NTNU – Trondheim
Norwegian University of
Science and Technology

Valuation of a Combined Cycle Gas Turbine

under price uncertainty and operational constraints

Øystein Arvesen
Vegard Gjelsvik Medbø

Industrial Economics and Technology Management
Submission date: June 2012
Supervisor: Sjur Westgaard, IØT
Co-supervisor: Stein Erik Fleten, IØT
Dipeng Chen, Centrica

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

NTNU
Det skapende universitet

MASTERKONTRAKT
- uttak av masteroppgave

1. Studentens personalia

Etternavn, fornavn	Fødselsdato
Arvesen, Øystein	**11. jun 1987**
E-post	Telefon
oystein.arvesen@gmail.com	**92251628**

2. Studieopplysninger

Fakultet	
Fakultet for Samfunnsvitenskap og teknologiledelse	
Institutt	
Institutt for industriell økonomi og teknologiledelse	
Studieprogram	Hovedprofil
Industriell økonomi og teknologiledelse	**Investering, finans og økonomistyring**

3. Masteroppgave

Oppstartsdato	Innleveringsfrist
16. jan 2012	**11. jun 2012**
Oppgavens (foreløpige) tittel	
Applications of spark spread models in valuation and decision making	
Oppgavetekst/Problembeskrivelse In this thesis, we aim to develop a model for the spark spread and use the model to value real options and/or financial options with the spark spread as the underlying price process. The thesis will include estimation and testing of different models on the spark spread, as well comparison and selection of the preferred model. We will illustrate how real and/or financial options can be used to make investment decisions for a player exposed to spark spread risk.	
Hovedveileder ved institutt	Medveileder(e) ved institutt
Førsteamanuensis Sjur Westgaard	**Stein Erik Fleten**
Ekstern bedrift/institusjon	Ekstern veileder ved bedrift/instutisjon
Centrica	**Dipeng Chen**
Merknader	
1 uke ekstra p.g.a påske.	

4. Underskrift

Student: Jeg erklærer herved at jeg har satt meg inn i gjeldende bestemmelser for mastergradsstudiet og at jeg oppfyller kravene for adgang til å påbegynne oppgaven, herunder eventuelle praksiskrav.

Partene er gjort kjent med avtalens vilkår, samt kapitlene i studiehåndboken om generelle regler og aktuell studieplan for masterstudiet.

Trondheim 13.01.2012
Sted og dato

Student

Hovedveileder

Originalen lagres i NTNUs elektroniske arkiv. Kopi av avtalen sendes til instituttet og studenten.

NTNU
Det skapende universitet

MASTERKONTRAKT
- uttak av masteroppgave

1. Studentens personalia

Etternavn, fornavn **Medbø, Vegard Gjelsvik**	Fødselsdato **16. jun 1986**
E-post **vegard.medbo@gmail.com**	Telefon **99512638**

2. Studieopplysninger

Fakultet **Fakultet for Samfunnsvitenskap og teknologiledelse**	
Institutt **Institutt for industriell økonomi og teknologiledelse**	
Studieprogram **Industriell økonomi og teknologiledelse**	Hovedprofil **Investering, finans og økonomistyring**

3. Masteroppgave

Oppstartsdato **16. jan 2012**	Innleveringsfrist **11. jun 2012**
Oppgavens (foreløpige) tittel **Applications of spark spread models in valuation and decision making**	
Oppgavetekst/Problembeskrivelse In this thesis, we aim to develop a model for the spark spread and use the model to value real options and/or financial options with the spark spread as the underlying price process. The thesis will include estimation and testing of different models on the spark spread, as well comparison and selection of the preferred model. We will illustrate how real and/or financial options can be used to make investment decisions for a player exposed to spark spread risk.	
Hovedveileder ved institutt **Førsteamanuensis Sjur Westgaard**	Medveileder(e) ved institutt **Stein Erik Fleten**
Ekstern bedrift/institusjon **Centrica**	Ekstern veileder ved bedrift/instutisjon **Dipeng Chen**
Merknader **1 uke ekstra p.g.a påske.**	

4. Underskrift

Student: Jeg erklærer herved at jeg har satt meg inn i gjeldende bestemmelser for mastergradsstudiet og at jeg oppfyller kravene for adgang til å påbegynne oppgaven, herunder eventuelle praksiskrav.

Partene er gjort kjent med avtalens vilkår, samt kapitlene i studiehåndboken om generelle regler og aktuell studieplan for masterstudiet.

Trondheim 13.01.2012
Sted og dato

Student **Hovedveileder**

Originalen lagres i NTNUs elektroniske arkiv. Kopi av avtalen sendes til instituttet og studenten.

Valuation of a Combined Cycle Gas Turbine

under price uncertainty and operational constraints

Vegard Medbø
Stud.Techn.

Øystein Arvesen
Stud.Techn.

6. juni 2012

Preface

This master thesis is conducted as the final part of a five year program at the Norwegian University of Science and Technology (NTNU), leading to a Master of Science degree in Industrial Economics and Technology Management. The work was completed in the spring of 2012, and is the shared work of two students. The thesis is written as a report combining the research fields of econometric time-series modelling and real option analysis. We call attention to the fact that the main title of the thesis has been altered from its original version stated in the master contract.

We wish to express our gratitude to our academic supervisors, Associate Professor Sjur Westgaard and Professor Stein-Erik Fleten. They have given us valuable feedback and good counsel. We would also like to thank Quantitative Analyst Dipeng Chen at Centrica Energy for providing us with price data and plant specifications, as well as taking time to answer questions on short notice. Further, we would like to thank Professor Fred Espen Benth at the University of Oslo and Kjersti Aas at Norsk Regnesentral for helping us with an in-depth understanding of the multivariate NIG distribution.

Trondheim, June 6$^{\text{th}}$, 2012

Øystein Arvesen

Vegard Gjelsvik Medbø

Abstract

In this thesis we combine multivariate time series modelling with real options theory to value a combined cycle gas turbine. We propose a novel price model with co-integrated power, gas and carbon prices, with multivariate stochastic volatility and MNIG distributed errors. The estimated model is found to outperform competing specifications in terms of higher likelihood and lower information criteria. We implement a Least Squares Monte Carlo (LSM) simulation to value the plant, incorporating ramp times, startup costs and variable plant efficiency. We take into account that day-ahead prices are settled the day before prices take effect, which is often overlooked in related literature. We find that ignoring this leads to suboptimal choices and a lower value estimate. An analysis of the regressions in the LSM algorithm reveals that the choice of basis functions has a significant effect on the estimated value of the plant. Particularly, for a low-efficiency plant, a regression on the spark spread underestimates the value by 20% compared to a regression on both the electricity price and the fuel cost components. This implies that in spread option valuations where the LSM is applicable, simulating all asset or commodity prices may be advantageous over simulating the spread alone.

Sammendrag

I denne masteroppgaven kombinerer vi multivariat tidsrekkemodellering med realopsjonsteori for å estimere verdien av et kombinert gasskraftverk. Vi foreslår en modell for strøm-, gass- og karbonpriser som hensyntar sesongvariasjoner, kointegrasjon, stokastisk volatilitet og ikke-Gaussisk oppførsel. Parameterestimering og testing viser at en CCC-GARCH-modell med MNIG-fordelte residualer fanger opp prisdynamikken vesentlig bedre enn de alternativene vi tar for oss. Videre bruker vi Least Squares Monte Carlo-simulering (LSM), og verdsetter kraftverket under fysiske beskrankninger som inkluderer variabel effektivitet, startkostnader og ledetid i produksjonsendring. Vi tar også hensyn til at prisene i et day ahead-marked er kjent dagen før de blir gjeldende, et faktum som ofte oversees i litteraturen, og viser at å ignorere dette fører til suboptimal produksjonsstyring og underestimering av kraftverksverdien. En analyse av regresjonsform i LSM-algoritmen avdekker at å simulere alle råvareprisene i stedet for en spread alene gir et mer realistisk verdiestimat. Spesifikt viser vi at en univariat regresjon estimerer verdien 1% for lavt i et scenario med høy effektivitet, og 20% for lavt ved lav effektivitet. Dette resultatet kan generaliseres til verdsetting av andre spreadopsjoner hvor LSM-algoritmen kan brukes.

Contents

List of Figures

List of Tables

Table 1: Commonly used abbreviations

ADF	Augmented Dickey-Fuller, a statistical test
AIC	The Akaike Information Criterion
BEKK	Baba, Engle, Kraft and Kroner, a class of volatility models
BIC	The Bayesian Information Criterion
c	Carbon, EUA or CO_2 quota
CCC	Constant Conditional Correlation, a volatility model
CCGT	Combined Cycle Gas Turbine
CO_2	Carbon Dioxide, a greenhouse gas
CSS_t	Clean spark spread on day t, the profit contribution of the plant
DCC	Dynamic Conditional Correlation, a class of volatility models
el	Electricity or power
EM	Expectation Maximization, a parameter estimation algorithm
EUA	EU Allowance, a permission to emit CO_2
EU ETS	European Union Emission Trading Scheme
g	Natural gas
GARCH	Generalized autoregressive conditional heteroskedasticity, a term for time varying volatility models
HRSG	Heat recovery steam generator, a component in a CCGT
HQ	Hannan-Quinn, an information criterion
KPSS	Kwiatkowski-Phillips-Schmidt-Shin, a statistical test
ICE	Intercontinental Exchange, a commodity exchange in the UK
LR	Likelihood Ratio, a test statistic
LSM	Least Squares Monte Carlo, a valuation algorithm
MNIG	Multivariate normal-inverse Gaussian, a probability distribution
MWh	Megawatt hour, an energy unit
NBP	National Balancing Point, a virtual gas hub in the UK
NPV	Net Present Value, a financial value estimate
OLS	Ordinary least squares, a type of regression
SC	Schwarz Criterion, an information criterion
VAR	Vector autoregressive, a time series model
VECM	Vector error-correction model, a time series model
QML	Quasi-maximum likelihood, a parameter estimation framework

Table 2: Commonly used symbols

χ	Steepness parameter in the MNIG distribution
$\boldsymbol{\alpha}$	Speed of error correction
α_{GARCH}	ARCH-coefficient in conditional volatility
b	Coefficient in LSM regression
$\boldsymbol{\beta}$	Co-integrating vector
β_{GARCH}	GARCH-coefficient in conditional volatility
c^{var}	Variable cost in the power plant, EUR/MWh
d	Dimensionality of model/number of variables
δ	Scale parameter in MNIG distribution
E	Overall efficiency of CCGT
ϵ_t	Unspecified residual, used for illustration purposes
Δc_F	Change in fuel cost
$\boldsymbol{\gamma}$	Skewness vector in MNIG distribution
$\boldsymbol{\Gamma}$	Covariance structure matrix in MNIG distribution
$\mathbf{H}_t$	Conditional covariance matrix
h	Univariate conditional volatility in CCC model
I_C	Carbon intensity of gas, in tonnes/MWh
κ	Coefficients in seasonal component regression
L	Heat rate for a specific plant
$L_{M,t}$	Market heat rate
$L^*_{M,t}$	Market heat rate accounting for plant-specific variable cost
m	Scenario index in LSM algorithm
μ_t	Short-hand for co-integrated mean equation
$\boldsymbol{\mu}$	Location vector in MNIG distribution
$\mathbf{P}_i$	Matrix of coefficients for endogenous lags in VECM model
p_t	Price at time t, in Euros, of electricity, gas or carbon
$\boldsymbol{\Pi}$	Matrix of coefficients in error correction term in VECM model
Φ	A constant in the VECM model
$\boldsymbol{\Psi}$	Coefficients for weekdays
$q^{el,gas}_{ij,t}$	Optimal consumption of gas or production of electricity in a state transition
$\mathbf{R}$	Coefficient matrix in CCC model
$r_{ij,t}$	Profit accrued by going from production state i to j
ρ	Risk-free discount rate
$\mathbf{s}(t)$	Seasonal component in commodity prices
t	Time index (occasionaly appears in the name of the Student's t distribution)
$\boldsymbol{\Theta}$	Dummy variable matrix for weekdays
$\boldsymbol{\theta}$	Adjustment to $\boldsymbol{\gamma}$ to get risk neutral MNIG distribution
$\boldsymbol{\Upsilon}$	Vector of residuals in X_t
$V_{i,t}$	Value of the CCGT being in state i at time t
ω	Constant in conditional volatility
$\mathbf{X}_t$	Vector containing the natural logarithms of the commodity prices
$\bar{\mathbf{X}}_t$	Vector of de-seasonalized log prices
$\mathbf{Z}_t$	Vector of i.i.d. errors

1 Introduction

Large investments are needed in European power generation capacity in the coming decades. Around 60% of the generation capacity was built more than 20 years ago, and 40% more than 30 years ago. The European Union emission trading scheme (EU ETS), designed to meet greenhouse gas reduction targets, further makes low-efficiency fossil fuel plants less profitable. Aging coal and low-efficiency gas power plants are likely to be replaced by modern combined cycle gas turbines, CCGTs. A CCGT can achieve efficiencies of around 60%, and is flexible enough to shut down in low demand periods, making it an ideal intermediate and base load power plant. Since 1990, most of the capacity additions in the UK has been gas-fired power plants, and CCGT capacity increased 28% from 2005 to 2010.

For a power plant operator, a credible valuation of generation assets is needed to make investment decisions. Increasing amounts of efficient coal and gas fired power plants, combined with feed in tariffs for renewable energy, has changed much of the business case for gas fired power plants. Traditionally, operators of such plants would hedge revenue with futures contract and run base load, while today the investment value is more dependent on a plant's ability to alter its output when prices change rapidly (see e.g. Timera Energy (2011)). This increases the complexity of the valuation. One needs to take into account both the variability of cash flows and the operational constraints of the specific power plant. In addition, a realistic assumption on price information arrival is necessary to not underestimate the plant's flexibility in production planning. In the literature, there is a vast amount of research on both price models and operational constraints, but there is little research on how the discrepancy between the arrival time of prices and the arrival time of price information affects the value of a power plant.

Power generators receive cash flows from the spread between electricity prices and the cost of fuel, which under the EU ETS also includes the cost of CO_2 allowances. Power, gas and carbon prices display both non-Gaussian behaviour and heteroscedastic volatility. The prices are bound by a co-integrating relationship, as the supply and demand of each commodity to varying degrees are affected by the prices of the other two. In this thesis we value a CCGT in the UK, proposing an accurate price model while taking operational constraints and information timing into account. We provide an extensive analysis of price dynamics, and incorporate the aforementioned price attributes in a vector error correction model (VECM) with CCC-GARCH volatility and MNIG distributed residuals. Our specification outperforms all the alternatives considered in terms of likelihood and information criteria. Taking a real options approach, we apply a Least Squares Monte Carlo (LSM) simulation to value a plant, whose actions are restricted by physical operational constraints. Assuming the power plant operates in the day-ahead market, we also take into account that information about the prices for a specific day arrives the day before the prices take effect. This gives the power plant operator more time to plan the next day's output than if one assumes that price information arrives the moment prices take effect.

Although the CCGT valuation problem has been studied in various forms, the authors have seen no related research with the same level of detail on *both*

the price dynamics and the observation of operational constraints. With our flexible valuation model, we analyze whether perfect foresight in prices is a valid assumption[1]. We find that the answer is strongly dependent on the efficiency of the plant, and that only high-efficiency plants justify this assumption. Traditional simulation-based valuation techniques normally make a myopic assumption with regards to price foresight, meaning that it is assumed that prices are only known from the instant they take effect. This is a too strict assumption when prices are determined on a day-ahead basis. We show that the implementation of a realistic assumption on the arrival of price information increases the value of the plant compared to the alternative.

Further, we find that the choice of regression form in the LSM algorithm impacts the value estimate to a high degree. Specifically, we show that multivariate regressions using price information on all three commodity prices outperform univariate regressions on the clean spark spread alone. We therefore conclude that simulating all commodity prices, taking their fundamental price relationship into account, gives a more realistic value estimate than simulating the spread by itself. This result can be generalized to other multi-asset valuation settings where LSM simulation is applicable, for example American style crack-spread options commonly used in oil price hedging.

1.1 Price model literature

A vast amount of literature focuses on modelling electricity and gas prices. To a varying degree, both commodities display stochastic volatility, regime changes and spikes, depending on data frequency and the market under analysis. Volatility clustering has been modeled by e.g. Escribano et al (2011) and Mu (2007), who both find strong evidence of stochastic volatility in electricity and natural gas prices, respectively. Geman and Roncoroni (2006) and de Jong (2006) model power price volatility by focusing on sudden price spikes, and suggest two different models: the former involving jumps such that if the price is above a threshold price it jumps back down to normal levels, and the latter a regime-switching model. Different GARCH models augmented with jumps are also frequently proposed for gas prices, see e.g. Bermejo-Aparicio et al. (2010) or Chan (2009). Most of the literature focuses on univariate modelling or consider the different commodities exogenous to each other - assumptions that we find to be over-simplified in this thesis. Johansen (1991) introduced the concept of vector error correction models (VECM), and Bunn and Fezzi (2009) employ the Johansen framework to estimate a structural VECM for power, gas and carbon prices. They show that carbon prices are affected by gas prices, and that both power and gas prices are influenced by shocks to the carbon price.

While Bunn and Fezzi assume a multivariate Gaussian distribution for the errors, a lot of recent literature has been concerned with the non-Gaussian behavior of both equities and commodities. Benth and Šaltytė-Benth (2004) employ the normal-inverse Gaussian distribution as the noise term in oil and natural gas spot prices, and Andresen et al. (2010) fit the distribution's multivariate form to

[1]Perfect foresight means that the decision maker has perfect knowledge of all future prices, giving him the opportunity to make the optimal decision at all times.

forward price returns at Nord Pool. Both papers find that the fit is superior to the Gaussian counterparts.

Daskalakis et al. (2009) study the prices of carbon allowances on different exchanges and conclude that a Geometrical Browinian Motion process augmented by jumps exhibits higher likelihood than other processes examined in the analysis. They also show that the different markets for carbon allowances do not provide arbitrage opportunities. Urig-Homburg and Wagner (2009) find that carbon spot and futures prices are linked by the cost-of-carry approach, and that due to cheap storability, the convenience yields in carbon prices are close to or equal to zero. Their results indicate that the price discovery of carbon allowances is led by the futures market. Bataller et al. (2007) study several variables considered to influence the price of carbon prices, including weather data. They find that the most important factors in determining prices are the prices of emission intensive sources, such as coal, crude oil and gas. Weather data showed little significance except in the cases of extreme temperatures. For an extensive literature review on the subject of carbon prices, see Chevallier (2011).

1.2 Valuation literature

The literature on power plant valuation can broadly be categorized in two sets, according to methodology. The papers in the first category use real options theory, while papers in the second category rely on operations research techniques, stochastic or deterministic, to estimate a plant's value. In the first category, Abadie and Chamorro (2008) assume gas prices follow a geometric mean-reverting process with normally distributed errors, while electricity prices are assumed deterministic. They derive closed form solutions for the option value of investing in a CCGT, and a CCGT that can switch to coal as its input fuel. Näsäkkäläand Fleten (2005) extend the spark spread model to have two stochastic, correlated factors, incorporating both uncertainty in equilibrium level and short term variations. They derive expressions for investment thresholds, in terms of spark spread levels, and for upgrading a base load plant to a peak load plant. In Fleten and Näsäkkälä (2010) they use the same methodology and data, but analyze the effect of an abandonment option, the effect of carbon prices and a plant's upper and lower value bound. Cassano and Sick (2009) use the Least Squares Monte Carlo (LSM) algorithm, which is a dynamic programming simulation approach where continuation value is approximated by a least squares regression. They estimate a regime switching model for the market heat rate (the ratio of power prices to fuel prices), and use the Margrabe approach to valuing a spread option. The LSM regressions are on a fourth degree power function, and they find that the market heat rate has good explanatory power for the plant value. Deng and Oren (2003) use the LSM algorithm on a gas power plant, with special emphasis on correct modelling of operational constraints. Deng and Xia (2006) extend the LSM approach to value a tolling agreement in which the buyer has the right to either operate a power plant or to receive its output under certain restrictions. The power and gas prices are assumed to be correlated mean-reverting processes augmented with jumps, and they find that the plant value increases with jumps in the power price. Los, de Jong and van Dijken

(2009) argue that correlated price processes is not sufficient to model power and gas prices' fundamental relationship over a longer period. Thus, for a power plant valuation, a co-integrating relationship should be estimated and be a part of the valuation.

In the second category, Tseng and Barz (2002) employ a Monte Carlo simulation to a multi-stage stochastic optimization model. They incorporate operational constraints such as minimum on and off times and start-up costs. They conclude that overlooking these constraints under-estimates the risks in the cash flows, but they do not incorporate ramp times. Zhu (2004) presents the valuation problem as a self-scheduling problem, and solves first a deterministic case via network optimization. He then generates price scenarios, and solves the deterministic model in each case. A regression based method is then proposed to combine the solutions to an optimal operations strategy under price uncertainty. Bjerksund, Stensland and Vagstad (2008) compare valuation techniques for gas storages. They apply a simple decision rule to a complex price process and a simple price process to a more complex dynamic programming approach, and compare both of these to a perfect foresight scenario. When benchmarked against the perfect foresight scenario they conclude that the model with the more advanced price process performs best, and that it is therefore more important to focus on the price process than the optimisation technique.

The papers of Deng and Oren (2003), Tseng and Barz (2002) and Zhu (2004) are especially concerned with modelling operational constraints, but neither of them incorporate a difference in time between the arrivals of price information and the arrival of prices, implicitly assuming zero foresight in prices.

This thesis uses some of the methods and results from the first category of papers, but extends the methodology to combine the LSM approach with daily granularity with deterministic dynamic programming within 24 hour periods. Our valuation setting is very similar to what Bjerksund, Stensland and Vagstad (2008) discuss, but we argue that it is possible, and even necessary, to combine an accurate optimization with an advanced price process to obtain a realistic value estimate.

The rest of the thesis is organized as follows: In Section 2, we describe the development of de-regulated electricity and gas markets, as well as the EU emission allowance markets. We also present some basic facts about combined cycle gas turbines. We analyze the price data and present the vector error-correction model in Section 3. In Section 4 we describe the valuation model, as well as some improvements that can be made when using daily decision points. The results of the valuation, along with a sensitivity analysis and discussion, is presented in Section 5, before we conclude in Section 6.

2 Institutional background

In this section, we will introduce the relevant markets in which a power plant operates, with a special consideration of the UK market. We first introduce the markets for power, natural gas and carbon allowances, and then give a short introduction to combined cycle gas turbines.

2.1 Market for electric power

The UK was the first country in Europe to liberalize the power market in 1989 under the Electricity Act. The Electricity Act was followed up with the privatization of the power sector in 1990. The *Transmission System Operator (TSO)*, who distributes the physical power and maintains infrastructure, is still a state-controlled entity as its function is crucial for the safety of electricity supply. The national generator and retail company, Central Electric Generating Board (CEGB), was privatized and split into three independent power-producing entities. The reason for privatization was to allow for competition to create a more efficient market both for consumers and suppliers of power. After the privatization, a wealth of generating and retail electricity companies have been founded, leading to a more complete and efficient market.

