factor analytic models cannot be interpolated or extrapolated.

6. Model estimation and evaluation

To estimate the models proposed in Section 5, we will use the maximum likelihood (ML) and Un-weighted least squares (ULS) methods, both available under these names in software packages like SAS.

As a pre-processing step, we compute standardised returns from actual returns using expressions (2.2) to (2.4). Here the parameters μ_m, α_m and σ_m are computed by maximum likelihood estimation. Finally, the sample correlation matrix $\mathbf{S}$ is computed from these standardised returns. By this two-step procedure, we ensure that the marginal distribution of a particular futures contract is not affected by whether the contract is considered alone or as part of a portfolio.

By Jöreskog [5], the ML estimator $\hat{\theta}$, provided it exists, is the vector $\hat{\theta}$ that minimises

$$L(\theta) = \log|\mathbf{C}(\theta)| + \mathrm{tr}\{\mathbf{S}\mathbf{C}^{-1}(\theta)\}, \tag{6.1}$$

where $|\cdot|$ and tr denote the determinant and trace of a square matrix respectively. The ULS estimator $\bar{\theta}$ on the other hand, if it exists, minimises

$$\begin{aligned} F(\theta) &= \frac{1}{2}\mathrm{tr}\{(\mathbf{S} - \mathbf{C}(\theta))^2\} \qquad (6.2) \\ &= \frac{1}{2}\sum_{m=1}^{M}\sum_{n=1}^{M}|\mathbf{S}_{mn} - \mathbf{C}_{mn}(\theta)|^2, \end{aligned}$$

which is half of the squared the Fröbenius norm of $\mathbf{S} - \mathbf{C}$. In (6.1) and (6.1), the correlation matrix $\mathbf{C}(\theta)$ may be replaced by the covariance matrix $\mathbf{\Sigma}(\theta)$ if a factor analytic model is to be estimated.

If $\mathbf{C}(\theta)$ is a reasonably good description of the data, $\hat{\theta}$ and $\bar{\theta}$ should be close. Otherwise the functional forms (6.1) and (6.2) of $L(\theta)$ and $F(\theta)$ suggest that the ULS method could be more robust if $\mathbf{C}(\theta)$ is close to being singular.

For the ML method, Jöreskog [5] gives asymptotic confidence intervals of $\hat{\theta}$ based on the Fisher information matrix. Moreover, an incremental likelihood ratio (LR) test may be used to test whether adding another parameter improves the fit significantly or not. More specifically, let H_A be the hypothesis that the data comes from a model with an N_A-dimensional parameter vector θ_A, and H_0 that it comes from a sub-model of H_A with $N_0 < N_A$ free parameters θ_0. Then under H_0, with T being the sample size,

$$Z_{ML} = T\left[L(\hat{\theta}_0) - L(\hat{\theta}_A)\right]$$

follows a χ^2 distribution with $N_A - N_0$ degrees of freedom.

To the best of our knowledge, there are no known asymptotic confidence intervals or distributions available for

$$Z_{ULS} = F(\bar{\theta}_A) - F(\bar{\theta}_0)$$

obtained from the ULS method. Instead, we use Monte-Carlo simulation according to Algorithm 1 below.

Algorithm 1. *Simulation based confidence intervals and hypothesis tests.*

(1) Estimate $\bar{\theta}_0$ from data by minimising (6.2).

(2) Simulate T independent samples of a normally distributed return vector $\mathbf{x}_t$ with mean zero and covariance matrix $\mathbf{C}(\bar{\theta}_0)$.
(3) Re-estimate $\bar{\theta}_0$ and store the value as $\bar{\theta}_{0,n}$.
(4) Estimate $\bar{\theta}_{A,n}$ of the complex model, using the same simulated data. Compute $Z_{ULS}(n) = F(\bar{\theta}_{0,n}) - F(\bar{\theta}_{A,n})$, and store the value.
(5) Repeat steps 2 to 4 N_{sim} times in order to obtain an empirical distribution of $\bar{\theta}_0$ and Z_{ULS}. This distribution is finally used to compute confidence intervals and perform tests.

The actual minimisation of (6.1) and (6.2) is done numerically by Matlab's `fminsearch` routine, which uses a Nelder-Mead algorithm for multi-dimensional unconstrained optimisation.

7. Results

In this section, estimation and hypothesis test results are presented for the correlation function and factor analytic models.

For the factor-analytic models, the ML and ULS estimates are very similar, so we only report the former. Unfortunately, this is not the case for the correlation function models, where the ML method performs poorly compared to the ULS method, when used with the data sets of Table 1. Since the ML-method is capable of estimating these models with simulated data, we are led to think that this failure is caused by a combination of model mis-specification and $\mathbf{C}$ being close to singular. Replacing the ML-method by the Generalised Least Squares (GLS) method by Lee [6] does not improve matters. With these observations in mind, we only report ULS-estimates for the correlation function models. To obtain confidence intervals and perform hypothesis testing, Algorithm 1 is used with $N_{\text{sim}} = 10^4$ simulations.

First we test if we could reject the hypothesis that the data is generated by any of our models. This is done by performing Test 1, described below.

Test 1:
H_0: The data is generated by our model with N_0 parameters.
H_A: The data is generated by the sample correlation matrix.

For the factor analytic models, Test 1 is performed before the normalisation to a correlation matrix.

For the correlation function models CM1, CM2 and CM3, Test 1 is repeated for the data sets DS1 to DS9. The test statistic is Z_{ULS} and Algorithm 1 is used with $N_{\text{sim}} = 10^4$ simulations in order to simulate its distribution given H_0. For the factor analytic models FM2, FM3 and FM4, the test is repeated only for the data sets DS5 to DS9, since DS1 to DS4 are used to compute the historical factors. Here we employ a standard LR-test, which means that the test statistic Z_{ML} is approximately χ^2 distributed with $\frac{1}{2}M(M-1) - (M_f + 1)$ degrees of freedom. Here $M = 18$ and $M_f = 2$, 3 or 4 depending on the order of the model.

Second we investigate if expanding a model by adding a parameter improves the fit by performing Test 2 below.

Test 2:

H_0: The data is generated by one of our models with N_0 parameters.

H_A: The data is generated by one of our models with $N_A = N_0 + 1$ parameters, such that H_0 is a sub-model of H_A.

For the correlation function models this means that we test whether adding θ_3 to CM1 and θ_2 to CM3, thereby creating CM2 and CM3 from CM1 and CM2, improves the fit significantly or not. The test statistic is again Z_{ULS} and Algorithm 1 with $N_{\text{sim}} = 10^4$ simulations is used to simulate its distribution given H_0. These tests are repeated for the data sets DS1 to DS9. Given the particular functional form of the correlation function models of Definition 5.1, Test 2 is equivalent to a test where H_0 is the hypothesis that the added parameter is zero, and H_A that it is non-zero.

