Spot Price Models for Natural Gas

-

Robustness of the Convenience Yield Approach

Thomas Volmer

March 31, 2011

Dissertation
zur Erlangung des Doktorgrades
an der
Fakultät Sozial- und Wirtschaftswissenschaften
der Otto-Friedrich-Universität Bamberg

Vorgelegt von:
THOMAS VOLMER

Erstprüfer: Prof. Dr. Matthias Muck, Otto-Friedrich-Universität Bamberg
Zweitprüferin: Prof. Dr. Susanne Rässler, Otto-Friedrich-Universität Bamberg

Tag der mündlichen Prüfung: 21. Dezember 2011

Acknowledgements

I am particularly grateful to Matthias Muck, my supervisor, for his invaluable support during my dissertation project. Without our numerous discussions and his advice I would not have been able to compose this work. I would also like to thank Susanne Rässler who provided me with helpful statistical guidance several times. She also kindly agreed to be part of my dissertation committee.

Furthermore, my work has strongly profited from fruitful discussions at the University of Bamberg with my colleagues from the Chair of Financial Control. I am especially grateful to Jan Marckhoff and Sebastian Paik for their enriching feedback and suggestions in many difficult situations. I would also like to thank the team of the Chair for the enjoyable working atmosphere during my time in Bamberg.

Moreover, many thanks go to Bagher Modjtahedi and Richard Quandt from the US as well as Scott Chaput and Martin Hill from New Zealand and Australia, respectively, for their helpful support and comments regarding my first paper. Additionally, I thank participants at the 13th Conference of the Swiss Society for Financial Markets Research (SGF) in Zurich and at the Conference on Energy Finance at the University of Agder, Kristiansand, Norway. Many thanks also go to APX-ENDEX, the British Atmospheric Data Center (BADC), the Intercontinental Exchange (ICE), the International Energy Agency (IEA) and National Grid UK for kindly providing the data used in my dissertation. I would like to highlight the personal support of Vlad Kaltenieks from APX-ENDEX and Rajesh Kukadia from National Grid who dealt with several of my individual queries.

Finally, I am greatly indebted to Kathrin, my family and friends for their emotional support and encouragement as well as their enduring belief in my dissertation project.

To my family

Contents

List of Tables

List of Figures

List of Abbreviations

APX	Amsterdam Power Exchange
BADC	British Atmospheric Data Centre
bcm	billion cubic meters
BOM	balance of the month
CIR	Cox, Ingersoll and Ross
ch.	chapter
c.p.	ceteris paribus
EEX	European Energy Exchange
EIA	Energy Information Administration
eq.	equation
FERC	Federal Energy Regulatory Commission
FGLS	feasible generalized least squares
fn.	footnote

GBM	geometric Brownian motion
ICE	Intercontinental Exchange
IEA	International Energy Agency
MAE	mean absolute error
mcm	million cubic meters
ME	mean error
ML	maximum likelihood
MMA	money market account
MWh	megawatt hours
mmBtu	million British thermal units
NBP	National Balancing Point
OCM	On-the-Day Commodity Market (UK)
OTC	over the counter
p.	page
PDE	partial differential equation
RMSE	root mean squared error
SCY	stochastic convenience yield
vs.	versus
SDE	stochastic differential equation
sec.	section
w.r.t.	with respect to

List of Symbols

Latin letters

$A(T)$	deterministic, parametric function of maturity T of Schwartz (1997)
B	value of money market account
C	amount of storage cost or amount of consumption good available (sec. 4.4)
D	basis differential (ch. 2) or D-value of Goldfeld and Quandt (1972)
F	forward or futures price
H	covariance matrix of measurement errors (Kalman filter)
I	inventory level
J	value function
K	option strike price
L	value of the likelihood function

Mi	month future with maturity in the ith-next month
N	length of time series
O	stock of explored oil
P	value of replicating portfolio (ch. 3, ch. 6) or inverse demand function (ch. 4) or marginal amount of storage risk premium (ch. 5)
$\mathcal{P}$	empirical probability measure
Q	marginal amount of convenience yield (ch. 5) or state noise covariance matrix (Kalman filter, ch. 6)
$\mathcal{Q}$	risk-neutral probability measure
R	marginal amount of interest (ch. 5) or error sensitivity matrix (Kalman filter, ch. 6)
R^2	coefficient of determination
S	spot price
T	transition matrix (Kalman filter)
U	jump size or utility function (sec. 4.4)
X	"amount on hand" or exploration quantity (sec. 4.4) or log-spot price
Z	exogenous production quantity or design matrix (Kalman filter)
c	storage cost rate (% of spot price) or vector of constants (Kalman filter)

d	differential operator or vector of constants (Kalman filter)
dN	increment of Poisson process
dW	increment of Wiener process
f	function
i	index (integer)
k	risky rate of return
m	net cost of storage (ch. 5) or state size (Kalman filter)
n	number of elements/observations in subsample or cross-section
r	(riskless) interest rate
t	time index or current period
u	marginal revenue of storage (sec. 5.1) or autocorrelated residual
v	portfolio weight
v^*	(risk-neutral) spot price drift
x	dependent variable
y	independent variable

Greek letters

Δ	difference operator
Ω	parameter set

α	mean-reversion level (stochast. processes) or state vector (Kalman filter)
β	regression coefficient
γ	rate of storage depreciation
δ	convenience yield rate (% of spot price)
ϵ	residual or measurement noise (Kalman filter)
η	residual or state noise (Kalman filter)
θ	mean-reversion level (ch. 4) or deseas. convenience yield rate (ch. 6)
κ	mean-reversion coefficient
λ	market risk premium
λ_P	Poisson intensity of jumps
μ	deterministic drift rate
ξ	long-term risk factor of Schwartz and Smith (2000) or measurement error standard deviation (ch. 6)
ρ	coefficient of correlation or rate of storage risk premium (ch. 5)
σ	standard deviation
χ	short-term risk factor of Schwartz and Smith (2000) or portfolio weight (ch. 6)
ϕ	coefficient of autocorrelation
ψ	portfolio weight

Decorations and superscripts

$\hat{}$	estimate of theoretical counterpart
$\tilde{}$	alteration of a quantity previously defined
*	risk-neutral parameter or FGLS-transormed coefficients (ch. 5)
^{e}	estimate of the theoretical counterpart
^{s}	deseasonalized variable

Chapter 1

Introduction

In recent years exchange traded natural gas contracts have gained great importance. This especially applies to continental Europe where increased liberalization and transparency are now on the agenda. One fact underpinning the rising interest in standardized contracts is the recent start of gas trading at several energy exchanges in this region. In July 2007 the German EEX (European Energy Exchange) launched its gas trading platform. In May and November 2008, respectively, the Scandinavian Nord Pool and the French Powernext followed suit. Finally, in December 2009 and 2010, respectively, Central European Gas Hub launched the Gas Spot and Gas Futures segments at Wiener Börse. Further support can be found by looking at the exchange traded volumes at the more established continental hubs: APX-ENDEX recently reported a year-on-year volume increase of 130% in gas trading at the Dutch TTF to 210 TWh in 2010 and a 22% increase for the Belgian hub Zeebrugge to 203 MWh in 2010.[1]

Along with the importance of exchange traded gas grows the necessity to find accurate pricing models for the different contract types and this

[1] Cf. APX-ENDEX (2011), p. 20.

regularly requires a future spot price. That said, one notable difference between commodities and stocks or bonds is that the current spot price is not directly observable in the market. During the last few decades a number of general commodity pricing models have been developed and tested. However, the application to natural gas pricing was predominantly tested using reduced-form models such as the well-known two-factor models by Gibson and Schwartz (1990), Schwartz (1997) and Schwartz and Smith (2000). As opposed to structural models, these models build on an exogenously specified stochastic process rather than supply and demand conditions founded in microeconomic theory. They are mostly uncomplex and also capture typical phenomena of the forward curve (e.g. the Samuelson effect[2]), but they do not offer fundamental economic explanations for the predicted prices.

The lack of insight into the real price drivers is a downside of pure reduced-form models because the risk of misspecification is high and the out-of-sample performance can be poor. With respect to the Schwartz and Smith model, Carlson et al. (2007) show that given frictional production adjustments in the oil and gas market the model will systematically overestimate the prices for oil and gas options. Similarly, Ribeiro and Hodges (2004) find that "...the use of current reduced-form models in the literature to price energy contingent claims has not been effective. In particular, the convenience yield process seems to be misspecified since its specification ignores some crucial properties of commodity price behavior such as the dependency of prices' variability on inventory levels" (p. 3).

Beyond the general criticism of reduced-form commodity price models, the latter statement points at difficulties caused by a specific price com-

---

[2] The *Samuelson effect* describes the fact that the volatility of forward prices increases with decreasing time to maturity and is due to the mean reverting property of many commodity spot prices.

ponent: the *convenience yield*. The concept of convenience yield originates from Kaldor (1939), the groundwork of the '*theory of storage'*. Kaldor states that goods in stock which are not yet sold forward have an unobservable flexibility yield because market participants owning these goods have the 'convenience' to make use of them whenever wanted.[3] For this reason, according to Kaldor, the observable *cost of carry* of a stored commodity, that is foregone interest and the outlay on physical storage, must be reduced by this availability premium. Brennan (1958) and Telser (1958) have further established that this premium varies with the level of storage in the economy. The concept of convenience yield is explained in detail in chapter 5.

The idea that such an availability premium impacts the spot price has also been implemented in a number of reduced-form commodity models. However, Schwartz and Smith (2000) abandon the concept of convenience yield in their generalized two-factor model and simply use a 'long-term' and a 'short-term' price component, one reason being that "[...] many find the notion of convenience yield elusive [. . .]" (p. 894). Given that the basic theory of the convenience yield is widely accepted, it seems questionable to neglect this variable in up-to-date pricing models since it contains fundamental economic information. Apparently, economists are simply lacking considerable knowledge about the dynamic behavior of the convenience yield in specific markets. Possession of such information would enable the economist to generate fundamental forecasts and to integrate them into a reduced-form model as an exogenous component. This could help to reduce the uncertainty in spot price forecasting significantly. At the same time, it

[3] This flexibility is of value since contrary to the owner of shares or bonds a commodity owner can put this good to additional use, for exapmle in industrial production. If there is uncertainty about the time when the input factor is needed, the producing company is more flexible by holding the commodity upfront than by entering a long postion in a forward contract with a fixed delivery date.

might constitute a viable compromise between the complexity of structural models and the danger of misspecification in reduced-form approaches.

The following study is based on this motivation and the aforementioned increase in demand for exchange-traded risk management instruments in the natural gas market, particularly in continental Europe. Therefore, in the first half of the main part, we present an econometric analysis of the drivers of convenience yield for natural gas. As a first contribution, it is shown that, in addition to national gas storage levels, air temperature is a highly relevant and robust determinant of the convenience yield. Moreover, a regime-switching model, although not optimal for predictions, makes explicit that basis variability in gas markets rises with inventories as long as the average observed inventory is sufficiently high. This contrasts findings of Fama and French (1987) for non-energy commodities, and it will be argued that the opposing behavior is due to the capacity constraints of underground gas storage. Kogan et al. (2009) document similar findings for crude oil futures prices. In addition, to the best of our knowledge, we are the first to investigate the robustness of a convenience yield model for pricing applications. A forecasting exercise, which identifies the squared storage model as the most appropriate specification, is presented afterwards. It is shown that the model keeps its explanatory power when the measurement interval is varied.

In the second half of the main part, we use the convenience yield model to develop a new model for the gas spot price based on the stochastic convenience yield model (*SCY model*) of Schwartz (1997). First, to estimate this *extended model*, futures prices are netted of those deviations from the equilibrium price which are explained by the fundamental variables above. Next, the model is estimated with the "netted" prices. Then, the prediction results are compared to those generated with the unmodified futures price data. It is shown that the extended model improves the out-of-sample forecast as the forecast horizon increases. At the same time, the in-sample and the cross-sectional fit are at least as good as for the benchmark. Nevertheless, conceptual and also numerical reasons lead to the fact that, irrespective of the model version, parameter stability along the cross-section of futures contracts leaves room for further amendments. For this reason, the study concludes with an in-depth analysis of alternative estimation methods. Yet, we find that these alternatives are not favorable to the standard method, i.e. a maximum likelihood estimation with the Kalman filter. Since the gas market is regionally fragmented[4] and for some additional reasons[5], the first part of the study is conducted on the liberalized UK gas market. Additionally, the scope is then extended to cover the US market.

The structure of the thesis is as follows: Firstly, in chapter 2, we explain why increased demand for financial risk management products has been arising in the gas market for a number of years. It is argued that this makes it worthwhile and necessary to find accurate pricing models for these emerging financial instruments. Next, the study discusses why gas needs to be studied separately from other commodities that have already received more research attention. It is also addressed why the UK and the US are suitable markets to study. The chapter closes with an overview of the most important risk management products for natural gas and leads over to the reason for the particular importance of the future spot price (and hence spot price models) in risk management. Chapter 3 elaborates on the two most important analytical methods for derivative pricing in financial theory and gives a brief overview of more advanced techniques. Chapter 4 presents and classifies existing models of spot (and forward)

---

[4] Geman (2006), for instance, distinguishes between the 3 regional markets: America, Europe and Asia.

[5] One practical advantage over the even more developed US market is that climate conditions are sufficiently uniform within this market which facilitates our econometric analysis. Further details are provided during the course of the study.

prices for energy commodities. The models are briefly evaluated according to a set of defined requirements. In addition, we explain why we choose the SCY model as the starting point for our model extension. In the main part, the determinants of convenience yield (chapter 5) and a hybrid spot price model with fundamental convenience yield forecasts (chapter 6) are investigated. The last chapter concludes.

Chapter 2

The natural gas market

In this chapter we illustrate, first of all, why financial risk management in natural gas has become an interesting and relevant field of research. Next, we explain why natural gas has to be modelled individually, i.e. why the existing research on the pricing of energy derivatives is not sufficient to build a reliable gas price model. Afterwards, we show that the gas markets of the UK and the US are most suitable for an empirical study. More precisely, we illustrate the difference of trading in liberalized vs. non-liberalized markets and classify both the British and the American market as rather liberalized. The last section of this chapter gives an overview about the derivatives which are actually traded in the natural gas market. It will be concluded that their payoffs regularly depend (directly or indirectly) on the future spot price. This is the ultimate reason for our objective to search for the most adequate spot price model.

2.1 Liberalization of natural gas markets

Open access to the distribution and sales segment of the US gas market was effectuated by the FERC orders 436 and 500 in 1985. They allowed end customers to buy directly from producers by reserving capacity on interstate pipelines. In the aftermath, different tariff structures and service levels evolved, such as firm and interruptible service of different degrees. In 1992, FERC final order 636 enforced the unbundling of pipeline services. It was now permitted to transfer unused firm transportation capacity to a third party. These orders together are responsible for the strong increase in business activity during the 1990s (cf. Figure 2.1).[1] In the years after the collapse of Enron in late 2001, US gas trading volumes entered into a phase of stagnation. However, since 2006, the massive influx of capital from financial investors, notably hedge funds, has led to a tripling of turnover and a corresponding push in market liquidity until today. In fact, this recent development in the world's biggest natural gas market best illustrates the rising importance of the commodity.

In Europe, the UK started deregulation in the mid-eighties. The Natural Gas Act of 1986 lead to the privatization of the formerly state-owned British Gas. The *tariff market*, the market segment for small-scale consumers, became regulated through a price cap while the *contract market* for large-scale clients remained, in fact, a monopoly until the mid-nineties. In 1995, the reform of the Natural Gas Act opened the whole residential market for competition until 1998. The threshold consumption quantity for clients' access to the contract market was gradually lowered until 1997, squeezing out the monopolistic tariff segment. The milestone of the liberalization process was the ultimate unwinding of transportation (BG Transco)

[1] Cf. Sturm (1997), ch. 1.

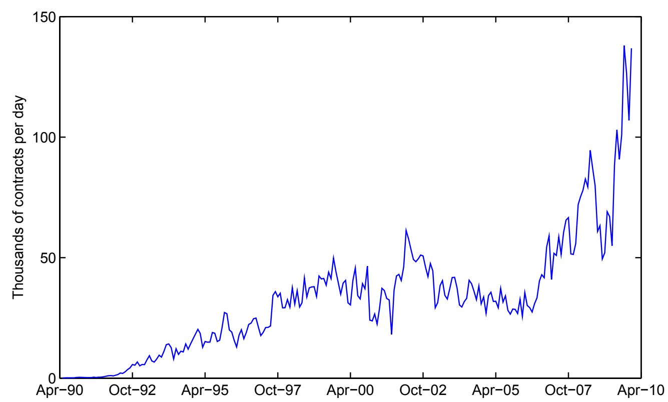


Figure 2.1: Daily volumes of natural gas month futures traded at the NYMEX (monthly averages, front month contract).

and storage from the exploration business (Centrica) in 1997.[2]

Deregulation in Continental Europe has been lagging behind. In 1998 the European Parliament and Council formulated the Directive on the Internal Market in Natural Gas I (1998/30/EC), but the speed with which member states promoted the national market opening varied considerably. Italy, the Netherlands and Spain, for instance, had already shown progressive developments until the publication of the second benchmarking report in October 2002, whereas Germany had not implemented the above Directive even by mid-year 2003. The second directive (2003/55/EC) and Regulation 1775/2003 set more ambitious targets in many respects. Free choice of supplier had to be established for all B2B customers by mid-year 2004 and for residential customers by mid-year 2007 in all member states.