In the liberalized power markets of the EU and the Nordics, the price is determined by the principles of supply and demand; meaning that participants in an auction agree on prices for a certain power measured in MW, or energy measured in MWh. The non-storability of electricity prevents any stocking, creating the need for a market that can regulate the supply on short time horizons. The power generators use different technologies to produce power, with different marginal cost of operations and production rates. For example, a run-of-the-river hydro power plant has low operating costs and will produce almost all the time; it serves as a base load plant. The marginal cost of coal fired plants is higher than for hydro power plants, and in general gas fired plants' costs are higher still. The generators will produce power only when their marginal costs are covered by the electricity price. As a result of the different marginal prices, a merit order curve can be created, where generators are ranked by their marginal prices and production capacity (Pöyry, 2010). See Figure 1 for an example. Because demand for power is almost inelastic, the production patterns differ for generators placed on different points on the merit order curve. A generator with low marginal cost will usually have a continuous production, while generators with high marginal cost, like gas-fired power plants, will have to shut down production when prices drop. An inexpected supply increase on the lower end of the curve will give a drop in prices. The entire merit order curve will then shift to the right, and the supply curve will cross the demand curve at lower prices. Before the privatization of the power market, CEGB used the merit order to set the price of electricity. Today, the price is settled in the market, but the merit order concept is still the driving force of electricity prices.

Power is traded on different markets, whos main difference is the time horizon of the participants. *The spot market* is a physical market where day-ahead and intraday contracts are traded for the next day. The *futures and forwards market*

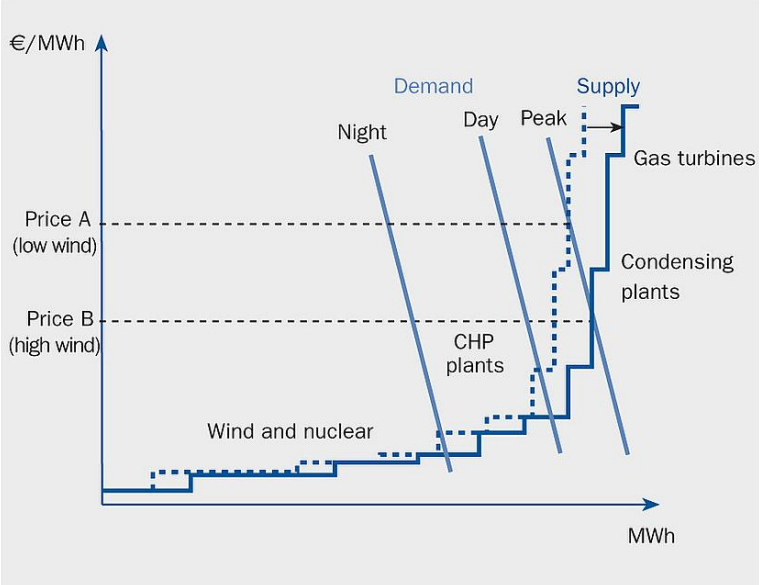

Figure 1: Illustration of the merit order curve (Pöyry, 2010).

serve longer-term hedging purposes and offer monthly, quarterly, semi-annual and annual contracts. The *Over-the-Counter market* consists of private and bilateral trades outside the power exchange.

In this thesis we will assume that the power plant is engaged in the day-ahead market only. The day-ahead market is settled at midday using auctions in which the price for the next day is settled by comparing total demand and supply. The settlement price for each auction is defined as the marginal price that balances the auction. This is the price that will be paid or received by all participants in the auction regardless of their initial bid, as long as the bid is above the settlement price for buyers, and below for sellers. The day-ahead structure implies an important difference between the trading date and the delivery date. The quoted date is the date of the transaction, whereas the delivery date is the date during which the seller of a contract generates the stipulated amount of power.

2.2 Market for natural gas

The UK market for natural gas was gradually liberalized during the period from 1986 through 1996, beginning with the privatization of British Gas in 1986. The 1995 Gas Act demanded that ownership of production and transmission of natural gas should be split up as the transmission function is a natural monopoly while production and sales would benefit from competition. British Gas was split into a producing entinty, Centrica plc, that was privatized, and a state-controlled transsmission entity, BG plc. In 1996, the market for natural gas was opened for all participants, meaning that spot and futures contracts could be traded freely on commodity exchanges. In the UK as in all liberalized gas markets, natural gas is traded on the basis of either virtual or physical trading points, called market hubs, where gas contracts are standardized to secure liquidity. Physical trading hubs, like for example Zeebrügge in Belgium, are usually situated at entry points

of gas pipelines or large storage facilities, and any commodity contract has a designated point of delivery. Virtual trading hubs however, like the National Balancing Point (NBP) in the UK, have no specific point of delivery; the gas is considered delivered when it is injected into the pipeline grid. NBP is the largest and most liquid gas trading hub in Europe. Other leading European trading hubs include TTF in Holland, Zeebrügge in Belgium and Netconnect Germany (NCG), formerly known as E.ON Gastransport (EGT). The electricity and gas markets are closely connected in the UK. The power sector accounts for 34% of the natural gas demand (DECC, 2011a), and gas-fired power plants provide 44% of the country's electricity production (DECC, 2011b). The relation implies both that the electricity price might be a driver for the natural gas price, and vice versa.

The standard natural gas contract in UK is traded on a day-ahead basis, using contracts that are quite similar to the day-ahead contracts for electricity. As the day-ahead contracts are only traded Monday to Friday, deliveries for Saturdays, Sundays and Mondays will all be settled on Fridays (the same is valid for holidays). The day-ahead contracts are essentially futures with a one-day maturity, but are usually referred to as spot contracts. The contracts are settled via midday auctions. Longer-term futures for monthly, quarterly, semiannual and annual delivery are also frequently traded, as well as options on these. In addition, there exist contracts for within-day delivery and other structures, but the trading volumes in these contracts are only a fraction of the day-ahead market's volume.

2.3 Market for carbon allowances

To comply with its obligations under the Kyoto protocol, the European Union (EU) introduced the *EU emissions trading scheme (EU ETS)* in 2005. The EU ETS is a cap-and-trade scheme, placing a fixed cap on the annual CO_2 emissions from the sectors that fall within the ETS. The cap is to be reduced each year to conform with the Kyoto targets. One certificate, or EU allowance (EUA), allows the holder of the certificate to emit 1 ton of CO_2. At the end of each year, companies must document their CO_2 emissions and submit the according number of EUAs to their national government. If they are left with a surplus of EUAs these can be sold on one of the markets for carbon emissions trading. If the company has a shortfall of EUAs, they must either buy the permissions on the market or pay a fine of EUR 100/ton CO_2 (which is not favourable compared to a market price that has ranged between 6–32 EUR/ton CO_2). Another way to comply with the EU ETS is to buy *certified emission reductions (CERs)*, wich are certificates that document a UN-approved reduction of greenhouse gases in a non-industrialized country.

The EUAs are organised in different *commitment periods*, where the first period, from January 1st, 2005 to December 31st, 2007, was described as a test phase. The second commitment period covers 2008-2012. Intra-phase *banking* and *borrowing* is allowed, meaning that an EUA bought in one year can be banked and used in the next year, or one can use it to cover an overemission during the previous year, but no banking or borrowing was allowed between the first and the second commitment period. As the Kyoto Protocol took effect in 2008, any

banking from previous periods would put the EU in risk of incompliance with Kyoto targets (Houser et al., 2008). Another reason for the banking ban may be that the market imperfections from the test phase should not be transferred to the second phase through banking (Chevallier and Alberola, 2009). The banking ban put severe pressure on spot prices for EUAs from 2006 to 2008, as it gradually became clear that the cap on emissions was set too high, resulting in an oversupply of certificates in that period. As a result, the prices of certificates expiring before 2008 declined to zero, while those expiring after 2008 remained in the range of 15-30 EUR/ton CO_2. From 2008 onwards, banking and borrowing is also allowed between commitment periods, eliminating the problem of prices going to zero. The fact that EUAs has no cost of storage other than the cost of capital creates an equilibrium between futures and spot prices based on the cost-of-carry approach (Urig-Homburg and Wagner, 2009) with a convenience yield close to zero.

EUAs are traded on several different commodity exchanges, the biggest of which is ICE in the UK, with a market share of around 91%. Other significant exchanges include Green Exchange, Nordpool, EEX and Bluenext. The members of the EU ETS include the 27 EU countries, as well as Iceland, Norway and Lichtenstein. The largest sectors include power production and industrials. The aviation sector is included as of January 1st, 2012. In the first two phases, emission allowances was distributed to companies according to their historical emissions, while from the third phase, starting in 2013, the distribution will mainly be done by auction.

2.4 Combined Cycle Gas Turbines

A gas-fired power plant converts natural gas to electricity by using the gas' combustion energy to heat and expand air. Air is compressed, mixed with natural gas, and burnt in a combustion chamber. The hot exhaust gas is then expanded through a gas turbine, making the turbine shaft rotate while the gas pressure decreases. The shaft often drives the compressor as well as an electricity-producing generator. Several configurations of multi-stage compression and expansion are available and in use in the industry. Modern gas turbines often have inter-cooling and re-compression between compressor and turbine stages to enhance efficiency, and a typical, modern gas turbine has an overall efficiency of 35–40%. However, the gases leaving the turbine often has considerable heat energy which can be used to power yet another power production cycle. A heat recovery steam generator (HRSG), or boiler, uses the hot exhaust gas to boil water. The steam is then expanded on a steam turbine. The two power cycles together are called a combined cycle power plant, or a combined cycle gas turbine (CCGT). Utilizing the spare heat in the gas turbine exhaust thereby improves the overall plant efficiency E, which is in the range of 55-60% for modern CCGTs. A simplified illustration of a CCGT is shown in Figure 2.

For a CCGT, the heat rate L is defined to be the units of gas required to produce one unit of power. Its relation to the overall plant efficiency is $L = \frac{1}{E}$, and it varies with the temperature in the plant's boiler. Thermodynamic equations including gas flow rate, compressor and turbine pressure ratios, ambient

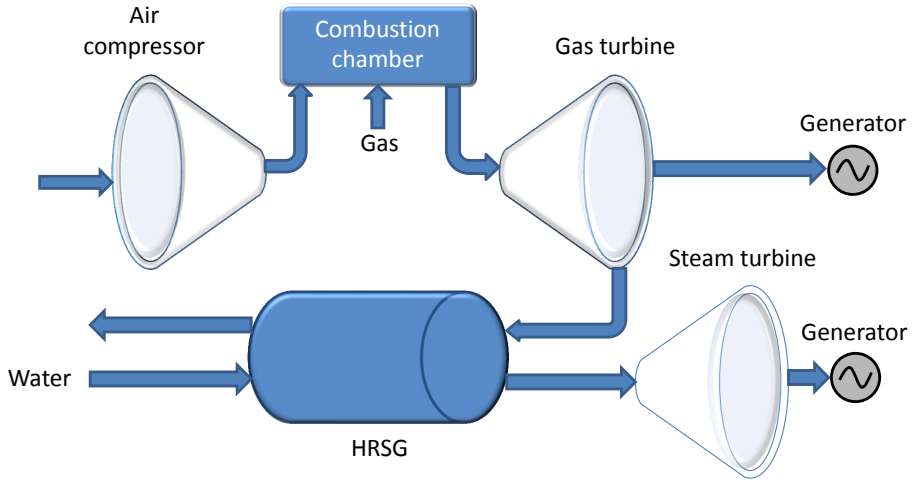

Figure 2: Schematic illustration of a combined cycle gas turbine. The HRSG utilizes spare heat in the exhaust gas to power a second power cycle, increasing overall efficiency.

temperature and heat exchanger temperatures are needed for exact calculations, but in general the plant will have a constant heat rate once it reaches stable operating conditions at its design capacity. At outputs of higher or lower power than the design capacity, and during ramp-up and ramp-down, the efficiency will be slightly impaired.

The plant's revenue stems from selling power in the market, while the direct costs are purchases of natural gas and carbon allowances. We denote the carbon intensity of the gas I_C, measured in tons of CO_2 per MWh of gas. We refer to the plant's direct contribution margin as the clean spark spread, $CSS_t = p_t^{El} - L_t p_t^G - I_C L_t p_t^C$, where p_t is the price of the plant's input or output and L_t is the plant's heat rate at time t. There are several markets in which the power plant can trade natural gas, electricity and CO_2 quotas. In this thesis, we will assume that the plant purchases natural gas and sells electric power in the day ahead market (spot market), and that it is bound by the European emission trading scheme for CO_2 quotas. As the CSS will vary for every power plant, the *market heat rate*, L_M is a more useful measure of the capacity utilization in the market. L_M is the ratio of power to gas prices, $L_M = p^{el}/p^g$, or for markets under an ETS; $L_M = p^{el}/(p^g + I_C p^c)$. If the marginal power producer is a gas-fired power plant, the market heat rate gives an indication of the marginal plant's efficiency $E_M = \frac{1}{L_M}$. A positive change of the market heat rate indicates that demand is shifted to the right in Figure 1, or that supply is shifted to the left. Authors including Cassano and Sick (2009) have used the market heat rate to value a gas-fired power plant.

When the clean spark spread becomes negative, meaning that the market heat rate falls below the heat rate of the power plant, it is no longer profitable to produce power and the plant should ideally be switched off. However, once the plant is switched off it will start to cool down. Switching the plant on again requires a certain amount of gas to reheat the boiler, which takes a certain amount of time (see Figure 3). The decision maker must include this consideration when deciding to switch off; for example, if he expects the CSS to become positive

again shortly, it might be more profitable to produce with a loss for a short period than to cease production and incur natural gas and carbon costs when restarting. Another consideration is that CCGTs can not ramp production up or down instantaneously. It will normally have a minimum output and a maximum output, and ramping up or down within these limits will take some time. As discussed previously, this will affect the heat rate and thereby the plant's CSS.

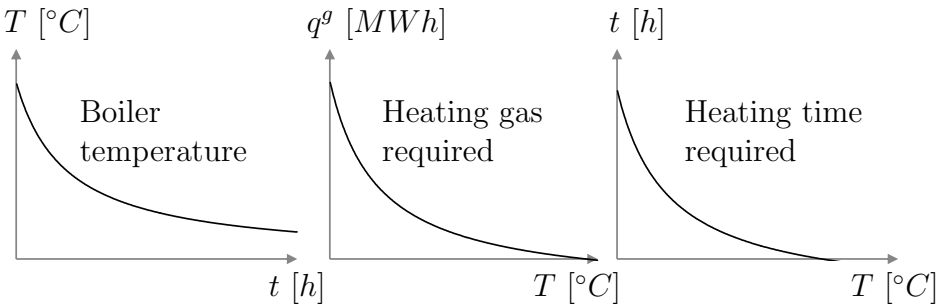

Figure 3: Simplified illustrations of what happens after a plant is switched off. Left: Boiler temperature as a function of time. Middle: Amount of heating gas, q^g needed to start plant, when the boiler temperature is $T\ [°C]$. Right: Startup time t given initial temperature T.

3 Data analysis and price modelling

This section describes the data series used in the analyses and discusses their most important properties with respect to price modelling. The prices we study are the UK prices for electricity, natural gas and carbon allowances. A common trait shown by all three data series is a gradual development from high volatility and fat tailed returns distributions, to lower volatility and fewer spikes in the later years of the data set. The shift is especially pronounced when comparing the periods before and after 2009. We expect this shift to be a result of tighter integration of power and gas markets, a more efficient market and, for carbon, a result of the learning curve for market participants engaging in the relatively new carbon market.

Electric power and natural gas prices are thought to be bound by a strong relationship, because both a large part of UK's electricity production uses gas as its primary energy input, and a large part of the natural gas consumption is consumed by power plants. In econometric terms, this would lead to an assumption that the prices of the two commodities are cointegrated. With regards to carbon allowances, the price of carbon affects the profitability of all fossil fuel power plants, as well as any industrial plant that uses natural gas. The drivers that affect the need for power and gas will also affect the need for carbon allowances. We therefore see a potential for all of the three commodities to be co-integrated, a hypothesis that we will devote part of this section to investigate.

In this section, we choose an appropriate price model, test its validity and estimate its parameters. We develop a co-integrated model for the power, gas and carbon prices with heteroscedastic volatility. We discuss several potential models for volatility and residuals before concluding that a VECM model with CCC-GARCH volatility and MNIG distributed error terms exhibits a superior fit to competing models. We assume that the price model is on the form $ln[\mathbf{P}(t)] = \mathbf{s}(t)+\boldsymbol{\mu}_t+\boldsymbol{\Upsilon}_t$, where $\mathbf{P}(t)$ is the column vector of prices at time t, $\mathbf{s}(t)$ is a function of time that captures seasonality effects, $\boldsymbol{\mu}_t$ is a deterministic mean equation and $\boldsymbol{\Upsilon}_t$ is a stochastic residual vector. Before analysing the data, we show a plot of the complete data set and the descriptive statistics of the series, in Figure 4 and Table 3 respectively.

Table 3: Descriptive statistics for power, gas and carbon prices. The data sets run from October 15$^{\text{th}}$, 2001 for electricity and gas and from April 22$^{\text{nd}}$, 2005 for carbon. All three sets end April 24$^{\text{th}}$, 2012.

	Power	Gas	Carbon
Mean	48.907	16.400	17.130
Median	44.405	14.784	16.020
Standard Deviation	23.648	7.848	5.075
Skewness	1.822	1.935	0.187
Excess Kurtosis	5.082	9.778	-0.446
Minimum	13.952	0.703	6.080
Maximum	228.095	96.724	32.250
Count	3163	3163	2135

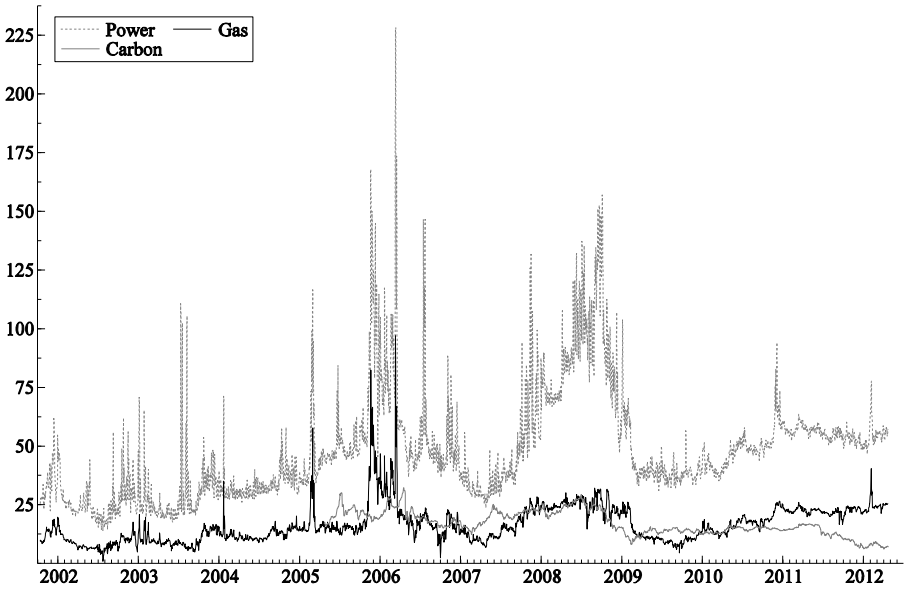

Figure 4: Plot of the complete data set for power, gas and carbon, 2002-2012. The carbon series begins in 2005. Prices in EUR/MWh for electricity and natural gas, and EUR/ton CO_2 for carbon.

3.1 Power price data

The data set on power prices consists of day-ahead prices for the UK Power Exchange. The data series runs from October 15$^{\text{th}}$ 2001 to April 24$^{\text{th}}$ 2012, a total of 10.5 years. Weekend prices are available from June 2002 onwards. The power prices are originally quoted in GBP/MWh, but all power and gas prices in this thesis will be presented in EUR/MWh as this is easily comparable to other European markets. The CO_2 prices are also quoted in EUR. To remove seasonal patterns in the power price, the logarithm of the prices is regressed on a set of sine-cosine combinations (as used by, e.g. Heydari and Siddiqui (2010) and Lucia and Schwartz (2002)) with periods of 1 year, half a year and three months, as well as a trend term. Sine and cosine combinations captured more of the variation in both power and gas prices than monthly or weekly dummies. Eq. (1) shows the full regression, while Table 4 displays the significant coefficients from its estimation.

$$\begin{aligned}\ln p_t^{el} = \kappa_0^{el} + \kappa_1^{el} t + \kappa_2 \sin(\frac{2\pi t}{365}) + \kappa_3 \cos(\frac{2\pi t}{365}) + \kappa_4 \sin(\frac{2\pi t}{183}) + \kappa_5 \cos(\frac{2\pi t}{183}) \\ + \kappa_6 \sin(\frac{2\pi t}{91}) + \kappa_7 \cos(\frac{2\pi t}{91}) + \epsilon_t^{el}\end{aligned} \tag{1}$$

Figure 5 shows the log power price series $x_t^{el} = \ln p_t^{el}$ as well as the fitted price and resulting residuals. A deterministic trend is included in the regression to improve the fit of the sine/cosine terms, but we do not assume a deterministic trend in the series. A stochastic trend, implying that the series are integrated rather than trend stationary, is a more probable model for financial prices. We will explore this further in Section 3.5.

Table 4: Estimated coefficients in Eq. 1, describing the deterministic, seasonal component of the power prices.

Variable	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	3.388	0.013	260.0	0.000	0.956
Time	2.08e-4	0.000	35.6	0.000	0.287
cos365	0.073	0.009	8.1	0.000	0.020
sin365	0.055	0.009	6.1	0.000	0.012
cos183	-0.025	0.009	-2.7	0.007	0.002
sin183	0.041	0.009	4.5	0.000	0.007

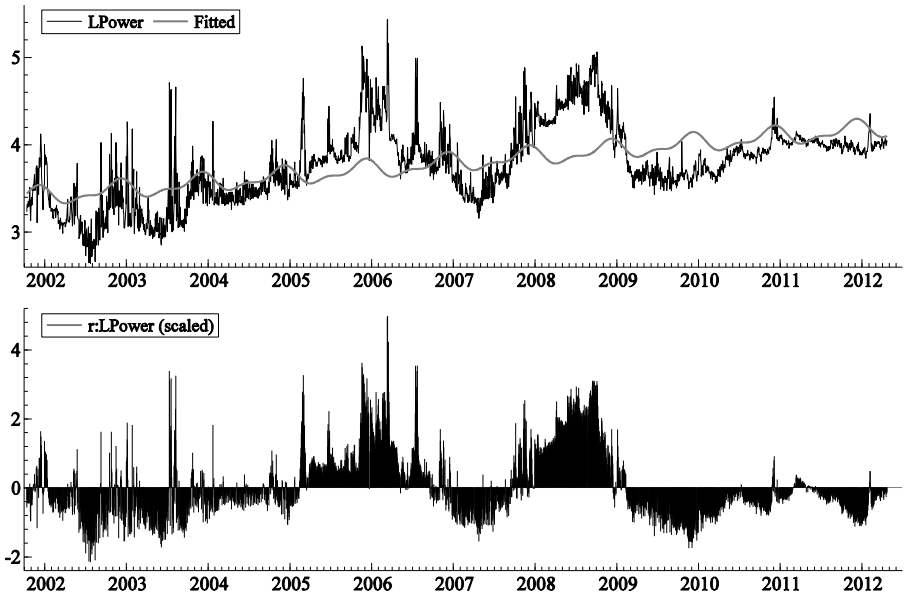

Figure 5: Top: Logarithm of power prices and its estimated seasonal pattern. Bottom: Residuals after removing seasonal effects

Day-of-the-week effects are also present, but the coefficients change significantly throughout the data set (see Figure 6), rendering an estimation of dummy variables for the whole data set rather useless. There may be several explanations to the declining day-of-the-week effects, but we expect that the increase of natural gas power in the UK fuel mix at the cost of less flexible power sources like nuclear and coal (see Figure 7) is a large part of the explanation, as natural gas power plants to a higher degree can adjust production to market demand. The deregulation and integration of both UK and neighbouring electricity markets, as well as increased transmission capacity to other countries, might also be part of the explanation.