Test 2 applied to the factor analytic models tests if adding a factor improves the fit significantly, or alternatively if the added parameter θ_{M_f+1} is non-zero. This is repeated for the expansion from FM2 to FM3 and from FM3 to FM4, for the data sets DS5 to DS9 separately. We employ a standard LR-test, which means that the test statistic is Z_{ML} is approximately χ^2 distributed with one degree of freedom.

Test 1 reports $p-$values less than 0.02 for all models and data sets. In other words, none of the data sets is likely to have been generated by any of our models.

Test 2 reports $p-$values less than 0.01 for all models and data sets, rejecting H_0 at the 99% level. Consequently, all of our added model parameters are significant, and this result is backed up by visual inspection of the fitted correlation models in Figures 7 and 12 below. The more complex the model is, the better the visual fit.

Next we display the estimated parameter values for the correlation function (CM1, CM2 and CM3) and factor analytic (FM2, FM3 and FM4) models together with 95% confidence intervals. For the factor analytic models, we choose to present the diagonal elements of $\mathbf{\Psi}$, θ_m^2, rather than θ_m.

DS	$\bar{\theta}_1$	
DS1	0.0118	[0.0097,0.0143]
DS2	0.0127	[0.0103,0.0157]
DS3	0.0183	[0.0148,0.0225]
DS4	0.0213	[0.0173,0.0262]
DS5	0.0242	[0.0197,0.0297]
DS6	0.0209	[0.0169,0.0259]
DS7	0.0154	[0.0124,0.0189]
DS8	0.0155	[0.0125,0.0192]
DS9	0.0109	[0.0088,0.0136]

Table 3: The estimated model parameter $\bar{\theta}_1$ of CM1 (introduced in Definition 5.1). Method of estimation: ULS. Two sided 95% confidence intervals computed by Algorithm 1 with $N_{\text{sim}} = 10^4$ simulations. Data sets: DS1 to DS9.

DS	$\overline{\theta}_1$	$\overline{\theta}_3$
DS1	0.0084 [0.0067,0.0106]	0.0364 [0.0261,0.0479]
DS2	0.0109 [0.0085,0.0135]	0.0240 [0.0114,0.0396]
DS3	0.0133 [0.0107,0.0163]	0.0590 [0.0384,0.0830]
DS4	0.0181 [0.0143,0.0227]	0.0434 [0.0195,0.0708]
DS5	0.0170 [0.0134,0.0211]	0.0790 [0.0575,0.1037]
DS6	0.0156 [0.0125,0.0191]	0.0714 [0.0473,0.1051]
DS7	0.0104 [0.0083,0.0127]	0.0599 [0.0418,0.0812]
DS8	0.0102 [0.0080,0.0126]	0.0714 [0.0500,0.0963]
DS9	0.0057 [0.0045,0.0071]	0.0567 [0.0439,0.0731]

Table 4: The estimated model parameters $\bar{\theta}_1$ and $\bar{\theta}_3$ of CM2 (introduced in Definition 5.1). Method of estimation: ULS. Two sided 95% confidence intervals computed by Algorithm 1 with $N_{\text{sim}} = 10^4$ simulations. Data sets: DS1 to DS9.

DS	$\overline{\theta}_1$	$\overline{\theta}_2$	$\overline{\theta}_3$
DS1	0.0005 [0.0001,0.0015]	0.00088 [0.00062,0.00110]	0.0300 [0.0217,0.0393]
DS2	0.0049 [0.0038,0.0064]	0.00057 [0.00038,0.00081]	0.0249 [0.0123,0.0402]
DS3	0.0072 [0.0058,0.0089]	0.00057 [0.00032,0.00087]	0.0615 [0.0416,0.0842]
DS4	0.0114 [0.0089,0.0142]	0.00066 [0.00034,0.00107]	0.0438 [0.0215,0.0694]
DS5	0.0070 [0.0043,0.0099]	0.00106 [0.00061,0.00159]	0.0717 [0.0524,0.0939]
DS6	0.0063 [0.0044,0.0085]	0.00092 [0.00059,0.00133]	0.0710 [0.0456,0.1007]
DS7	0.0049 [0.0039,0.0061]	0.00051 [0.00030,0.00074]	0.0624 [0.0441,0.0831]
DS8	0.0036 [0.0023,0.0050]	0.00064 [0.00039,0.00091]	0.0712 [0.0483,0.0990]
DS9	0.0016 [0.0005,0.0028]	0.00043 [0.00042,0.00062]	0.0531 [0.0411,0.0681]

Table 5: The estimated model parameters $\bar{\theta}_1$, $\bar{\theta}_2$ and $\bar{\theta}_3$ of CM3 (introduced in Definition 5.1). Method of estimation: ULS. Two sided 95% confidence intervals computed by Algorithm 1 with $N_{\text{sim}} = 10^4$ simulations. Data sets: DS1 to DS9.

DS	$\hat{\theta}_1^2$	$\hat{\theta}_2^2$	ϵ
DS5	15.84 [13.85,17.81]	1.41 [1.23,1.59]	0.214 [0.211,0.217]
DS6	16.14 [14.12,18.16]	1.48 [1.29,1.66]	0.159 [0.156,0.161]
DS7	16.61 [14.53,18.68]	1.18 [1.03,1.33]	0.126 [0.124,0.128]
DS8	16.60 [14.52,18.68]	1.09 [0.95,1.23]	0.147 [0.146,0.148]
DS9	16.99 [14.86,19.11]	0.67 [0.58,0.76]	0.150 [0.148,0.153]

Table 6: The estimated model parameters $\hat{\theta}_1^2$, $\hat{\theta}_2^2$ and ϵ of FM2 (introduced in Definition 5.2). Method of estimation: ML. Two sided 95% confidence intervals from asymptotic ML-theory. Data sets: DS5 to DS9.