---

[2] Cf. Price (1997).

Unused capacity had to be offered to third parties. Integrated gas undertakings had to legally and organizationally separate their transportation and distribution from their exploration and production businesses. However, legal unbundling of the distribution system has still been subject to a postponement option. For this reason, the liberalization process has continued to progress with different speed in the member states.[3]

On a macroeconomic level, market structures in Continental Europe remain inefficient to date. One reason is that roughly 90% of volumes contracted in these markets are long-term take-or-pay contracts. Hence, short-term volume and price risk might not always be optimally distributed between supply and demand at present. Yet, it is envisioned by a number of experts that long-term contracts will loose some importance and that a greater fraction of gas will soon be traded on exchanges. For instance, Neumann and von Hirschhausen (2004) point out that *"[. . .] the empirical evidence from the US and the UK suggests an inverse relationship between gas sector liberalization and contract length, although long-term contracts do not entirely disappear with market liberalization"* (p. 177). The demand for risk management in Continental Europe will hence continue to rise and efforts to deepen the understanding of the underlying price risk may constitute an important research contribution.

2.2 Properties of gas as a commodity

According to the IEA, natural gas accounted for 21% of global primary energy demand in 2008, and its share will increase to 25% by 2035, while oil is projected to drop from 33% to 27% in the same period. The most important rise in gas consumption in absolute terms will come from elec-

[3] Cf. Haase (2008).

tricity generation, whose share will increase from 21% of total gas demand in 2008 to 24% in 2035.[4] The main drivers of the projected growth are global climate concerns and the corresponding national commitments to reduce greenhouse gas emissions. Since gas-fired power plants operate with much lower CO_2 emissions than coal plants, some important CO_2 producers like China plan to replace an important part of their current generation capacity.

Natural gas is a very specific commodity. As opposed to many other energy carriers, pipelines are the predominant medium for transportation. Long-distance transportation necessitates repressurizing the gas after a certain distance which is an energy-intensive process. This adds to the maintenance of metering stations and pipelines such that the total transportation cost per unit of calorific value is five times higher than for oil. Since the pipeline system has a limited reach and supply and demand centers are not evenly distributed over the world, regional markets instead of a single world market have developed. This is contrary to the oil market, in which transportation by tankers is an economical alternative to pipelines and a single contract per sort is traded worldwide.[5]. Geman (2006) distinguishes three different demand regions for gas: North America, Europe and Asia. Even within one of these markets, prices can differ substantially depending on the capacity to transport gas to the location of maximum scarcity. This can lead to persistent demand-supply imbalances and hence to intra-regional price differences.

The gas price reflects these facts in different forms. If short-term imbalances in the intra-regional flows of gas occur, price jumps or spikes are a common consequence. They can especially be observed on an intraday basis. This type of imbalance is mainly due to the prominent role of gas in

---

[4] Cf. IEA (2011), GAS scenario.

[5] Cf. Geman (2006), p. 236, Shively and Ferrare (2007), ch. 4.

power generation, but can also be caused by disruptions in production or pipeline operations.[6] Medium-term intra-regional price imbalances last for a day or longer and can be due to insufficient pipeline capacity between different market areas in the considered region. A prominent example is the price difference between gas at the National Balancing Point (NBP), the main hub in UK, and the hub Zeebrugge (Belgium). Such a situation can be due to overbooking of the interconnector pipeline which links the two regions. Maintenance operations and time lags in the process of reversing the flow direction do also play an important role. Figure 2.2 shows the daily price differential between the two market areas in percent of the mean of the two prices. As can be seen, the differential is substantial overall. Moreover, the daily differentials are obviously not independently and symmetrically distributed, but skewed with clustered volatility. These inefficiencies are due to the grid-boundedness of the commodity.

Apart from intra-regional disequilibria, there are repeating patterns of price changes caused by annual seasonality in demand. In the winter, space heating demand drives up prices. In the US, prices also rise in the summer due to the prominent role of air conditioning which increases gas demand from power stations. Figure 2.3 shows the daily quotes of the UK 3-month ahead (M3) gas future traded on the Intercontinental Exchange (ICE). One can observe an annual price peak at the beginning of the winter season when the deliveries for the coldest months (January and February) are traded.

Furthermore, storage does not even out the seasonality and short-term fluctuations significantly. In fact, storage technologies in the industry are very specific because gas is stored underground. Since this is very costly, it takes a considerable intertemporal price difference to have agents use

[6] The works of Douglas and Popova (2008) and Routledge et al. (2001) look at the price relationship between gas and power in detail.

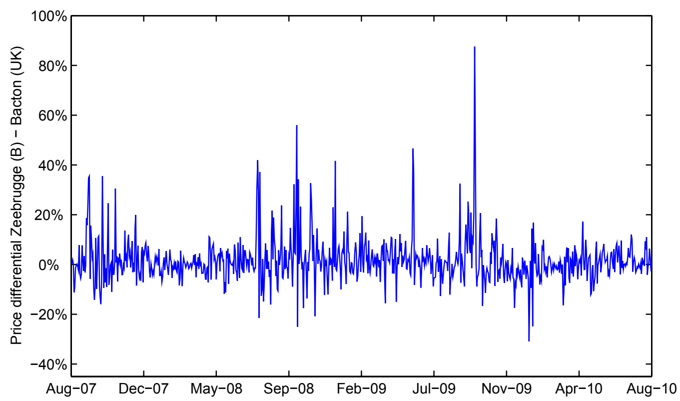


Figure 2.2: Daily price differential of day-ahead gas along the interconnector (Bacton - Zeebrugge) in % of the mean price.

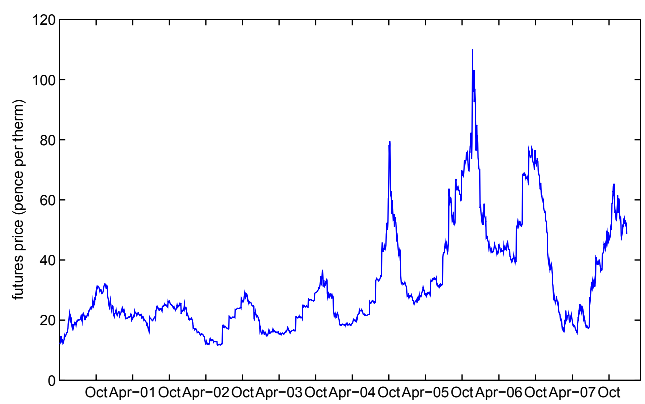


Figure 2.3: Daily prices of the 3-month ahead future in UK.

storage. In addition, the exploration of new storage space can take a long time, and the required amount of capital can be substantial. The geological formations which are commonly exploited for gas storage are depleted gas and oil fields, salt caverns and aquifers. The first type of formation is normally the largest and the least costly to develop. However, the storage cycle time is the lowest such that these facilities can only offer seasonal storage. Aquifers have higher deliverability rates and gas can be cycled more than once per season. However, they have high base gas requirements[7] and bear a high degree of geological risk. Salt caverns, in turn, have very low base gas requirements, an even higher deliverability and almost no geological risk. In turn, the exploration cost is extremely high.[8] Briefly, the high investment and operating expenses of gas storage can lead to structural shortages of available storage space in the market. If the growth of a market's storage capacity does not keep pace with the growth of demand, price volatility will increase. The same will happen if supply uncertainty rises for which the UK market is a prominent example.[9]

To summarize, this section has shown that gas is an economically relevant commodity and that some of its fundamental properties suggest an individual modelling. The most important properties in comparison to other energy carriers are the grid-boundedness, the pronounced daily and seasonal fluctuations of demand and prices as well as the costly storage technology. Next, it is explained how trading in gas markets works in order to justify the choice of markets for the empirical analysis.

---

[7] The gas which remains in a storage facility to keep pressure at a sufficient level (also called *cushion gas*).

[8] Cf. Dietert and Pursell (2000).

[9] See section 5.3, p. 64 for details on the case of the UK.

2.3 Trading in liberalized vs. regulated markets

A "commodity" in the sense of financial theory is not only defined by the nature of the good and its time of delivery, but also by its quality and delivery location. The last mentioned aspect is especially relevant for grid-bound commodities such as natural gas. This means that, for instance, natural gas somewhere in California is a different commodity than natural gas in New York City. Regarding quality, a given volume of natural gas can have a varying chemical composition. In Western Europe, there are currently two different networks for gas with high calorific value (H-gas) and low calorific value (L-gas). In the US, there is no such distinction. In addition to these properties, natural gas is sold for delivery over a period of time rather than at a point in time. This is due to the fact that the possibility of storing gas at the location of demand is usually very limited. These features are relevant for any type of supply contract presented in the following.

Given the importance of the delivery location, it becomes immediately clear that standardized and efficient trading operations are only possible if the commodity is priced with respect to a single reference point. In the natural gas market, the so-called *hubs*, i.e. interconnects of a number of long-distance pipelines, serve as such point. The reason for this is that, in any case, most of the gas traded in the respective region or market area has to pass the hub in order to be shipped from the supplier to the customer. Hubs do not necessarily exist physically, but can be virtual as well. A virtual hub is an area rather than a specific point in the network. Henry Hub in Louisiana is an example of a physical gas hub whereas the NBP in the UK is a virtual hub.

The mentioned hubs are examples from liberalized markets in which network access to deliver gas to the hub is *non-discriminatory* and *non-*

frictional. Non-discriminatory means that every market agent who wants to offer gas is allowed and able to book pipeline capacity from any entry point of the pipeline system to the hub. Non-frictional means that the cost of shipping is not excessive. The following example illustrates this criterion: Until 2007, the German transportation system was fragmented into multiple pipeline sections. The owners of these sections each required a separate shipping contract for capacity booking (called *transportation path model*). The high coordination effort in this system prevented the establishment of (a limited number of) hubs with sufficient liquidity. On the contrary, shippers in the current UK transportation system only close one contract for entry and/or exit from the whole long-distance network, hence network access in the UK can be characterized as non-frictional in contrast to the former German model.

In this study, we consider both the US and the UK as liberalized markets from which we can, with some limitations, draw conclusions about the prospective economic relationships driving a reference gas price in continental European markets.

2.4 Traded products

In this section, it is demonstrated why gas spot price models are important for the pricing of derivatives. It is shown that the common feature of all introduced products is their (at least indirect) dependence on the gas spot price estimate. In stock and bond markets the spot product is regularly the most frequently traded contract. Opposingly, virtually no commodity is delivered immediately after purchase and a true spot price does not exist. This makes the empirical exercise of pricing commodity derivatives especially challenging. Regarding natural gas, the shortest-term (standardized) product available is the so-called Within-Day contract which delivers a specified calorific amount of gas per hour for the remainder of the current gas day.[10] Its specification details can vary greatly, but one common feature is that it does not start to deliver immediately after purchase. For example, the EEX offers Within-Day gas contracts delivering from three hours after trading until the end of the gas day. Such a contract is, in fact, a futures contract, a tradable document which stipulates delivery of a given quantity of gas at a future point in time to the holder of the contract.

However, even in liberalized markets the majority of trades is currently closed over the counter (OTC). One reason for this is the fact that the era of liberalized markets is still young and liquidity in many specific contracts is insufficient for exchanges to be profitable with these products. Yet, exchanges frequently offer the clearing of contracts which are traded OTC. Hence, OTC contracts in energy markets are not necessarily free from margin requirements. Moreover, in contrast to electricity markets, there is a similar expiry specification for forwards and futures contracts in gas markets. Both contract types cease trading before the delivery period. Instead, within-day and balance-of-the-month (BOM) futures are traded as separate contracts equivalent to the already delivering futures. For this reason, the study does not make an explicit distinction between forwards and futures unless otherwise stated.[11]

Month, quarter and BOM contracts are those contracts most common on exchange. They all deliver a predefined daily quantity over the respective contract period. In terms of traded volumes, month contracts are the most important products. In fact, the highest volumes are recorded in the last week of the month, the so-called “bid week”. In this time, market

[10] A gas day is different from a calendar day and is commonly defined from 6 a.m. on the same calendar day to 6 a.m. on the following day.

[11] This simplification requires the assumption of deterministic interest rates (cf. Hull (2009) for details).

participants trade the majority of their gas requirements or available volumes for the next month on a firm basis. The payoff of a month future (or forward) with start of delivery on day T_1 and end of delivery on day T_2 is

$$\text{Payoff}_{Future} = \frac{1}{T_2 - T_1} \sum_{t=T_1}^{T_2} (S_t - F). \tag{2.1}$$

wherein F stands for the preagreed futures (or forward) price and S_t for the spot price.[12] Contracts with financial as well as physical settlement exist.

Additional gas derivatives are day contracts, basis swaps, index swaps or different options such as futures options, calendar spread options and swing options. Compared to month contracts, these products are less often traded on exchange. Day contracts are single day strips of a month futures or forward contract. A very common swap contract in the US is the futures look-alike swap (or simply "futures swap"). It exchanges a floating price shortly before expiry of the front-month future contract (e.g. simple average of the last three days' futures settlement prices, L3D) against a fixed price which is similar to the fixed price to pay for the front-month future. The only difference to the futures contract is that the settlement takes place financially, not physically (i.e. by delivering the underlying). A buyer in a (fixed-float) futures swap pays the fixed price and receives the floating price. To calculate the buyer's payoff, (2.1) has to be modified by replacing S_t above by L3D. The payoff received during the delivery period is, hence, known up-front, after the L3D price has been computed.[13]

A basis swap in natural gas markets means to exchange the value of

[12] Cf. Marckhoff and Muck (2008).

[13] Cf. Sturm (1997), p. 43ff.

gas at two different delivery points. Thus, one could also call this product a “locational swap”. Usual delivery is taken or made monthly and payoffs can be replicated by a long position in one location’s contract and a short position in the other location’s contract (both monthly with financial settlement). Given that the final settlement prices of the two forwards are fixed, only the difference D between them (the basis differential for the month) remains in the payoff equation:

$$\text{Payoff}_{Basis\,Swap} = \frac{1}{T_1 - T_2} \sum_{t=T_1}^{T_2} (S_t^A - S_t^B) + D. \tag{2.2}$$

S_t^A denotes the spot price at location A and S_t^B the spot price at location B. Since there may be many possible smaller trading points in a network or market area with centralized trading, liquidity in many of these single products is often not sufficient to allow for exchange trading.

Fixed-float index swaps (or simply index swaps) are a combination of futures swaps and basis swaps. An index is an indicative fixed price for gas in the current delivery month published at the start of the month. Usually it is the median price at which the contract traded during the bid week. The buyer of an index swap receives the index price of a certain location and pays a fixed price to the counterparty. The fixed price is usually negotiated in such a way that the initial value of the swap agreement is zero. Buying a futures swap gives exposure to the risk of the month contract at the main hub and buying a basis swap in addition hedges the price differential risk to the regional trading point under consideration.

In contrast to swaps (standard) options are instruments which entitle but do not oblige the owner to buy or sell a predefined quantity of the underlying at an agreed-upon price at a future point of time (European option) or until that point of time (American option). The underlying in a futures option is the futures contract for a particular month. The payoff of this contract at execution time t is

$$\text{Payoff}_{Futures\,Option} = max\left\{F_t - K; 0\right\}$$

with K denoting the strike price of the option and F_t the month future's price at the execution date. Natural gas futures options are traded on exchange in the US (NYMEX, American style options), but not in Europe. A common application is the insurance of fixed price risk for industrial clients. Being long in a futures call option agents can still participate in favorable decreases of the monthly futures price while capping the appreciation risk.

Calendar spread and swing options are exotic contracts. Intuitively, a calendar spread option permits the holder to exchange gas in one month (e.g. the front month, M1) with gas in a later month (e.g. the 4th next month ahead, M4). This is especially of relevance in markets which have strong seasonality such as the natural gas market. The payoff for a calendar spread option is

$$\text{Payoff}_{Calendar\,Spread\,Option} = \max\left\{F_t^{T_1} - F_t^{T_2}; 0\right\}.$$

$F_t^{T_1}$ denotes the time t-price of the future with delivery in the more recent period T_1 (M1 in the above example) whereas $F_t^{T_2}$ denotes the one with longer term (e.g. M4). Hence, the payoff of this product is again dependent on the futures price.

Swing options are special in that they ensure volume flexibility for the holder in addition to an agreed-upon price. These options are, for instance, implied in the long-term take-or-pay (TOP) contracts between gas marketers and producers. On each day or in each month during the lifetime of the option, the holder has the right to recall a quantity q_t with $m \leq q_t \leq M$. m and M are the minimum and the maximum quantity respectively. The total quantity withdrawn during the option's lifetime is often bounded by a lower limit A and an upper limit B. If the total quantity withdrawn turns out to be below A, a penalty has to be paid to the producer. The entire option payoff depends on an optimal path which defines what quantities should be withdrawn each following day until maturity. Since the optimal path is contingent on the path of the underlying spot price, the valuation of such a contract is cumbersome. The interested reader is referred to Jaillet et al. (2004) for an illustrative example.