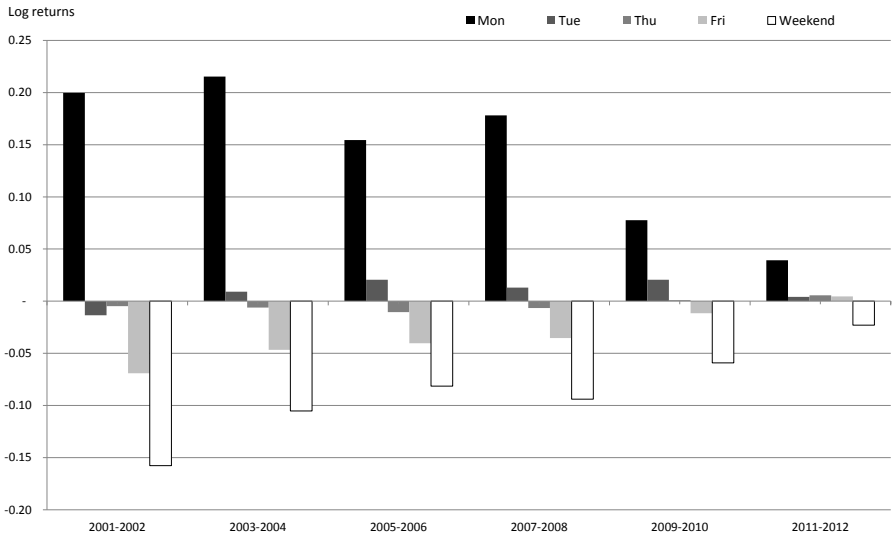

Figure 6: Estimated day-of-the-week effect in electricity log returns over six different two-year periods. The effect is clearly diminishing throughout the 10 years of data.

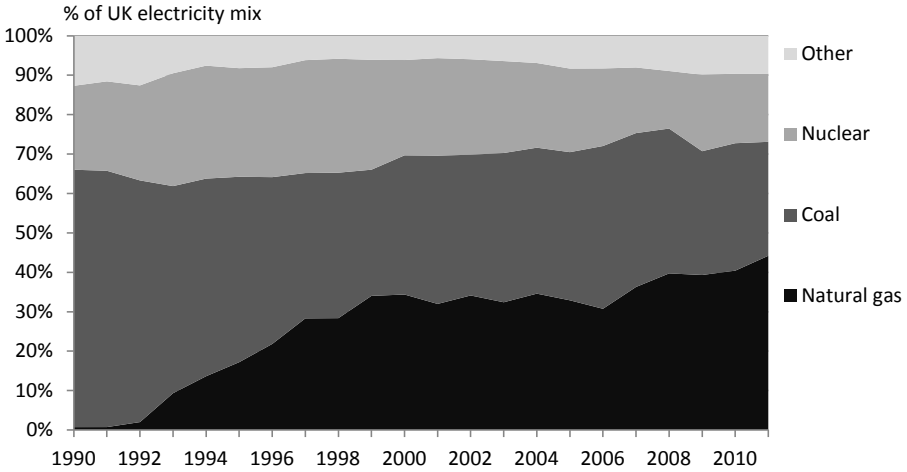

Figure 7: Natural gas as a percentage of the fuel mix used in electricity production in the UK (DECC, 2011b).

A visual inspection of the prices and the fitted price of Figure 5 reveals that prices are expected to be at their highest during December and at their lowest

in April. We leave the estimation of weekday-effects for Section 3.5, where they will be estimated for all three variables simultaneously. Inspecting the residuals, we see that their volatility does not seem constant. Spikes in the price series are also visible, but diminish throughout the period.

3.2 Gas price data

Day-ahead prices for the UK National Balancing Point (NBP) have been used in the modelling of gas prices. As for the power price series, the data series runs from October 15th, 2001 to April 24th, 2012, a total of 10.5 years. The prices are originally quoted in GBP/therm, but are presented in EUR/MWh. Performing the same regression as in Eq. (1) with the log prices of natural gas as the endogenous variable, reveals a pattern similar to that of power prices. The regression results are summarized in Table 5. The main difference is that the lowest price for the year is expected in July rather than in April. The same diminishing spike behavior can be observed in Figure 8.

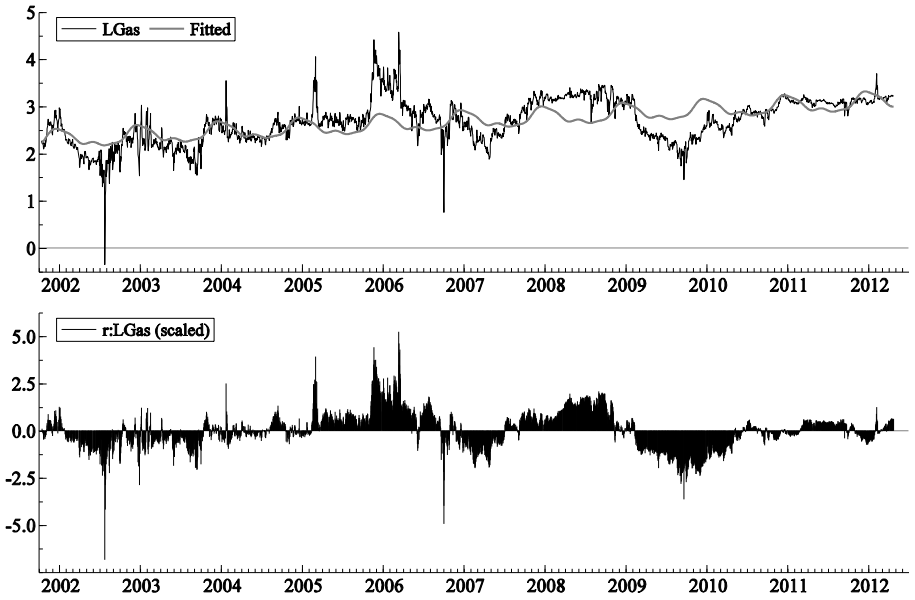

Figure 8: Top: Log gas prices and its estimated seasonal pattern. Bottom: Residuals after removing seasonal effects

Table 5: Estimated coefficients in Eq. 1 performed on the log gas prices.

Variable	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	2.270	0.013	170.0	0.000	0.902
Time	2.16e-4	0.000	36.2	0.000	0.294
cos365	0.031	0.009	3.3	0.001	0.003
sin365	0.173	0.009	18.6	0.000	0.099
cos183	-0.047	0.009	-5.1	0.000	0.008
sin183	0.048	0.009	5.2	0.000	0.008
cos91	-0.025	0.009	-2.7	0.007	0.002

3.3 Carbon price data

In modeling carbon price data, we use first position futures prices from ICE. We use futures prices that are valid in the second phase of the EU ETS, thereby avoiding the problem of spot prices declining to zero during 2007 (see Section 2.3). The series of log carbon prices is shown in Figure 9. The data set runs from April 22$^{\text{nd}}$, 2005 to April 24$^{\text{th}}$, 2012, a total of seven years, and is quoted in EUR/ton CO_2.

For the sake of completeness, we also run the least squares regression in Eq. (1) for carbon prices. This results in significant coefficients on annual and semi-annual sine functions, as well as a negative trend. However, due to the low cost of carry of carbon certificates (Urig-Homburg and Wagner, 2009) and the ability of banking and borrowing certificates from adjacent periods, there is no economic reason that seasonal patterns should exist. Neither do we see a reason for a negative trend in the price of a commodity that, by its cap-and-trade construction, has a declining supply through time. The significant coefficients can be explained by carbon certificates being a very young commodity. As market participants have learned and adjusted during the first few years, prices have risen from levels of EUR 6/ton to EUR 30/ton, and fallen back to EUR 7/ton. With only seven years of data, each of these movements can give a seasonal pattern a high significance. As the market for EUAs is open also to pure financial participants, any emerging seasonal pattern should be neutralized by arbitrageurs. We therefore choose to model the carbon price without trend or seasonal patterns, as the seasonal patterns probably result from a spurious regression.

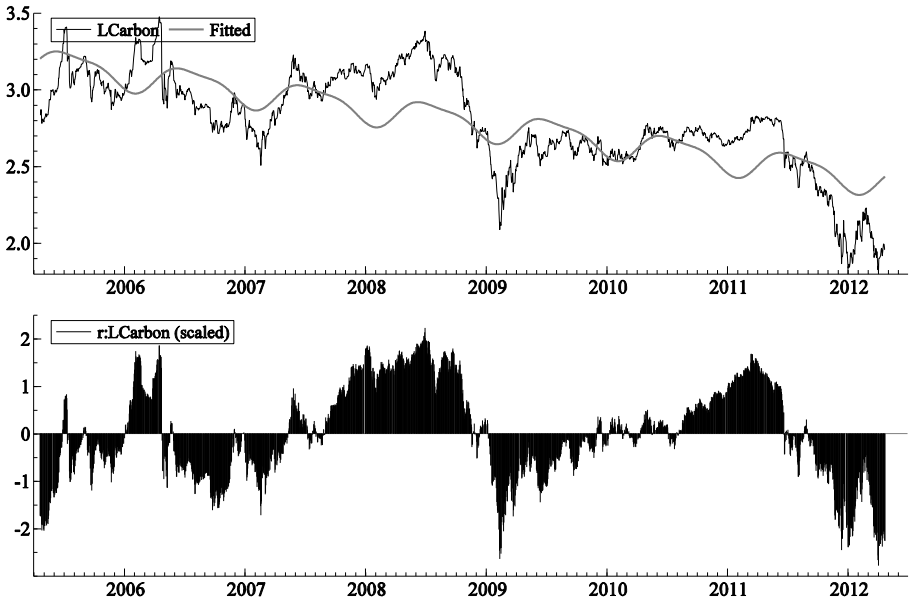

Figure 9: Top: Log carbon prices and its estimated seasonal pattern. Bottom: Residuals after removing seasonal effects

3.4 Stationarity

After removing the seasonal effects $\mathbf{s}(t)$, but not the deterministic time trends or the constants, we are left with the de-seasonalized series $\bar{\mathbf{X}}_t = \mathbf{X}_t - \mathbf{s}(t)$. We now investigate whether the three de-seasonalized time series contain unit roots, i.e., whether they are $I(1)$. A time series being $I(1)$ means that they have a stochastic trend that may or may not have a zero mean. We perform the augmented Dickey-Fuller (ADF) unit root test (Said and Dickey, 1984) as well as the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) stationarity test (Kwiatkowski et al., 1992). The null hypothesis in the ADF test is that the series is $I(1)$, meaning that they have one unit root $\phi = 1$, and in the KPSS test that the series is stationary.

The result of the tests, for both the log price series and the log return series, are shown in Table 6. As expected, the series for power, gas and carbon all seem to be $I(1)$. The ADF t-statistic for gas prices is significant on the 5% level, indicating stationarity, but stationarity is strongly rejected by the KPSS statistic. Non-stationarity in gas spot prices is consistent with research by e.g. Modjtahedi and Movassagh (2005).

Table 6: Stationarity tests for power, gas and carbon prices and returns. * and ** mean rejection of null hypotheses on the 5% and 1% level, respectively. The null hypothesis of the ADF test is an I(1) series, and the null of the KPSS test is an I(0) series, so we conclude that all three series are I(1).

Variable	Sample size	ϕ	ADF statistic	KPSS statistic
$\bar{x}^{el}$	3163	0.987	-2.688	36.723**
$\bar{x}^{g}$	3163	0.988	-2.889*	32.785**
$\bar{x}^{c}$	2135	0.998	-1.146	37.736**
$\Delta\bar{x}^{el}$	3162	-1.925	-13.798**	0.005
$\Delta\bar{x}^{g}$	3162	-1.410	-14.030**	0.006
$\Delta\bar{x}^{c}$	2134	0.072	-42.978**	0.133
Residuals of VECM model, Eq. 2a				
υ^{el}	2135	-0.329	-13.456**	0.329
υ^{g}	2135	-0.417	-13.791**	0.172
υ^{c}	2135	-0.086	-12.791**	0.064
Significance levels				
10%			-2.57	0.347
5%			-2.86	0.463
1%			-3.43	0.739

3.5 Co-integration modelling

In Figure 10, we show the log price series after correcting the power and gas series for seasonality. It is clear that power and gas prices follow each other's movements quite closely, while carbon carbon prices exhibit the same behaviour more weakly. When two or more integrated variables share a common stochastic drift, they are said to be co-integrated (Engle and Granger, 1987). In this section

we test formally for co-integration, choose the specifications of the co-integration model and estimate its parameters.

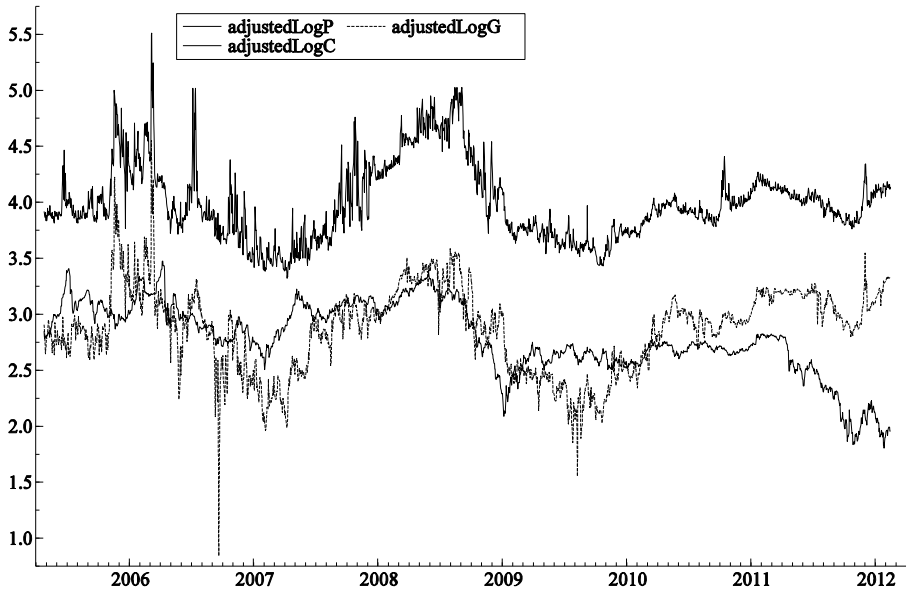

Figure 10: The logarithms of electricity, natural gas and carbon prices after correcting for seasonal effects.

Within the Johansen framework (Johansen, 1991), a VAR(k) model may be tested for co-integration with the trace test. The model contains k lags, chosen by minimizing an information criterion, and can be rewritten as follows:

$$\bar{\mathbf{X}}_t = \Phi + \sum_{i=1}^{k} \mathbf{\Lambda}_i \bar{\mathbf{X}}_{t-i} + \mathbf{\Upsilon}_t \tag{2a}$$

$$\Delta\bar{\mathbf{X}}_t = \Phi + (\mathbf{\Lambda}_1 + \ldots + \mathbf{\Lambda}_k)\bar{\mathbf{X}}_{t-1} - (\mathbf{\Lambda}_2 + \ldots + \mathbf{\Lambda}_k)\Delta\bar{\mathbf{X}}_{t-1} \tag{2b}$$

$$- (\mathbf{\Lambda}_3 + \ldots \mathbf{\Lambda}_k)\Delta\bar{\mathbf{X}}_{t-2} - \ldots - \mathbf{\Lambda}_{k-1}\Delta\bar{\mathbf{X}}_{t-k+1} \tag{2c}$$

$$= \Phi + \mathbf{\Pi}\bar{\mathbf{X}}_{t-1} + \sum_{i=1}^{k-1} \mathbf{P}_i \Delta\bar{\mathbf{X}}_{t-i} + \mathbf{\Upsilon}_t \tag{2d}$$

where $\bar{\mathbf{X}}_t$ is the vector of de-seasonalized log prices, $\Phi \in \mathbb{R}$ is a constant, $\{\mathbf{\Lambda}_i, \mathbf{\Pi}, \mathbf{P}_i\} \in \mathbb{R}^{d\times d}$ are matrices of coefficients, and $\mathbf{\Upsilon}_t \in \mathbb{R}^d$ is the vector of residuals. When all d variables are $I(1)$, $\Delta\bar{\mathbf{X}}_t$ is stationary, and the linear combination(s) $\mathbf{\Pi}\bar{\mathbf{X}}_{t-1}$ must also be stationary. The trace test is a test on the rank of the matrix $\mathbf{\Pi}$ with the null hypothesis that the rank $r \leq r_0$. If $0 < r < d$, then $\bar{\mathbf{X}}_t$ is said to have r co-integrating relations. When modelling the logarithms of electricity, natural gas and carbon prices, we get $d = 3$ and $\bar{\mathbf{X}}_t = [\bar{x}_t^{el}, \bar{x}t^g, \bar{x}_t^c]^\intercal$. Both the Akaike Information Criterion and Akaike's Final Prediction Error recommend $k = 8$ endogenous lags (note that we only use the data from April 22$^{\text{nd}}$, 2005, as this is where the carbon series starts).

From the trace test reported in Table 7, it is clear that we reject the hypothesis of the rank being zero, but we cannot reject it being less than, or equal to,

Table 7: Johansen's trace test. The rank under H_0 is r_0, and LR refers to the likelihood ratio. We reject the hypothesis that the rank is equal to zero. We can not reject that it is less than, or equal to, one.

r_0	LR	p-value
0	59.74	0.00
1	15.92	0.18
2	1.71	0.82

one. We conclude that there exists one co-integrating vector between the logarithms of electricity, natural gas and carbon prices. The long term matrix $\mathbf{\Pi}$ can be decomposed into $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}^\intercal$, such that the term $\boldsymbol{\alpha}\boldsymbol{\beta}^\intercal\bar{\mathbf{X}}_{t-1}$ can be interpreted as the correction caused by deviations from the co-integrating relationship. The co-integrating vector is $\boldsymbol{\beta}^\intercal$ and the coefficients in $\boldsymbol{\alpha}$ are the speeds of correction for each variable. For this reason, Eq. (2d) is known as a Vector Error Correction Model (VECM). In our setting, the constant term Φ should be a part of the long term relationship between the variables. The long term relation $\boldsymbol{\beta}^\intercal\bar{\mathbf{X}}$ has an equilibrium level $\Phi \neq 0$ that the prices revert towards. Although there does not exist an exact mathematical relationship between the average CSS and e^Φ, the two values are closely connected. For example, when the fuel costs are constant, $\mathbb{E}[CSS] = \vartheta_0 + \vartheta_1 e^\Phi$. The economic interpretation of the vector $\boldsymbol{\beta}$ is the price elasticities of $\mathbf{P}_t$. We estimate $\boldsymbol{\alpha}$ under the assumption that $\Phi \neq 0$. As discussed in Section 3.1, the weekday effects have changed considerably between 2005 and 2012. To model this phenomenon while keeping the model as parsimonious as possible, we include different dummy variables for weekdays before and after January 1$^{\text{st}}$, 2009. This should filter out the effects of less liberalized and interconnected markets, as well as the lower percentage of natural gas in the electricity mix, in the years before 2009. The VECM is therefore:

$$\Delta\bar{\mathbf{X}}_t = \boldsymbol{\alpha}(\boldsymbol{\beta}^\intercal\bar{\mathbf{X}}_{t-1} - \Phi) + \sum_{i=1}^{l}\mathbf{P}_i\Delta\bar{\mathbf{X}}_{t-i} + \mathbf{\Psi}\mathbf{\Theta}_t + \mathbf{\Upsilon}_t \tag{3}$$

where l is not necessarily equal to k, but chosen to minimize an information criterion while removing significant autocorrelation in the residuals $\mathbf{\Upsilon}_t$. The dummy variables for weekdays are contained in the vector $\mathbf{\Theta}_t$, and their associated coefficients in the matrix $\mathbf{\Psi}$. Guided by the Hannan Quinn criterion, we choose $l = 2$ and estimate Eq. (3). The estimation was performed in the JMuLTi software (Lütkepohl and Krätzig, 2004). The coefficients $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\mathbf{P}$ that are significant at the 5% level are shown in Table 8, while insignificant parameters have been set to zero. The full list of parameters, weekday coefficients and corresponding t-values can be found in Appendix A.

Table 8: Significant coefficients in the VECM of Eq. (3). The model has significant coefficients in the two first lags, and one co-integrating vector $\boldsymbol{\beta}$.

Coefficient	$\Delta\bar{x}_t^{el}$	$\Delta\bar{x}_t^{g}$	$\Delta\bar{x}_t^{c}$
$\boldsymbol{\alpha}^\intercal$	-0.089	0.035	0.000
$\boldsymbol{\beta}$	1.000	-0.805	-0.307
Φ	0.842	0.842	0.842
$\Delta\bar{x}_{t-1}^{el}$	-0.200	0.000	0.000
$\Delta\bar{x}_{t-1}^{g}$	0.085	-0.054	0.000
$\Delta\bar{x}_{t-1}^{c}$	0.000	0.249	0.072
$\Delta\bar{x}_{t-2}^{el}$	-0.131	0.000	0.000
$\Delta\bar{x}_{t-2}^{g}$	0.000	-0.086	0.000
$\Delta\bar{x}_{t-2}^{c}$	0.000	0.000	0.000

In Table 8, rows 4, 5 and 6 constitute $\mathbf{P}_1$ and rows 7, 8 and 9 constitute $\mathbf{P}_2$. Each column correponds to the effect on a different variable. We see that the coefficients have reasonable signs, e.g., the correction to $\Delta\bar{x}_t^{el}$ is negative if $\bar{x}_{t-1}^{el} - 0.805\bar{x}_{t-1}^{g} - 0.307\bar{x}_{t-1}^{c} - 0.842 > 0$. It is also evident from the table that carbon prices are not affected by deviations from the long term equilibrium (its $\boldsymbol{\alpha}$ coefficient is zero), even though it has a significant effect on both electricity and gas prices (its $\boldsymbol{\beta}$ coefficient is not zero). This is in line with the results of Bunn and Fezzi (2009).

3.6 Residual analysis

Examining the autocorrelation function of the residuals $\boldsymbol{\Upsilon}_t$, shown in Figure 11, we see that there is very little autocorrelation left.