DS	$\hat{\theta}_1^2$	$\hat{\theta}_2^2$	$\hat{\theta}_3^2$
DS5	15.84 [13.85,17.81]	1.41 [1.23,1.59]	0.351 [0.305,0.398]
DS6	16.14 [14.12,18.16]	1.48 [1.29,1.66]	0.185 [0.160,0.209]
DS7	16.61 [14.53,18.68]	1.18 [1.03,1.33]	0.080 [0.062,0.084]
DS8	16.60 [14.52,18.68]	1.09 [0.95,1.23]	0.125 [0.107,0.142]
DS9	16.99 [14.86,19.11]	0.67 [0.58,0.76]	0.203 [0.176,0.229]

DS	ϵ
DS5	0.155 [0.152,0.157]
DS6	0.117 [0.115,0.119]
DS7	0.106 [0.104,0.108]
DS8	0.117 [0.116,0.119]
DS9	0.100 [0.099,0.101]

Table 7: The estimated model parameters $\hat{\theta}_1^2$, $\hat{\theta}_2^2$, $\hat{\theta}_3^2$ and ϵ of FM3 (introduced in Definition 5.2). Method of estimation: ML. Two sided 95% confidence intervals from asymptotic ML-theory. Data sets: DS5 to DS9.

DS	$\hat{\theta}_1^2$	$\hat{\theta}_2^2$	$\hat{\theta}_3^2$	$\hat{\theta}_4$
DS5	15.84 [13.85,17.81]	1.41 [1.23,1.59]	0.351 [0.305,0.398]	0.164 [0.142,0.187]
DS6	16.14 [14.12,18.16]	1.48 [1.29,1.66]	0.185 [0.160,0.209]	0.049 [0.041,0.056]
DS7	16.61 [14.53,18.68]	1.18 [1.03,1.33]	0.080 [0.062,0.084]	0.075 [0.065,0.085]
DS8	16.60 [14.52,18.68]	1.09 [0.95,1.23]	0.125 [0.107,0.142]	0.033 [0.027,0.038]
DS9	16.99 [14.86,19.11]	0.67 [0.58,0.76]	0.203 [0.176,0.229]	0.084 [0.073,0.095]

DS	ϵ
DS5	0.114 [0.112,0.116]
DS6	0.102 [0.100,0.104]
DS7	0.079 [0.078,0.080]
DS8	0.108 [0.106,0.110]
DS9	0.066 [0.065,0.067]

Table 8: The estimated model parameters $\hat{\theta}_1^2$, $\hat{\theta}_2^2$, $\hat{\theta}_3^2$, $\hat{\theta}_4^2$ and ϵ of FM4 (introduced in Definition 5.2). Method of estimation: ML. Two sided 95% confidence intervals from asymptotic ML-theory. Data sets: DS5 to DS9.

The parameters of the correlation function models presented in Tables 3 to 5 are not particularly stable across data sets. For the factor analytic models, Tables 6 to 8 show that the parameter $\hat{\theta}_1$ related to the level factor is quite stable, but that the stability of the other parameters deteriorates gradually as the order of the factor increases.

Tables 6 to 8 also show that expanding a given factor analytic model FMi by adding a factor does not change the values of $(\hat{\theta}_1, \ldots, \hat{\theta}_i)$. This is expected given the orthogonality of the columns of $\mathbf{V}_h$, the least squares form (6.2) of $F(\theta)$ and that the ML- and and ULS estimates are very similar for the factor analytic models.

In order to compare the methods, the least squares errors $F(\bar{\theta})$ defined in (6.2) are given in Table 9 below for DS5 to DS9. Recall that the ML and ULS estimates of the factor analytic models are very similar. For the factor analytic models, we compute $F(\bar{\theta})$ after we have normalised the estimated covariance matrix in order to make it a correlation matrix.

	DS5	DS6	DS7	DS8	DS9
CM1	0.257	0.187	0.103	0.166	0.120
CM2	0.110	0.095	0.043	0.067	0.030
CM3	0.035	0.033	0.022	0.033	0.014
FM2	0.107	0.040	0.013	0.084	0.055
FM3	0.044	0.027	0.008	0.083	0.029
FM4	0.027	0.027	0.004	0.083	0.027

Table 9: Fitting error $F(\bar{\theta})$ for the models CM1, CM2, CM3, FM2, FM3 and FM4 (introduced in Definitions 5.1 and 5.2). Data sets: DS5 to DS9.

Table 9 reveals that for both classes of methods, adding an extra parameter often reduces the error by about 50%. Moreover FM4 is the best performing model followed by FM3 and CM3 on a shared second place. When comparing models with a given number of parameters, CM3 has the smallest errors. For some data sets FM3 and FM4 have the same least square error, apparently contradicting the results of Test 2. This is caused by the fact that Test 2 is performed on the fitted covariance matrix $\mathbf{\Sigma}$ before the normalisation, whereas the fitting errors of Table 9 are computed from $\mathbf{C}$ after the normalisation. Obviously, the post-normalisation procedure needed for the factor analytic models could make the fit sub-optimal.

When comparing models, visual inspection of the fits is a good complement to the least squares errors of Table 9. For this purpose, parts of the sample and estimated correlation matrices of DS5 are shown in Figures 7 to 12 below.