Looking at the entirety of the above presented contracts, one sees that the payoff of many of them directly depends on the expected future spot price. Yet, even those which depend on the futures price of the commodity can be valued up-front by defining this futures price in terms of the expected spot price. Details on this methodology will be provided in the following chapter on valuation models. For now, the point to make is that *the future spot price plays a central role in valuing virtually all the derivative contracts in the natural gas market*. In this respect, the gas market is similar to other commodity markets such as the one for crude oil, for example. Together with the increasing importance of risk management in liberalized gas markets this shows that investigating the robustness and forecasting quality of gas spot price models is an important research topic.

Chapter 3

Contingent claim valuation

This chapter explains why the spot price of a commodity has an essential role for the pricing of the derivatives presented in the previous chapter. This explanation is necessary to understand the intention of the following analysis, which aims to find the most appropriate spot price model for natural gas.[1] In addition, this chapter explains how pricing formulas for derivatives are linked to the spot price to obtain an analytical solution where possible. This is one of the basics to understand the setup of any closed-form pricing formula for derivatives which is discussed throughout this study.

Derivatives on storable assets, such as natural gas, can be valued by building a replicating portfolio with the same payoff structure as the claim and known price dynamics. This portfolio normally contains the (spot) asset, i.e. the underlying. For instance, a replicating strategy for a long

[1] A discussion of what "appropriate" means in this context is part of the next chapter.

position in a forward contract is to lend the money needed to invest in the asset today and buy it on the spot market. After the expected price change over time, this portfolio must have the same value as the position in the forward contract.[2] The amount to be lent for this strategy today is the current market price of the forward contract discounted by the riskless interest rate: $F_{t,T}e^{-r(T-t)}$. $T-t$ denotes the time to maturity of the forward. The amount required to invest into the asset depends on the expectation of the future spot price at T and the risk-adjusted return k. The total amount to be invested in order to have one unit of the asset at expiry is hence $E(S_T)e^{-k(T-t)}$. If the market is efficient, current costs and benefits of the replication are equal, i.e.

$$E(S_T)e^{-k(T-t)} = F_{t,T}e^{-r(T-t)}. \tag{3.1}$$

To determine the present value of the forward contract one needs to rearrange the equation to

$$F_{t,T} = E_t(S_T)e^{(r-k)(T-t)}. \tag{3.2}$$

The forward price has now been replicated in terms of the expected future spot price and a money market account (MMA). Hence, the importance of the spot price for the valuation of derivatives is based on the fact that its dynamics are known or easier to conceive than those of the derivative itself.[3] These dynamics are "hidden behind" the expectation operator of the future spot price. To be able to calculate the futures price, we therefore need to solve the right hand side of equation (3.2) given the known or assumed dynamics of the spot price. In fact, there are two different ba-

[2] Since forward and futures prices are equal if interest rates are deterministic (cf. Hull (2009), p. 109f.), the distinction between the two prices will not matter in the following as long as this is not explicitly stated.

[3] Cf. Seppi (2002), p. 9 and Hull (2009), p. 120f.

sic methods to obtain an analytical solution for a derivative price. These methods are presented in the following two sections of the chapter. Finally, in the last section, we give an overview of more advanced methods and the associated literature.

3.1 Risk-neutral valuation method

The *risk neutral valuation principle* (or *martingale method*) goes back to Cox and Ross (1976). It changes the dynamics of the underlying such that the discounted spot price becomes a martingale under a new probability measure. This new probability measure is used to price any contingent claim. We now illustrate this method in more detail in a continuous time framework. In the example above, the risk-adjusted required return k of an investor depends on the amount of systematic risk of the asset and the individual risk attitude of the investors. This means that an objective value cannot be determined from this formula right away. However, the risk-neutral valuation principle makes it possible to value any derivative by pretending that investors be risk-neutral. That is, their individual risk attitude does not matter.

Let us assume the spot price S in the example above is driven by a one-factor stochastic process as suggested by Schwartz (1997) (one-factor model):

$$dS = \kappa(\mu - \ln S)Sdt + \sigma SdW. \tag{3.3}$$

This is a trend-stationary process wherein κ denotes the speed of mean reversion to trend μ. The deterministic drift of this process is $\kappa(\mu - \ln S)dt$. Furthermore, dW is the increment of a Brownian motion and σ the volatility. Applying Ito's Lemma and setting $X = \ln S$ yields[4]

$$\begin{aligned} dX &= \kappa(\alpha - X)dt + \sigma dW \\ \alpha &= \mu - \frac{\sigma^2}{2\kappa}. \end{aligned} \tag{3.4}$$

Parameter α now plays the role of the mean-reversion level of the asset's log-return. To value a derivative on this asset with the risk-neutral approach, the expected log-return α has to be reduced by the market risk premium λ. This means that the asset price only grows with log-return $\alpha^* = \alpha - \lambda$, which is equal to the riskless interest rate. In compensation for this reduction of drift in the deterministic part, one has to apply *Girsanov's transformation* to the stochastic part of the process: Firstly, the risk premium is added again in the stochastic part such that the underlying process remains the same. The new process is $d\tilde{W} = \frac{\lambda}{\sigma}dt + dW$, which contains a drift. Secondly, one applies a different probability measure under which the new process $\tilde{W}$ becomes driftless again. The probability measure which satisfies this condition is often called the $\mathcal{Q}$-measure as opposed to the empirical measure $\mathcal{P}$. Harrison and Kreps (1979) and Harrison and Pliska (1981) have shown that in an arbitrage-free market at least one risk-neutral measure $\mathcal{Q}$ exists. $\mathcal{Q}$ can be obtained from $\mathcal{P}$ by the *Radon-Nikodym derivative*. Formally, for the new stochastic increment, it holds that $E^{\mathcal{Q}}(d\tilde{W}) = 0$ as opposed to $E^{\mathcal{P}}(d\tilde{W}) = \frac{\lambda}{\sigma}dt$. In other words, the measure change simply relocates the probability mass such that $d\tilde{W}$ becomes a martingale and $\tilde{W}$ a standard Brownian motion under $\mathcal{Q}$. The $\mathcal{Q}$-measure can then be applied to the pricing of contingent claims without explicit knowledge of the investors' risk attitude and their required return k and all future cash flows can be discounted by the risk-free interest rate.

Once the risk-neutral measure is known, it is easy to derive a price

[4] See e.g. Wiersema (2008), p. 110f. for a detailed derivation of this solution.

for the forward contract (3.2), given the assumptions about the spot price dynamics (3.3). Under the risk-neutral measure, the required return k will be equal to the risk-free rate. The exponent in (3.2) therefore cancels out so the forward price must simply equal the expected spot price under $\mathcal{Q}$, i.e.

$$F_{t,T} = E_t^Q(S_T).$$

Since the mean-reversion process (3.4) is defined for the log-spot price, one obtains

$$E_t^Q\left(S_T\right) = E_t^Q\left(e^{X_T}\right) = e^{E_t^Q(X_T)+\frac{1}{2}Var_t^Q(X_T)}.$$

Now, only the first two moments of the log-spot price under the risk-neutral measure are required. The change of measure does not affect the variance. This means that the variance is the same under $\mathcal{P}$ and $\mathcal{Q}$. Mean and variance of the stochastic differential equation (SDE) with mean reversion are known[5] to be

$$E_t^Q(X_T) = e^{-\kappa(T-t)}X_t + (1-e^{-\kappa(T-t)})\alpha^* \tag{3.5a}$$

$$Var_t(X_T) = \frac{\sigma^2}{2\kappa}(1-e^{-2\kappa(T-t)}). \tag{3.5b}$$

Therefore, the value of the forward contract results as follows:

$$F_{t,T} = \exp\left(e^{-\kappa(T-t)}X_t + (1-e^{-\kappa(T-t)})\alpha^* + \frac{\sigma^2}{4\kappa}(1-e^{-2\kappa(T-t)})\right). \tag{3.6}$$

3.2 Partial differential equation method

This method was developed by Black and Scholes (1973), using key insights by Merton (1973). For this method Ito's Lemma is directly applied to the

[5] For the derivation see Wiersema (2008), p. 110f.

SDE of the contingent claim. The general idea is that the claim and the underlying on which it is written must depend on the same risk factors. The result is a partial differential equation (PDE) which has to be solved to obtain the claim's price in closed form.

Assuming again that spot prices follow process (3.3), the change in the forward price results from changes in time and in the spot price and can, hence, be written as

$$dF = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S}dS + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}(dS)^2. \tag{3.7}$$

By substituting dS and dS^2 one obtains the SDE

$$dF = \left(\frac{\partial F}{\partial t} + \frac{\partial F}{\partial S}\left[\kappa(\mu - \ln S)S\right] + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S^2\right)dt + \frac{\partial F}{\partial S}\sigma S dW \tag{3.8}$$

wherein $\kappa\mu$ can be interpreted as the total asset return and $\kappa \ln S$ as an "external" yield component, which does not result in a price appreciation of the asset.[6] Furthermore, we know that borrowing on the MMA and investing in the asset simultaneously can be used as a hedging strategy for the forward position. This information can be used to derive the PDE. Let B be the value of the MMA with

$$B_t = \exp\left\{\int_0^t r_s ds\right\} \tag{3.9}$$

and ϕ as well as ψ portfolio weights. Since, by assumption, the external

[6] The "external" yield can be, for instance, a convenience yield (explained in chapter 5) or a dividend yield.

yield can be reinvested, the value change of portfolio P is given by

$$\begin{aligned} dP &= \phi dS + \psi dB \qquad (3.10)\\ &= \phi\left[\kappa(\mu - \ln S)Sdt + \sigma SdW + \kappa \ln S S dt\right] + \psi rBdt. \end{aligned}$$

The initial net investment of the replication strategy equals zero, i.e. $\phi S + \psi B \stackrel{!}{=} 0$. Therefore, we can substitute $\psi = \frac{-\phi S}{B}$ and simplify (3.10) to

$$dP = \phi[\kappa\mu S - rS]dt + \phi\sigma SdW.$$

The replicating portfolio must have the same deterministic and stochastic components as the forward contract, i.e.

$$\begin{aligned} \phi[\kappa\mu S - rS] &= \frac{\partial F}{\partial t} + \frac{\partial F}{\partial S}\kappa(\mu - \ln S)S \qquad (3.11a)\\ &\quad + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S^2 \\ \phi\sigma SdW &= \frac{\partial F}{\partial S}\sigma SdW \qquad (3.11b) \end{aligned}$$

Substituting $\phi = \frac{\partial F}{\partial S}$ from (3.11b) into (3.11a) yields the following PDE:

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial S}(r - \kappa \ln S)S + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma^2 S^2 = 0. \qquad (3.12)$$

This is the fundamental PDE of the forward price, which can be solved with the boundary condition $F_{T,T} = S$ to obtain the closed-form price formula (3.6). Yet in general, solving PDEs is only possible under restrictive conditions and there is no comprehensive solution scheme available. Solutions for some common types of PDEs, including (3.12), can be found in Evans (1998).

We have now presented two standard analytical methods to derive a closed-form pricing formula for a derivative contract. More advanced methods are not discussed here in detail, since they are not of immediate relevance for the main analysis. Nevertheless, the following final section of the chapter gives a brief overview of them and names sources for more detailed information.

3.3 Advanced techniques

Only a limited number of models can be solved by the standard methods described in the previous two sections. Especially when more complicated (e.g. non-linear, path-dependent) derivatives are to be priced or coupled SDEs are involved, these techniques are often insufficient to yield a solution. A more general analytical method for affine models is described in Duffie et al. (2000). It is somewhat more involved since it makes use of Fourier transforms. The idea is to transform the expectation, compute the transform function explicitly and apply an inversion formula in which only a single integral has to be computed numerically. For an intuitive description of the method and an application to option pricing, we refer to Muck (2006).

If an analytical solution is inefficient or impossible, one has to revert to numerical methods. One possibility is to use *Monte Carlo Simulation*, pioneered for this purpose by Boyle (1977) and later applied, among others, by Johnson and Shanno (1987), Hull and White (1987) and Duffie and Glynn (1996). The idea is to simulate paths of the risk-factors which determine the payoff of the derivative contract under the risk-neutral measure. From the realizations one can then compute the corresponding payoffs at maturity. In the simplest case, the expected payoff under the risk-neutral measure is simply the unweighted arithmetic mean of all payoff realizations.[7] From this quantity the price is easily computed by discounting by the risk-free interest rate where applicable.

Another possible numerical method is *finite differences*. This method is applied to the pricing of options, for instance, by Schwartz (1977), Brennan and Schwartz (1978) and Courtadon (1982). It makes use of the partial differential equation of the model, e.g. (3.12), and of the corresponding terminal condition, e.g. $F_{T,T} = S$. A grid with as many dimensions as risk factors and the time dimension is set up. Starting from the nodes of the terminal point in time, the value of the derivative at each node of the grid is computed by working backwards in time. The information set to do so contains the values of the derivative one time-step ahead and the difference quotients of the partial derivatives at the current node. By letting the differences become very small, i.e. by subdividing the mesh, the price at the current state of the derivative will approach the analytical solution.

Finally, *binomial or multinomial trees* can be used. They have been first studied by Cox et al. (1979) and Boyle (1986) for derivative pricing. The process of the underlying is described by a tree consisting of nodes and branches. Nodes are pricing points as in the finite-difference schemes. Branches describe the dynamics of the underlying over time. Figure 3.1 shows the simplest form, a binomial tree with recombining branches. With the risk-neutral probabilities (*here*: p and $1-p$) and the possible payoffs at maturity on the lowest line of nodes, one can recursively determine prices at antecedent nodes until the upmost node is reached and the current price is known. Again, by letting the time steps become very small, the discrete distribution of the risk-factor approaches the dynamics of the continuous process. Applications of (trinomial) tree procedures are found, for example,

[7] More recent studies such as Avellaneda et al. (2001) and Glasserman and Yu (2005) also allow to weight the paths (Weighted Monte Carlo) which offers more flexibility.

in Hull and White (1994a), Hull and White (1994b) and Muck and Rudolf (2005).

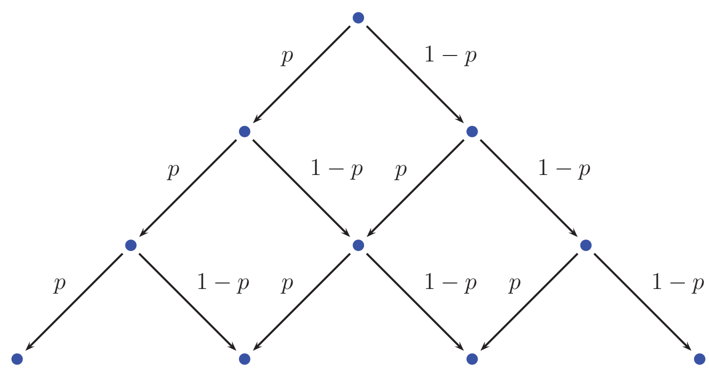


Figure 3.1: Binomial tree scheme.

Briefly, a number of advanced techniques exists to circumvent complications arising with the methods described in the previous sections. For more detailed information on these numerical methods we recommend, among others, Hull (2009), chapter 19 and Epps (2007), chapter 8.

In this chapter, we have demonstrated the importance of the expected future spot price and the spot price dynamics for the valuation of (commodity and other) derivatives. Previously, in chapter 2, we have explained that natural gas is a commodity which deserves individual attention for modelling. Consequently, the following chapter describes the existing spot price models which might be used for natural gas. It also explains the choice of the particular model which will be adapted to the gas market.

Chapter 4

Models of spot and forward prices

Numerous models for commodity prices can already be found in literature. Therefore, the first objective in this chapter is to set out criteria for an appropriate model. Next, existing models are classified according to these criteria and the most influential models are reviewed, whereby the research gap regarding gas spot price models becomes evident. Finally, conclusions from this review for the appropriate design of a gas price model are drawn which lays the groundwork for the main part, i.e. chapters 5 and 6.

4.1 Requirements for an appropriate model

We start the analysis of existing models by defining a list of model selection criteria. The aim is to capture the most important points to consider whenever a pricing model has to be chosen for valuation purposes. The first criterion is the *completeness* of information input. That is, a model should take account of existing insights into the price-building mechanism from previous work. For example, it has been established that the price of a flow commodity depends negatively on the rate of production. If a model takes account of such an economic relationship, it is less likely to attribute a price movement to the wrong cause and, thereby, less prone to misspecification. Another important quality feature is the *(in-sample) goodness of fit*. A model should always be able to adequately reproduce observable patterns in historical prices. For example, the price volatility of futures contracts on storable commodities is negatively related to the time to maturity of the future which is known as the *"Samuelson effect"*.

Furthermore, an important model requirement is *robustness*. This means that the model has to perform well in forecasting exercises, possibly over different time horizons (e.g. from one to three periods ahead). Robust also means that the fit remains viable if the frequency of input data is altered. Data frequency for model estimation can range from daily to monthly or even yearly time periods. Eventually, the insensitivity to outliers or stress scenarios is another desirable aspect of robustness. The last criterion for an appropriate model is just as important as robustness: The *ease of implementation*. Very complex models have the disadvantage of causing numerical estimation problems ("curse of dimensionality") and long computation times. Evidently, a model has to be updated regularly to take account of the most recent market information. If the applicable estimation algorithm consumes excessively much time or even breaks down, new forecasts cannot be produced in due time. Besides estimation time, another critical issue can be that the produced estimates are not sufficiently reliable due to numerical inaccuracy.