However, a multivariate ARCH-LM test (Engle, 1982) with one lag on the residuals has a test statistic of 112.66. This means that the null hypothesis of no multivariate ARCH effects is rejected, even at the 1% level. We get the same result when performing univariate ARCH-LM tests on each residual υ_t. We therefore need to model the heteroscedasticity, and look for candidates that can incorporate multivariate effects.

3.6.1 Stochastic volatility modelling

To model the ARCH effects, we estimate five different multivariate GARCH(1,1) models and select one by three different information criteria. All models assume that $\boldsymbol{\Upsilon}_t = \mathbf{H}_t^{1/2}\mathbf{Z}_t$ where $\mathbf{H}_t \in \mathbb{R}^{d\times d}$ is a conditional covariance matrix, and $\mathbf{Z}_t \in \mathbb{R}^d$ is a vector of i.i.d. errors. The estimation is performed in a Quasi Maximum Likelihood (QML) setting, where we assume a multivariate Gaussian error distribution for $\mathbf{Z}_t$. Bollerslev and Wooldridge (1992) show that QML estimates are consistent, even if the assumption of normality in the i.i.d. errors is violated. As we shall see below, a better fit is indeed obtained with a more flexible distribution, known as the multivariate normal-inverse Gaussian (MNIG) distribution.

The two first multivariate GARCH models are the Scalar BEKK(1,1) and Diagonal BEKK(1,1) models (Baba, Engle, Kraft and Kroner, 1991), in which

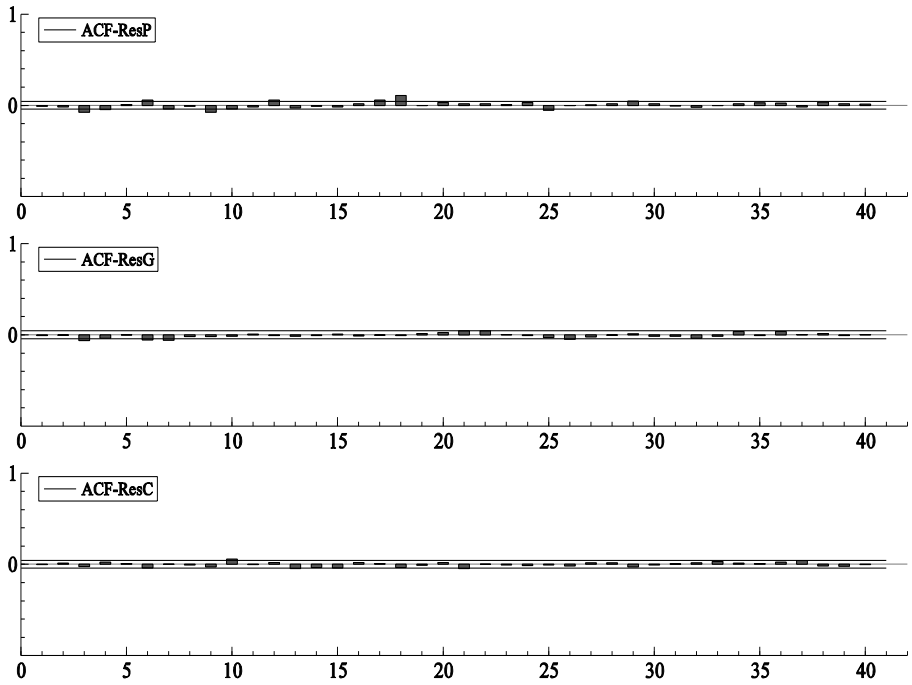

Figure 11: The autocorrelation functions of the residuals of the VECM model. The graph shows the first 42 lags of power, gas and carbon prices from the top to the bottom respectively.

the conditional covariance matrix is assumed to follow the process:

$$\mathbf{H}_t = \mathbf{C}^\intercal \mathbf{C} + \mathbf{A}^\intercal \boldsymbol{\Upsilon}_{t-1} \boldsymbol{\Upsilon}_{t-1}^\intercal \mathbf{A} + \mathbf{B}^\intercal \mathbf{H}_{t-1} \mathbf{B} \tag{4}$$

In the Scalar BEKK model, $\mathbf{A}$ and $\mathbf{B}$ are scalars, while in the Diagonal BEKK, they are diagonal matrices of coefficients. For more details on these models, see (Baba, Engle, Kraft and Kroner, 1991).

The third and fourth models are the Dynamic Conditional Correlation model specifications of Engle (2002) and Tse and Tsui (2002), respectively, while the fifth model is the Constant Conditional Correlation model. The last three models assume that the conditional covariance $\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t$, where the elements in $\mathbf{D}_t = diag(h_{11}^{1/2}, h_{22}^{1/2}, \cdots, h_{dd}^{1/2})$ are univariate GARCH models, and $\mathbf{R}_t$ is a conditional correlation matrix. While model three and four assume two different, lag-dependent specifications for the matrix $\mathbf{R}_t$, model five restrics $\mathbf{R}_t$ to be constant. We estimate the models in OxMetrics, but only on a subset of the data, from January 1$^{\text{st}}$, 2009 and onwards. This is because of the structural changes in the markets in 2007 and 2008, as described in Sections 3.1 and 3.2. We believe an estimation on this subset is more likely to reflect the future volatility behavior than the entire data set. The logarithm of the QML, number of parameters, Schwarz Criterion, Hannan-Quinn criterion and Akaike Information Criterion of the models are reported in Table 9.

Table 9: Log Quasi Maximum Likelihood and information criteria for the MGARCH model candidates. The model with best fit is highlighted in bold figures.

Model	$log(QML)$	Parameters	SC	HQ	AIC
Scalar BEKK	6129.18	9	-12.099	-12.127	-12.143
Diagonal BEKK	6148.47	13	-12.110	-12.149	-12.174
DCC, Engle	6163.87	6	-12.189	-12.207	-12.218
DCC, Tse and Tsui	6163.96	6	-12.189	-12.207	**-12.218**
CCC	6160.30	4	**-12.195**	**-12.207**	-12.215

Considering the results in Table 9, it is not quite clear which model to adopt. The smallest information criterion is highlighted, and it seems that the DCC or the CCC models provide the best fit. By principle of parsimony, we proceed with the CCC model, thereby assuming the following:

$$\boldsymbol{\Upsilon}_t = \mathbf{H}_t^{1/2}\mathbf{Z}_t \tag{5a}$$

$$\mathbf{H}_t = \mathbf{D}_t\mathbf{R}\mathbf{D}_t, \quad \mathbf{D}_t = diag(h_{11}^{1/2}, h_{22}^{1/2}, \cdots, h_{dd}^{1/2}) \tag{5b}$$

$$h_{ii,t} = \omega_i + \alpha_{i,garch}\upsilon_{i,t}^2 + \beta_{i,garch}h_{ii,t-1}, \quad i \in \{1, \ldots, d\} \tag{5c}$$

$$\mathbf{Z}_t \sim i.i.d(\mathbf{0}, \mathbf{I}) \tag{5d}$$

Note that adopting the CCC model is not the same as assuming that the observed correlation is constant - it only means that we restrict the coefficients in the matrix $\mathbf{R}$ to be constants. As the univariate volatility terms h_{ii} vary individually, a simulation of the residuals do indeed produce time varying correlation. The estimated parameters in the CCC model are shown in Table 10.

Table 10: Estimated parameters of the CCC-GARCH model in equations (5a)-(5d).

	Coefficient	Value	p-value
Power	ω_1	1.46e-4	0.011
	$\alpha_{1,garch}$	0.126	0
	$\beta_{1,garch}$	0.795	0
Gas	ω_2	8.42e-5	0.008
	$\alpha_{2,garch}$	0.180	0
	$\beta_{2,garch}$	0.789	0
Carbon	ω_3	3.68e-6	0.201
	$\alpha_{3,garch}$	0.082	0
	$\beta_{3,garch}$	0.919	0
$\mathbf{R}$ matrix	$R_{2,1}$	0.280	0
	$R_{3,1}$	-0.045	0.137
	$R_{3,2}$	-0.128	0

The standardized residuals $\mathbf{Z}$, are shown in Figure 12. Normality tests confirm what is obvious from this figure: the null hypothesis of a Gaussian error distribution is rejected at any reasonable significance level. The univariate Jarque-Bera statistics are 2609, 3146 and 271 for power, gas and carbon respectively, while the

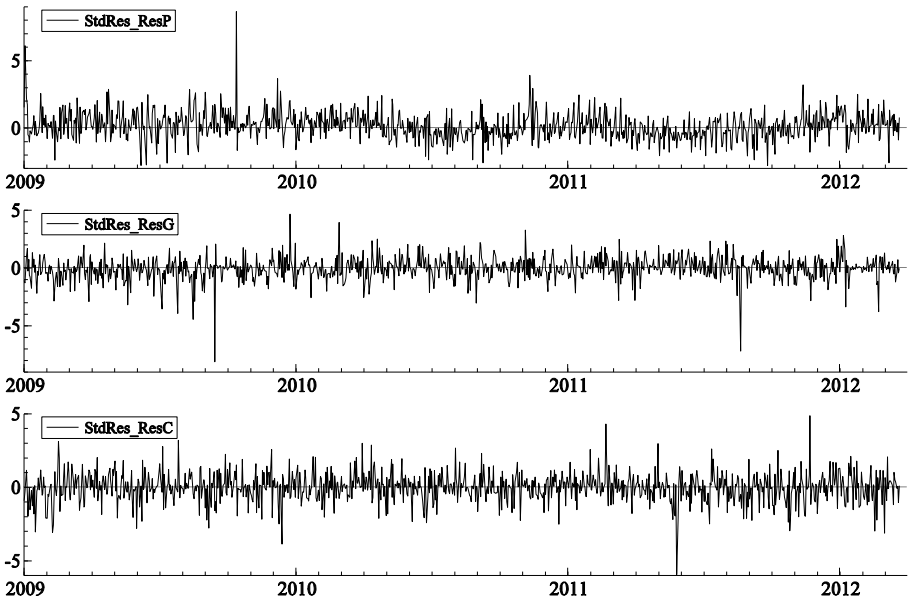

Figure 12: Standardized residuals $\mathbf{Z}$ from the Constant Conditional Correlation model from 2009-2012. High kurtosis indicates a non-normal distribution.

vector normality test statistic is 943. We therefore need to look for other distribution candidates to model the standardized residuals, and we consider here the multivariate normal-inverse Gaussian distribution (MNIG) and the multivariate skewed t distribution, which both have been used to model semi heavy-tailed distributions in financial applications before (see for example Aas, Haff and Dimakos (2006), Benth and Henriksen (2011) and Andresen et al. (2010)).

3.6.2 The Multivariate Normal-Inverse Gaussian distribution

The univariate NIG distribution, and indications of its multivariate form, was described by Barndorff-Nielsen (1977). Let $\mathbf{Z}$ be MNIG distributed with parameter vectors $\{\boldsymbol{\mu}, \boldsymbol{\gamma}\} \in \mathbb{R}^d$, a dispersion matrix $\boldsymbol{\Gamma} \in \mathbb{R}^{d \times d}$ and scalar parameters χ and δ, such that $\mathbf{Z}_t \sim MNIG(\boldsymbol{\mu}, \delta, \boldsymbol{\Gamma}, \chi, \boldsymbol{\gamma})$. Note that not all parameter symbols are consistent with what is common in the literature, for examle χ is commonly denoted α, but we reserve α for its use in the co-integration model to avoid confusion. The probability density function $f_Z(\mathbf{Z})$ may then be written as:

$$f_Z(\mathbf{Z}) = \frac{\delta}{2^{\frac{d-1}{2}}} \left(\frac{\chi}{\pi q(\mathbf{Z})}\right)^{\frac{d+1}{2}} e^{p(\mathbf{Z})} K_{\frac{d+1}{2}}(\chi q(\mathbf{Z})) \tag{6a}$$

$$q(\mathbf{Z}) = \sqrt{\delta^2 + (\mathbf{Z} - \boldsymbol{\mu})^\intercal \Gamma^{-1} (\mathbf{Z} - \boldsymbol{\mu})} \tag{6b}$$

$$p(\mathbf{Z}) = \delta \sqrt{\chi^2 - \boldsymbol{\gamma}^\intercal \boldsymbol{\Gamma} \boldsymbol{\gamma}} + \boldsymbol{\gamma}(\mathbf{Z} - \boldsymbol{\mu}) \tag{6c}$$

where $K_i(\cdot)$ is the modified Bessel function of second kind with index i. The vector $\boldsymbol{\mu}$ is a location parameter, δ is a scale parameter, χ indicates the steepness of the distribution and the vector $\boldsymbol{\gamma}$ indicates the skewness in d dimensions. Even though it is possible to estimate the parameters via maximum likelihood estimation when $d = 3$, the likelihood surface becomes very flat. The number

of parameters makes the Expectation Maximization algorithm (Dempster *et al.*, 1977) a more tractable method. Øigard and Hanssen (2002) derive the EM algorithm for the MNIG distribution, and Aas, Haff and Dimakos (2006) modify it to fit the restrictions that $\mathbb{E}(\mathbf{Z}) = \mathbf{0}$ and $\text{cov}(\mathbf{Z}) = \mathbf{I}$.

An EM algorithm is guaranteed to converge, and when implementing it in Matlab, we find that it converges quickly. With a convergence criterion that the maximum change in any parameter is less than 10^{-6} in two consecutive iterations, it only takes about 50 iterations to find stable parameter estimates with around three years of daily observations. The EM algorithm for the MNIG distribution and the parameter estimates are shown in Appendix B.1.

3.6.3 The multivariate skewed t distribution

There are several version and parametrizations of the skewed t distribution. It is in the family of generalized hyperbolic distributions, and the parametrization we use is described by e.g. McNeil et al. (2005). It has parameters to incorporate both skewness in all dimensions and tail heaviness. We give the probability density function and its estimation procedure in Appendix B.2, where we also derive the equations necessary for estimating the parameters of its standardized version.

3.6.4 Selecting the distribution

We fit the standardized MNIG and the standardized skewed t distributions to the standardized residuals, and compute their log-likelihoods. We do the same for the multivariate Gaussian and Student's t distributions for comparison. Whether we select by the highest log-likelihood or the lowest information criterion, we come to the conclusion that the MNIG distribution provides the best fit and and the skewed t the second best fit. Note that the multivariate Student's t is a restricted version of the multivariate skewed t distribution, with skewness vector equal to zero, and the multivariate Gaussian is a special case of the MNIG distribution. Based on the results, we assume that the standardized residuals $\mathbf{Z}$ are MNIG distributed for the rest of this thesis. The log-likelihood, number of parameters, Akaike's information criterion and the Bayesian information criterion are summarized in Table 11.

Table 11: Results of fitting standardized multivariate distributions to standardized residuals. The best fit according to each test is outlined with bold figures.

Distribution	Log-Likelihood	Parameters	AIC	BIC
MNIG	**-4326**	17	**8686**	**8782**
Skewed t	-8624	16	17280	17370
Student t	-8803	1	17608	17613
Gaussian	-9044	0	18088	18088

3.7 Risk-free approximation

When discounting the cash flows of our LSM simulation, we can either discount the real cash flows using a risk-adjusted discount rate, or we can adjust the cashflows to a risk-neutral measure and discount these using the risk-free rate. As the volatility of each day's cash flows depends on the level of the spark spread, the risk-adjusted discount rate will differ for each point in the LSM simulation. The calculation of such a discount rate is cumbersome and in the LSM literature most authors use a risk-neutral measure of the simulated cash flows.

To perform a risk-neutral valuation, we implicitly assume that the market is complete and that the cash flows of our power plant can be replicated. This may be a bold assumption, but with the existence of a well-developed spot and futures market for both electricity, gas and carbon, the variance of power plant cash flows can to a large extent be hedged, resulting in a low-volatility cash flow that, assuming no-arbitrage conditions, resembles a risk-free return.

For our price model to be utile in a risk-free valuation setting, we must simulate the prices so that their expected return equals the risk-free interest rate ρ. Benth and Henriksen (2011) show that for a number d of MNIG-distributed variables to all have an expected return of ρ, the only adjustment needed is to add a vector $\boldsymbol{\theta}$ to the existing vector $\boldsymbol{\gamma}$ of Eq. (6c), thereby changing the skewness of the distribution. The vector $\boldsymbol{\theta}$ is calculated by numerically solving a system of d equations. In our model, however, the prices are in addition affected by the CCC-GARCH structure and the vector error-correction. With these models added to the residuals, providing an analytical solution to a risk-free distribution is beyond the scope of this thesis. Rather, we exchange $\boldsymbol{\gamma}$ for $\boldsymbol{\gamma}+\boldsymbol{\theta}$ in our MNIG model and use a heuristic to find a $\boldsymbol{\theta}$ that ensures that $\mathbb{E}[ln(p_{t+\Delta t}/p_t)] = \rho\Delta t$. The optimization is computationally heavy. A large amount of simulations per iteration is needed due to the fat tails of the distributions. The surface of the problem is neither convex nor concave, with many local optima. Assuming a continuously compounded risk-free rate of $\rho = 2.5\%$, we find the best approximation to be $\boldsymbol{\theta} = [-0.0010 \quad 0.0019 \quad 0.0002]^\intercal$. The average return of 13 runs with 800 simulations per run is $[0.0245 \quad 0.0254 \quad 0.0254]^\intercal$ with sample standard deviations for the 13 runs of $[0.0055 \quad 0.0056 \quad 0.0075]^\intercal$ for power, gas and carbon, respectively. The returns are taken on a one-year simulation of prices after allowing for 20 days of initial simulations.

To sum up, the price model has the following form:

$$\mathbf{X}_t = \bar{\mathbf{X}}_{t-1} + \mathbf{s}(t) + \Delta\bar{\mathbf{X}}_t + \boldsymbol{\Upsilon}_t \tag{7a}$$

$$\Delta\bar{\mathbf{X}}_t = \boldsymbol{\alpha}(\boldsymbol{\beta}^\intercal\bar{\mathbf{X}}_{t-1} - \Phi) + \sum_{i=1}^{2}\mathbf{P}_i\Delta\bar{\mathbf{X}}_{t-i} + \boldsymbol{\Psi}\boldsymbol{\Theta}_t + \mathbf{H}_t^{1/2}\mathbf{Z}_t \tag{7b}$$

$$\mathbf{H}_t = \mathbf{D}_t\mathbf{R}\mathbf{D}_t, \quad \mathbf{D}_t = diag(h_{11}^{1/2}, h_{22}^{1/2}, \ldots, h_{dd}^{1/2}) \tag{7c}$$

$$h_{ii,t} = \omega_i + \alpha_{i,garch}\upsilon_{i,t}^2 + \beta_{i,garch}h_{ii,t-1}, \quad i \in \{1,\ldots,d\} \tag{7d}$$

$$\mathbf{Z}_t \sim MNIG(\boldsymbol{\mu}, \delta, \boldsymbol{\Gamma}, \chi, \boldsymbol{\gamma}+\boldsymbol{\theta}) \tag{7e}$$

In this chapter we have shown a step-by-step approach to building a credible price model for the three commodities power, gas and carbon allowances.

The model is composed of a seasonal component, a VECM model on the deseasonalized price series and a CCC-GARCH model with MNIG-distributed residuals. The whole model is in essence only driven by three factors, the residuals $\mathbf{Z}_t$ of the MNIG distribution. We might have added other price drivers such as e.g. weather and temperature variables, but as such variables are hard to simulate without an extensive weather model, they have been omitted. During the rest of this thesis, the price model will be used in simulations of prices to estimate the value of a power plant trading in the three commodities.

4 Valuing the combined cycle gas turbine

In this section we describe the simulation-based technique we use to value the CCGT. We use the Least Squares Monte Carlo (LSM) algorithm, proposed by Longstaff and Schwartz (2001). The LSM algorithm is a simulation based dynamic programming technique where the continuation value of an asset is approximated with a least squares regression. During backward induction, the values in the next time period are regressed on a set of problem specific variables in the current time period. This regression approximates the continuation value, implying that we do not assume perfect foresight. As start-up costs and ramp times are present, the cash flows earned in each period depends on the choice of production in the period(s) before. This makes our problem a path-dependent problem, for which the LSM algorithm has proved effectful in related research (see, e.g., Cassano and Sick (2009) or Boogert and de Jong (2008)). The algorithm is also flexible enough to incorporate the operational constraints. We define the state of the power plant as its power output for producing states, and the boiler temperature for non-producing states.

A power plant may pursue different production strategies in terms of hedging and marketing of electricity. For example, a large plant with little operational flexibility may purchase gas and sell power on monthly forward contracts, to decrease exposure to spot price volatility. A smaller, highly flexible plant, could rather trade in the hourly market and ramp down when the CSS becomes negative. In this thesis we assume a plant trading in the day-ahead market when modelling the value of the power plant, but the valuation approach may easiliy be generalized to any time granularity.

The auctions for gas and power day-ahead prices are settled at noon on day t, meaning that at noon on any day, we know the price valid from midnight and the next 24 hours, p_{t+1}. From the day before we also know p_t, the price valid for the rest of the day. As new information arrives at noon, we define this to be the moment where the power plant makes the decision on how much to produce for the next 24 hours. Carbon allowances are not structured as day-ahead contracts, but for simplicity we shall assume that the power plant operator also buys carbon allowances at noon.

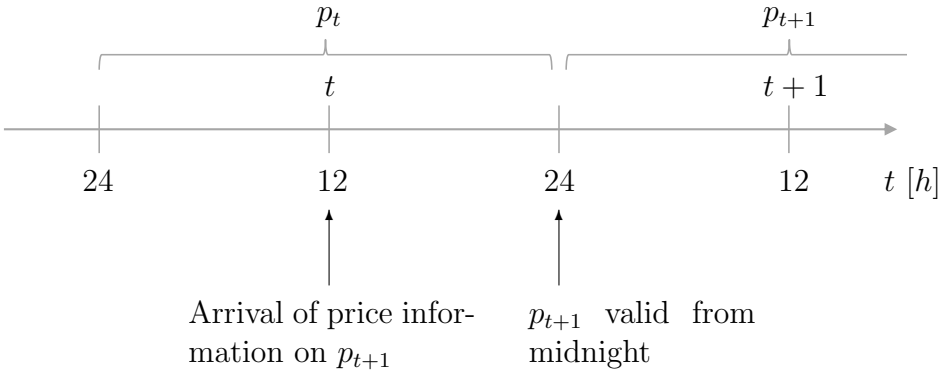

Figure 13: Illustration of decision timing. Price information for the next day arrives at noon. The prices are valid from midnight to midnight.

When planning the production for the next 24-hour period, the power plant operator needs to take the following elements into consideration:

- The power plant state at the arrival of new information (state i)
- The prices of power, gas and carbon for the rest of the day, P_t
- The prices of power, gas and carbon for the next day, P_{t+1}
- His expectations about the optimal state in which to end period t (state j)

All of these elements will affect the decision of whether and when to produce power. Operational constraints also affect the decision: Assume, for example, that the power plant at time t is idle when information about tomorrow's prices arrives. Based on the new information, the highest expected continuation value is achieved by being in the maximum production state at time $t+1$. The prices of electricity, gas and carbon will affect the optimal production profile. When in the idle state, the plant has to consume gas to pre-heat the machinery before reaching the minimum production state. Next, it has to ramp up from minimum to maximum production. If the price of gas is high on day one and lower on day two, it may be optimal to postpone the heating until the next day. However, the CSS_{t+1} might also be so high that it is worthwile to pre-heat during day one so that maximum production is achieved for every hour of day two.