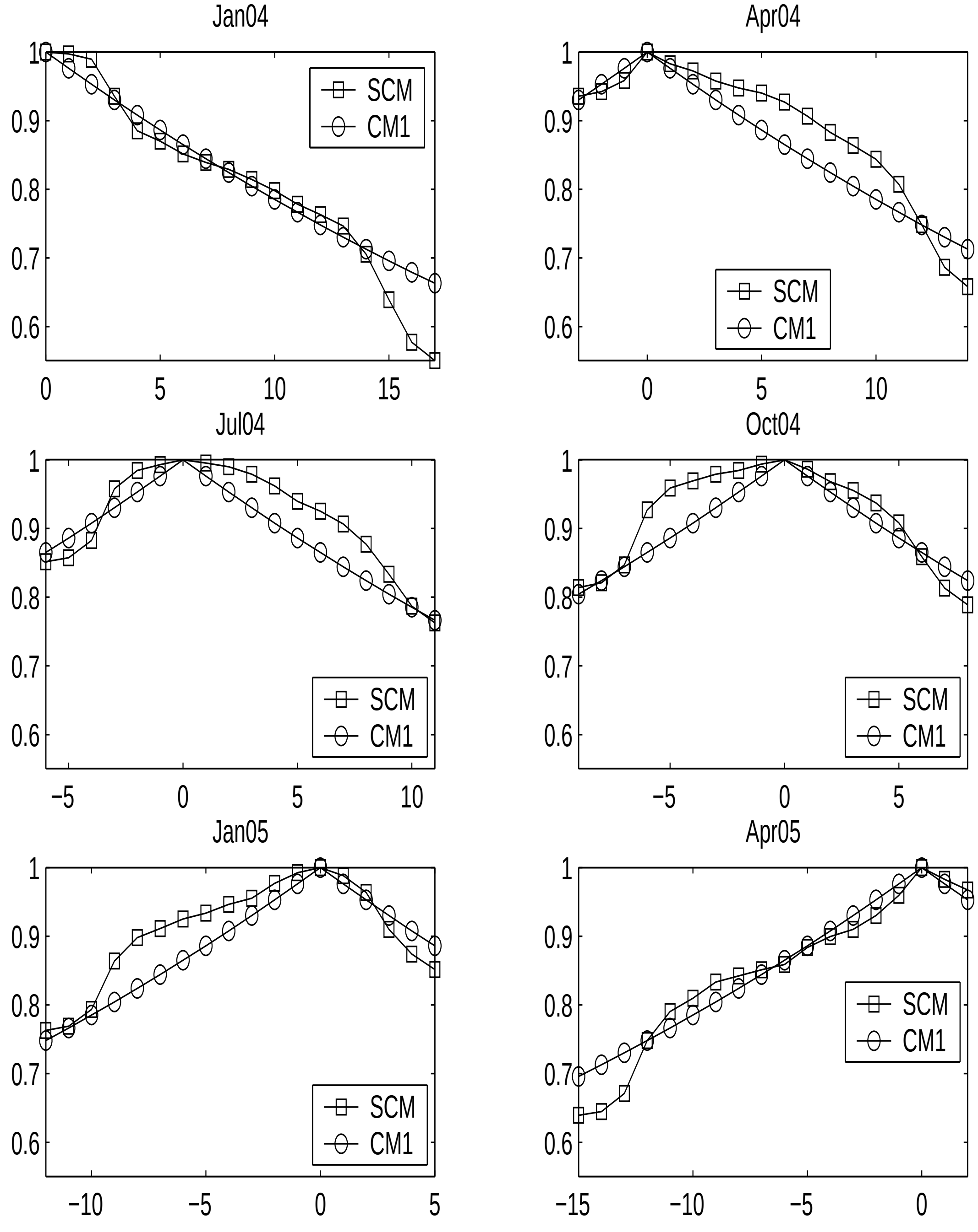

Figure 7: Some sample correlations (SCM) of standardised returns plotted together with fitted model correlations for CM1 (Introduced in Definition 5.1). Estimation method: ULS. Data: DS5. Correlation pairs: All contracts of DS5 with Jan04 (top left), Apr04 (top right), Jul04 (middle left), Oct04 (middle right), Jan05 (bottom left) and Apr05 (bottom right) contracts. The x-axis shows the lag in months between the contract in the title and the other contract in the correlation pair.

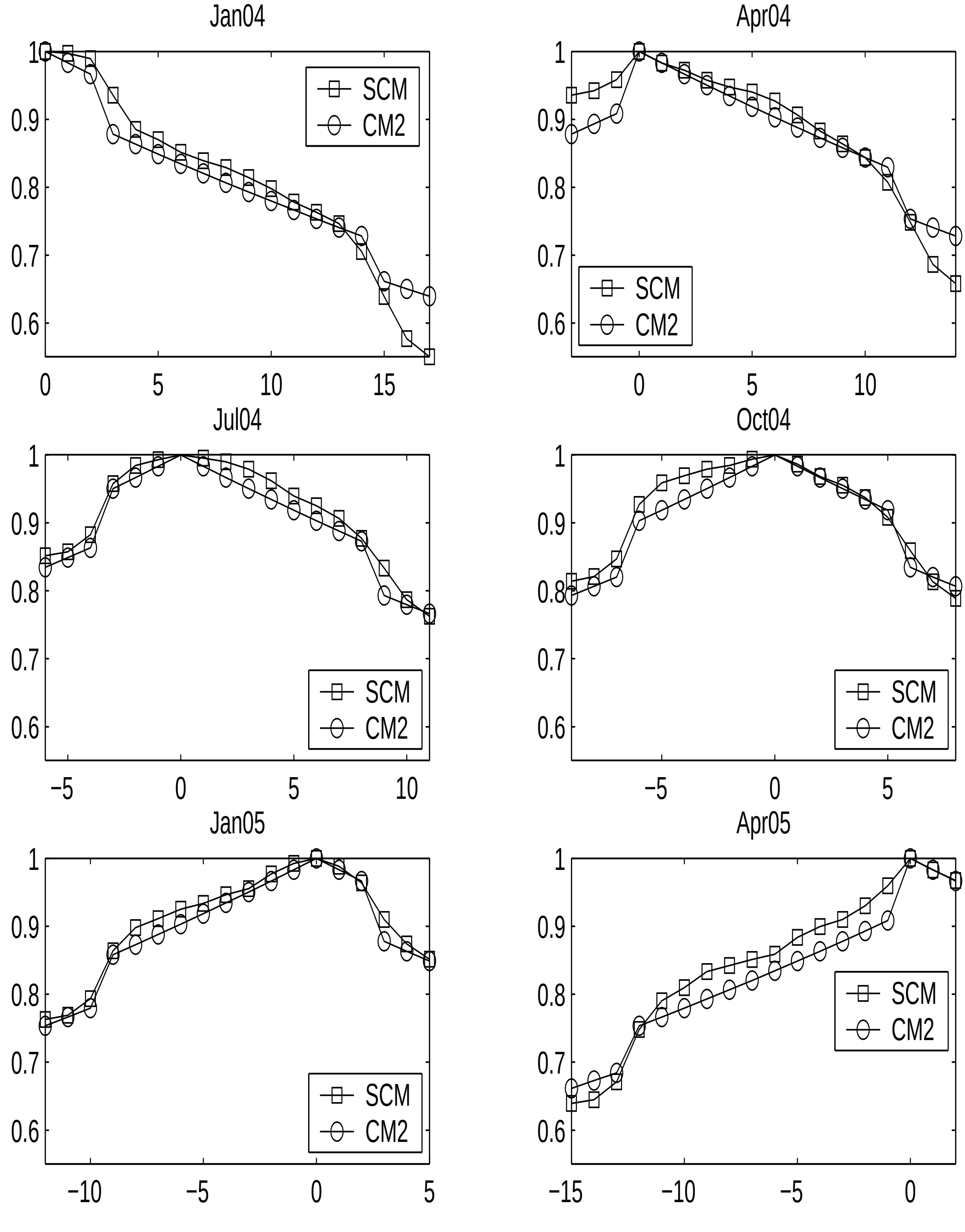

Figure 8: Some sample correlations (SCM) of standardised returns plotted together with fitted model correlations for CM2 (Introduced in Definition 5.1). Estimation method: ULS. Data: DS5. Correlation pairs: All contracts of DS5 with Jan04 (top left), Apr04 (top right), Jul04 (middle left), Oct04 (middle right), Jan05 (bottom left) and Apr05 (bottom right) contracts. The x-axis shows the lag in months between the contract in the title and the other contract in the correlation pair.