These criteria, besides of their possible incompleteness, are naturally interrelated. For instance, the ease of implementation is often negatively related to completeness and positively related to robustness. Nevertheless, we believe that they constitute a first guideline for the clarification of the most important strengths and weaknesses of the existing pricing models. These models will be presented in the remainder of the chapter.

4.2 Classification of existing models

Econometric models can be distinguished along two dimensions. The first one defines the formal layout: Some models are derived directly from microeconomic quantities such as the supply and demand of individual agents. These models are called *structural-form* models. In contrast, *reduced-form* models state relationships between the variables of interest directly. The simplest model of the latter type is a linear regression, formally given by $y = Xb$ with X denoting the data matrix and b the coefficient vector.[1] The main advantage of structural models is their comprehensive picture of the economy which reduces the risk of misspecification. In turn, reduced-form models tend to be less complex and, thereby, easier to calibrate with empirical data. In fact, certain structural models can be converted to the reduced form.[2]

The second dimension determines the implied assumption about the efficiency of markets. The first type, which is called *equilibrium model*, borrows ideas from Neokeynesian theory. Market equilibrium conditions in the model, but short-run deviations from equilibrium due to temporary shocks and frictional readjustments are allowed. Due to this possibility the model does not necessarily need to be in equilibrium at current market

---

[1] A linear regression is a particular case in which the model is not only *reduced*, but also *solved*, i.e. no derivatives of the endogenous variables remain in the model.

[2] Seppi (2002), p.48f.

prices, which is considered as a disadvantage by some practitioners.[3] The second type is known as *no-arbitrage model*. It enforces consistency with current market prices by using them as an input for parameter estimation. These models revert to neoclassical theory by assuming frictionless and efficient markets. Equilibrium deviations are corrected by market forces instantaneously. Again, some simple equilibrium models can be converted to no-arbitrage models by including a function of time into the stochastic process for the uncertain economic quantity (e.g. for the instantaneous market interest rate). Yet, enforcing fit to current prices comes at the cost of time inconsistency. That is, if the model is recalibrated, the new parameter estimates might significantly differ from the previous ones. Another inherent risk is the enforcement of fit to contaminated current prices.

The following presentation of commodity spot price models will distinguish along the first dimension and present the relevant models in their order of complexity. It has to be noted that applying the following models to the gas market might necessitate certain adjustments, such as removing seasonality (or possibly jumps) in prices up-front. In how far this applies depends on the type of contract and the frequency of observations considered. Details are provided later on in the analysis.

4.3 Reduced-form models

The earliest reduced-form commodity price model was designed for contingent claims on futures by Black (1976). It is conceptually similar to the well-known model for stock options by Black and Scholes (1973), but specifies a process for the futures price, not the spot price (under the $\mathcal{P}$-

[3] Cf. Hull (2009), p. 678.

measure):

$$\frac{dF_{t,T}}{F_{t,T}} = (\mu - r)dt + \sigma dW_t,\ t \leqslant T. \tag{4.1}$$

$F_{t,T}$ denotes the futures price at time t with maturity in T. μ defines the total required return for an investment in the asset.[4] dW is the increment of a Wiener process and σ a scaling factor. The above process is called a geometric Brownian motion (GBM), since price changes are always proportional to the level the price. The model is not limited to applications in commodity markets, but has also been used for bond prices, for instance. As can be seen from the process equation, the change in the futures price is modelled as a geometric Brownian motion.The problem is that under this assumption the volatility of futures returns converges to infinity with increasing time to maturity. This feature is inconsistent with empirical observations of many commodities including natural gas. While it is possible to adjust the portfolio growth rate for a constant convenience yield[5], futures and spot price movements are assumed to be perfectly correlated and only a single maturity of futures is considered. Therefore, the model has a particular weakness with respect to the goodness-of-fit criterion.

Brennan and Schwartz (1985) develop a model for extractable resources which contains an explicit relationship between futures and spot prices. They also introduce the convenience yield as an additional model parameter to explain changes in the basis, i.e. differential of futures and spot prices. The general relationship between the futures price F and the spot price S originates from Ross (1978):

$$F(S, T-t) = Se^{(r-\delta)(T-t)}. \tag{4.2}$$

---

[4] This convention is kept throughout the thesis. Note that in the discussed papers μ sometimes denotes the required return *net of the implicit convenience yield*!

[5] A convenience yield would be conceptionally identical to a dividend yield as in Ross (1978).

r is the constant riskless interest rate of an investor and δ is the constant (net) convenience yield in % of the spot price which accrues to the holder of the commodity. The proportional relationship of the (absolute) convenience yield to S is a shortcut to the assumption that this availability premium depends negatively on inventories and that the spot price will be high whenever inventories are low.[6] The stochastic process for the spot price is a geometric Brownian motion:

$$\frac{dS}{S} = (\mu - \delta)dt + \sigma_1 dW_1. \tag{4.3}$$

According to the risk neutral valuation principle, the log-futures price can be derived in closed-form:

$$\ln F_{t,T} = \ln S_t + (r - \delta + \frac{1}{2}\sigma_1^2)(T - t). \tag{4.4}$$

However, no closed-form solution for any contingent claim is given by the authors. Despite the enhancement of this model concerning the relationship of spot and forward prices, the term structure of futures volatilities is again misspecified as in Black (1976). In addition, changes in the curvature of the term structure are still not reflected in the model.

These issues are resolved in Gibson and Schwartz (1990) who develop a valuation model for crude oil investments. The convenience yield in this model is mean-reverting and stochastic. Therefore, it is referred to as the *stochastic convenience yield model* (SCY model). The authors illustrate that the convenience yield empirically has a mean-reverting pattern, most likely because inventories influence the value of convenience.[7] The spot

[6] See e.g. Brennan (1958). More information on the convenience yield's relationship to inventories is provided later in the analysis.

[7] Details on the theory of storage are provided in chapter 5.

price dynamics are taken from (4.3) and the convenience yield therein is specified as an Ornstein-Uhlenbeck process:

$$d\delta = \kappa(\alpha - \delta)dt + \sigma_2 dW_2. \tag{4.5}$$

κ denotes the mean-reversion speed, α the mean-reversion level and σ_2 the scale factor of Brownian increment dW_2. The increment is correlated with the spot price increment with coefficient of correlation ρ. Despite the greater flexibility of this model compared to the previous one, the simplifying assumption of a constant market price of risk for the convenience yield still constitutes a restriction in the proper pricing of energy futures such as crude oil as shown by the authors. While the model fits short-term futures prices well, pronounced mispricing occurs for long-term maturities. Nevertheless, it seems to offer a reasonable trade-off between the goodness-of-fit and the robustness criterion established above.

Schwartz (1997) provides the closed-form solution for the valuation of futures contracts in this two-factor model. He also shows that futures price volatility returns in this model converges to a finite value with increasing time to maturity. The model is estimated for datasets of crude oil, copper and gold futures. Estimation problems appear regarding the market premium for convenience yield risk and the empirical drift of the spot price. These parameters cannot be properly identified since only the risk-neutral dynamics are directly observed. That is, the empirical parameters can only be estimated implicitly from futures prices. Another problem specific to the oil futures data is the high (estimated) correlation of the stochastic increments dW_1 and dW_2 of the spot price and the convenience yield. Relating this fact to our model selection criteria shows that the trade-off between goodness of fit and robustness in the SCY model should be enhanced.

For this reason, the model is restated in a more general mathematical form by Schwartz and Smith (2000). This version does not use coupled

SDEs, but two independent additive stochastic components for the (log-)spot price:

$$\begin{aligned} \ln S_t &= \xi_t + \chi_t \\ d\xi &= \mu_\xi dt + \sigma_\xi dW_\xi \\ d\chi &= -\kappa\chi + \sigma_\chi dW_\chi . \end{aligned} \tag{4.6}$$

ξ denotes the long-term drift component whereas χ denotes the short-term mean reverting component of the spot price. This model is mathematically equivalent to the Schwartz (1997) model, but has one parameter less. The authors also provide a closed-form solution for European options on futures contracts. In the empirical part of the paper, it can be shown that the correlation of the stochastic increments is lower than in the SCY model with the same oil futures dataset. However, the parameter identification problems turn out to be the same as before. The authors admit that this will affect the quality of price forecasts. It is explained that whether or not the convenience yield is explicitly specified, it takes very long time series of futures prices to get reliable estimates for the empirical dynamics of the spot price. The authors propose a different estimation method to deal with this problem. We will return to this issue during the course of our study. Furthermore, both the Schwartz (1997) and the Schwartz and Smith (2000) model have been criticized for overpricing the value of option contracts with long maturities systematically (see, for instance, Carlson et al. (2007)). According to the critics, the reason for this overpricing is that, in reality, price volatility is not proportional to the spot price but conditional on a number of fundamental variables.

Models with stochastic volatility respond, among other things, to this criticism. For instance, Eydeland and Geman (1999) combine a one-factor mean-reversion model discussed in Schwartz (1997) with the Heston (1993) model to create the following two-factor model for gas and electricity prices:

$$
\begin{aligned}
dS &= \kappa(\alpha - \ln S)Sdt + \sqrt{V}SdW_1 && (4.7\text{a}) \\
dV &= b(\theta - V)dt + \sigma_V\sqrt{V}dW_2,\ E(dW_1 dW_2) = \rho dt. && (4.7\text{b})
\end{aligned}
$$

Again, the spot price reverts to level α with speed κ. However, stronger movements in the spot price will be reflected in the stochastic volatility term $\sqrt{V}$. The volatility itself is a mean reverting stochastic process and the Feller condition ensures that $V > 0$ in (4.7b). Yet, mathematically, square root diffusion processes are complicated to handle since V is no longer normally distributed. An even more serious concern is, again, the incompleteness of certain energy markets. Liquid options to hedge the volatility risk are not always available. For example, natural gas futures options are exchange-traded only in the US, and liquidity in longer-term contracts is often insufficient. Another problem is the negligence of fundamental information on the futures basis. As shown by Fama and French (1987, 1988) and others, there is a strong relationship between the level of volatility and the size of the basis. Briefly, the model might perform well in-sample due to its high flexibility, but the completeness criterion might not be sufficiently met to ensure reliable estimates of the volatility risk premium.

Yet a different source of randomness has been considered already a long time before by Merton (1976). The author introduces jumps in the stochastic process:

$$\frac{dS}{S} = \mu dt + \sigma dW + UdN. \tag{4.8}$$

This famous jump-diffusion model features random price jumps with Poisson intensity λ_P. Whenever the Poisson increment dN_t takes the value of 1, a jump with size U_t occurs. The latter variable is normally distributed such that jumps can be positive as well as negative. The model was originally designed for the stock market only, but jumps also occur in electricity and gas prices in the intraday and day market (cf. chapter 2). Yet, apart from the above mentioned problems of the GBM process for energy markets, a key problem of this model is the necessary assumption of risk neutrality with respect to jump risk. In Merton's setting, jump risk cannot be hedged because the stock and a risk-free bond are the only traded assets. In fact, empirical evidence is not in line with the assumption of risk neutrality. An additional problem is that all jumps in the model are followed by a pure diffusion period. Frequently in reality though, up-jumps are immediately followed by reverse jumps, a phenomenon which is called a *spike*. More recent models along the same strand of literature account for this phenomenon, but these models are again more complicated such that robustness becomes an even more critical issue. Whether or not it pays to include jumps into the stochastic process will generally depend on the desired data frequency, since many commodity prices revert fairly quickly after a jump. In the gas market, the majority of currently traded derivatives has monthly delivery periods such that especially intradaily time series are less frequently needed for model calibration. Therefore, the modelling of jumps is not in the primary focus of this study. For further information on this strand, the reader is referred to recent studies for the electricity market, e.g. by Huisman and Mahieu (2003), Geman and Roncoroni (2006), Seifert and Uhrig-Homburg (2007) and Marckhoff and Muck (2009), as well as to a bond market study by, among others, Björk et al. (1997).

Extensions to the two-factor models are presented, among others, by Cortazar and Schwartz (2003). The authors develop a three-factor model which is an upgrade to the SCY model with the trend factor as the third stochastic variable. The trend also has mean-reversion properties and a mean value $\bar{\mu}$. Its contemporaneous value in is related to the total required asset return μ of the SCY model by $\tilde{\mu} = \mu - \alpha$ with α as the equilibrium convenience yield. In compensation, the modified convenience yield variable $\tilde{\delta}$ reverts to zero ($\tilde{\delta} = \delta - \alpha$).

$$\begin{aligned} dS &= (\tilde{\mu} - \tilde{\delta})Sdt + \sigma_1 SdW_1 & \text{(4.9a)} \\ d\tilde{\delta} &= -\kappa\tilde{\delta}dt + \sigma_2 dW_2 & \text{(4.9b)} \\ d\tilde{\mu} &= a(\overline{\mu} - \tilde{\mu})dt + \sigma_3 dW_3 & \text{(4.9c)} \end{aligned}$$

The model parameters are estimated from the risk-neutral version of these equations with further simplifying modifications in order to reduce the number of parameters. The paper provides and compares in-sample and out-of-sample estimates of the model. Interestingly, it is not benchmarked against the simpler two-factor model hence it remains unclear whether this more complicated version of the Schwartz (1997) model is also more robust. What makes the study interesting instead is the simplified estimation algorithm, termed *"implied methods"* compared to the more elaborate *Kalman filtering* procedure used in Schwartz (1997) and Schwartz and Smith (2000). We will investigate both algorithms in more detail in chapter 6.

Another three-factor model by Cassasus and Collin-Dufresne (2005) releases the constraint of time-invariant risk-premia which is implied in the two-factor models presented. The risk factors are the spot price, the convenience yield and the short rate. The short rate process is mean-reverting as in Vasicek (1977) with mean-reversion speed a and equilibrium rate m:

$$dr = a(m - r)dt + \sigma_3 dW_3. \tag{4.10}$$

Empirically, the authors find a positive relationship of both the short rate and the spot price to the convenience yield. Higher interest rates make storage more costly which c.p. drives down inventories. The theory of storage, discussed extensively in the next chapter, predicts that convenience yields rise in this case. While a likelihood-ratio test shows that the coefficients for time variation in risk-premia and for the endogenous influence on convenience yield are significant, an explicit comparison of model performance with the two-factor models is not provided. Model 3 in Schwartz (1997) also includes the short rate as an additional stochastic factor. It is shown that the improvement in goodness of fit compared to the two-factor model is very small since convenience yield variation is much more important. In turn, numerical implementation of these models is tedious and parameter estimates are probably not very stable over time.

Finally, some reduced-form models also contain regime-switches of the spot price. For instance, Ribeiro and Hodges (2004) implement a model in which the spot price can follow two distinct processes depending on the state of inventories. The key criticism of these authors towards the Schwartz (1997) model is that it might temporarily violate the cost-of-carry constraint because in a risk-neutral world the discounted futures price net of carrying cost should always be smaller or equal to the current spot price. In their model, they take this restriction into account: As long as the drift of the spot price remains below a critical level S^*, investors buy and hold inventory and the spot price has a drift equal to the marginal cost of carry, $r + c$:

$$dS_t = (r + c)S_t dt + \sigma S_t dW,\ S_t \leqslant S^*. \tag{4.11}$$

Once the price surpasses the critical level, all inventory is sold and the price starts to follow a mean reverting pattern which is equivalent to the one-factor model in Schwartz (1997)

$$dS_t = \kappa(\alpha - \ln S_t)S_t dt + \sigma S_t dW,\ S_t > S^*. \tag{4.12}$$

The model is implemented with a trinomial-tree procedure. It turns out that some of the stylized facts of commodity forward curves are more accurately represented by this model than by a one-factor model. One of these facts is the high kurtosis in spot price distributions. However the authors admit that some misspecifications become apparent as well. For instance, when the initial spot price is very low and the market is in contango[8], it takes an excessively long time until the market price reaches the equilibrium long-run price (in the mean-reverting regime). The reason for this is the cap imposed on the growth rate in the GBM regime. Another problem is that the model produces left-skewed price distributions which is contrary to empirical observations for most commodities.

Other regime-switching models are usually more complicated and are used in the gas market mainly if storage decisions are the major focus of the exercise such as in storage facility valuation problems.[9]

4.4 Structural models

Structural models give a micro-founded picture of the economy and derive commodity prices from the relevant supply and demand conditions. Early structural commodity models are developed by, among others, Williams and Wright (1991), Deaton and Laroque (1992), Chambers and Bailey (1996) and Deaton and Laroque (1996). Unlike most reduced-form models, these models take into account the role of inventory levels in determining the convenience yield and specify the latter one as an endogenous variable. This allows for a non-linear relationship of inventories to spot and forward prices and also to the basis, such that inventory changes become the driver for heteroscedastic price variations, which are empirically observed.

For example, in Deaton and Laroque (1992), the economy is charac-

[8] A market is in *contango* whenever the term structure of futures prices is increasing in the time to maturity whereas it is in *backwardation* when the term structure is falling.