While the value of being in state j at time $t+1$ is unknown, the optimal production between time t and $t+1$, going from state i to j, is a deterministic problem. Because we know p_t and p_{t+1}, it can easily be solved via dynamic programming. The value $V_{i,t}$ of the power plant at time t being in state i, conditioned on choosing the optimal next state j, is expressed via the Bellmann equation:

$$V_{i,t} = \arg\max_j \left\{ r_{i,j,t} + e^{-\rho \Delta t} \mathbb{E}[V_{j,t+1}] \right\} \tag{8}$$

where $r_{i,j,t}$ is the profit from choosing the optimal production profile, and ρ is the risk-free interest rate. We implement the LSM algorithm with a daily granularity, where the expected value of being in state j at time $t+1$ is approximated with an OLS regression, and the optimal production planning between state i and j is decided via deterministic dynamic programming. We first go through the deterministic production planning and then the LSM algorithm. Finally we propose an approach to reduce computational time drastically. This is accomplished by proving that the optimal production profile only depends on three ratios of prices, such that we can pre-solve the dynamic program once for all ratios.

4.1 Deterministic production planning

The function to maximize in the deterministic dynamic program is the profit $r_{i,j,t}$ in Eq. (8):

$$r_{i,j,t} = q^{el}_{i,j,t}(p^{el}_t - c^{var}) + q^{el}_{i,j,t+1}(p^{el}_{t+1} - c^{var}) - q^{g}_{i,j,t}(p^{g}_t + p^{c}_t \cdot I_C) - q^{g}_{i,j,t+1}(p^{g}_{t+1} + p^{c}_{t+1} \cdot I_C) \tag{9}$$

where q^{el}_{ijt}, q^{el}_{ijt+1}, q^{g}_{ijt} and q^{g}_{ijt+1} are the optimal production volumes of power and consumption of gas for the periods before and after midnight. We define p^{el}_t, p^g_t and p^c_t to be the corresponding prices of power, gas and carbon, c^{var} to be the variable cost per MWh of power, and I_C to be the carbon intensity of natural gas as defined in Section 2.4.

The intraday production profile is optimized with an hourly granularity, taking the operational constraints of the power plant into consideration. First, we discretize the state space of the power plant such that all transitions between states take 1 hour. The characteristics of each state are the boiler temperature for non-producing states, and the generator output for producing states. When the total number of states is S, the matrix $\mathbf{L} \in \mathbb{R}^{d\times d}$ defines feasible transitions such that:

$$l_{ij} = \begin{cases} 0 & \text{, if the transition from } i \text{ to } j \text{ is feasible} \\ -1 & \text{, otherwise} \end{cases} \tag{10}$$

The dynamic program we need to solve is

$$\bar{V}_h(i, p^{el}_h, p^g_h, p^c_h) = \arg\max_j \left\{ r_h(i, j, p^{el}_h, p^g_h, p^c_h) + \bar{V}_{h+1}(j) + l_{ij}B \right\} \tag{11}$$

where $\bar{V}_h$ is the value of being in state i at hour h, r_h is the revenue from staying one hour in state i and B is a penalty imposed for infeasible transitions. We also define $r_h = \bar{q}^{el}_h(i,j)p^{el}_h - \bar{q}^g_h(i,j)(p^g_h + I_C p^c_h)$ such that $\bar{q}_h(i,j)$ is the hourly production of electricity or consumption of gas when going from state i to j. These are predefined for each state transition. This problem is deterministic, because $p_h = p_t$ for $h \in [1, \ldots, 12]$ and $p_h = p_{t+1}$ for $h \in [13, \ldots, 24]$, or in other words, both prices and production/consumption for all transitions are known with certainty.

Note that the $r_{i,j,t}$ that we use in the LSM algorithm and Eq. (8) is just the sum of the hourly profits, $r_{i,j,t} = \sum_{h=1}^{24} r_h$, conditioned on a sequence of optimal state transitions. We obtain this sequence, and the resulting daily profit, through traditional dynamic programming. In the LSM algorithm this intraday production scheduling is calculated between a known initial state and end-state, and we therefore impose the penalty B for not ending up in the desired state, as a boundary condition.

4.2 The LSM valuation algorithm

Starting at time $T-1$ and working backwards, the LSM algorithm solves Eq. (8). We need boundary conditions for time $t = T$ and $t = 1$ to start and terminate the algorithm, and we therefore define $V_{i,T} = r_{ii,t}$ for all states i, and that the plant is idle in $t = 1$. The LSM algorithm gives, for each scenario $m \in [1, \ldots, M]$, time period $t \in [1, \ldots, T]$ and state $j \in [1, \ldots, S]$, an expected continuation value of ending time period t in state j. The expected continuation value is based on a least squares regression where the actual values in $t+1$ are regressed on different

variables in t, for example the clean spark spread.

$$\hat{V}_{j,t+1,m} = \sum_{k=1}^{l} b_k f_k(\cdot) \tag{12}$$

where b_k are the coefficient estimates in the regression. The basic functions $f_k(\cdot)$ can have several forms. Longstaff and Schwartz (2001) suggest power functions, Laguerre polynomials and several other functions of the relevant variables, including their cross products. In our setting, the variables are the prices of the commodities and the state of the power plant. The inclusion of the power plant state in the regression can be avoided by structuring the problem as a three-dimensional grid (time × scenarios × states). The expected value is calculated for each state, and will then only depend on the prices. See Boogert and de Jong (2008) for more details on reducing the dimensionality of the regressions.

At time $T-1$ we compute the expected values $\hat{V}_{j,T,m}$ for all states j and in each scenario m. For this, we use the boundary condition and a regression. We proceed to pick the next state j that maximizes $\{r_{i,j,t} + e^{-\rho t}\hat{V}_{j,T,m}\}$ where the profit $r_{i,j,t}$ is calculated as described in Section 4.1. The *actual* values $V_{i,T-1,m}$ resulting from this decision are computed, and these steps are repeated backwards in time until $t = 1$. The Monte Carlo simulated value of the power plant is defined as $\frac{1}{M}\sum_{m=1}^{M} V_{1,1,m}$.

4.3 Improving the computational speed

Solving a deterministic dynamic program for every 24 hour period in all scenarios and for all feasible state transitions is computationally heavy. We rather exploit the fact that the only input we need for solving the valuation problem with a daily granularity is the amount of electricity produced ($q^{el}_{i,j,t}$ and $q^{el}_{i,j,t+1}$) and gas consumed ($q^{g}_{i,j,t}$ and $q^{g}_{i,j,t+1}$) during the 24-hour period. The actual production profile, meaning the specific choice in each hour, is irrelevant for the daily profit calculation. We also use the result that the optimal intraday production profiles only depend on three ratios of prices rather than the actual prices[2], namely:

$$L^*_{M,t} = \frac{p^{el}_t - c^{var}}{p^g_t + I_C \cdot p^c_t} \tag{13}$$

$$L^*_{M,t+1} = \frac{p^{el}_{t+1} - c^{var}}{p^g_{t+1} + I_C \cdot p^c_{t+1}} \tag{14}$$

$$\Delta c_{F,t} = \frac{p^g_{t+1} + I_C \cdot p^c_{t+1}}{p^g_t + I_C \cdot p^c_t} \tag{15}$$

where $L^*_{M,t}$ is the market heat rate when taking the specific plant's variable cost into consideration, and $\Delta c_{F,t}$ is the relative change in fuel costs from time

[2]Much of the related literature make a myopic assumption, ignoring that the price is known before it takes effect. When making this assumption, the optimal production profile only depends on the market heat rate $L^*_{M,t}$. To see this, assume that price information arrives at midnight and that the 24-hour period has only one price. In this setting, $L^*_{M,t} = L^*_{M,t+1}$ and $\Delta c_{F,t} = 1$.

t to $t+1$. A proof of this is given in Appendix D. We can therefore solve the deterministic production planning program in Section 4.1 for a range of $L^*_{M,t}$'s, $L^*_{M,t+1}$'s and $\Delta c_{F,t}$'s that span their minimum and maximum values, and store the optimal production and consumption values for all start states i and end states j. These values, $q^{el}_{i,j,t}$ and $q^{g}_{i,j,t}$, are the sum of the $\bar{q}_h(i,j)$ conditioned on optimal transition choices. Thus, instead of calculating the optimal production profile for each day and each scenario in the LSM algorithm, we calculate the ratios $L^*_{M,t}$, $L^*_{M,t+1}$ and $\Delta c_{F,t}$ from the prices at any given time, and extract the relevant q's.

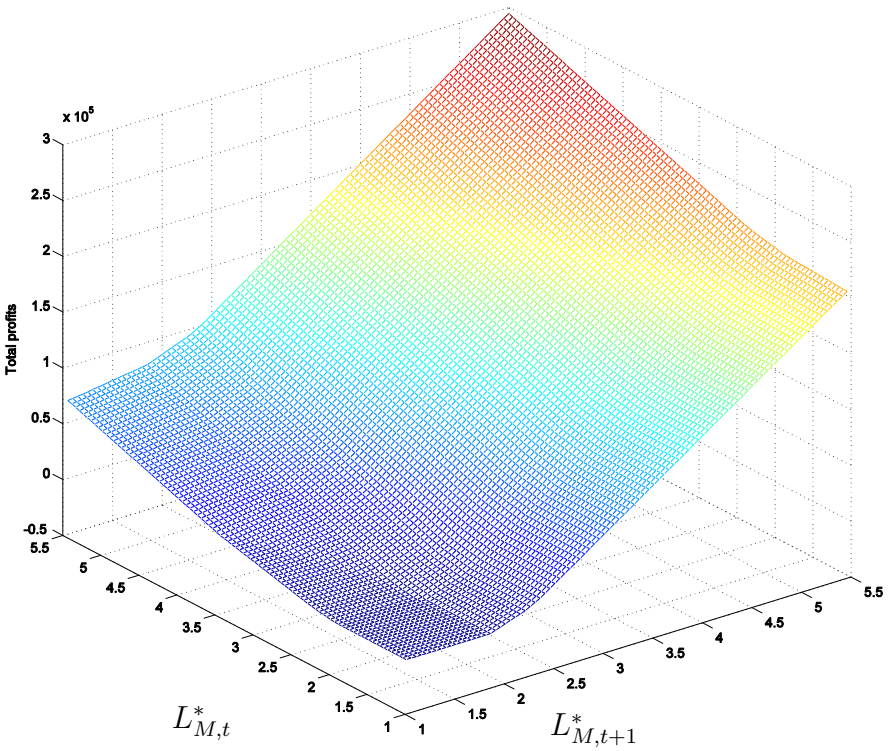

Figure 14: Example of 24 hour profits incurred from optimal intraday production profile. Shown here for $\Delta c_{F,t} = 1$ (fuel cost at $t+1$ equal to fuel cost at t) and fuel costs $p^g_1 + I_C \cdot p^c_1 = 20$.

Because we must make finite matrices of pre-stored qs, we may make a small rounding error when computing the $(L^*_{M,t}, L^*_{M,t+1}, \Delta c_{F,t})$ in simulation and rounding the numbers down to the nearest pre-defined ratio. However, at the threshold between two (or more) production profiles, the *profit* will be the same regardless of profile. Figure 15 shows the total production of power and consumption of gas. We see that the total production of power and gas may change dramatically for a small increment in e.g. $L^*_{M,t}$, while the accrued profit (Figure 14) is a continuous function.

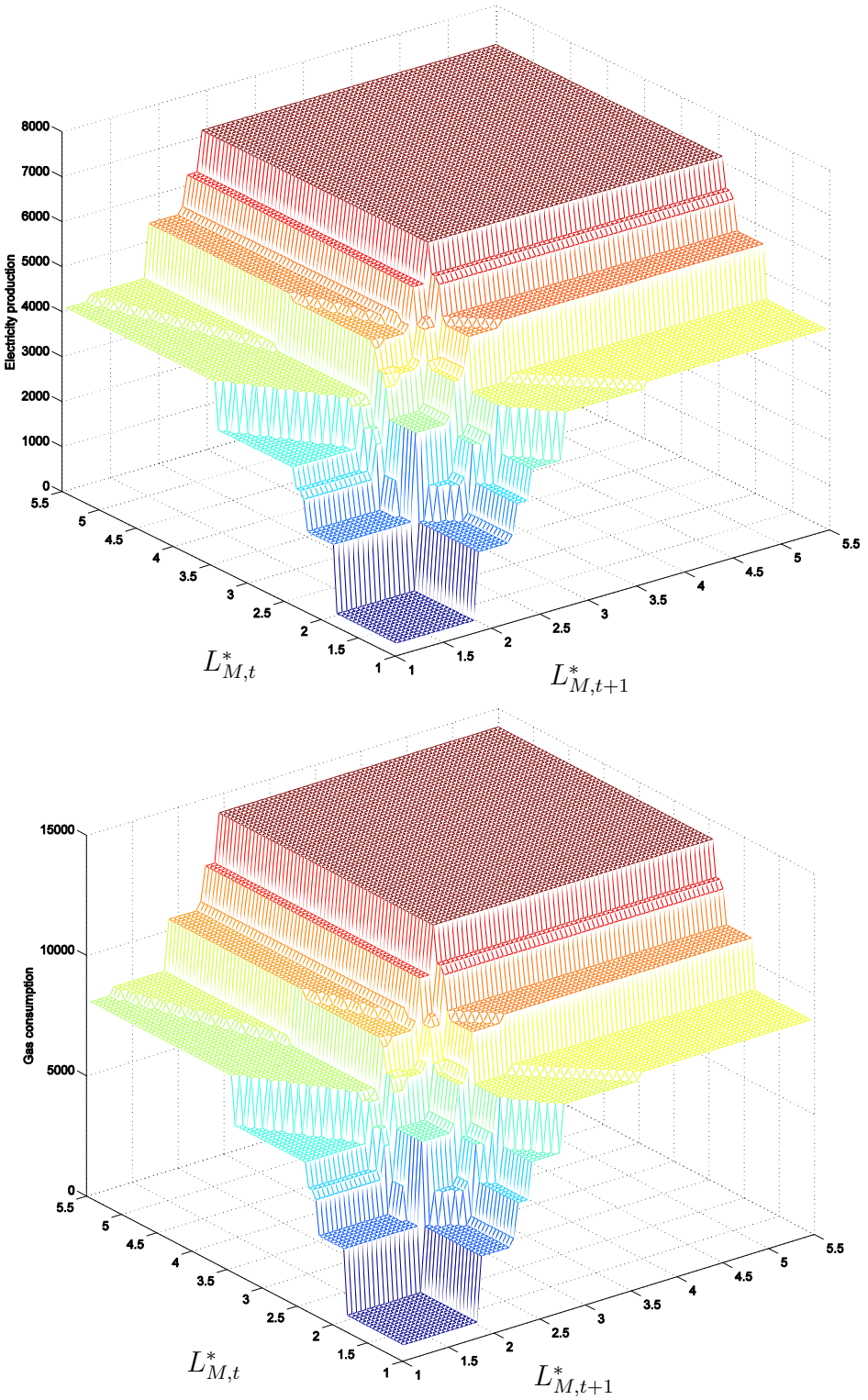

Figure 15: Optimal production of electricity and consumption of gas from deterministic dynamic program. Top: 24-hour production of electricity, $q^{el}_{4,5}(L^*_{M,t}, L^*_{M,t+1})$ for $\Delta c_{F,t} = 1$. Bottom: 24-hour consumption of gas, $q^{g}_{4,5}(L^*_{M,t}, L^*_{M,t+1})$ for $\Delta c_{F,t} = 1$.

5 Results and discussion

5.1 Power plant specifications

We will here value the cash flows of a power plant with realistic specifications, using the methods described in Sections 3 and 4. With the possible plant states defined by its power output and temperature, we discretize the state space to five states; three of which are non-producing and two of which are producing states. We refer to the idle states as the Cold state (idle for a long time), the Warm state (intermediate state in cooling down) and the hot state (recently switched off). The Max state produces at maximum capacity, while the Min state produces at minimum capacity. For an illustration of how we discretize the state-space, see Figure 16. The CCGT has the following operational characteristics[3]:

- Minimum output of 150 MW at 55% efficiency
- Maximum output of 300 MW at 50% efficiency
- Gas used for pre-heating from the cold state: 200 MWh
- Gas used for pre-heating heating from the warm state: 100 MWh
- Gas used for pre-heating heating from the hot state: 50 MWh
- State transition times, as summarized in Figure (17)

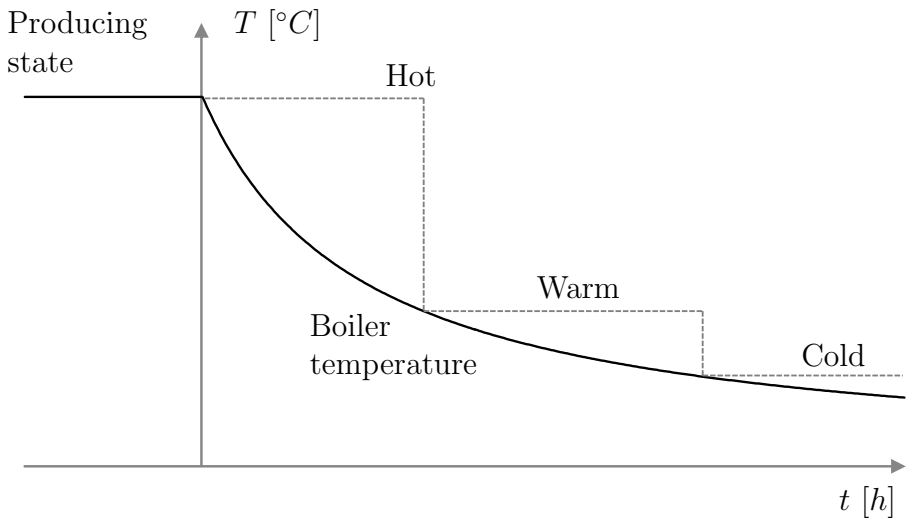

Figure 16: The plant cools continuously down from producing temperature to ambient temperature. In the valuation model, we discretize the states as shown in this graph.

In the intraday production optimization we discretize the transitions further into $S = 23$ states to ensure that all transitions take exactly 1 hour. This implies that between, for example, the warm state and the hot state, we create 11 intermediate states. Cooling down from hot to warm, which takes 12 hours, will

[3]Kindly provided by Centrica, plc.

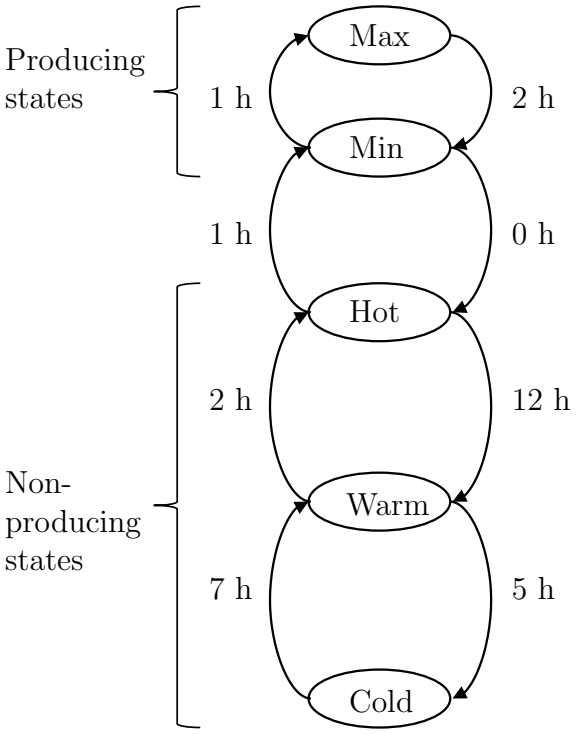

Figure 17: State transition times in the discrete state-space.

be modelled as doing 12 downward jumps from 20 to 19, 19 to 18, et cetera. The opposite transition, from warm to hot, takes only two hours and is therefore be modeled as taking two jumps of six states each.

Transition between maximum production and minimum production, i.e. ramping up or down, are modelled with $q_{ij}^{el} = (q_{ii}^{el} + q_{jj}^{el})/2$ for $i, j \in \{\text{Min,Max}\}$. We assume the lowest plant efficiency while ramping.

We further assume that the plant's production does not influence market prices of neither electricity, gas nor EUAs, and that the plant is only active in the spot market. We use risk-neutral valuation, and assume a risk-free rate of 2.5% p.a.

5.2 Price Simulation

To simulate prices from Eq. (7), we start with drawing the random numbers $\mathbf{Z}_t$ from the risk-neutral MNIG distribution. We use the fact that the MNIG distribution is a normal mean-variance mixture with an inverse-Gaussian mixing distribution (see Appendix B), and sample $\mathbf{Y} \in \mathbb{R}^3$ from $N_3(\mathbf{0}, \mathbf{I})$ and $W \in \mathbb{R}$ from $IG(\delta, \sqrt{\chi^2 - (\boldsymbol{\gamma}^*)^\intercal \boldsymbol{\Gamma} (\boldsymbol{\gamma}^*)}$, such that

$$\mathbf{Z}_t = \boldsymbol{\mu} + W\boldsymbol{\Gamma}\boldsymbol{\gamma}^* + \sqrt{W}\boldsymbol{\Gamma}^{1/2}\mathbf{Y} \sim MNIG(\boldsymbol{\mu}, \delta, \boldsymbol{\Gamma}, \chi, \boldsymbol{\gamma}^*) \tag{16}$$

with $\boldsymbol{\gamma}^* = (\boldsymbol{\gamma} + \boldsymbol{\theta})$ as the risk-neutral adjustment. At each time t we calculate the three univariate GARCH volatilities $h_{ii,t}$ and the conditional covariance matrix $\mathbf{H}_t$. Then the co-integrated mean equation is computed, and finally we add the seasonal component $\mathbf{s}(t)$. Figure 19 shows the distribution of one-year price

simulations, while Table 12 shows the descriptive statistics of the simulated prices along with the descriptive statistics for power, gas and carbon prices during the period used for estimation of the price model.

Table 12: Descriptive statistics for power, gas and carbon prices from 2009-2012 against the descriptive statistics of 1000 one-year simulations. Means and medians are higher as the start prices of the simulations are higher than the mean in the 2009-2012 period.

	Actual prices			Simulated prices		
	Power	Gas	Carbon	Power	Gas	Carbon
Mean	49.13	17.60	12.96	56.96	19.79	13.28
Median	50.27	18.44	13.62	55.24	19.47	13.01
Standard Deviation	9.18	5.67	2.66	11.51	4.58	2.29
Skewness	0.62	-0.19	-0.94	0.49	0.23	0.26
Excess kurtosis	1.93	-0.95	-0.01	0.10	-0.40	-0.23
Minimum	30.81	4.24	6.08	37.59	11.64	9.47
Maximum	103.66	40.15	16.88	90.33	30.80	18.82
Count	1,008	1,008	1,008	365,000	365,000	365,000

We have tested the model specification with information criteria, but to further test its robustness we perform a backtest on the actual data. At each time step, we use the realized prices up untill time t to forecast 10,000 prices for $t+1$. We save the 2.5 and 97.5 percentiles of the 10,000 forecasts, and thereby calculate a 95% confidence interval for tomorrow's price forecast. The real prices and confidence intervals are shown in Figure 18, from January 1st and the next 400 observations. For 1,000 consecutive observations, the confidence intervals are breached less than 5% for the three commodity price forecasts, with an average of 4.6%. As a comparison, we perform the same analysis with a multivariate Gaussian distribution for the i.i.d. errors. The number of breaches averages 5.42% across the three series, with a maximum of 6.25% for the natural gas prices. This demonstrates the model's ability to capture the observed price dynamics, and to a certain degree forecast prices.