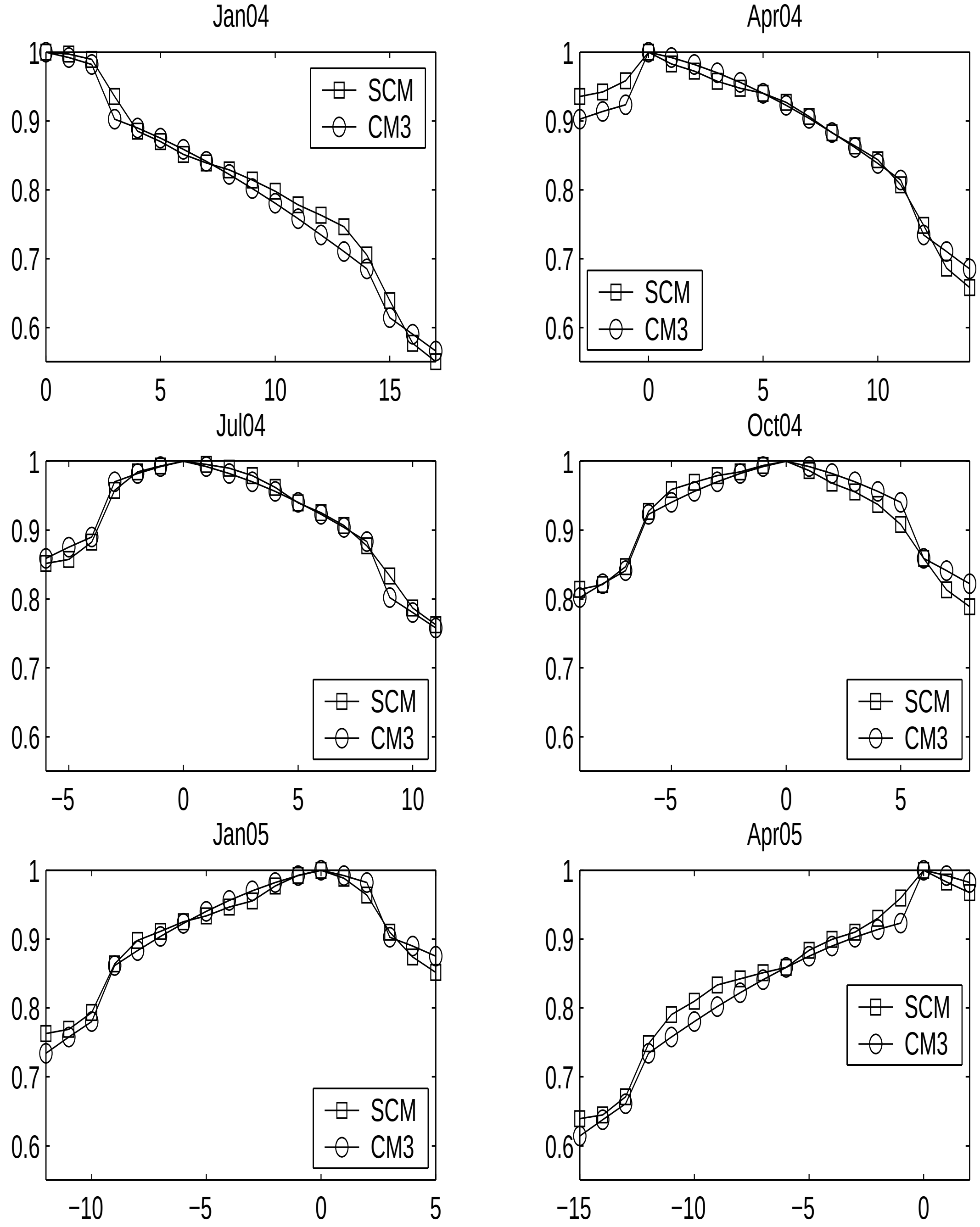

Figure 9: Some sample correlations (SCM) of standardised returns plotted together with fitted model correlations for CM3(Introduced in Definition 5.1). Estimation method: ULS. Data: DS5. Correlation pairs: All contracts of DS5 with Jan04 (top left), Apr04 (top right), Jul04 (middle left), Oct04 (middle right), Jan05 (bottom left) and Apr05 (bottom right) contracts. The x-axis shows the lag in months between the contract in the title and the other contract in the correlation pair.