[9] Cf. e.g. Chen and Forsyth (2006).

terized by two regimes: one in which contemporaneous stock levels I_t are greater than zero and one in which there is a stockout:

$$I_t \geqslant 0, \text{ if } \frac{1}{1+r}(1-\gamma)E_t(S_{t+1}) = S_t \tag{4.13a}$$

$$I_t = 0, \text{ if } \frac{1}{1+r}(1-\gamma)E_t(S_{t+1}) < S_t. \tag{4.13b}$$

r denotes the risk-free interest rate, γ the rate of storage depreciation (or equivalently storage cost) in percent of the stock levels I_t and E_t the expectation operator for time t. Hence, inventories are non-empty whenever the discounted expected value of the stored commodity equals the current spot price. While I_t is endogenous, production quantity Z_t is stochastic. The state variable which determines the current equilibrium price is the "amount on hand" $X_t = Z_t + (1-\gamma)I_{t-1}$.[10] Under rational expectations, it must hold that

$$S_t = f(X_t) = \max\left\{\frac{1}{1+r}(1-\gamma)E_t\left[f(X_{t+1})\right](S_{t+1}), P(x_t)\right\} \tag{4.14}$$

where P is the inverse demand function. Prices are then asymmetrically distributed due to the maximum function. A particular disadvantage of these early models is that they are specific to agricultural commodities which are harvested once a period. In addition, the model is not able to give account of term structure dynamics. Moreover, the high autocorrelation in empirical data cannot be explained.

Routledge et al. (2000) formulate a generalized model of this kind with the same basic model setup. An extension is the introduction of a forward price curve. Forward prices are assumed to revert to a long-run level, i.e.

[10] The market clearing condition enforces an intertemporal restriction on I_t. Precisely, current demand $D(S_t)$ must always be met by net supply: $Z_t + (1-\gamma)I_{t-1} + I_t = D(S_t)$.

long maturities are independent of the current demand state. While the above reduced-form models with mean reversion imply that the Samuelson effect holds irrespective of the economic conditions, violations of this effect can occur in Routledge et al. (2000). This is the case whenever the storage level is very high such that any demand or supply shock can be offset by storage reactions. In this situation, spot price volatility can be lower than futures price volatility. This conforms to empirical observations of many commodity prices.[11] That said, the empirical performance of this model with oil price data is mixed. In particular, the market is twice as often in backwardation as the model predicts. In fact, a study by Litzenberger and Rabinowitz (1995) shows that the frequent backwardation scenarios are a systematic pattern of the crude oil market (or, more generally, of commodities which serve as an input to production rather than an asset). Hence, the negligence of such a fact clearly constitutes a drawback.

Cassasus et al. (2005), among others, address this problem in a specific oil market model which defines an economy with oil supply and production of a consumption good (the numeraire). Another major change with respect to the previous models is that agents are assumed to be risk-averse instead of risk-neutral. Therefore, market risk premia of oil and their variation can be studied in the model. The production technology of the consumption good requires both the consumption good c itself and the commodity o as inputs. The (isoelastic) production function is

$$f(o,c) = \alpha o^n c^{1-n}. \tag{4.15}$$

Consumption demand is endogenous and the supply of oil is modelled as a resource extraction problem with costly and discretionary investments in additional extraction capacity. The dynamics of the stock of (explored)

[11] It is demonstrated later that this line of reasoning only holds for commodities for which the total storage capacity in the economy does not have a strict upper bound.

oil O_t and the amount of the consumption good C_t available at t are characterized by

$$\begin{aligned} dO_t &= -(\bar{\imath}+\gamma)O_t dt + \sigma_O O_t dW_{O,t} + X_t dN_t && (4.16a)\\ dC_t &= (f(\bar{\imath}O_t, C_t) - C_t)dt + \sigma_C C_t dW_{C,t} - \zeta dN_t. && (4.16b) \end{aligned}$$

$\bar{\imath}$ denotes consumption, γ the depreciation rate of the oil stock and N_t a Poisson process such that $dN_t = 1$ when an investment occurs at time t and $dN_t = 0$ otherwise. Furthermore, X_t is the amount of oil explored when an investment is made at t and $\zeta(X_t, O_t, C_t)$ is the cost incurred for an investment, measured in units of the consumption good. Variables σ_O and σ_C are standard deviations.

Two factors characterize the forward price term structure: the contemporaneous spot price and the "investment regime". The equilibrium spot price for oil S_t is determined by the ratio of the marginal value of oil wells to the marginal value of the consumption good

$$S_t = \frac{\partial J/\partial O}{\partial J/\partial C}. \quad (4.17)$$

O_t denotes the stock of explored oil and C_t the amount of the consumption good available at t. $J(t, C, O) = \sup_{\mathcal{C};\mathcal{A}} E_t\left[\int_t^\infty e^{-r(s-t)}U(C_s)ds\right]$ is the aggregate value function of the economy - the supremum of the present value of future utility over a set of allowable consumption policies $\mathcal{C}$ and admissible investment strategies $\mathcal{A}$. Each time, an investment in additional oil wells is undertaken, the marginal value of oil $\partial J/\partial O$ drops and, thereby, the spot price does as well. Discretionary investments in new wells are undertaken whenever the relation of explored oil to consumption good undershoots a certain threshold.

The second factor which impacts on the term structure of forward prices is the "investment regime". When the economy is in the "far-from-investment regime", the market expects the oil price to rise in the near future, since ongoing consumption will lower the available reserves of oil through production (contango or hump-shaped structure). In turn, when the economy is in the "near-investment regime", the market expects the price to drop soon, due to the opening of new wells (backwardation or trough-shaped structure).

The most important merit of the model is that it can be calibrated such that it features long backwardation periods as they are observed in the market. The authors also show that time-variation in the market risk premium of oil returns can be partly explained by the above mentioned investment regimes. However, the econometric performance is difficult to evaluate since the data are aggregated to quarterly frequency which is not useful for many practical risk management applications. In addition, only a subset of the parameters is estimated due to the computational burden of the chosen method (simulated method of moments and a numerical technique to solve a Hamilton-Jacobi-Bellman differential equation). Therefore, the model's practical applicability is limited.

Finally, a more recent structural model is presented by Carlson et al. (2007). One difference to the aforementioned study is that the assumption of unlimited resource availability is relaxed. Instead, the model applies the Hotelling (1931) rule for extraction of exhaustible resources, i.e. in equilibrium the resource is depleted such that prices grow with the risk-free interest rate. The economy has four state variables: demand, reserves, marginal extraction cost and historic production rate and one continuous choice variable, the current production rate. The paper analyses settings with different types of demand and technology innovations. The key feature, however, is the inclusion of a production adjustment cost. Due to this cost, small demand shocks are not followed by production adjustments. These shocks will subsequently die off with a time lag due to mean reversion in demand. If a shock is large, in turn, the production reaction will overshoot in order to keep adjustment costs small. This will dampen the effect of the shock on prices such that large temporary price shocks occur less frequently. The authors demonstrate empirically that the non-linear price sensitivity to demand is more realistic than the constant sensitivity which is understated, for instance, in the Schwartz and Smith (2000) reduced-form model. The effect is especially relevant for long-term derivatives which depend on price volatility. In fact, the authors show that long-term options are significantly overpriced if the Schwartz and Smith (2000) model is used. In turn, the computational burden of implementing such a multidimensional model for short-term applications does most likely not pay off.

4.5 Summary of review and contribution

The main insights from the survey of reduced-form and structural models can be summarized as follows: Early reduced-form models are easy to implement, but they make unrealistic assumptions concerning the sensitivity of the long-run price expectations to price changes at the short end of the term structure. The consequence is that these models will strongly misprice long-term derivatives whose value depends on the volatility of the underlying. More recent models, especially those with two risk factors, are more flexible, but it was found that the assumption of a constant market price of risk made in these models is still restrictive and long-term (real option-type) investments are still priced inaccurately. Models incorporating jumps are more complicated to derive and estimate. The right type of discontinuity of the model (jump with or without mean reversion or spike) might be crucial for the appropriateness. However, this type of model is primarily useful if the commodity is grid-bound and the price series considered has an intradaily or daily frequency of observations. Models with more than two risk factors regularly have even better in-sample results due to the greater flexibility they offer. In turn, their out-of-sample performance (robustness) is seldom superior to that of a two-factor model.

Structural models provide microeconomic justifications for the empirically observed price dynamics. They also relax some of the assumptions made in reduced-form models that are inconsistent with economic theory. In general, the completeness-of-information criterion suggested above is met to a higher extent in these models. The contrary holds, in turn, for the ease-of-implementation criterion. In fact, the specification of supply and demand functions is sometimes arbitrary since they are unobservable and the coefficients of influential factors cannot be estimated explicitly.

We have shown that no specific commodity price model for natural gas has been developed to date. In addition, a trade-off between the strengths of reduced-form and structural models has to be found. Reduced-form models tend to neglect some important fundamental properties of the commodity under consideration while structural models are often rather complex. It has been demonstrated above that one parameter appears in both reduced-form and structural models: the convenience yield. It builds a link between the financial and physical markets of an economy and is driven by fundamental factors. In addition, it has already received much research attention and is, hence, well founded in economic theory. Criticism about the high correlation between estimated convenience yields and spot prices in reduced-form models might be partly due to an incomplete model setting and the fact that the convenience yield has not been determined independently from empirical data. Additionally, we recall that we are not interested in finding the best model to fit a particular set of futures prices, but a model for the spot price that will ultimately serve for the pricing of a variety of derivatives in the natural gas market. The contribution of this thesis is, firstly, the identification of fundamental convenience yield drivers and, secondly, their incorporation into the SCY model which permits us to benchmark empirical results against those generated by the conventional SCY model. These two aspects is each dedicated one of the following chapters.

Chapter 5

Fundamental convenience yield model*

In the previous chapters of this thesis, we pointed out the economic importance of natural gas as a resource and the necessity to have a spot price model specifically for the gas market. The last chapter, in particular, looked at existing commodity models, taking account of their strengths and shortcomings. It turned out that no specific model has yet been proposed for natural gas. In addition, no study has attempted, so far, to combine the advantages of structural and reduced-form models in a hybrid model framework. In fact, the convenience yield links the financial market and the goods market and it appears in both model classes.

Therefore, this chapter analyses the fundamental drivers of convenience yield, with the ultimate objective to combine a fundamental convenience

*This chapter is based on Volmer (2011).

yield forecast with a reduced form model (see chapter 6). The first section of the chapter is a recap of the theory of storage. In the following sections, existing empirical findings for the convenience yield in the gas market are reviewed before testable additional drivers are discussed. Next, the data is described and prepared and, finally, the empirical results are presented and evaluated.

5.1 The theory of storage

As mentioned in the introduction, the seminal work for the theory of storage is Kaldor (1939). It says that due to the double role of many commodities as both a consumption good and an asset, the holder of a commodity can profit from the flexibility to put the commodity to use whenever needed. Since buying on the spot market and selling forward must be a riskless deal, a *'cash-and-carry trade'*, the marginal cost of carry net of the convenience yield must be equal to the difference between the forward price F and the spot price S

$$F - S = R + C - Q. \tag{5.1}$$

R, C and Q stand for the marginal cost of capital, the marginal cost of storage (incl. wastage) and the marginal convenience yield respectively. Now, intuition suggests that abundance of storage should lower the premium for immediate availability. This would imply a positive relationship between the marginal convenience yield and the spot price because stocks will be abundant when the commodity is cheaply available. This indirect relationship between the spot price and the convenience yield works, according to Brennan (1958), through the demand and supply in the market for storage. Brennan proposes an equilibrium model for commodity inventories in which the marginal convenience value of a good is a decreasing function of the aggregate stocks held in the economy. The demand for storage is implicitly given by

$$\begin{aligned} S_{t+1} - S_t &= f_{t+1}(D_{t+1}) - f_t(D_t) \\ &= f_{t+1}(I_t + Z_{t+1} - I_{t+1}) - f_t(I_{t-1} + Z_t - I_t). \end{aligned} \tag{5.2}$$

S_t denotes the spot price and D_t consumption demand at time t. $f_t(D_t)$ is the inverse demand function, I_t the commodity's inventory carried over to the next period $t+1$ and Z_t the exogenous production level. The price spread on the left-hand side is a decreasing function of I_t. The equilibrium condition in the economy is that the marginal revenue of storage $u_t^{'}$ equals the marginal cost $m_t^{'}$

$$u_t^{'}(I_t) = m_t^{'}(I_t). \tag{5.3}$$

In an atomistic competition, $u_t^{'}$ must equal the expected change in price given by (5.2). The net marginal cost of storage is given by

$$m_t^{'}(I_t) = C_t(I_t) - Q_t(I_t) + P_t(I_t) \tag{5.4}$$

where in addition to the quantities defined above P denotes the marginal inventory risk premium which is discussed later in the analysis. For now, note that $C_t^{'}(I_t) \leq 0$, i.e. the marginal convenience yield is assumed to be decreasing in the storage levels. This means that the first unit of inventory should have a very high convenience yield, but the yield will diminish for every additional unit.

A formal but intuitive illustration of this concept is found in a recent model by Ribeiro and Hodges (2004), which accentuates the buffering function of storage in the process of equating supply with demand. The authors assume an Ornstein-Uhlenbeck stochastic process for the production rate of a commodity and analyze both a monopolistic and a competitive market for storage. In this market stock owners buy up the whole production and decide what fraction to store and to sell later and what fraction to sell immediately to consumers. This decision depends on an optimal stocking plan which is updated dynamically at every point in time and wherein the new optimum depends on the current rate of production. The aim of this plan is to smooth the effective supply in order to maximize intertemporal profits from storage. Since the Gaussian supply process implies that higher deviations from the expectation become more and more unlikely the greater the deviation, the first few units of stock will be used much more frequently than the remainder. In result, every additional unit of stock must, in fact, have a decreasing convenience yield. Another insight from the paper is that shocks to production should generally be absorbed by the storage facilities, a fact that will also be relevant for this study.

The relationship of convenience yield to stock levels has been verified in numerous commodity markets since Brennan's article was published. A natural question is how these studies deal with the fact that the convenience yield is not directly observable, but implicitly results from the cost-of-carry equation. In addition, the marginal physical stocking cost is regularly not known for any storage level. In fact, it is most often assumed that C simply is linearly proportional to the asset's price S such that $C_{t,T} = c_{t,T}S(T-t)$ with storage duration $T-t$. Similarly, one can define $q_{t,T}$ to be the rate of convenience yield, a percentage of the asset price. Therefore, since Fama and French (1988) many econometric studies have used an equivalent of

$$c_{t,T} - q_{t,T} = \frac{F_{t,T} - S_t}{S_t} - r_{t,T} \tag{5.5}$$

or its negative to analyze the variation in convenience yield. (5.5) is simply (5.1) rearranged and restated with respect to a particular current date t and a maturity of the forward contract T. $r_{t,T}$ is the applicable percentage cost of capital (i.e. the interest rate). $c_{t,T} - q_{t,T}$ then denote what is commonly labeled the *'net storage cost'* in percent of the spot price. and is computed from the observable variables on the right hand side, the *'interest adjusted basis'* (IAB).

When we differentiate the body of empirical convenience yield literature by the type of commodity analyzed, we find that natural gas has so far received little attention. This is important to note since comparable studies for other commodities are only of limited value for the gas market. This is due to the unique characteristics of gas outlined in section 2.1, even in comparison to other energy commodities such as oil, coal or electricity. Hence, the insights of convenience yield studies in other industries are not directly transferable to the gas market .[12]

The determination of basis and convenience yield in the gas market is analyzed, with more or less focus, in Dincerler et al. (2005), Modjtahedi and Movassagh (2005), Wei and Zhu (2006), Cartea and Williams (2008), Haff et al. (2008) and Stronzik et al. (2008). Yet, these papers do not go beyond the traditional theory of storage outlined above. Dincerler et al. find that the convenience yield falls at a decreasing rate as national gas inventory levels rise. Similarly, Modjtahedi and Movassagh (2005) discover a positive relationship between gas storage levels and the futures' basis which is the main source of variation in the interest adjusted basis. The curvature of this functional relationship is not consistent over all the maturities analyzed. Wei and Zhu (2006) test whether deviations from 5-year average storage levels can explain changes in the convenience yield and find a moderate significance. In addition, spot price shocks turn out highly significant in explaining convenience yield so that the R^2-values of their model are strikingly high. However, this model is probably not apt for forecasting since the spot price shock is measured as the residual from an ARMA-process of the log spot price and this residual will be difficult to project. Moreover, we believe that the significant influence of the shock is

[12] Those who are interested will find a comprehensive list of these studies in Gao and Wang (2005).

not overly surprising because the spot price also entered into the calculation of the explained convenience yield variable itself.

Cartea and Williams (2008) price multiple interruptible supply contracts in the UK gas market with the Schwartz and Smith (2000) model. They present a short qualitative analysis of what drives the short-term price component of Schwartz and Smith which is directly related to the convenience yield. The focus of the paper is on the contract valuation, however. Haff et al. (2008) test the influence of storage levels and interest rates on the basis of forward contracts along the lines of Modjtahedi and Movassagh (2005). The paper confirms a positive influence of storage levels on the basis. Yet, we find some of their various model specifications not sufficiently founded on economic intuition and the authors do not state a clear preference for any of them. Besides this, the spot prices used to compute the basis of the futures contracts seem to conflict with the cash-and-carry strategy. The use of survey-based price estimates and a relatively short time series could be additional minor problems.