5.3 Valuation results

5.3.1 Convergence of values

In this section we present the most important results of our analyses. We first find how many scenarios we must simulate before the value converges in the LSM algorithm. Figure 20 shows that at least 1500 simulations are necessary to provide an accurate estimate of the plant value. Using 6000 simulations of 365 days, we create a reference case that serves as a benchmark for our further analyses. The reference case has an average present value of 29.66 million euros per year, with a standard deviation of 36.5%. For the rest of this section, every presented simulation is performed using 1500 simulations, unless otherwise is stated. One such simulation takes around 17.5 minutes to run in MATLAB R2011a on a 2.66 GHz Intel Core2 Duo computer running Windows XP.

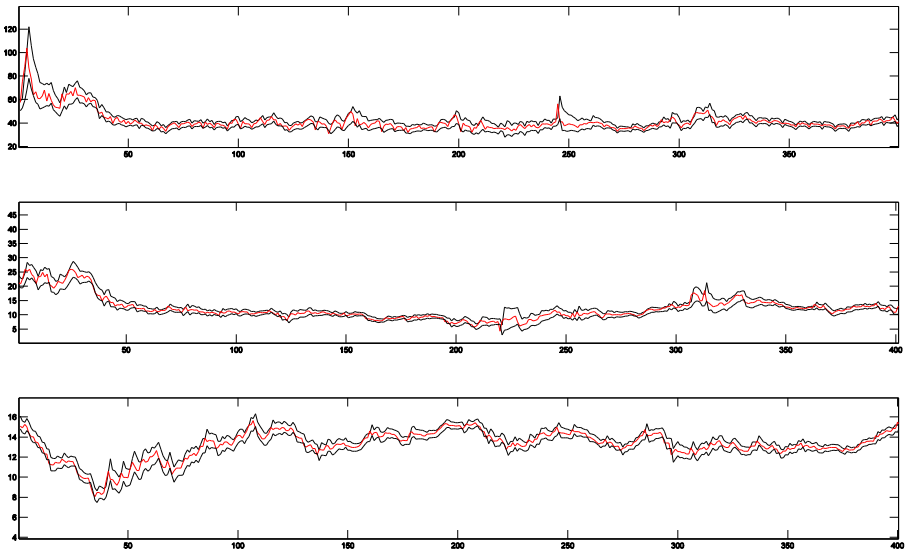

Figure 18: 400 consecutive forecasts from the price model. Actual prices in red, with 95% confidence interval for the forecast in black. Top: Power prices. Middle: Gas prices. Bottom: Carbon prices.

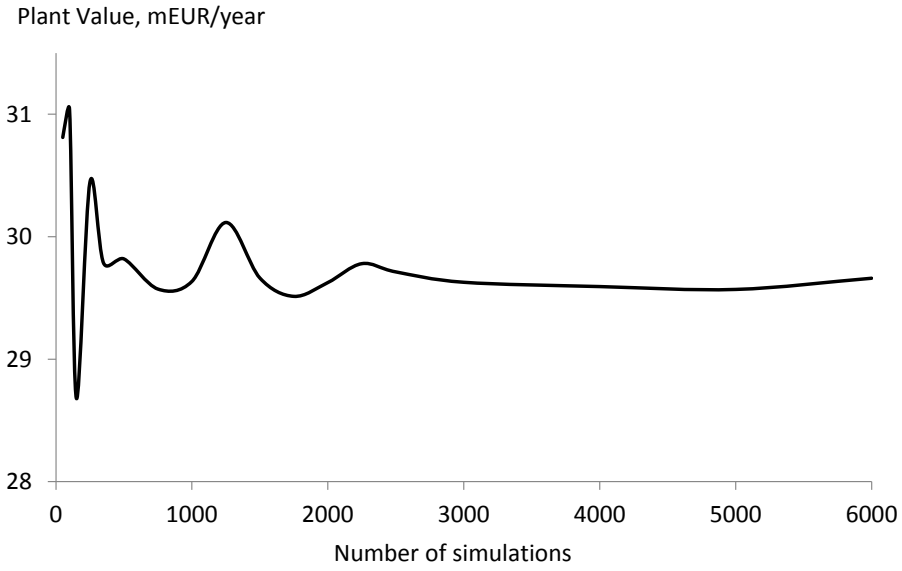

Figure 20: Convergence of LSM simulation. The value estimate converges at around 1500 simulations.

5.3.2 Comparison to a perfect foresight scenario

In the rest of this section, we will frequently compare the obtained value of a simulation with the value obtained assuming perfect foresight. We will refer to the ratio of estimated plant value to the value obtained through perfect foresight as the *relative value*. The perfect foresight simulation is based on the same price series as the tested simulations, but instead of estimating the continuation value of each state using least squares regression, the continuation value of the next state is known with certainty. The reference case obtains a relative value of 99.83% of the perfect foresight value, indicating that few errors are made in the

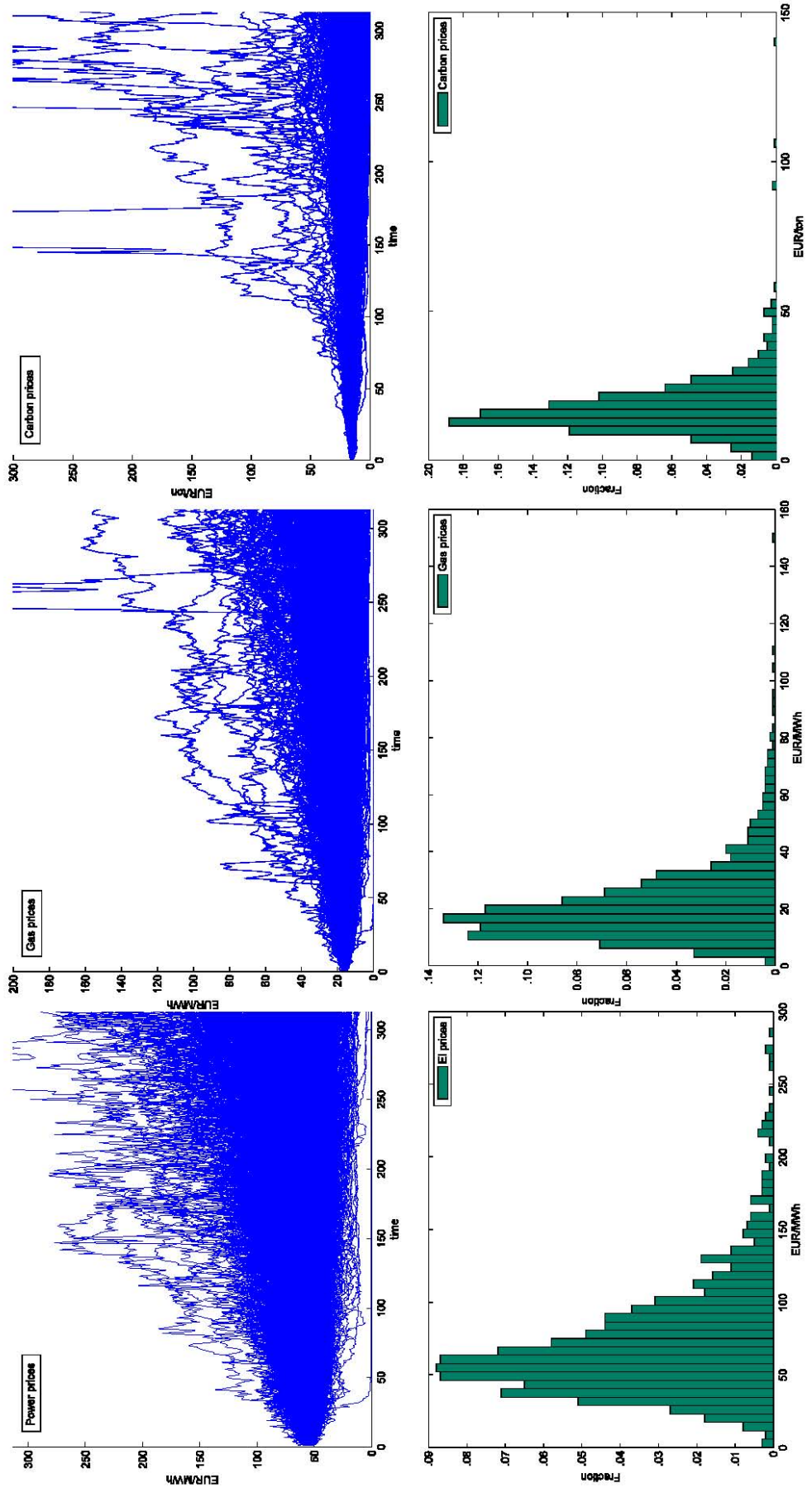

Figure 19: The top row shows 10,000 simulations of 365 days of power, gas and carbon prices. The bottom row shows the distribution of prices at the final date. From left to right, the columns correspond to power, gas and carbon respectively. We simulate 313 observations of six day weeks, that correspond to one year because sunday prices are equal to saturday prices.

state transition choices. The reason that this high relative value can be reached is because the efficiency of our reference case CCGT is so high that it is seldom optimal to operate at any other level than maximum production. In the reference case, the plant is switched off only 1.3 times per year, on average.

Figure 21 shows one run of price simulations with the corresponding power production. The depicted scenario is among the more volatile ones, and is presented in order to show the effects of the market heat rate falling below the specific minimum heat rate of our power plant. As the market heat rate is usually above the heat rate of our power plant, most simulations find it optimal to keep production at the maximum level during the whole simulation period.

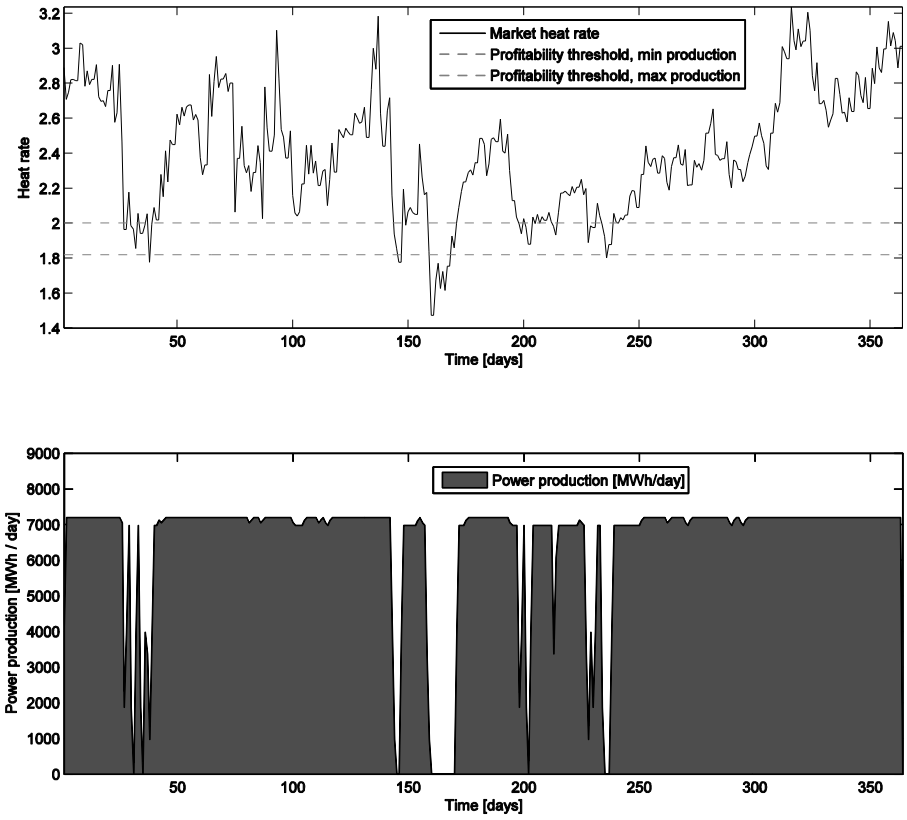

Figure 21: Results from an LSM simulation of 365 days. Top: Prices, converted to the market heat rate L_M. Bottom: The corresponding optimal production of electricity.

5.3.3 Impact of regression form

When choosing the optimal state transition, the LSM algorithm estimates the continuation value for each state by using an OLS regression. The endogenous variable is the set of simulated continuation values at time $t + 1$, while the regressors can be any set of variables based on the information available at time t. The choice of regressors and their functional form is therefore crucial to the value estimate, as it directly affects the state transition choices. While much of the related literature favours the spark spread or market heat rate as the regressor, we argue that using both power, gas and carbon prices provides a more realistic value estimate. As all three prices would be available to a plant manager, this

additional information should be used in the decision rule. Below we show that the multivariate regression indeed outperforms the univariate counterparts when benchmarked against a perfect foresight scenario.

Longstaff and Schwartz (2001) suggest several different polynomials as the basic functions in the LSM regressions. We consider here the power function and Laguerre polynomial of different orders, as they are popular specifications in the LSM literature. The purpose of this analysis is to examine whether our co-integrated model where all three prices are available is advantageous to a model that simulates the spark spread directly. We compare the power function and the Laguerre polynomial of the spark spread, to a multivariate regression where both electricity prices and fuel cost (the sum of gas and carbon costs) enter as regressors. The latter regression has the form:

$$\begin{aligned}\hat{V}_{j,t+1,m} = b_0 + b_1 p^{el}_{tm} + b_2(p^{g}_{tm} + I_C \cdot p^{c}_{tm}) + b_3(p^{el}_{tm})^2 \\ + b_4(p^{g}_{tm} + I_C \cdot p^{c}_{tm})^2 + b_5 p^{el}_{tm}(p^{g}_{t} + I_C \cdot p^{c}_{tm})\end{aligned} \tag{17}$$

Figures 22, 23 and 24 show the fit obtained by power functions, Laguerre polynomials and two-variable regressions, respectively.

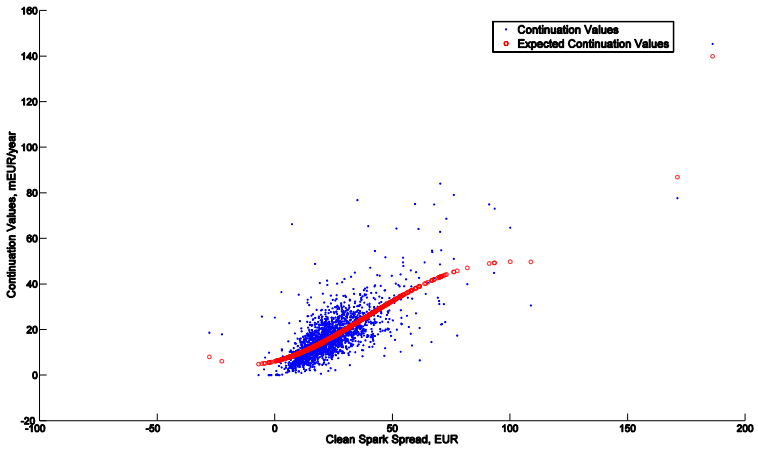

Figure 22: Actual (blue) and fitted (red) continuation values, based on a 6th order power function regression.

To test the impact of the regression form, we run the valuation algorithm with 1st to 5th order power functions and Laguerre polynomials, and compare the resulting relative value to the relative value obtained with Eq. (17). The results are shown in Figure 25.

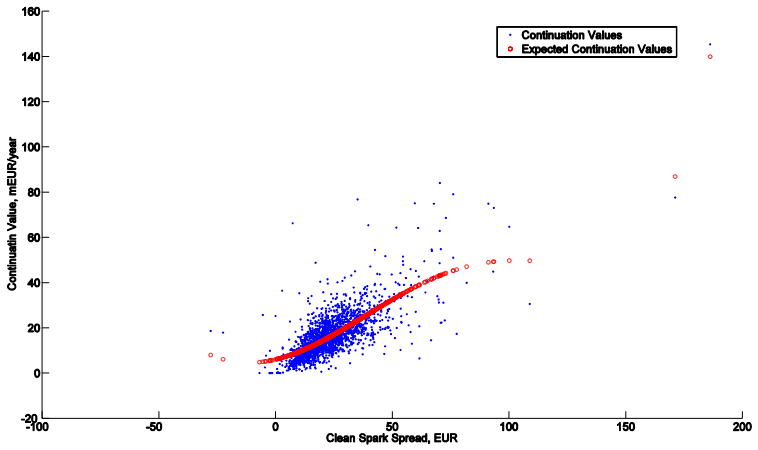

Figure 23: Actual (blue) and fitted (red) continuation values, based on a regression on a Laguerre polynomial of order 6

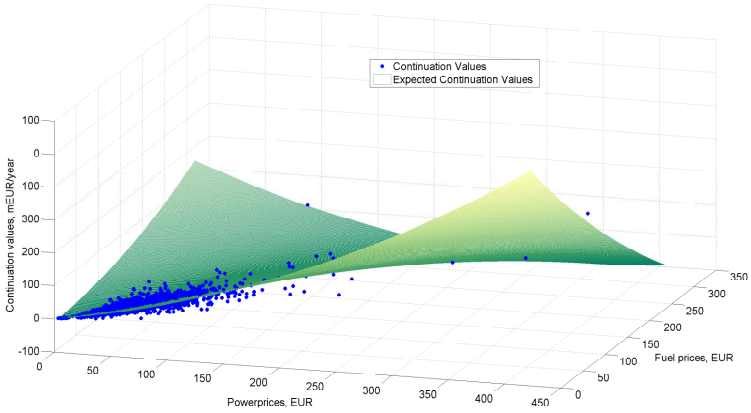

Figure 24: Actual (blue) and fitted (green surface) continuation values, based on the two-dimensional regression in Eq. (17)

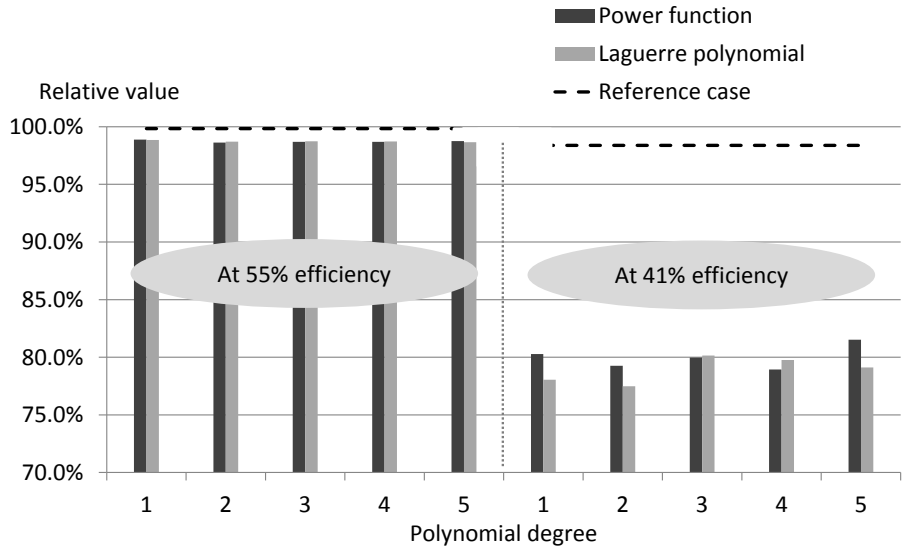

Figure 25: At the reference case efficiency, the univariate regressions estimate a relative value that is 1% lower than the multivariate regression. At an efficiency of 41% this effect is even more pronounced: The univariate regressions estimate the relative value to be 81.5%, while the multivariate regressions estimate 98.4%.

From Figure 25 it seems clear that using univariate regressions consistently underestimates plant value. This effect increases as the plant's efficiency is lowered. We also estimate the value assuming 41% efficiency. Clearly, additional information is contained in the full set of prices that consitutes a spread, compared to the spread itself. As the three prices are available to a decision maker, we argue that the multivariate regression provides a more accurate and realistic estimate of the power plant value than univariate power functions and Laguerre polynomials. This further implies that simulating the clean spark spread or market heat rate alone, as a stationary process, may underestimate the value of the CCGT. This result can be generalized to LSM valuation of other derivatives on several assets, such as crack spread options, rainbow options or basket options. For the rest of the thesis, we use Eq. (17) in the LSM algorithm unless otherwise is stated.

5.3.4 Is perfect foresight a valid assumption?

Perfect foresight is a commonly used comparison to evaluate an optimization technique, and it is occasionally assumed when valuing some producer with uncertain cash flows. Our results show that under the price conditions we consider, the perfect foresight assumption is a good assumption for base load power plants where turning the plant off is seldom or never optimal. When the plant's efficiency is close to the marginal producer's efficiency (whos contribution margin is barely positive), this assumption is rejected because the number of risky decisions on whether to switch on or off the plant increases. Figure 26 shows how the number of turn-offs per year and the relative value with different efficiencies for the CCGT. It seems clear that the estimation error is higher when the number of switch-offs is highest, and we see that at 41% efficiency, which is not uncommon

for single-cycle gas turbines, the perfect foresight assumption over-estimates the value by 2%. We argue that the perfect foresight assumption is neither wrong nor right, but its validity depends on the efficiency of the plant compared to current market conditions.

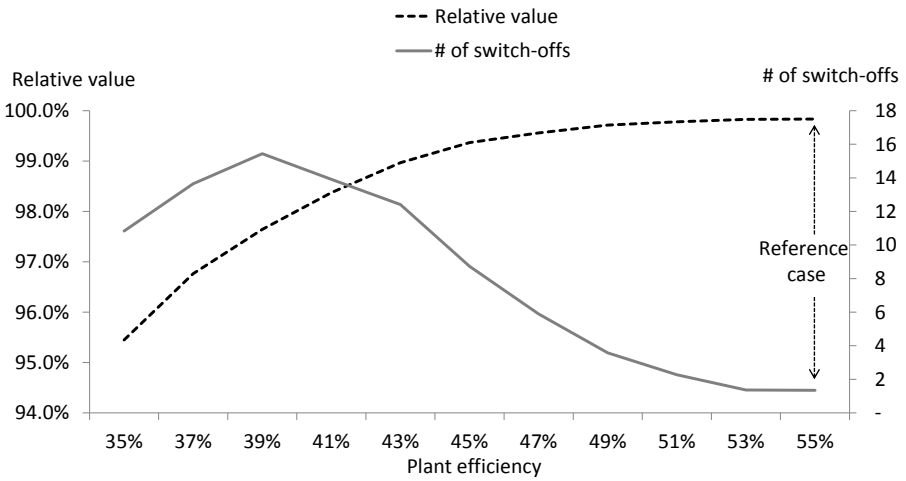

Figure 26: Simulation results using different levels of power plant efficiency. Relative values are shown on the left axis and average number of switch-offs per year on the right. The X-axis efficiency denotes the maximum plant efficiency, with the minimum efficiency being 5 percentage points lower in all scenarios.