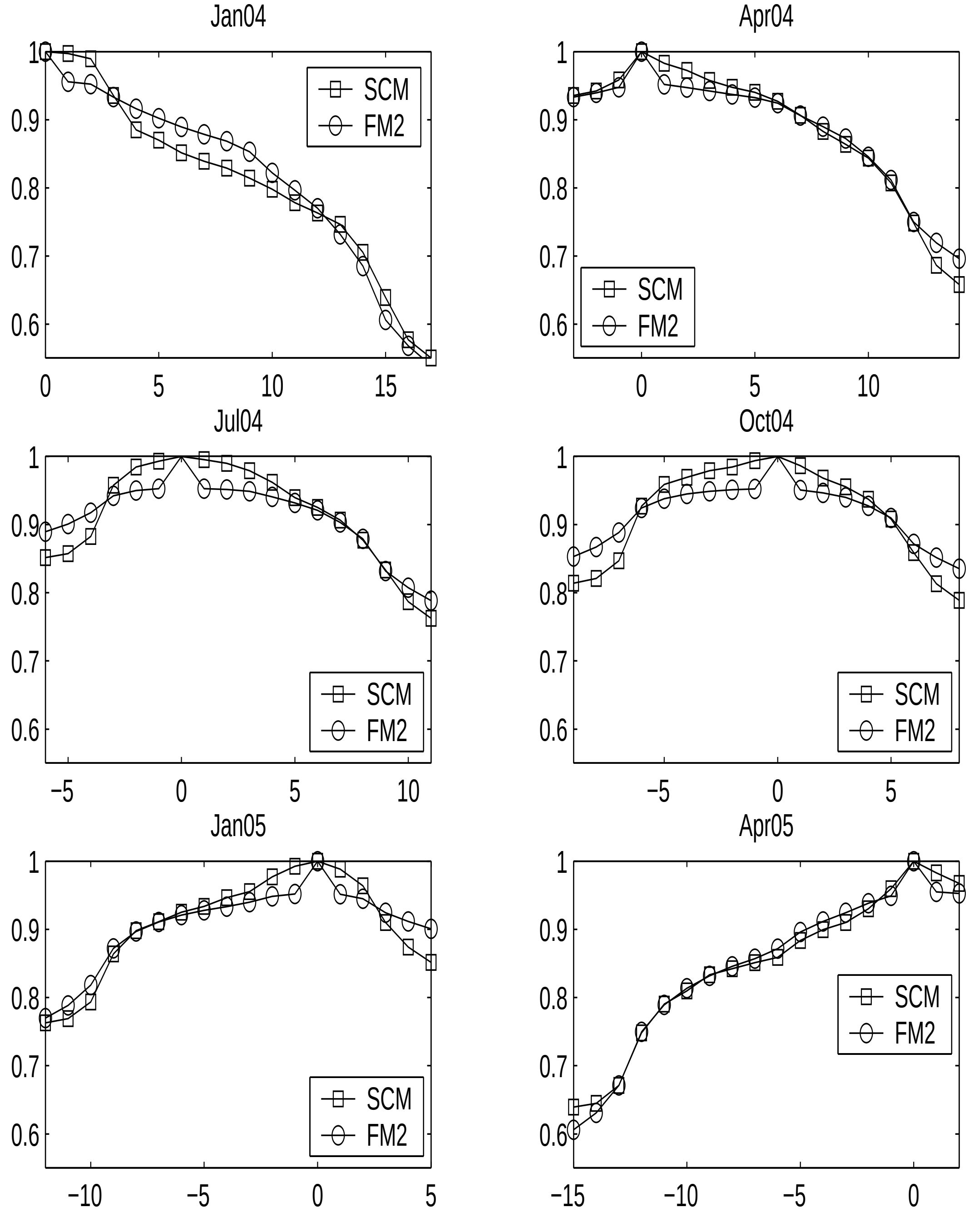

Figure 10: Some sample correlations (SCM) of standardised returns plotted together with fitted model correlations for FM2(Introduced in Definition 5.2). Estimation method: ML. Data: DS5. Correlation pairs: All contracts of DS5 with Jan04 (top left), Apr04 (top right), Jul04 (middle left), Oct04 (middle right), Jan05 (bottom left) and Apr05 (bottom right) contracts. The x-axis shows the lag in months between the contract in the title and the other contract in the correlation pair.

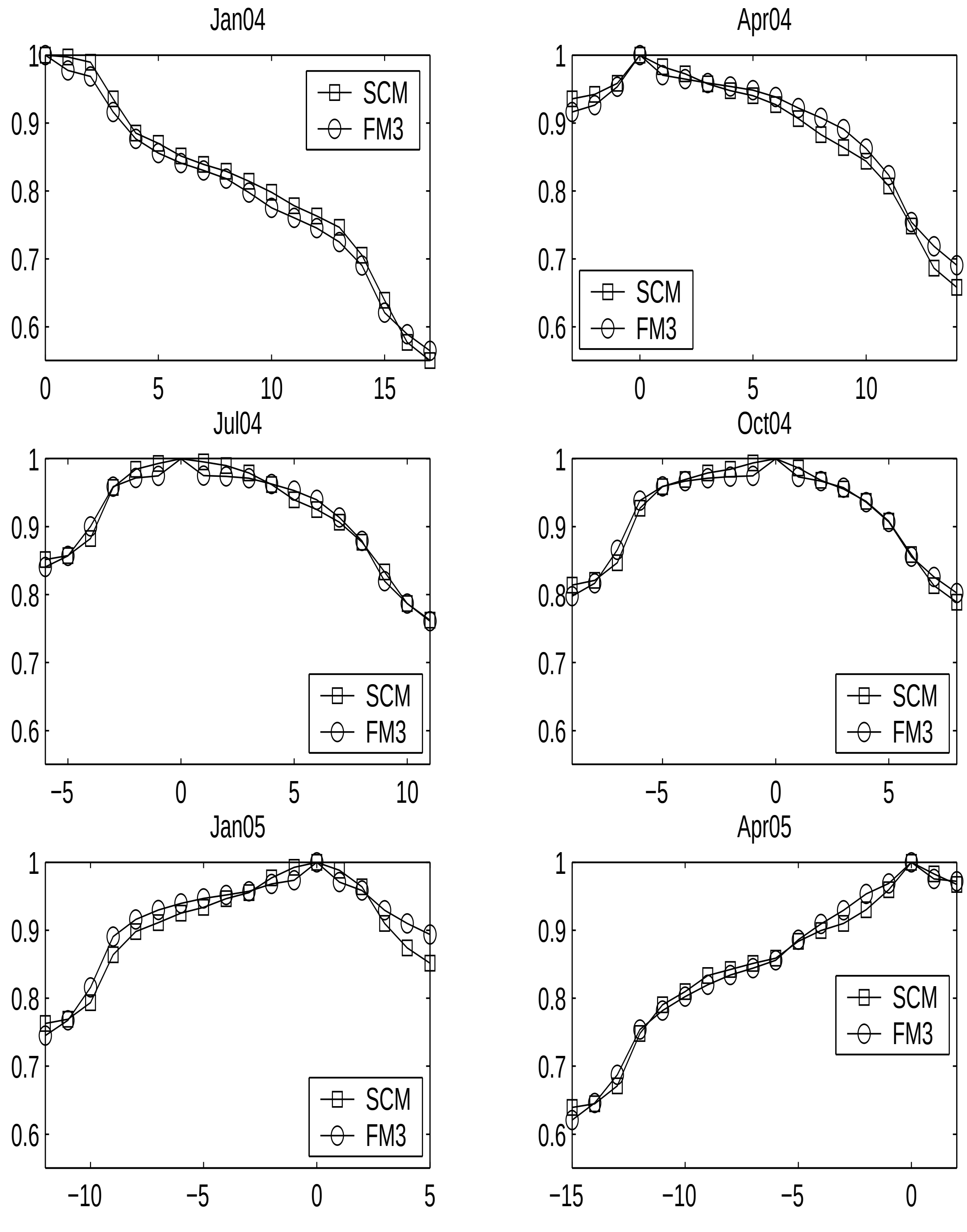

Figure 11: Some sample correlations (SCM) of standardised returns plotted together with fitted model correlations for FM3(Introduced in Definition 5.2). Estimation method: ML. Data: DS5. Correlation pairs: All contracts of DS5 with Jan04 (top left), Apr04 (top right), Jul04 (middle left), Oct04 (middle right), Jan05 (bottom left) and Apr05 (bottom right) contracts. The x-axis shows the lag in months between the contract in the title and the other contract in the correlation pair.