Finally, a study by Stronzik et al. (2008) investigates the validity of the cost-of-carry relationship in (5.1) with a methodology originally applied in Fama and French (1987). The actual drivers of q are not studied. The authors find that the seasonality of the basis is not entirely consistent with their expectation, and that the basis does not vary one for one with the nominal interest rate, which is contradictory to the theory of storage. In fact, two conceptual problems should have impacted on the result: Firstly, the study uses the 6-month and 12-month futures contract for the basis, where the latter one is extremely illiquid in the UK. A second problem is that daily observations of the basis are calculated with monthly maturing futures contracts whereby convenience yields of different maturities are mixed up in one sample. Even if the term structure of the convenience yield was flat, this procedure could not be absolutely accurate because the convenience yield used is actually not annualized, but specific to the time to maturity as is the interest adjusted basis in equation (5.5).

Overall, the existing studies in the gas market consider storage levels an important determinant and do also provide mixed evidence for a non-linear dependence of the convenience yield on this variable. Evidence for other fundamental drivers or for the robustness of the proposed relationship has not been looked at yet. This should make apparent that the mentioned contribution in the first part of this study clearly goes beyond the existing body of literature. We will continue by justifying the choice of variables for the following econometric analysis.

5.2 Testable drivers

The previous section already pointed out the empirical importance of stock levels as a determinant of the convenience yield. Hence, this driver should doubtlessly be part of a convenience yield model for gas, in particular due to the strong seasonality of gas demand. However, recently made assumptions about the precise way in which stock levels reveal information about the convenience yield are debatable. Ribeiro and Hodges (2004) and also Cartea and Williams (2008) assume that there is some dynamically updated optimal stocking policy made by the owner(s) of storage, coupled with the ability to smooth the marginal convenience value over time. In consequence, Cartea and Williams (2008) argue that the convenience yield will not react in response to the absolute storage level but only to deviations from the expected seasonal level because seasonal changes in stocks are already incorporated in the optimal plan. Capacity restrictions on stock inflow and outflow is the only mentioned reason for deviations from the plan and for the resulting fluctuations in convenience yield.

We argue against this assertion. Even with unconstrained operational capacity the storage owners do not have full control over the market price at the central hub. One should consider, for instance, that the decisions of all owners of inventory are made independently and that there is a time lag from withdrawal to delivery as well. In addition, not only inflow and outflow capacity but also the total storage capacity is restricted. This holds especially true in the UK where, until recently, stocking capacity only covered 14 days of average winter demand.[13] In whatever way the convenience yield will react when storage facilities become completely replenished or depleted: One should try to take account of these further restrictions which is not possible using deseasonalized data. For these reasons, our study first reverts to the more established theoretical work and uses absolute (not deseasonalized) stock levels in the main part of the analysis. We expect a negative relationship between convenience yield and stocks as it was already found in previous studies.

While storage impacts on the convenience value by affecting the aggregate supply, the scarcity of a good is a relative measure and is also influenced by demand changes. If short term demand rises, the convenience value of the gas in stock must increase because production rate is sticky in the short-run and the additional demand must be met from stock to keep the market in equilibrium. However, to the best of our knowledge, no study has yet tested any variable that immediately relates to gas demand. Instead, one has been contented to use indirect variables for the short term imbalance of supply and demand such as the spot price shock in Wei and Zhu (2006).

Therefore, we point out that heating accounts for the largest fraction of gas consumption in many countries. The most important share usually comes from domestic and commercial space heating activity. In the UK for example, this share accounts for roughly 40% of demand on a yearly

[13] Cf. POST (2004).

average (cf. Appendix 1). The high fraction of heating consumption makes aggregate demand highly seasonal and sensitive to temperature. Therefore, temperature should be among the most important variables influencing the convenience yield because a fall in temperature - from below a certain level - means increased heating demand. In US based studies, it has become common to measure this influence with the so-called heating degree days (HDDs). They represent the difference between a comfortable room temperature, e.g. 20°C, and the prevailing outside temperature on a daily basis. The opposite indicator, cooling degree days, also play a role in the US due to the widespread use of air-conditioners. They increase, particularly in the summer, the demand from gas power stations.[14] In the UK market, average day temperature does rarely surpass this threshold, so plain temperature observations will lead to the same result. For the econometric test of this temperature-convenience yield relationship, we exclusively look at the UK data at the outset. The hypothesis is that the relationship is negative since an increasing demand will increase the scarcity of gas in the short run before any production adjustment can ease the situation again.

Finally, an interesting question is whether the crude oil price or convenience yield should have a bearing on the convenience yield of gas. On the one hand, this might be considered especially because a recent cointegration study by Panagiotidis and Rutledge (2007) for UK crude oil and gas has shown that the two price series did not decouple after the UK gas market liberalization. In addition, Cassasus et al. (2009) find co-movements of convenience yields with the price of a related commodity in pairs of oil products. Hence, it is possible that this relationship includes natural gas. Bailey and Chan (1993) document common elements in the basis variability of different commodities other than gas, due to a shared sensitivity to stock and bond markets. On the other hand, the conclusion that a short

[14] Cf. Shively and Ferrare (2007), p. 27.

run relationship for natural gas exists is not directly possible. Empirical observations in the UK market, e.g. by ILEX (2004), show that the gas price has historically reacted to oil price changes with a time lag of about 9 months. Besides this, Panagiotidis and Rutledge (2007) cannot find a significant impulse response of the gas price on oil price shocks in a simulation exercise. In addition, the crude oil price is not seasonal as is the gas price. Finally, we do not look at the level of the gas price but at the (negative) interest adjusted basis. We do actually not have a strong reason to assume that the oil price should directly affect the scarcity of gas in the short run. Merely the demand from electricity plants might be substitutable on this horizon, but crude oil itself is not the immediate substitute in this case. Briefly, there are arguments to both views on this question. Regressions controlling for an impact of crude oil can help to verify whether or not such an influence is present in the market. Our a priori expectation is that the results will be negative and, hence, in line with the last two cited studies, which have specifically looked at natural gas.

5.3 Data description and preparation

The study of convenience yield drivers is, at the outset, limited to the UK market primarily because the generation of a representative "market temperature" is rather involved for geographically extended markets such as the US. We measure convenience yield as the negative of the interest adjusted basis in (5.5). Since there is no real spot market price for gas, as it is the case for many other commodities, the nearest-to-maturity futures can be used which are usually contracts with a delivery period of a single day (*'day contracts'*). The most liquid day contracts are traded in the On-the-Day Commodity Market (OCM) for delivery at the National Balancing

Point (NBP), the most active UK gas hub.[15] This official market segment is run by APX-ENDEX who provided daily weighted average prices for the NBP OCM Title Day contract for the time span between 04/00 and 12/07.[16] The most liquid UK futures are the monthly contracts traded at the Intercontinental Exchange (ICE). They stipulate constant daily delivery volumes of 1,000 therms deliverable at the NBP during a specified month. These contracts start trading up to one year ahead of delivery and expire on the second last business day before the delivery month. We choose this contract as the futures contract in the interest adjusted basis.

Moreover, the monthly contract immediately before expiry (the *'front month contract'*) can be used as an additional proxy for the spot price. This proxy is slightly less accurate than the OCM contract since it delivers gas during the whole next month, but it offers the possibility to extend our econometric analysis to additional time series. Since the contract expires at a monthly frequency, the sample time series will also have a monthly frequency in order to generate a series of convenience yields with an equal term. The data provided by the ICE extends from 01/97 to 01/08 and consists of daily weighted average prices and volumes. We compute six time series of the convenience yield wherein the OCM price serves as the spot price and five more time series wherein the front month price is used instead.

In October 1998 the Interconnector pipeline for gas exchange between Bacton (UK) and Zeebrugge (B) was opened and several studies hypoth-

[15] The OCM market is mainly used by National Grid, the network operator, for settling network balancing trades with the shippers of physical gas.

[16] The average is calculated per *gas day*. Each gas day (D) lasts from 6:00a.m. on the same calendar day to 6:00a.m. on the following day. Contracts start trading at 12:00 noon on the day before the start of delivery (*D-1 or D1*) and cease at 3:35a.m. on the calendar day after the start of delivery. This implies that the daily average quote contains the prices of 2 different contracts, the D-1 and the D contract.

esize that this has led to a structural break in the market. Although we expect the distortion caused in the basis to be small, we exclude observations before this date.[17] To calculate monthly 'spot' and forward prices, we pick the last five trading days before the maturity of the front month future. These days are considered as the *'bid week'*. It is well known that the majority of gas trades occurs in the last week before the month-end, so this procedure assures that quotes which enter into the monthly price are backed by sufficient volume and contain less noise. We exclude all monthly futures with maturities greater than the one of the 6th-month ahead contract (M6) because of insufficient volumes during the bid week. For the remaining time series, we compute a volume weighted average of the month contract prices and a simple average of the OCM prices. A full history of daily volumes was not obtainable in the latter case.

To complete the data necessary to calculate the negative interest-adjusted basis, sterling LIBOR interest rates with different maturities, dating back to 10/98, are used. The data are available on the homepage of the British Bankers Association. We recall from equation 5.1 that interest in the cost-of-carry relationship accounts for the opportunity cost of the tied-up capital. At the same time, we note that the cash-and-carry trade implied in our convenience yield calculation is actually done with a month future, meaning that an equal fraction of our invested capital is unwound every day throughout the delivery month. Strictly speaking we would, hence, need to account for this by calculating a different opportunity cost for each day of delivery. Instead, we simplify the problem: We compound interest for all delivery days of the monthly contract until the middle of the month and use the LIBOR rate with maturity closest to this date, in fact, the 1-month to 6-month rates. The error committed is economically negligible since only the interest on interest within a single

[17] A further discussion on the potential impact of the interconnector on our study is found in the concluding section of this chapter.

month is affected.

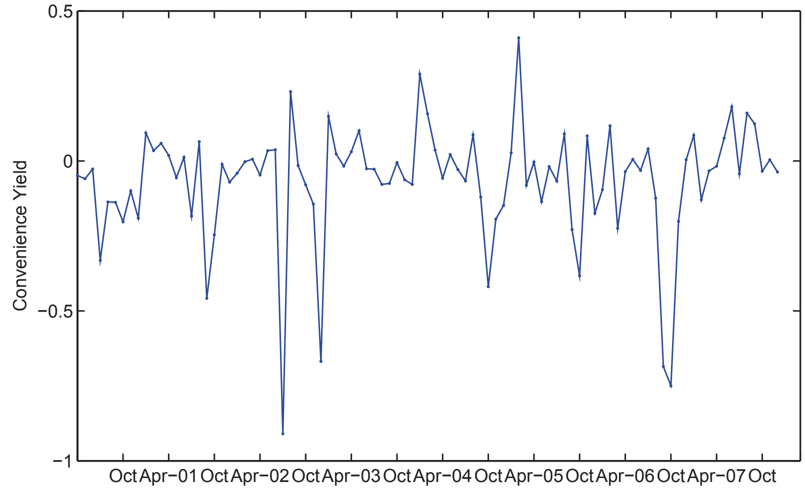
As for the inventory data, the International Energy Agency (IEA) provided monthly closing stock levels of working gas from 10/98 to 01/08. Working gas is the fraction of total inventory which can effectively be released and delivered. As explained in the last section, we refrain from deseasonalizing storage figures for now. Our monthly temperature history dates back to 10/98 and comes from the UK Office of Meteorology. In fact, determining a meaningful frequency of temperature observations is critical and involves a trade-off: On the one hand, the measured temperature on the single day when someone enters into a cash-and-carry trade should be the most relevant one for the spot gas demand and the convenience yield on that day. On the other hand, temperature on a daily basis is more difficult to forecast far ahead in time which would, however, be a desirable property of our explanatory variable. Monthly mean temperature observations spanning the start dates of the cash-and-carry trade explain demand at that particular date less precisely and a larger part of the model fit is explained seasonal variation. Yet, these values are certainly more stable and are predictable farther into the future. Hence, we start our analysis with monthly average temperature, i.e. the average in a certain month is used to explain the convenience yield measured in that month. In this set-up any seasonal variation in the convenience yield does not necessitate a separate treatment since it is directly explained by its fundamental causes, varying storage levels and temperature.

A preliminary look at the time series and unit root tests (augmented Dickey-Fuller test, Phillips-Perron test) indicate that the computed convenience yield and the temperature are stationary time series, whereas the stock levels have a linear time trend in addition to the expected seasonality pattern. This time trend is most likely due to a structural phenomenon of the market: Namely, it is apparent that for several years the domestic gas resources are shrinking and the UK is gradually becoming more dependent on imported gas. Since a market can react to a demand shock more quickly by increased domestic production than by increased imports, storage volumes need to grow over time if the share of home production in total supply diminishes.[18] This means that the annual increase in average stock should not lead to a lower convenience value. To solve this problem, we remove the linear trend from the series by an OLS regression of stock levels on time. The residuals from this regression compose a stationary and centered series which will be used in the analysis.

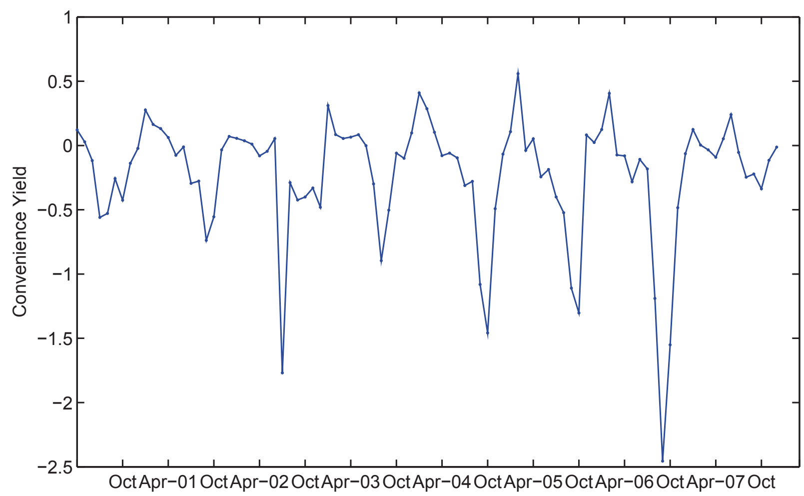
For a descriptive analysis of the convenience yield we select two of the eleven time series. The first series we look at contains the price of the front month contract (M1) as the future price and the OCM price as the spot price. In the second series M1 is replaced by the 3rd-month future (M3). The front month contract is the future which starts delivering right after the turn of the month, whereas the latter one starts two months further ahead. Both convenience yield series are shown in Figure 5.1. A look at the graphs reveals that sharp drops have been somewhat more frequent than strong increases. The annual seasonality pattern of convenience value is more clearly visible in the second time series. This is because market participants consider a fading of the spot price shocks within two month plus a few days more likely than the fading within only the few days until the M1-contract starts to deliver. Therefore, the M1 price moves more closely together with the spot price such that the overall variability of the shorter term convenience yield in Figure 5.1 is smaller.

Some of the most remarkable price reactions have been fostered by extraordinary events: Operational problems of the interconnector pipeline are likely to be responsible for the sharp decrease observable in July 2002. As

[18] This can be verified by comparing the total gas storage capacity of countries with (formerly) autonomous gas supply (e.g. UK, NL) with the capacity of highly import dependent countries (e.g. GER, FR, IT). See Appendix 2 for more details.



Basis: M1-OCM



Basis: M3-OCM

Figure 5.1: Selected time series plots of the convenience yield in UK.

the flow direction had to be reversed during the course of a cleaning action, gas was imported in the UK even though prices on the mainland had already been higher than in the UK. Consequently the UK spot price dipped down. In 2006, a very warm September and October temporarily brought the spot price under pressure. The pronounced drop in convenience yield at that time occurred when the new Langeled pipeline bringing Norwegian gas to the UK was tested at full capacity simultaneously.

The reason for the sharp rise in February 2005 does not seem to be caused by any extraordinary event. It has already been discussed by Cartea and Williams (2008) who observe an increased short-term risk premium at this time. They argue that the winter had initially been very mild such that the temperature drop in February came as a surprise. This meant that stock levels had already been depleted too much compared to the seasonal mean. Our data confirm the observation with regard to temperature, whereas the conclusion made for stock levels cannot be verified. Evidently though, in this situation as well as during the periods discussed before the storage facilities did not absorb the shocks to the scarcity of gas. This is visible from the remarkable reaction of the convenience yield. We will pick up this phenomenon during the course of the following analysis.

5.4 Basic model

To test for the drivers of convenience yield, we start with an OLS regression of the negative IAB from (5.5), here labelled y_t, on stock levels and temperature. The estimation equation thus reads

$$y_t = \beta_{cons} + \beta_{stock} x_{stock,t} + \beta_{temp} x_{temp,t} + \epsilon_t \tag{5.6}$$

with $x_{stock,t}$ denoting the detrended stock levels as explained on p. 65 (originally in bcm[19]) and $x_{temp,t}$ the monthly mean temperature, both demeaned. This regression is run separately for each of the eleven different bases. They are obtained by using the chosen futures contracts together with the two different spot price proxies, the price from the OCM and the front month future (M1).