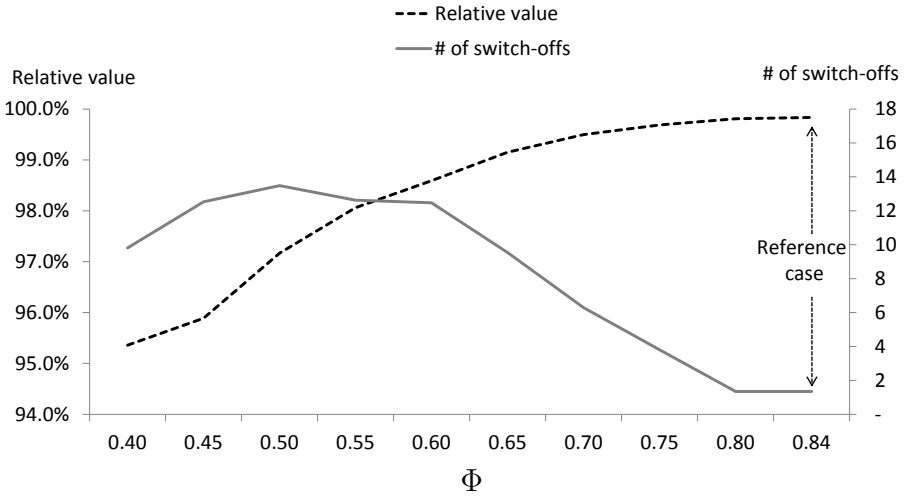

Figure 27: Simulation results with varying Φ. Φ is the equilibrium in the VECM model, and a lower Φ indicates a lower average CSS. Relative value is shown on the left axis and average number of switch-offs per year on the right. The parameter Φ is the constant in the price equilibrium, and is therefore related to the expected clean spark spread.

5.3.5 Future developments in the spark spread

As around 60% of today's European power generation was built more than 20 years ago, it is natural to assume that these will be replaced by newer, more

efficient plants in the future. One result of this may be that the marginal producer has a higher efficiency, thereby lowering the CSS and the market heat rate, given that the marginal producer is still a gas-fired power plant. Our reference case CCGT has a high efficiency, but as the average efficiency of the thermal power generation park increases, it will shift our plant further to the right end of the merit order curve. To analyse the implications of such a development, and to test our model's strength in an environment with when the CSS falls below zero more often, we analyse how the plant value and the relative value is affected by the parameter Φ in the price model. This parameter is the constant in the co-integrating vector, and represents an equilibrium level of the log prices. Recall that the average spark spread is closely related to the value of e^{Φ} and that $\mathbb{E}[CSS] = \vartheta_0 + \vartheta_1 e^{\Phi}$ when fuel cost is constant. Figure 27 shows how the number of swhitch-offs and the relative value varies with Φ. Notice its similarity with Figure 26. The effect of a lower Φ is much the same as assuming a lower efficiency of the plant. The result of either is that the plant efficiency moves closer to the average market efficiency, lowering the plant's achieved spark spread.

To test the effect of a changing Φ on the plant value, the regression $V_0 = \varrho_1 + \varrho_2 e^{\Phi} + \varrho_2 e^{2\Phi}$ expresses the estimated present value as a function of Φ. The regression has an R^2 of 99.94% for $\Phi \in [0, 0.84]$, meaning that the value of a power plant to a large degree can be expressed from an expectation of Φ. A value estimate based on this regression along with our simulated value estimate is shown in Figure 28.

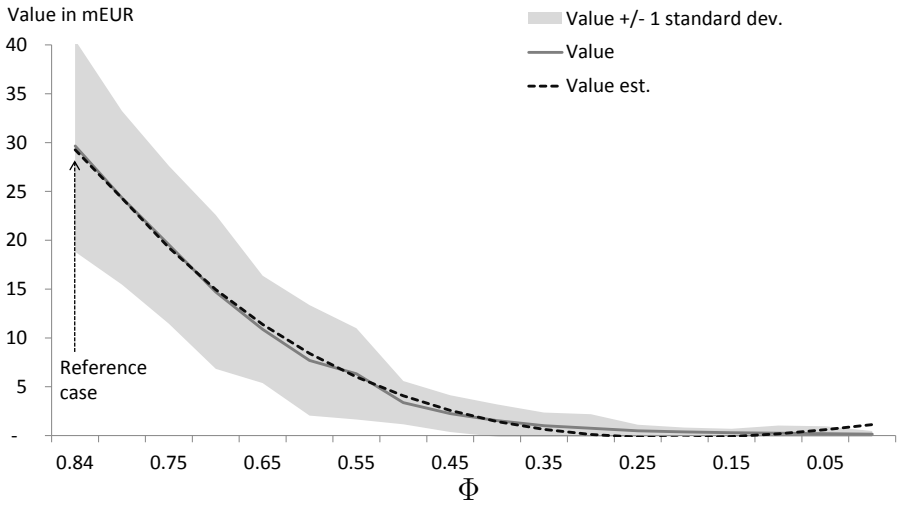

Figure 28: The value of the plant can be estimated with small errors using a polynomial on e^{Φ}. The grey line shows the value obtained by the LSM simulations. The dotted line shows the estimate based on the regression. The grey fan indicates the standard deviation of the LSM simulations. $R^2 = 99.7\%$.

5.3.6 The effect of a myopic assumption on price information

Our model takes into account the knowledge of prices 12 hours in advance of their arrival. Making a myopic assumption means that one assumes that price information arrives at the same time as prices take effect. We will refer to our

assumption of 12 hour knowledge of tomorrow's price as a day-ahead assumption. To measure the effect of the myopic assumption, we compare our valuation results to a simulation based on a myopic assumption. Whenever new prices arrive, the plant manager decides whether to switch states or not. If he decides to switch, the order to switch is given immediately, meaning that the ramp-up or ramp-down is always started in the first hour of the day. The resulting relative values are shown in Figure 29.

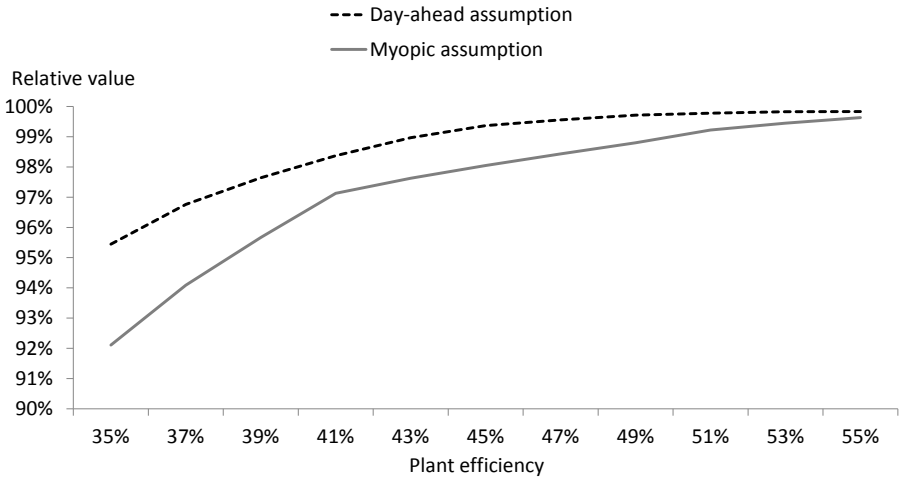

Figure 29: Value estimates using the myopic assumption vs. the day-ahead assumption for different plant efficiencies. Modelling the power plant on a myopic assumption leads to an undervaluation that increases with decreasing efficiencies.

The graph shows that ignoring the flexibility in intraday production scheduling, and assuming a simple model with ony daily decision points, under-estimates the plant value. At high efficiencies and present price conditions, the effect is very small because the plant would always be running. However, at an efficiency 10 percentage points lower than the reference case, the simplified model underestimates the value by 1.5%, and at 20 percentage points lower efficiency it underestimates by 3.5%. As the option to alter production at any hour is present at a real power plant, we conclude that over-simplifying the optimization algorithm can impair the accuracy of the value estimate quite significantly.

As a final analysis, we simulate the combined effect of both a myopic assumption and a 4^{th} degree Laguerre polynomial on the CSS, at the reference case efficiency. The two effects alone yield relative values of 98.7% and 99.6%, respectively. The simulation of both effects simultaneously yields a relative value of 84.8%. The interesting part here is that the combination provides an underestimation of a larger magnitude than the product of their respective relative values would indicate. See Figure 30. The simulation yielded identical results for 1500, 2000 and 4000 simulations, implying that the low relative value is not due to a higher number of simulations needed for values to converge.

Concluding this section, we find that making the wrong assumptions when valuing an asset whose cash flows originate from a price spread can have a large impact on the estimated value. Our procedure outperforms competing ones by up to 19% of the value obtained with perfect foresight, without imposing unrealistic

assumption about the information availible to the plant operator.

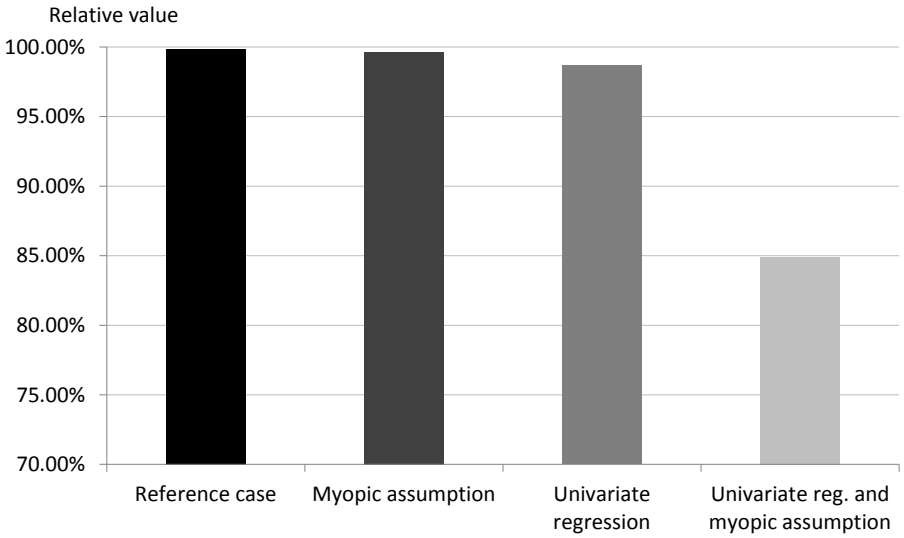

Figure 30: Relative value of the reference case, the zero foresight case, the univariate regression case and the combination of both zero foresight and univariate regression. Notice that the combination of the two yields a much lower relative value than the product of their stand-alone relative values.

6 Conclusion

Combining recent advances in both price modelling, volatility modelling and dynamic programming, we present a complete model for the valuation of a CCGT power plant with special emphasis on realistic modelling of both prices and the plant's operational constraints. We show that power, gas and carbon prices are co-integrated with heteroskedastic volatility, and that a CCC-GARCH model with MNIG distributed residuals captures the price variations in a better way than competing models. The valuation is performed using an LSM simulation where operational constraints are observed. We enhance computational efficiency by showing that the optimal intraday production profile only depends on three ratios of prices, and can be stored as deterministic values rather than being calculated for every scenario.

The modelled 300 MW power plant specified by Centrica plc. is a credible example of a real-life power plant. Our simulations average a first-year discounted cash flow of 29.7 million EUR. The results indicate that the LSM simulation is 99.83% of the perfect foresight value for high efficiencies, while for efficiencies of 30-35%, the relative value falls to around 95%. The relative value tends to fall as the number of switch-offs rises, as there are more occasions in which estimation errors result in a suboptimal choice. According to our simulations, there also exists additional information in the complete set of commodity prices compared to the clean spark spread. A regression on all prices provides a better fit of the continuation values in the LSM simulation than a regression on the spark spread only. This result generalizes to LSM valuations of other spread options such as the crack spread, and multi-asset derivativese such as rainbow options and basket options. Finally, we analyze the implication of a myopic assumption on day-ahead prices. We find that the common, but unrealistic myopic assumption foresight underestimates the value of a power plant. The underestimation is more pronounced in situations of low efficiency and thus multiple switch-offs.

References

Aas, K., Haff, I.H., Dimakos, X.K, 2006. Risk estimation using the multivariate normal inverse Gaussian distribution. The Journal of Risk 8(2), 39-60

Abadie L.M., Chamorro J.M., 2008. Valuing flexibility: The case of an Integrated Gasification Combined Cycle power plant. Energy Economics 30, 1850-1881

Andresen, A., Koekebakker, S., Westgaard, S., 2010. Modeling electricity forward prices using the multivariate normal inverse Gaussian distribution. The Journal of Energy Markets 3 (3), 3-25

Bermejo-Aparicio, R., Moreno, M., Villaplana, P., 2008. European natural gas spot markets: Volatility transmission and jumps modelling. Available at `https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=forofinanzas2008&paper_id=127`[Accessed 04.12.2012]

Baba, Y., Engle, R.F., Kraft, D.F., Kroner, K.F., 1991. Multivariate simultaneous generalised ARCH. Discussion paper 89-57, University of California, San Diego, Department of Economics

Barndorff-Nielsen, O.E., 1977. Exponentially decreasing distributions for the logarithm of particle size. Proceedings Royal Society London A353, 409–19

Mansanet Bataller, M., Pardo Tornero, Á., Valor, E., 2007. CO2 Prices, Energy and Weather. Energy Journal 28 (3), 73-92

Benth, F.E., Henriksen P.N., 2011. Pricing of Basket Options Using Univariate Normal Inverse Gaussian Approximations. Journal of forecasting 30, 355-376

Benth, F.E., Šaltytė-Benth, J., 2004. The Normal Inverse Gaussian Distribution and Spot Price Modelling in Energy Markets. International Journal of Theoretical and Applied Finance 7 (2), 177–192

Bjerksund, P., Stensland, G., Vagstad, F. Gas Storage Valuation: Price Modelling v. Optimization Methods. Working Paper 2008, Norwegian School of Economics and Business Administration

Bollerslev, T., Wooldridge, J.M., 1992. Quasi–Maximum Likelihood Estimation and Inference in Dynamic Models with Time–varying Covariances. Econometric Reviews 11, 143–172

Boogert, A., de Jong, C, 2008. Gas storage valuation using a Monte Carlo method. Journal of Derivatives 15 (3), 81-91

Bunn, D. W., Fezzi, C., 2009. Structural interactions of European carbon trading and energy prices. Journal of Energy Markets 2 (4), 53-69

Cassano, M., Sick, G., 2009. Valuation of a Spark Spread: an LM6000 Power Plant. European Journal of Finance (forthcoming)

Chan, W. H., Wang, G. H. K., Yang, L., 2009. Weather, inventory and common jump dynamics in natural gas futures and spot markets. Working Paper, Waterloo, Wilfrid Laurier University

Chevallier, J., 2011. Carbon Price Drivers: An Updated Literature Review. International Journal of Applied Logistics (forthcoming)

Chevallier, J., Alberola, E., 2009. Banking and Borrowing in the EU ETS: An Econometric Appraisal of the 2005-2007 Intertemporal Market. International Journal of Energy, Environment and Economics 17 (4), 441-457

Daskalakis, G., Psychoyios, D. and Markellos, R. N., 2009. Modeling CO2 Emission Allowance Prices and Derivatives: Evidence from the European Trading Scheme. Journal of Banking & Finance 33 (7), 1230-1241

de Jong C, 2006.The Nature of Power Spikes: A Regime Switch Approach. Studies in Nonlinear Dynamics and Econometric 10 (3), Article 3

Dempster A.P., Laird, N.M., Rubin, D., 1977. Maximum likelihood from incomplete data using the EM algorithm. Journal Royal Statistical Society B39, 1–38

Deng, S.J., Oren, S.S., 2003. Incorporating Operational Characteristics and Startup Costs in Option-Based Valuation of Power Generation Capacity. Probability in the Engineering and Informational Sciences 17 (2), 155-181

Deng, S.J., Xia, Z., 2006. Pricing and Hedging Electricity Supply Contracts: a Case with Tolling Agreements. International Journal of Information Technology and Decision Making 5 (3), 421-436

Department of Energy and Climate Change. Digest of UK energy statistics: Chapter 4 Natural gas, `http://www.decc.gov.uk/en/content/cms/statistics/energy_stats/source/gas/gas.aspx`[Accessed 03.12.2012]

Department of Energy and Climate Change. Digest of UK energy statistics: Chapter 5 Electricity, `http://www.decc.gov.uk/en/content/cms/statistics/energy_stats/source/electricity/electricity.aspx`[Accessed 03.12.2012]

Engle, R.F., 2002. Dynamic Conditional Correlation - a Simple Class of Multivariate GARCH Models. Journal of Business and Economic Statistics 20, 339-350

Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction: Representation, estimation and testing. Econometrica 55 (2), 251-276

Engle, R.F., 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica 50 (4), 987-1007

Eriksson, A., Forsberg, L., Ghysels, E., 2004. Approximating the Probability Distribution of Functions of Random Variables: A New Approach. Discussion paper CIRANO

Escribano, A., Peña, J.I., Villaplana, P., 2011. Modelling Electricity Prices: International Evidence. Oxford bulletin of economics and statistics 73 (5), 622-650

Fleten, S.-E., Näsäkkälä, E., 2010. Gas-fired power plants: Investment timing, operating flexivility and CO_2 capture. Energy Economics 32, 805-816

Geman, H., Roncoroni, A., 2006. Understanding the fine structure of electricity prices. Journal of Business 79, 1225-1261

Heydari, S., Siddiqui, A., 2010. Valuing a gas-fired power plant: A comparison of ordinary linear models, regime-switching approaches, and models with stochastic volatility. Energy Economics 32, 709-725

Houser T., Bradley R., Werksman J., Childs B., Heilamyr R., 2008. Leveling the Carbon Playing Field: International Competition and US Climate Policy Design. Peter G. Peterson Institute for International Economics and the World Resources Institute, Washington DC, USA, 18-20

Hu, W., Kercheval, A., 2008. The skewed t distribution for portfolio credit risk. Advances in Econometrics 22, 55-83

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59, 1551-1580

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., Shin, Y., 1992. Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. Journal of Econometrics 54, 159-178

Lucia, J., Schwartz, E.S., 2002. Electricity Prices and Power Derivartives - Evidence from the Nordic Power Exchange. Review of Derivatives Research 5, 5-50

Longstaff F.A., Schwartz E.S., 2001. Valuing American options by simulation: A simple least-squares approach. The Review of Financial Studies 14 (1), 113-47

Los, H.S., de Jong, C., van Dijken, H., 2009. Realistic Power Plant Valuations. How to Use Cointegrated Power and Fuel Prices. WorldPower 2009, 48-53

Lütkepoh, H., Krätzig, M. Applied Time Series Econometrics. Cambridge University Press, Cambridge; 2004

McNeil, A., R. Frey, P. Embrechts. Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press, Princeton; 2005

Modjtahedi, B., Movassagh, N., 2005. Natural-gas futures: Bias, predictive performance, and the theory of storage. Energy Economics 27, 617-637

Mu, X., 2007. Weather, storage, and naturalgas price dynamics: Fundamentals and volatility. Energy Economics 29 (1), 46-63

Näsäkkälä, E., Fleten, S.-E., 2005. Flexibility and technology choice in gas fired power plant investments. Review of Financial Economics 14, 371-393

Pöyry, 2010. Wind Energy and Electricity Prices: Exploring the 'Merit Order Effect. European Wind Energy Association litterature review, Brussels, Belgium

Said, S. E., Dickey, D. A., 1984. Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. Biometrika 71, 599–607

Timera Energy, 2011. CCGT investment in the UK and Germany. Available at `http://www.timera-energy.com/uk-power/ccgt-investment-in-the-uk-and-germany/`[Accessed 06.06.2012]

Tse, Y., Tsui, A., 2002. A Multivariate GARCH Model with Time-Varying Correlations. Journal of Business and Economic Statistics 20, 351-362

Tseng, C.L., Barz, G., 2002. Short-term generation asset valuation: A real options approach. Operations Research 50 (2), 297-310

Uhrig-Homburg M., Wagner M. Futures Price Dynamics of CO_2 Emission Certificates - An Empirical Analysis. Journal of Derivatives 17 (2), 73-88

Zhu, W., 2004. Thermal generation asset valuation problems in a competitive market. Ph.D. dissertation, University of Maryland, College Park, MD, USA. Available at `http://hdl.handle.net/1903/1771`

Øigard, T.A., Hanssen, A., 2002. The multivariate normal inverse Gaussian heavytailed distribution: Simulation and estimation. IEEE International Conference on Aucustics, Speech and Signal Processing, Orlando Florida, May 2002

A Estimation results from the VECM model

The VECM model is

$$\Delta\bar{\mathbf{X}}_t = \alpha(\beta^\intercal\bar{\mathbf{X}}_{t-1} - \Phi) + \sum_{i=1}^{q} P_i\Delta\bar{\mathbf{X}}_{t-i} + \Psi\Theta_t + \Upsilon_t \tag{18}$$

where $\bar{\mathbf{X}}_t \in \mathbb{R}^d$ is the vector of de-seasonalized log prices, Φ is a scalar, Λ_i, Π, $P_i \in \mathbb{R}^{d\times d}$ are matrices of coefficients, and $\Upsilon_t \in \mathbb{R}^d$ is the vector of residuals. There is a clear economic intepretation of the term $(\beta^\intercal\bar{\mathbf{X}}_{t-1} - \Phi)$: it is the long-term equilibrium between the three de-seasonalized log prices. The estimation results are shown in Table 13.