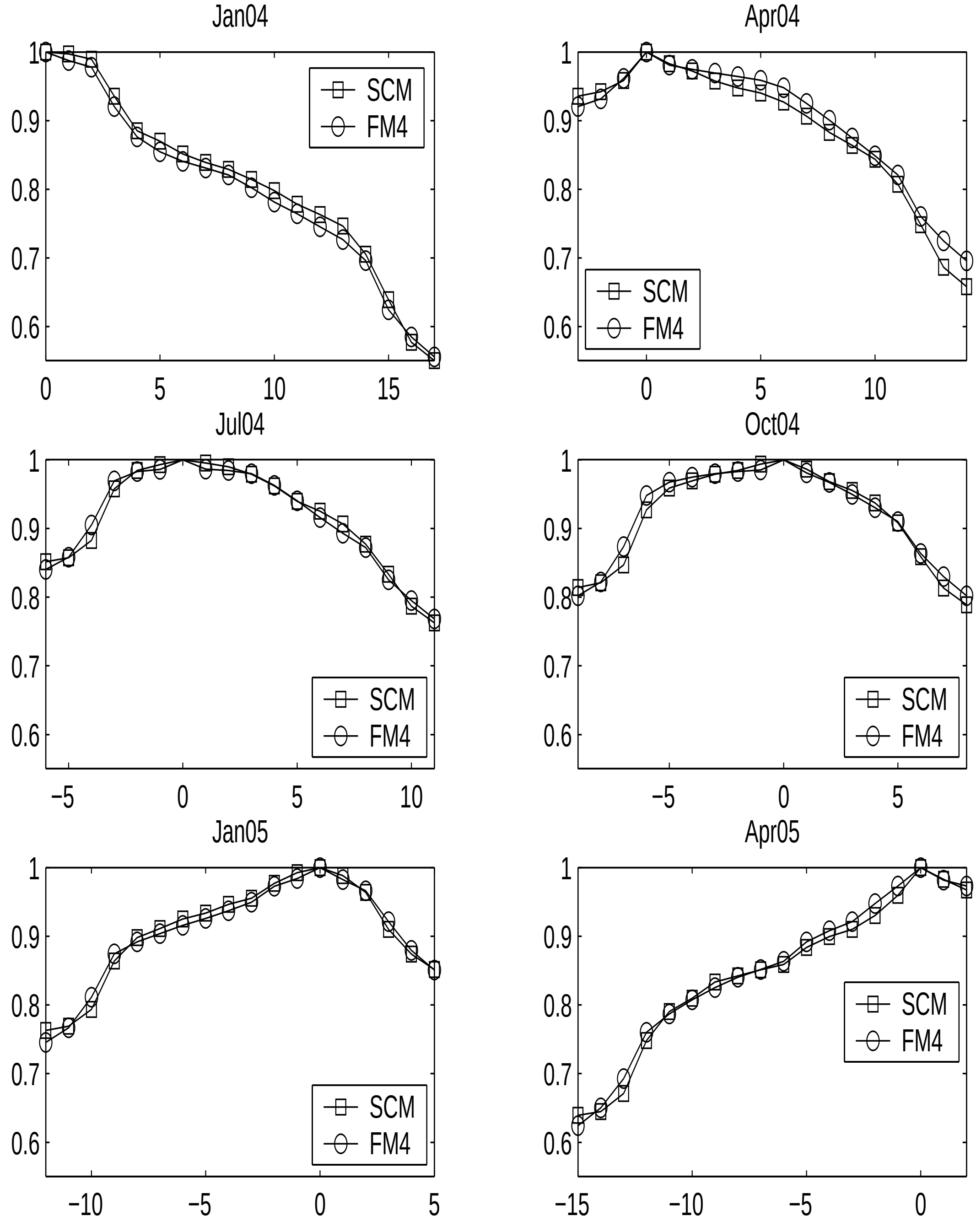

Figure 12: Some sample correlations (SCM) of standardised returns plotted together with fitted model correlations for FM4(Introduced in Definition 5.2). Estimation method: ML. Data: DS5. Correlation pairs: All contracts of DS5 with Jan04 (top left), Apr04 (top right), Jul04 (middle left), Oct04 (middle right), Jan05 (bottom left) and Apr05 (bottom right) contracts. The x-axis shows the lag in months between the contract in the title and the other contract in the correlation pair.

Figures 7 to 12 show that the visual fit is improved when an extra parameter is added. According to our opinion, FM4 offers the best visual fit, followed by CM3 and FM3 on a shared second place.

The most significant non-stationary feature of the sample correlation matrix displayed in Figures 3 and 7 to 12 is the drop in correlation between two contracts, which occurs for each April month located between the two maturities (*"the April drop"*). This effect could be caused by the typical shape of the futures curve (see Figure 1) and the cost of storage. More specifically, let $T_1 < T_2$ be two maturities, and assume that the demand and price of $F(t, T_2)$ increases. Under some conditions storage owners may find it profitable to go long $F(t, T_1)$, short $F(t, T_2)$, and hedge by storing the gas delivered from $F(t, T_1)$. In absence of arbitrage, the total effect is that the demand and price increase of $F(t, T_2)$ causes a demand and price increase of $F(t, T_1)$. Consequently, the prices move together, which creates correlation between the contracts. Considering the shape of the futures curve given in Figure 1, with the winter (December to March) contracts being more expensive than the other contracts, this kind of transaction is very unlikely to be profitable if T_1 and T_2 are on either side of an April month. No one buys expensive gas in the winter and sells it cheap in the Spring or Summer. Buying the gas even earlier does not make sense either, since it would still be more profitable to sell it during the winter. Besides that, the storage costs would be high. To conclude, the "extra" correlation caused by inter-temporal substitution as described above is absent if the two maturities of a correlation pair has an April month located between them. This explains the drop.

Test 2 for CM1 vs. CM2 has a $p-$value of 0.01, so explicitly modelling the April drop effect improves the fit significantly. Furthermore, Tables 4 and 5 show that the 95% confidence interval for the parameter θ_2 is well above zero for CM2 and CM3. The same tables also show that θ_2 has roughly the same value for CM2 and CM3, indicating that this parameter indeed captures an actual drop.

Figures 10 to 12 reveal that the factor analytic models reproduce the April drop effect automatically. This is not surprising since it is a cyclical effect and we use historical factors computed from contracts maturing one year earlier. Consequently it is crucial to compute the historical factors from contracts maturing the same calendar months as the contracts for which one wishes to model the correlation matrix.

8. Concluding remarks

We have specified and estimated some hierarchical correlation function and factor analytic models of the correlation matrix for natural gas futures returns. Statistical tests, least squares errors and visual inspection all show that the more complex the model, the better the fit.

The most important non-stationary effect present in the sample correlation matrix is the April drop effect, which is caused by the shape of the futures curve and the cost of storage. Most of our models are capable of reproducing this effect.