The regression results are in Table 5.1. The column entitled "Basis" denotes the contracts used to compute the negative interest adjusted basis, i.e. the convenience yield variable. For example, "M3-M1" means that y_t is given by $-\frac{F_t^{M3}-S_t^{M1}}{S_t^{M1}} + r_{t,t+\frac{2}{12}}$. The middle columns present the point estimates for the coefficients in equation (5.6) as well as their t-statistics and significance levels in brackets. *, ** and *** denote significance at the 10%, 5% and 1% level respectively. Since the residuals are first-order autocorrelated and heteroscedastic, Newey-West standard errors have been used. The second last column reports the p-value from White's heteroscedasticity test, and the last column shows the adjusted R^2-statistic. The number of observations is $N = 93$ in equations 1 to 6 and $N = 112$ in equations 7 to 11. Overall, the explanatory power of the model is strikingly high, and the coefficient values and their t-statistics are in line with the predictions. The regression coefficient for the stock variable has a negative sign, indicating that the marginal convenience yield diminishes when storage levels increase. As can be seen from the t-statistics, this relationship is strongest when the respective futures contracts are close to maturity, while it decreases when the time to maturity becomes higher. The intuition for this effect is that storage can be replenished or depleted in the long run, so current storage levels have a decreasing predictive power for the convenience yield on a longer horizon. This also justifies the insignificant β_{stock} in regressions 6 and 10, and only the significantly positive coefficient of

[19] Billion cubic meters.

No.	*Basis*	$\hat{\beta}_{cons}$	$\hat{\beta}_{stock}$	$\hat{\beta}_{temp}$	p_{White}	R^2_{adj}
1	M1-OCM	-.0686 (-3.90***)	-.0859 (-5.46***)	-.0033 (-1.05)	.094	14.8%
2	M2-OCM	-.1478 (-4.67***)	-.1855 (-5.45***)	-.0194 (-2.70***)	.017	35.2%
3	M3-OCM	-.2178 (-5.14***)	-.2287 (-5.14***)	-.0467 (-4.86***)	.012	42.7%
4	M4-OCM	-.2616 (-5.95***)	-.1907 (-4.31***)	-.0787 (-7.79***)	.022	52.7%
5	M5-OCM	-.2916 (-6.90***)	-.0860 (-2.16**)	-.1069 (-10.18***)	.086	60.6%
6	M6-OCM	-.3128 (-6.63***)	.0437 (1.00)	-.1175 (-10.41***)	.303	57.3%
7	M2-M1	-.0568 (-5.15***)	-.0786 (-7.33***)	-.0147 (-5.17***)	.000	49.5%
8	M3-M1	-.1104 (-5.83***)	-.1095 (-5.77***)	-.0388 (-8.25***)	.001	60.8%
9	M4-M1	-.1508 (-7.22***)	-.0733 (-3.52***)	-.0673 (-13.04***)	.000	74.2%
10	M5-M1	-.1867 (-7.51***)	.0167 (0.68)	-.0934 (-14.36***)	.000	76.1%
11	M6-M1	-.2153 (-6.25***)	.1268 (3.73***)	-.1038 (-12.88***)	.004	68.1%

Table 5.1: Estimation results of basic convenience yield model (5.6).

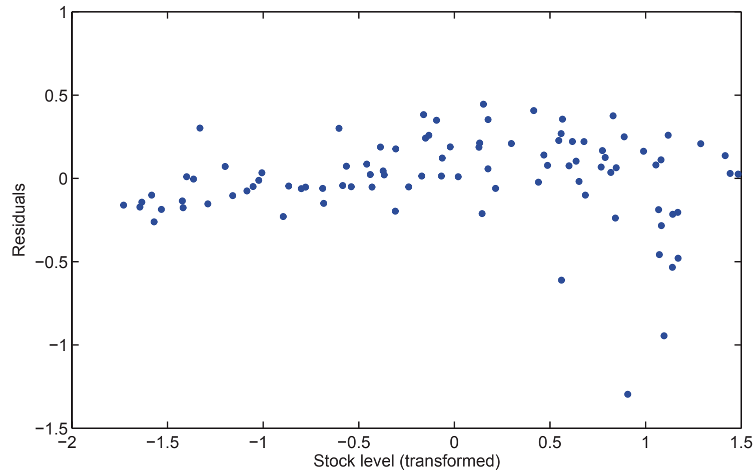
regression 11 remains unexplained. The absolute coefficient values of this variable start decreasing with longer maturities from regression 4 and 9 onwards. They increase from regressions 1 to 3 and from 7 to 8 because we measure the maturity matching convenience yield instead of the annualized one so the increase in the time dependent yield slightly counters the effect of diminishing predictive power of stocks over time.

The temperature coefficient is also in line with the posited hypotheses. The coefficient has a negative sign in all regressions and its absolute value increases when time to maturity becomes longer. The reasons for this phenomenon are straightforward: Firstly, while current temperature has a significant and negative effect on current demand and the spot price, its indication of the temperature at the futures' maturity becomes weaker the more the maturity date is ahead in time. To illustrate this effect, let us consider regression 1: Since the front month contract starts to deliver right after the turn of the month under consideration, the front month price reacts almost identically to temperature as does the spot price. Therefore, the differential between the two prices is not dependent on temperature and the respective coefficient is insignificant. Now compare this to the reaction on temperature in regression 6. The OCM price will again react negatively to a temperature increase, but the demand in 5 months plus a few days should hardly be influenced by current weather conditions, thereby making the M6 price insensitive to current temperature changes. Hence, the differential $\frac{-(F-S)}{S}$ will diminish when temperature increases. Secondly, the maturity matching convenience yield leads again to more pronounced absolute changes when the maturity of the considered futures increases.

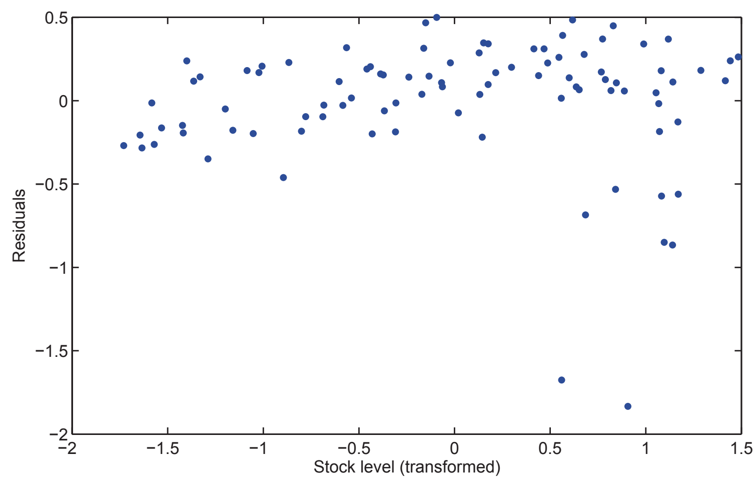
The intercept coefficient is significantly negative in all cases. In theory the linear relationship of the convenience yield to both regressors separately should have a positive intercept, considering that the convenience yield itself is strictly positive. However, the stock coefficient is only a residual term since we have made the series stationary and use also a demeaned temperature coefficient. Additionally, the cost of physical storage is still figured into the dependent variable y_t, and we do not know how sizeable it is compared to the convenience yield. Thus, the observed intercept is not contradictory to theory. Based on the coefficient of determination, the model seems to be very powerful overall, but it performs weaker when the short end of futures maturities is regarded and when OCM prices are used for the spot price (regressions 1 to 3). The p-values from the White test document the strong heteroscedasticity of the residuals. It is necessary to further investigate its source.

The graphical analysis of the model residuals shows that a number of severe outliers is present as could be expected from Figure 5.1 already. These outliers are most likely the cause of the changing error variance and can potentially bias parameter estimates. Interestingly, most of them occur in the upper third of stock level observations as is visible from Figure 5.2. Plotted against time, the residuals display the most remarkable outliers between late summer and late autumn, one exception being the previously mentioned incident in July 2002 (Figure 5.3). This is the period when storage levels are highest each year, i.e. the time before the heating season starts and stocks are withdrawn.

The simultaneous occurrence of high storage levels and drops in the residual could be a misspecification that arises from the simplified measurement of convenience yield according to equation (5.5). Taking a more differentiated view on the net storage cost, Brennan (1958) provides two possible explanations for the observed phenomenon. Firstly, he hypothesizes that the marginal outlay on physical storage might start increasing once inventories are sufficiently close to the capacity limit of storage. That is, the scarcity of available storage space can potentially drive up the net storage cost. Secondly, Brennan suggests that besides the cost of physical storage and convenience yield, a risk premium is figured into the marginal

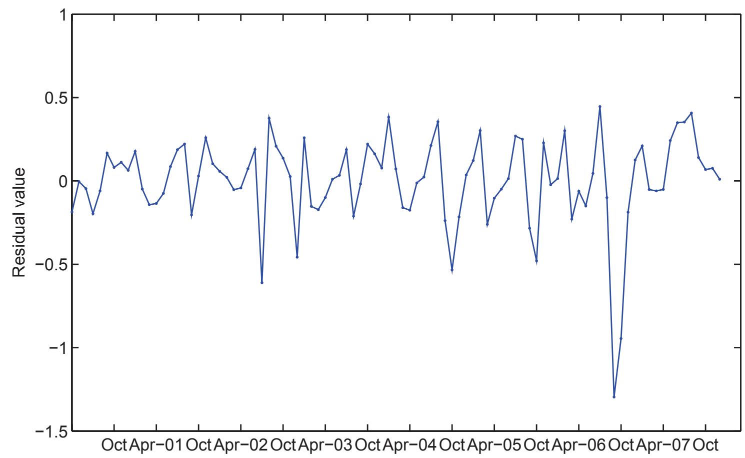


Basis: M2-OCM

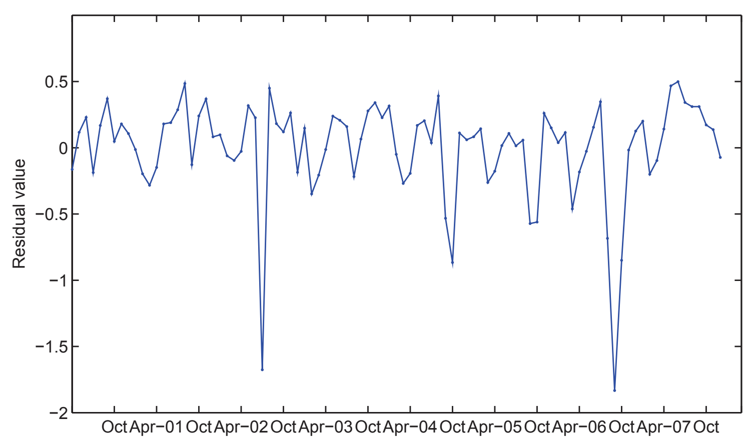


Basis: M4-OCM

Figure 5.2: Selected plots of residuals from basic convenience yield model against detrended stock levels.



Basis: M2-OCM



Basis: M4-OCM

Figure 5.3: Selected plots of residuals from basic convenience yield model against time.

cost of storage, and this premium increases strongly when inventories approach the capacity limit. It is interpreted as a compensation for the risk of price depreciations of goods in stock.

In the natural gas market this storage risk premium can equally be regarded as a compensation for foregone arbitrage opportunities as already noted by Cartea and Williams (2008). We recall that intertemporal arbitrage is done by buying at a low spot price and selling at a higher forward price at the same time, thereby making use of available storage capacity until the forward contract matures. Now, by putting the respective gas into storage one will occupy some or all of the remaining capacity and thus eliminate the possibility of even better arbitrage opportunities in the future. Therefore, the risk premium which the owner of storage capacity should earn starts to rise and does so at an increasing rate as soon as storage levels approach the total capacity limit. Note that this interpretation is economically similar to the cause of an increase of the physical storage charge above. To illustrate the effect of the two additional theoretical phenomena, Figure 5.4 shows the total net storage cost and the convenience yield as a function of the storage level. The net storage cost function also contains the risk premium. The figure tells us that if one observed data points only in the lower range of possible stock levels, the convexity of the convenience yield function leads to a concave shape of the net storage cost. Conversely, if stock levels were in the upper range, the net storage cost would appear as a convex increasing function of stocks since the storage charge or the mentioned risk premium increase non-linearly. The lowest storage level in our time series is observed in 04/00 when the remaining volume was 1,550 mcm[20], which is still far away from complete depletion. Accordingly, since we use the negative of the net storage cost from Figure 5.4, we should observe a dependent variable which is a concave decreasing

[20] Million cubic meters.

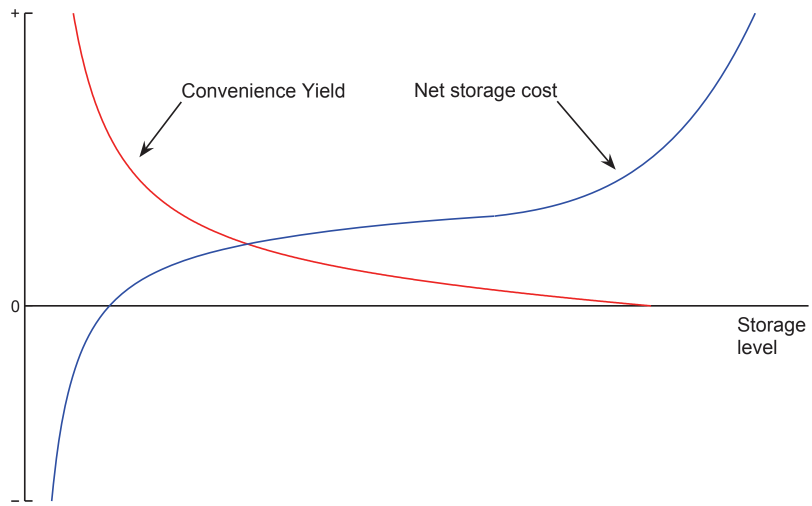


Figure 5.4: Theoretical shape of the marginal cost of storage (continuous line) according to Figure 2 in Brennan (1958).

function of the stock levels. The marginal risk premium, denoted by ρ, or the marginal storage charge c should be responsible for the concavity. Our dependent variable y_t, the negative net storage cost from (5.6), would, hence, be composed of the marginal cost of physical storage, the marginal risk premium and the marginal convenience yield:

$$y_t := -(c_{t,T} + \rho_{t,T} - q_{t,T}). \tag{5.7}$$

In other words, if there were excess gas on the market and stocks were already filled up to a certain extent, market participants would want the basis to increase more strongly before they were willing to put additional gas in stock. Consequently, we would see the negative net storage cost plunging more strongly as well, and this cannot be captured by our basic model with only a linear storage regressor.

Unfortunately, we cannot measure the components of the negative net storage cost individually. Both risk premium ρ and convenience yield q are implicit. Moreover, measuring the cost of physical storage is difficult because cost data is regularly not available, but proprietary to the storage operators. In the UK, APX-ENDEX and Centrica have recently launched a secondary trading platform for storage capacity in the Rough reservoir. Yet, a price index is not provided so far because of insufficient liquidity in this market. In Appendix 3, we explain briefly the components of storage cost and we also show the long-term development of season-ahead storage prices. However, the yearly frequency of these data does not permit us to include it in the econometric analysis.

This does not pose a general problem. Changes in the risk premium and the cost of physical storage should be taken account of in a spot price model as well so one can simply regard y reduced by a constant cost of storage as the *'net convenience yield'*. In turn, whether variations in ρ or c can explain most of the changing error variance remains an open question. This means that we may need an additional heteroscedasticity correction. Two of the extended model specifications in the following section are designed to address this problem.

5.5 Extended Specifications

Based on Brennan (1958) and the previous residual analysis, we have found that the basic model does not seem to be optimally specified yet, and that our convenience yield variable (the negative net storage cost) could be falling in the stock levels at an increasing rate. The question is how we could potentially capture this curvature and, hence, improve the model. This will be dealt with in the first part of this section. In the second part, we investigate whether the crude oil price can add further explanatory power to our model.

One possibility to deal with a non-linear relationship between the convenience yield and the storage levels is to include the squared storage observation as an additional regressor. Dincerler et al. (2005) and Modjtahedi and Movassagh (2005) have already included such a squared term in their models with ambiguous results. Findings of the former study correspond to the left part of Figure 5.4 and, hence, provide support for a convex decreasing convenience yield. The latter study finds that the corresponding coefficients change sign: Those for the short maturities correspond to the right end of the net storage cost curve, but the longer ones have an opposing sign again. We also test a linear model with the square of stock levels as an additional regressor (*'squared stock model'*) since this seems economically justified and since it could diminish the heteroscedasticity to some extent. The corresponding regression equation is

$$y_t = \beta_{cons} + \beta_{stock} x_{stock,t} + \beta_{stock^2} x_{stock^2,t} + \beta_{temp} x_{temp,t} + \epsilon_t \qquad (5.8)$$

In addition, we test a switching regression model (Goldfeld and Quandt (1972)) as an alternative specification. This model will allow the convenience yield to be governed by two different regimes while the storage levels shall be the trigger for the regime switch in addition to its role as a regressor. This model has three advantages: Firstly, we point out that this specification yields a regression which is piecewise linear in the storage levels but keeps the time series properties of the data and endogenously determines the cut-off point between the 'pieces'. This permits us to test for a changing inventory risk premium or a changing storage charge with a piecewise linear relationship instead of a quadratic one. Quadratic regressors can make the model more sensitive to noise in the respective explanatory variable, so a piecewise linear specification might stabilize the estimate. If one of the hypothesized effects is present in the cost of storage, the stock coefficient will be more strongly negative in the second regime since we are trying to fit the negative of the curve in the right part of Figure 4. Secondly, since the model is estimated with a smooth transition between the two regimes, each observation will be attributed an individual error variance. This means, a more acceptable weighting of observations within the estimation is achieved. No simple transformation could do a similar job in the single regime models.[21] Thirdly, since the model implies a separate error variance for each of the two regimes, we will be able to test the theory established by Fama and French (1987), stating that the variability of the basis will be relatively higher when stock levels are low and vice versa.