B Probability distributions for the standardized errors

B.1 Estimation of the MNIG distribution

The MNIG distribution can be represented as a normal mean-variance mixture, with the inverse Gaussian (IG) distribution as the mixing distribution, such that when $\mathbf{Y} \sim N_d(\mathbf{0}, \mathbf{I})$ and $W \sim IG(\delta, \sqrt{\chi^2 - \boldsymbol{\gamma}^\intercal\boldsymbol{\Gamma}\boldsymbol{\gamma}})$, then $\mathbf{Z}_t = \boldsymbol{\mu} + W\boldsymbol{\Gamma}\boldsymbol{\gamma} + \sqrt{W}\boldsymbol{\Gamma}^{1/2}\mathbf{Y}$ is MNIG-distributed $\mathbf{Z}_t \sim MNIG(\boldsymbol{\mu}, \delta, \boldsymbol{\Gamma}, \chi, \boldsymbol{\gamma})$. The probability density function is given below in Eq. (19a).

$$f_Z(\mathbf{Z}) = \frac{\delta}{2^{\frac{d-1}{2}}}\left(\frac{\chi}{\pi q(\mathbf{Z})}\right)^{\frac{d+1}{2}} e^{p(\mathbf{Z})} K_{\frac{d+1}{2}}(\chi q(\mathbf{Z})) \tag{19a}$$

$$q(\mathbf{Z}) = \sqrt{\delta^2 + (\mathbf{Z} - \boldsymbol{\mu})^\intercal\Gamma^{-1}(\mathbf{Z} - \boldsymbol{\mu})} \tag{19b}$$

$$p(\mathbf{Z}) = \delta\sqrt{\chi^2 - \boldsymbol{\gamma}^\intercal\boldsymbol{\Gamma}\boldsymbol{\gamma}} + \boldsymbol{\gamma}(\mathbf{Z} - \boldsymbol{\mu}) \tag{19c}$$

where $K_i(\cdot)$ is defined as in Section 3.6.2. Its expectation and covariance are given in Eqs. (20a) and (20b).

$$\mathbb{E}[\mathbf{Z}] = \boldsymbol{\mu} + \frac{\delta\boldsymbol{\Gamma}\boldsymbol{\gamma}}{\sqrt{\chi^2 - \boldsymbol{\gamma}^\intercal\boldsymbol{\Gamma}\boldsymbol{\gamma}}} \tag{20a}$$

$$cov[\mathbf{Z}] = \frac{\delta\left[\boldsymbol{\Gamma} + (\chi^2 - \boldsymbol{\gamma}^\intercal\boldsymbol{\Gamma}\boldsymbol{\gamma})^{-1}\boldsymbol{\Gamma}\boldsymbol{\gamma}\boldsymbol{\gamma}^\intercal\boldsymbol{\Gamma}^\intercal\right]}{\sqrt{\chi^2 - \boldsymbol{\gamma}^\intercal\boldsymbol{\Gamma}\boldsymbol{\gamma}}} \tag{20b}$$

When dimensionality is low, parameter estimation is feasible with traditional maximum likelihood estimation. However, even with $d = 3$ dimensions, the EM algorithm (Dempster *et al.*, 1977) is faster and seems, to the authors, less dependent on starting values. With three dimensions, we must estimate 17 parameters, and the flat likelihood surface increases the chance of ending up in a local optimum instead of the global optimum.

The EM algorithm for the MNIG distribution was derived by Øigard and Hanssen (2002), and Aas, Haff and Dimakos (2006) modified the algorithm to

Table 13: Estimation results of the VECM model. The values in each column is the effect on the log-return of each variable (see the top row). T-values are shown in parantheses. Wednesday dummy coefficients have been set to zero, and β for log power has been set to one.

Coefficient	$\Delta\bar{x}_t^{El}$	$\Delta\bar{x}_t^{G}$	$\Delta\bar{x}_t^{C}$
$\boldsymbol{\alpha}^\intercal$	-0.089	0.035	0.002
	(-7.618)	(3.076)	(0.453)
$\boldsymbol{\beta}^\intercal$	1	-0.805	-0.307
	(N/A)	(17.589)	(-4.168)
ϕ	-0.842	-0.842	-0.842
	(-3.709)	(-3.709)	(-3.709)
$\Delta\bar{x}_{t-1}^{el}$	-0.2	-0.017	-0.009
	(-8.636)	(-0.754)	(-1.274)
$\Delta\bar{x}_{t-1}^{g}$	0.085	-0.054	-0.008
	(3.488)	(-2.321)	(-1.196)
$\Delta\bar{x}_{t-1}^{c}$	0.133	0.249	0.072
	(1.787)	(3.477)	(3.327)
$\Delta\bar{x}_{t-2}^{el}$	-0.131	0.021	-0.009
	(-5.940)	(1.005)	(-1.339)
$\Delta\bar{x}_{t-2}^{g}$	0.008	-0.086	0
	(0.339)	(-3.713)	(-0.061)
$\Delta\bar{x}_{t-2}^{c}$	0.069	0.049	0.015
	(0.931)	(0.682)	(0.693)
Mon Pre 2009	0.074	0.058	0.003
	(10.805)	(8.777)	(1.480)
Tue Pre 2009	0.014	0.011	-0.001
	(2.102)	(1.779)	(-0.422)
Wed Pre 2009	0	0	0
	(N/A)	(N/A)	(N/A)
Thu Pre 2009	-0.017	-0.009	0
	(-2.724)	(-1.529)	(0.093)
Fri Pre 2009	-0.054	-0.008	0
	(-8.454)	(-1.251)	(0.021)
Weekend Pre 2009	-0.033	-0.034	-0.001
	(-5.087)	(-5.435)	(-0.288)
Mon After 2009	-0.009	0.014	-0.003
	(-1.243)	(2.004)	(-1.311)
Tue After 2009	0.002	-0.001	0.001
	(0.270)	(-0.193)	(0.677)
Wed After 2009	0	0	0
	(N/A)	(N/A)	(N/A)
Thu After 2009	-0.002	-0.002	0
	(-0.247)	(-0.356)	(0.011)
Fri After 2009	-0.01	-0.004	-0.001
	(-1.390)	(-0.594)	(-0.367)
Weekend After 2009	0.024	-0.018	0
	(3.641)	(-2.783)	(-0.080)

fit the restrictions that the expected value is the zero vector and that the covariance matrix is the identity matrix. We follow the recommendation of Aas *et al.*, and estimate three univariate NIG distributions to each of the data sets, to use as input for the starting values. A univariate NIG distribution has parameters $(\mu_i, \delta_i, \chi_i, \gamma_i)$, and we estimate these parameters by the method of moments (Eriksson et al., 2004). We then set as starting values:

$$\boldsymbol{\mu}_0 = [\mu_1 \mu_2 \dots \mu_i]^\intercal \tag{21}$$

$$\delta_0 = \frac{1}{d} \sum_{i=1}^{d} \delta_i \tag{22}$$

$$\boldsymbol{\gamma}_0 = [\gamma_1 \gamma_2 \dots \gamma_d]^\intercal \tag{23}$$

$$\boldsymbol{\Gamma}_0 = \mathbf{I}_d \tag{24}$$

$$\chi_0 = \frac{1}{d} \sum_{i=1}^{d} \sqrt{\chi_i^2 - \gamma_i^2 - \boldsymbol{\gamma}_0^\intercal \boldsymbol{\Gamma} \boldsymbol{\gamma}_0} \tag{25}$$

$$\tag{26}$$

The EM algorithm consists of two steps; the expectation step and the maximization step. In the first step, we compute the expected values of the inverse-Gaussian distributed variable W and its inverse:

$$\zeta_t = \mathbf{E}[W_t | Z_t = z_t] = \frac{q(z_t)}{\chi} \frac{K_{\frac{d-1}{2}}(\chi q(z_t))}{K_{\frac{d+1}{2}}(\chi q(z_t))} \tag{27}$$

$$\phi_t = \mathbf{E}[W_t^{-1} | Z_t = z_t] = \frac{\chi}{q(z_t)} \frac{K_{\frac{d+3}{2}}(\chi q(z_t))}{K_{\frac{d+1}{2}}(\chi q(z_t))} \tag{28}$$

$$\tag{29}$$

In the second step, we use these expected values to update parameter estimates. For ease of notation, we compute the temporary variables $\lambda = \sum_{t=1}^{T} \zeta_t / T$, $\xi = \sum_{t=1}^{T} \phi_t / T$ and $\upsilon = T(\sum_{t=1}^{T} (\phi_t / - 1/\lambda))^{-1}$, where T is the number of observations. We then proceed by updating the parameter estimates according to the

following equations:

$$\delta = \sqrt{\upsilon} \tag{30a}$$
$$\eta = \frac{\delta}{\lambda} \tag{30b}$$
$$\boldsymbol{\gamma} = \arg\min_{\gamma} \left[\sum_{t=1}^{T} (\phi_t \mathbf{Z}_t^{\intercal} \boldsymbol{\Gamma}^{-1} \mathbf{Z}_t) \right. \tag{30c}$$
$$\left. -2 \sum_{t=1}^{T} ((1 - \lambda\phi_t) \mathbf{Z}_t^{\intercal} \boldsymbol{\gamma}) + T(-\lambda + \lambda^2 \xi) \boldsymbol{\gamma}^{\intercal} \boldsymbol{\Gamma} \boldsymbol{\gamma} \right] \tag{30d}$$
$$\bar{\boldsymbol{\Gamma}} = \eta \left[\frac{\mathbf{I}}{\delta} + \frac{1}{\boldsymbol{\gamma}^{\intercal}\boldsymbol{\gamma}} \left(\frac{\eta}{2\boldsymbol{\gamma}^{\intercal}\boldsymbol{\gamma}} \left(\sqrt{1 + \frac{4\boldsymbol{\gamma}^{\intercal}\boldsymbol{\gamma}}{\delta\gamma}} - 1 \right) - \frac{1}{\delta} \right) \boldsymbol{\gamma}\boldsymbol{\gamma}^{\intercal} \right] \tag{30e}$$
$$\boldsymbol{\Gamma} = \frac{\bar{\boldsymbol{\Gamma}}}{\left|\bar{\boldsymbol{\Gamma}}\right|^{\frac{1}{d}}} \tag{30f}$$
$$\chi = \sqrt{\eta^2 + \boldsymbol{\gamma}^{\intercal} \Gamma \boldsymbol{\gamma}} \tag{30g}$$
$$\boldsymbol{\mu} = -\lambda \boldsymbol{\Gamma} \boldsymbol{\gamma} \tag{30h}$$
$$\tag{30i}$$

where Eqs. (30e) and (30h) are the result of solving $\mathbb{E}[Z] = 0$ and $cov[Z] = \mathbf{I}$. The EM algorithm was written in Matlab. After the two steps, we save the parameter estimates, and repeat the expectation and maximization. When the largest difference in any parameter estimate is smaller than some limit in two consecutive iterations, we say stop and save the parameter estimates. These are given in Table 14.

Table 14: MNIG parameter estimates from the EM algortihm, using a tolerance limit of 1e-6

$\boldsymbol{\mu}^{\intercal}$	-0.1333	0.0127	0.1409
δ	1.1323		
$\boldsymbol{\Gamma}$	0.9962	0.0013	0.0142
	0.0013	1.0095	-0.0014
	0.0142	-0.0014	0.9946
χ	1.1287		
$\boldsymbol{\gamma}^{\intercal}$	0.1334	-0.0127	-0.1411

B.2 The skewed t distribution and its estimation procedure

There are several skewed t distributions, and several parametrizations one can use. We use the parametrization given by Hu and Kercheval (2008). Let $\mathbf{Z}$ be skewed t distributed with $(d \times 1)$ parameter vectors $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$, a $(d \times d)$ matrix $\boldsymbol{\Sigma}$

and a scalar parameter ν, such that $\mathbf{Z}_t \sim SkewedT_d(\boldsymbol{\mu}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \nu)$. The probability density function $f_Z(\mathbf{Z})$ may then be written as:

$$f_Z(\mathbf{Z}) = c\frac{K_{\frac{\nu+d}{2}}\left(\sqrt{(\nu+\rho(\mathbf{z}))(\boldsymbol{\gamma}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma})}\right)e^{(\mathbf{z}-\boldsymbol{\mu})^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}}}{\left(\sqrt{(\nu+\rho(\mathbf{z}))(\boldsymbol{\gamma}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma})}\right)^{-\frac{\nu+d}{2}}(1+\frac{\rho(\mathbf{z})}{\nu})^{\frac{\nu+d}{2}}} \tag{31a}$$

$$c = \frac{2^{1-\frac{\nu+d}{2}}}{\Gamma(\frac{\nu}{2})(\pi\nu)^{\frac{d}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \tag{31b}$$

$$\rho(\mathbf{x}) = (\mathbf{x}-\boldsymbol{\mu})^\intercal\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \tag{31c}$$

where $K_i(\cdot)$ is defined as in Section 3.6.2. The parameter ν is the degrees of freedom in the distribution, while the vector $\boldsymbol{\gamma}$ represents the skewness in each of the d dimensions. The expectation and covariance are defined if $\nu > 2$ and $\nu > 4$ respectively, and are given by Eqs. (32a) and (32b).

$$\mathbb{E}(\mathbf{Z}) = \boldsymbol{\mu} + \boldsymbol{\gamma}\frac{\nu}{\nu-2} \tag{32a}$$

$$cov(\mathbf{Z}) = \frac{\nu}{\nu-2}\boldsymbol{\Sigma} + \boldsymbol{\gamma}\boldsymbol{\gamma}^\intercal\frac{2\nu^2}{(\nu-2)^2(\nu-4)} \tag{32b}$$

The EM algorithm for the skewed t is very similar to the one for the MNIG distribution in Appendix B.1, as they are both generalized hyperbolic distributions. The skewed t distribution can also be written as a normal mean-variance mixture, but with the inverse gamma distribution as the mixing distribution. The setup of the algorithm is very similar to the one in Appendix B.1, and we will therefore only give the necessary equations to estimate the parameters. To initialize the algorithm, we set $\boldsymbol{\mu}_0 = \bar{Z}$ and $\Sigma_0 = \mathbf{I}$, $\boldsymbol{\gamma}_0$ equal to the sample skewness and $\nu = 20$. In each iteration we compute the auxilliary variables

$$\rho_t = (\boldsymbol{z}_t-\boldsymbol{\mu})^\intercal\boldsymbol{\Sigma}^{-1}(\boldsymbol{z}_t-\boldsymbol{\mu}) \tag{33a}$$

$$\theta_t = \left(\frac{\rho_t+\nu}{\boldsymbol{\gamma}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}}\right)^{-\frac{1}{2}}\frac{K_{\frac{\nu+d+2}{2}}\left(\sqrt{(\rho_t+\nu)(\boldsymbol{\gamma}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma})}\right)}{K_{\frac{\nu+d}{2}}\left(\sqrt{(\rho_t+\nu)(\boldsymbol{\gamma}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma})}\right)} \tag{33b}$$

$$\eta_t = \left(\frac{\rho_t+\nu}{\boldsymbol{\gamma}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}}\right)^{\frac{1}{2}}\frac{K_{\frac{\nu+d-2}{2}}\left(\sqrt{(\rho_t+\nu)(\boldsymbol{\gamma}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma})}\right)}{K_{\frac{\nu+d}{2}}\left(\sqrt{(\rho_t+\nu)(\boldsymbol{\gamma}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma})}\right)} \tag{33c}$$

$$\xi_t = \frac{1}{2}log\left(\frac{\rho_t+\nu}{\boldsymbol{\gamma}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}}\right) + \frac{\frac{\partial K_{-\frac{\nu+d}{2}+\alpha}\left(\sqrt{(\rho_t+\nu)(\boldsymbol{\gamma}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma})}\right)}{\partial\alpha}|_{\alpha=0}}{K_{\frac{\nu+d}{2}}\left(\sqrt{(\rho_t+\nu)(\boldsymbol{\gamma}^\intercal\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma})}\right)} \tag{33d}$$

The partial derivative of the bessel function ($K_i(\cdot)$) is cumbersome to both derive and compute, and we therefore use the numerical approximation:

$$\frac{\partial K_{-\frac{\nu+d}{2}+\alpha}(\cdot)}{\partial\alpha}|_{\alpha=0} = \frac{K_{-\frac{\nu+d}{2}+h}(\cdot) - K_{-\frac{\nu+d}{2}-h}(\cdot)}{2h} \tag{34}$$

We also compute the means of the auxiliary variables $\bar{\theta}, \bar{\eta}$ and $\bar{\xi}$. We then proceed to update the parameter estimates, according to the following equations:

$$\gamma = \frac{T^{-1}\sum_{t=1}^{T}\theta_t(\bar{Z} - Z_t)}{\bar{\theta}\bar{\eta} - 1} \tag{35}$$

ν solves the equation

$$-\psi(\frac{\nu}{2}) + log(\frac{\nu}{2}) + 1 - \bar{\xi} - \bar{\theta} = 0 \tag{36}$$

where $\psi(\cdot)$ is the polygamma function, and

$$\begin{aligned} \mu &= -\gamma\frac{\nu}{\nu - 2} \\ \Sigma &= \left(\frac{\nu - 2}{\nu}\right)\left(\mathbf{I} - \boldsymbol{\gamma}\boldsymbol{\gamma}^{\intercal}\frac{2\nu^2}{(\nu-2)^2(\nu-4)}\right) \end{aligned} \tag{37}$$

where we have used Eqs. (32a) and (32b) to impose zero expectation and identity covariance. When the maximum change in any parameter between two iterations is smaller than some convergence limit, we terminate the algorithm and keep the parameters. These are given in Table 15.

Table 15: Skewed t parameter estimates from the EM algortihm, using a tolerance limit of 1e-6

$\boldsymbol{\mu}^{\intercal}$	-0.1621	0.1717	0.1326
$\boldsymbol{\gamma}^{\intercal}$	0.1073	-0.1137	-0.0878
$\boldsymbol{\Sigma}$	0.6440	0.0192	0.0148
	0.0192	0.6418	-0.0157
	0.0148	-0.0157	0.6500
ν	5.9192		

C The deterministic production scheduling

C.1 Further description of the intraday optimization

This appendix goes a little more in detail on the generation of the intraday production profiles presented in Section 4.3. We define the $S \times S$ matrices $\bar{\mathbf{Q}}^g$ and $\bar{\mathbf{Q}}^{el}$ such that $\bar{q}^g_{ij}$ is the gas burnt during the hour in which the power plant goes from state i to state j, and $\bar{q}^{el}_{ij}$ is the electricity produced from i to j. If both i and j are states in which the power plant is turned off, there will be no electricity production, and thus $\bar{q}^{el}_{ij} = 0$. We define the numbering of states such that if $j < i$, then the power plant is in a cool-down phase in which no gas is burnt. If $i > j$, then the transition implies that the plant is warming up, and $\bar{q}^g_{ij} > 0$. Further, the $S \times S$ profit matrices $\bar{\mathbf{R}}_h$ contain the elements $\bar{r}_{ijh}$ that defines the intraday profit of going from state i to state j during hour h. The prices p^{el}_h and p^g_h are defined as the price valid for the hour h. These will only

change at midnight, meaning that $p_h = p_t$ for $h \in [1, 12]$ and $p_h = p_{t+1}$ for $h \in [13, 24]$. The profit matrix for hour h is defined as follows:

$$\bar{\mathbf{R}}_h = (p_h^{el} - c^{var})\bar{\mathbf{Q}}^{el} - (p_h^g + p_h^c \cdot I_C)\bar{\mathbf{Q}}^g + B\mathbf{L} \tag{38}$$

Where B is a penalty imposed on infeasible transitions, and is set to a value so high that infeasible transitions are never optimal. Using the $\mathbf{R}$ matrix, we can now calculate the continuation values for each hour of the day. Through traditional dynamic programming, we get the optimal profit from starting in state i and ending in state j 24 hours later. The total production of power and consumption of gas are stored for all feasible values of $L_{M,t}$, $L_{M,t+1}$ and $\Delta c_{F,t}$ for both the period before and after midnight, and for all possible start states and end states (i, j). Using these databases, we can accurately compute the profits of going from i to j without performing the intraday dynamic programming for all days.

The $S \times 25$ matrix $\mathbf{V}_{s_{end}}$ denotes the values of being in state i in the beginning of hour h, when the last state is set to be s_{end}. The matrix contains 25 columns, as the last column denotes the value of being in state i at the end of the 24-hour period. For a given end state s_{end} the elements $v_{hi,s_{end}}$ of $\mathbf{V}_{s_{end}}$ are determined by computing backwards from $h = 25$ to $h = 1$:

$$v_{25,i,s_{end}} = \begin{cases} 0, & \text{if i} = s_{end} \\ -B, & \text{otherwise} \end{cases} \tag{39}$$

$$v_{hi,s_{end}} = \arg\max_j (r_{hij} + B(l_{ij} - 1) + v_{h+1,j}) \quad \forall i \in (1, S), h \in (1, 24) \tag{40}$$

Where B is considered a penalty imposed on infeasible or undesired states or transitions. B must be set so high that performing an infeasible or undesired transition in Eq. 40 will never be optimal. The dynamic program consisting of Eqs. 38 - 40 is repeated for all feasible end states. When the matrices $\mathbf{V}_{s_{end}}$ are completed for all end states, the total profit for the 24-hour period when going from state s_{start} to s_{end} is $v_{1,s_{start},s_{end}}$, which can be used in the LSM algorithm.

D Proof of dimensionality reduction

When the power plant operator is choosing the optimal production profile, he is in maximizing the profit realized during the next 24-hour period. Let us define the hourly productions of electricity and consumptions of gas as the vectors N^{el} and N^{g}, having elements n_h^{el} and n_h^{g}, $h \in \{1, \ldots, 24\}$. The prices of electricity, gas and carbon are p_t^{el}, p_t^{g} and p_t^{c}. The profit arising from one day of production is

$$\begin{aligned} r_t = &\sum_{h=1}^{12} n_h^{el}(p_t^{el} - c^{var}) + \sum_{h=13}^{24} n_h^{el}(p_{t+1}^{el} - c^{var}) - \\ &\sum_{h=1}^{12} n_h^{g}(p_t^{g} + I_C \cdot p_t^{c}) - \sum_{h=13}^{24} n_h^{g}(p_{t+1}^{g} + I_C \cdot p_{t+1}^{c}) \end{aligned} \tag{41}$$

When choosing between two (or more) different production profiles N^{el1}, N^{g1} and N^{el2}, N^{g2}, the profile giving the highest total profit is chosen. For simplicity, we will denote the net revenue $p_t^{el} - c^{var}$ as p_t and the fuel cost $p_t^{g} + I_C \cdot p_t^{c}$ as $c_{F,t}$. Production profile 1 is chosen if and only if

$$\begin{aligned} &\sum_{h=1}^{12} n_h^{el1} p_t + \sum_{h=13}^{24} n_h^{el1} p_{t+1} - \sum_{h=1}^{12} n_h^{g1} c_{F,t} - \sum_{h=13}^{24} n_h^{g1} c_{F,t+1} \\ \geq &\sum_{h=1}^{12} n_h^{el2} p_t + \sum_{h=13}^{24} n_h^{el2} p_{t+1} - \sum_{h=1}^{12} n_h^{g2} c_{F,t} - \sum_{h=13}^{24} n_h^{g2} c_{F,t+1} \end{aligned} \tag{42}$$

To show that the choice between the two distinct profiles are independent of the prices and only depends on the relative prices, $L_{M,t} = \frac{p_t}{c_{F,t}}$, $L_{M,t+1} = \frac{p_{t+1}}{c_{F,t+1}}$ and $\Delta c_{F,t} = \frac{c_{F,t+1}}{c_{F,t}}$, we divide by $c_{F,t}$ and substitute for $L_{M,t}$, $L_{M,t+1}$ and $\Delta c_{F,t}$:

$$\begin{aligned} &L_{M,t}^{*} \sum_{h=1}^{12} n_h^{el1} + L_{M,t+1}^{*} \Delta c_{F,t} \sum_{h=13}^{24} n_h^{el1} - \sum_{h=1}^{12} n_h^{g1} - \Delta c_{F,t} \sum_{h=13}^{24} n_h^{g1} \\ \geq &L_{M,t}^{*} \sum_{h=1}^{12} n_h^{el2} + L_{M,t+1}^{*} \Delta c_{F,t} \sum_{h=13}^{24} n_h^{el2} - \sum_{h=1}^{12} n_h^{g2} - \Delta c_{F,t} \sum_{h=13}^{24} n_h^{g2} \end{aligned} \tag{43}$$

As the choice between two sets of production profiles here only depends on the relations $L_{M,t}$, $L_{M,t+1}$ and $\Delta c_{F,t}$, we can generate a set of universal rules for the choice of profile regardless of the actual price levels of the underlying commodities.