Whether to choose a correlation function or a factor analytic model is to some extent a matter of personal taste. Our correlation function models are very simple and may be extrapolated outside the range of the maturities used for their estimation. They also perform better than factor analytic models for a given number of model parameters. Due to the requirement of positive semi-definiteness however, it is not straightforward to find a valid extension of CM3 which could compete with FM4 in terms of mean squared error. Moreover, correlation function models could prove difficult to apply to other commodities. First, the stylised facts of the sample correlations to be modelled must be identified. Second, a suitable functional form that reproduce these stylised facts must be found. Third, these functional forms must prove to yield a positive semi-definite matrix. If the stylised facts are more complex than the April drop effect, this may be difficult.

Factor analytic models on the other hand, accomplish this three step procedure automatically, provided that the stylised facts are either stationary or seasonal. The fit of the covariance matrix may always be improved by adding another factor, although this may occasionally not improve the fit of the correlation matrix after normalisation. Drawbacks of the factor analytic models include that they are maybe not as simple as the correlation function models and cannot be extrapolated.

To conclude, we believe that the correlations of the fitted CM3, FM3 and FM4 models are close enough to the sample correlations in order to be useful. Economic agents who seek a parsimonious way of calculating Value-at-Risk (VAR) of a portfolio or to price a derivative with many underlying futures contracts could benefit from using the results presented in this paper.

One extension of this paper would be to model cross-commodity correlations, like natural gas and crude oil. Another would be to use copulas to investigate whether correlation is a good way to model dependence or not.

References

[1] F. Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3:167–179, March 1976.
[2] E. Buell. Correlation functions for wind and geopotential on isobaric surfaces. *Journal of Applied Meteorology*, 11(1):51–59, 1972.
[3] A. Cortazar and E. Schwartz. The valuation of commodity-contingent claims. *The Journal of Derivatives*, 4(1):27–39, Summer 1994.
[4] W. Dunsmuir and D. Nott. Estimation of nonstationary spatial covariance structure. *Biometrika*, 89(4):819–829, 2002.
[5] K. Jöreskog. A general method for analysis of covariance structures. *Biometrika*, 57(2):239–251, 1970.
[6] S. Lee. Constrained estimation in covariance structures. *Biometrika*, 66(3):539–545, 1979.
[7] F. Longstaff, P. Santa-Clara, and E. Schwartz. The relative valuation of caps and swaptions: Theory and empirical evidence. *Journal of Finance*, 56(6):2167–2109, 2001.
[8] R. Nielsen. *An Introduction to Copulas*. Springer, New York, 1999.
[9] M. Pourahmadi. Maximum likelihhod estimation of generalised linear models for multivarate normal covariance matrix. *Biometrika*, 87(2):425–435, 2000.
[10] R. Rebonato. *Volatility and Correlation in the Pricing of Equity, FX and Interest Rate Options*. John Wiley & Sons, Chichester, 1999.
[11] E. Schwartz. The stochastic behavior of commodity prices: implications for valuation and hedging. *Journal of Finance*, 52(3):923–972, July 1997.
[12] M. Stein. *Interpolation of Spatial Data*. Springer, New York, 1999.

Appendix A. Proofs

In this appendix we prove that the functions CM1 to CM3 introduced in Definition 5.1 are proper correlation functions. This result is given in Proposition A.2 below, which relies on Lemma A.1. This lemma is also of independent interest when proving positive semi-definiteness of candidate correlation functions other than the ones used in this paper.

Lemma A.1. *The following statements hold.*

(1) *Let* $\theta \geq 0$ *and* $0 < \alpha \leq 2$*. Then*

$$C(m,n) = \exp(-\theta|m-n|^{\alpha})$$

is a correlation function.

(2) *Let* $\{C_k(m,n)\}_{k=1}^{K}$ *be a set of correlation functions, and* $\{\alpha_k\}_{k=1}^{K}$ *a set of non-negative weights summing to one. Then*

$$C(m,n) = \sum_{k=1}^{K} \alpha_k C_k(m,n)$$

and

$$C(m,n) = \prod_{k=1}^{K} C_k(m,n)$$

are correlation functions.

(3) *Let* $C_1(m,n)$*,* $C_2(m,n)$ *be correlation functions and* g *be a deterministic function such that* $g(n) \in [0,1]$*. Then*

$$C(m,n) = [g(m)g(n)]^{1/2}\, C_1(m,n) + \{[1-g(m)]\,[1-g(n)]\}^{1/2}\, C_2(m,n)$$

is a correlation function.

(4) *Let* $C(m,n) = \rho(|m-n|)$ *be a stationary correlation function and* h *be a function such that* $h : \mathbb{N} \to \mathbb{N}$*. Then*

$$\overline{C}(m,n) = \rho\,[|h(m) - h(n)|]$$

is a correlation function.

Proof. Part 1: see Stein [12]. Part 2, Claim 1: This is immediate from the definition of a correlation function. Part 2, Claim 2: Let X_n and Y_n be two independent and standardised stochastic processes. Then compute the correlation function of the process $Z(n) = X(n)Y(n)$. Part 3: Again consider two independent and standardised processes X_n and Y_n, with correlation functions $C_1(m,n)$ and $C_2(m,n)$ respectively. Then $Z_n = \sqrt{g(n)}X_n + \sqrt{1-g(n)}Y_n$ has the desired correlation function. Part 4: Let $X(n)$ be a stationary process with correlation function $C(m,n) = \rho(|m-n|)$ and define $Y(n) = X(h(n))$. Then $Y(n)$ is a (not necessarily uniform or monotone!) re-sampling of $X(n)$, and has the desired correlation function. □

Apart from the exponential family of stationary correlation functions introduced in Part 1 of Lemma A.1, several other families of stationary correlation functions exist. These include, but are not limited to, the Buell and Matérn families described in Buell [2] and Stein [12] respectively.

Proposition A.2. *The models CM1 to CM3 introduced in Definition 5.1, are correlation functions.*

Proof. This Proposition follows almost directly from Lemma A.1, in particular the second claim of Part 2. To prove that $C(m,n) = e^{-\theta N_A(m,n)}$ is a correlation function, apply of Part 3 of Lemma A.1 with $C(m,n) = \exp(-\theta|m-n|)$ and let $h(n)$ be an integer valued step function that increases by one each April. □

Department of Economics, School of Business, Economics and Law, Göteborg University, Box 600, S-405 30 Göteborg, Sweden
E-mail address: mats.kjaer@economics.gu.se

Department of Finance, McCombs School of Business, 1 University Station, B6600, University of Texas at Austin, Austin, TX. 78712-1179, U. S. A.
E-mail address: eronn@mail.utexas.edu