It might also be asked whether a regression model with jumps could be a more straightforward representation to account for the extreme values of residuals we have observed above. Indeed, jumps in the gas spot price occur frequently on an intraday and interday basis, partly due to the prominent role of gas in short term power generation. However, we recall that we are using 5-day averages of the spot price and a monthly interval of observations, so the outliers we observe cannot have the same causes as the short term spot price movements. In addition, many of our drops exhibit sluggishness, that is they regularly evolve and revert over an interval of several observations. While it is probable that capacity restrictions similarly play a role in our case (e.g. with respect to operations and space of the storage system), we have no single observable trigger variable proposing a stable estimation of a jumping regression. Finally, the number of extreme observations would most likely be too small.

The characteristic feature of our switching model, which Goldfeld and

[21] Log-transformation of the model is not possible due to the non-negativity constraint for the observations. Conditioning the residual standard deviation on the square root of stock levels through GLS is neither possible (for the same reason) nor a meaningful representation of the observed heteroscedasticity pattern.

Quandt labeled the 'D-method', is that the discontinuous step function governing the change between two possible regimes is replaced by a continuous approximation through a standard normal integral, the D-value. To estimate the model parameters, we have to maximize the following log-likelihood function:

$$\begin{aligned}\log L &= -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=2}^{n}\log[\sigma_1^2(1-D_t)^2+\sigma_2^2D_t^2] \\ &\quad -\frac{1}{2}\sum_{t=2}^{n}\frac{(y_t-f_t)^2}{\sigma_1^2(1-D_t)^2+\sigma_2^2D_t^2}\end{aligned} \tag{5.9}$$

wherein the combined regression equation f_t reads

$$f_t = (1-D_t)(x_{1,t}\beta_1+\rho_1\epsilon_{1,t-1})+D_t(x_{2,t}\beta_2+\rho_2\epsilon_{2,t-1}).$$

D_t denotes the weighting factor which lies between 0 and 1. The cut-off point between the two regimes is contained in the expression of the D-value (not shown for ease of exposition) and is determined endogenously by the model. X_1.and X_2 are the data matrices for the two regimes, σ_1 and σ_2 are the error variances, β_1 and β_2 are the vectors of partial regression coefficients and ρ_1 and ρ_2 are the coefficients of first-order autocorrelation. Finally, ϵ_1 and ϵ_2 are matrices of residuals. The last term of the log-likelihood function (5.9) is the equivalent of a weighted least squares expression. We have estimated the model parameters using the quadratic hill-climbing algorithm (Goldfeld and Quandt (1972)).

The results obtained with the squared stock model (5.8) are displayed in Table 5.2. We can see that the previously analyzed coefficients are unchanged with respect to sign and are still significant overall. Only the intercept becomes less significant for short maturities now. The squared stock coefficient is statistically significant, but it changes the sign in both sets of regressions. For longer term futures contracts, i.e. regressions 6, 10

No.	*Basis*	$\hat{\beta}_{cons}$		$\hat{\beta}_{stock}$		$\hat{\beta}_{stock^2}$		$\hat{\beta}_{temp}$		p_{White}	R^2_{adj}
1	M1-OCM	-.0228	(-1.06)	-.1008	(-4.16***)	-.0574	(-2.52**)	-.0052	(-1.18)	.166	18.4%
2	M2-OCM	-.0487	(-1.72*)	-.2180	(-5.31***)	-.1244	(-3.34***)	-.0235	(-3.38***)	.019	42.0%
3	M3-OCM	-.0929	(-2.44**)	-.2696	(-4.95***)	-.1565	(-3.18***)	-.0518	(-5.52***)	.022	47.9%
4	M4-OCM	-.1676	(-3.91***)	-.2215	(-4.09***)	-.1178	(-2.41**)	-.0825	(-8.06***)	.054	54.6%
5	M5-OCM	-.2726	(-5.52***)	-.0922	(-2.01**)	-.0239	(-.56)	-.1077	(-10.03***)	.172	60.2%
6	M6-OCM	-.3719	(-5.79***)	.0630	(1.39)	.0741	(1.88*)	-.1151	(-10.21***)	.453	57.5%
7	M2-M1	-.0252	(-1.98**)	-.0866	(-7.25***)	-.0393	(-3.48***)	-.0159	(-6.03***)	.000	53.0%
8	M3-M1	-.0617	(-3.06***)	-.1218	(-5.71***)	-.0605	(-3.15***)	-.0406	(-8.98***)	.001	63.3%
9	M4-M1	-.1300	(-5.75***)	-.0786	(-3.40***)	-.0258	(-1.22)	-.0681	(-12.96***)	.000	74.3%
10	M5-M1	-.2284	(-7.17***)	.0273	(1.13)	.0518	(2.17**)	-.0919	(-14.13***)	.001	76.7%
11	M6-M1	-.3182	(-6.53***)	.1529	(4.80***)	.1278	(4.29***)	-.0999	(-12.82***)	.016	71.3%

Table 5.2: Estimation results of squared stock model for the convenience yield (5.8). Details on the notation are given in the explanation for Table 5.1 on p. 68.

and 11, the sign is not in line with Brennan's theory. This finding is similar to that in Modjtahedi and Movassagh (2005) and can partly be attributed to the different pricing of long term and short term storage service charge. As the latter authors explain, capacity limits might put an increasing upward pressure on the charge for short term storage, while the one for long term storage should remain more or less constant. Indeed, storage facilities tend to sell (primarily firm) storage service once a year for the entire next storage year, so well ahead in time. In turn, there is also trading for the short term, either via the facility owner or another secondary market, and the corresponding prices will most likely increase when storage space becomes scarce.[22] This supports the view that changes in the cost of physical storage are the main cause of the non-linearity in the net storage cost function. Unfortunately, the reduction in heteroscedasticity with respect to the basic model is only moderate as shown by the p-values of the White-test. An additional test for ARCH-effects does not yield consistent results across the time series so including such effects does not seem to be meaningful for monthly data. In turn, with respect to the explanatory power of the model, the squared stock model seems to outperform the basic version at the short end of maturities while being equally powerful at the long end.

The results of the switching model (5.9) are found in Table 5.3. The subscripts denote the number of the regime and regime 2 contains those observations with higher stock levels. In addition to the parameter estimates and their z-statistics, the F-test statistic for homogeneity of the standard deviations of residuals and the number of observations in each regime are displayed. The column labeled R^2_{ML} reports the generalized coefficient of determination as defined by Nagelkerke (1991). Overall the significance of the coefficients is reduced compared to the single regime models. Neither

[22] Cf. e.g. the terms of purchase and the general information of Centrica Storage and Scottisch Southern Energy. These are the operators of the largest UK storage facilities, the Rough field and the Hornsea salt caverns.

$No.$	$\hat\beta_{1cons}$	$\hat\beta_{1stock}$	$\hat\beta_{1temp}$	$\hat\beta_{2cons}$	$\hat\beta_{2stock}$	$\hat\beta_{2temp}$	$\hat\rho_1$	$\hat\rho_2$	$\hat\sigma_1$	$\hat\sigma_2$	F	$\hat n_1$	$\hat n_2$	R^2_{ML}
1	-.1288	-.0958	-.0086	.1263	-.3181	-.0044	-.16	-.01	.11	.24	.21***	45	48	10.4%
	(-1.14)	(-1.15)	(-1.70*)	(1.15)	(-2.67***)	(-.55)	(-1.00)	(-.06)	(7.26***)	(6.21***)				
2	-.1572	-.1270	-.0213	.1985	-.6109	-.0265	-.21	.35	.12	.34	.13***	46	47	21.7%
	(-1.32)	(-1.45)	(-3.75***)	(1.19)	(-3.22***)	(-2.02**)	(-1.24)	(2.19**)	(7.45***)	(5.99***)				
3	-.1879	-.1479	-.0261	.1978	-.7393	-.0715	-.20	.30	.13	.45	.09***	45	48	24.5%
	(-1.70*)	(-1.69*)	(-3.86***)	(1.23)	(-3.78***)	(-4.28***)	(-.93)	(2.02**)	(7.34***)	(6.98***)				
4	-.2236	-.1461	-.0409	.0317	-.5370	-.1212	-.03	.20	.14	.51	.08***	46	47	30.7%
	(-1.89*)	(-1.50)	(-4.10***)	(.15)	(-2.26**)	(-5.43***)	(-.17)	(1.26)	(7.22***)	(6.34***)				
5	-.4229	-.2271	-.0630	-.1490	-.2228	-.1369	.10	.14	.18	.48	.14***	41	52	34.8%
	(-2.18**)	(-1.30)	(-2.44**)	(.85)	(-1.08)	(-7.37***)	(.30)	(.92)	(5.89***)	(8.14***)				
6	-.3517	.0001	-.1172	-.1244	-.1872	-.1082	.65	.13	.26	.51	.28***	50	43	33.4%
	(-3.30***)	(.00)	(-7.57***)	(.65)	(-.84)	(-5.12***)	(4.49***)	(.83)	(9.30***)	(8.69***)				
7	-.0037	-.0149	-.0148	-.0066	-.1517	-.0166	.15	.18	.06	.12	.25***	53	59	28.6%
	(-.19)	(-.76)	(-5.57***)	(-.20)	(-3.73***)	(-4.10***)	(1.05)	(1.46)	(9.89***)	(10.6***)				
8	.0032	-.0070	-.0220	-.0384	-.2056	-.0494	.27	.40	.08	.17	.19***	39	73	46.1%
	(.15)	(-.30)	(-6.39***)	(-.61)	(-2.75***)	(-7.26***)	(1.95*)	(3.08***)	(10.5***)	(10.6***)				
9	1.3879	.4375	.0097	-1.8124	.5998	-.1383	.16	.68	.00	.26	.00***	57	55	42.2%
	(-.83)	(-.89)	(-.72)	(.54)	(-.57)	(-1.70*)	(.62)	(1.97**)	(.00)	(1.92*)				
10	-.1686	-.0278	-.0644	-.2413	.0973	-.0942	.46	.45	.14	.21	.43***	37	75	39.8%
	(-2.18**)	(-1.30)	(-2.44**)	(-.85)	(1.08)	(-7.37***)	(.30)	(.92)	(5.89***)	(8.14***)				
11	.2546	.0843	-.0308	.1047	-.1010	-.0814	1.76	.54	.19	.23	.65	34	78	29.7%
	(1.59)	(1.15)	(-1.59)	(1.52)	(-1.63)	(-7.54***)	(10.1***)	(5.53***)	(7.75***)	(12.0***)				

Table 5.3: Estimation results of regime-switching model for the convenience yield (5.9).

does the stock level coefficient in regime 1 have any explanatory power nor the constant term in regime 2. Even the constant in regime 1 is only significant in 5 of 11 cases. As we know from the previous analyses that a constant does exist in the model, we assume that the number of parameters is excessive. An additional problem might be that maximum likelihood estimation is a large sample method so significance is evaluated based on the asymptotic normal distribution. The number of observations per regime is evenly distributed in most cases, but in regression 8, 10 and 11 the second regime contains significantly more observations. This is somewhat counter-intuitive given the hypothesis that the storage charge or the risk premium should not increase remarkably until storage levels are sufficiently close to the capacity limit.

However, the model reveals that the residual standard deviation in regime 2 is significantly higher than in regime 1 as evidenced by the significant F-statistics. This is in line with the residual plots in Figure 2 from the previous section. Fama and French (1987) have found for several classes of commodities that the basis variability is declining in stock levels. They argued that larger inventories mean extended possibilities to dampen (positive) demand and (negative) supply shocks. The authors reason that if inventories are depleted (or expected to be depleted shortly), spot price variability decouples from futures price variability since any short term demand or supply shock will transmit one for one to the spot price, but not to the futures price. Yet, this theory about the behavior of the basis variability cannot generally hold for the gas market, given that available storage capacity is a tight restriction. In fact, storage space, which is mostly underground, cannot be enlarged in the short run. Figure 5.5 demonstrates this fact for the period from 10/05 to 01/08 for both long range and total storage levels.[23]

[23] Long range storage as classified by National Grid UK is based on the storage facility used (Rough, North Sea). Unlike our notation 'long term storage' it does not refer to

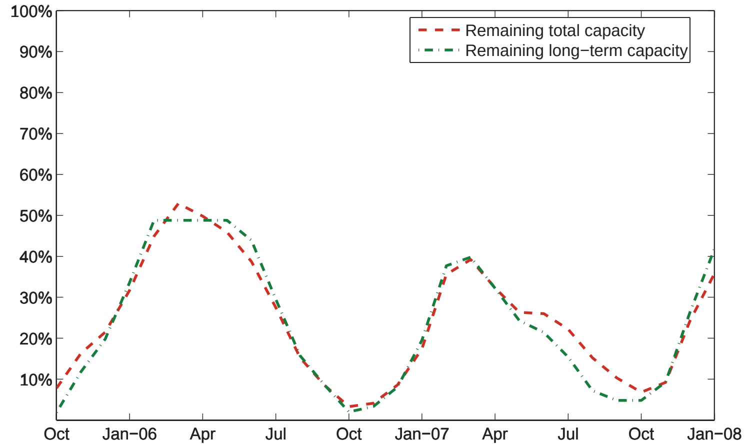


Figure 5.5: Relation between residuals of basic model and remaining storage space in UK.

It is apparent from the graph that capacity limits have been closely approached in the winters during this time period. But since the upper bound of the storage facilities is the binding one, we have an effect inverse to the one described by Fama and French. That is, spot price volatility increases now because positive supply and negative demand shocks can no longer be offset by storage injections. This effect has not been documented in the literature on the theory of storage so far. With respect to the overall model fit though, we conclude that a switching model based on stock levels does not enhance the explanatory power.

After having investigated alternative specifications with regard to the storage variable, we now study the influence of the crude oil market as suggested in the discussion of testable drivers in section 5.2. Two effects of

the actual holding time.

oil seem particularly meaningful to be tested. First, changes in the oil price could directly influence the scarcity of gas through a substitution effect in the spirit of Cassasus et al. (2009). In this case, a drop in the price of oil should (possibly by a change in the price of its chemical derivatives) lead to a decrease of gas demand. We test for such a positive relationship by including the front month contract of the brent future traded on the ICE into the linear regression model. This time, the model is stated in first-difference form since the brent price series is integrated of order 1:

$$\begin{aligned} \triangle y_t &= \beta_{cons} + \beta_{\triangle stock} \triangle x_{stock,t} + \beta_{\triangle (stock^2)} \triangle x_{stock^2,t} \\ &\quad + \beta_{\triangle temp} \triangle x_{temp,t} + \beta_{\triangle P_{oil}} \triangle x_{P_{oil},t} + \epsilon_t. \end{aligned} \tag{5.10}$$

As is clearly visible from Table 5.4, which shows the regression results, the oil price itself does not have an impact on the determination of the convenience yield in the gas market. The t-statistics of the corresponding regressor is insignificant in all cases. Since the dependent variable has already been stationary in the level form, the regression constant has turned insignificant now. The ratio of all other t-statistics remains nearly unchanged compared to the results in Tables 5.1 and 5.2.

Second, it is possible that the convenience yields of crude oil and gas are directly correlated because they share exposure to the same macroeconomic factors as proposed by Bailey and Chan (1993). Therefore, we also estimate a regression which includes the convenience yield of crude oil as an additional regressor:

$$\begin{aligned} y_t &= \beta_{cons} + \beta_{stock} x_{stock,t} + \beta_{stock^2} x_{stock^2,t} \\ &\quad + \beta_{temp} x_{temp,t} + \beta_{CY_{oil}} x_{CY_{oil},t} + \epsilon_t. \end{aligned} \tag{5.11}$$

Table 5.5 shows the results of this regression. Since we use a front month convenience yield for oil in all regressions, both convenience yields are

No.	*Basis*	$\hat{\beta}_{cons}$		$\hat{\beta}_{\Delta stock}$		$\hat{\beta}_{\Delta stock^2}$		$\hat{\beta}_{\Delta temp}$		$\hat{\beta}_{\Delta P_{oil}}$		R^2_{adj}
1	M1-OCM	-.0002	(-.01)	-.1568	(-3.49)***	-.0856	(-2.33)**	-.0013	(-.14)	.0687	(.92)	7.3%
2	M2-OCM	.0028	(.09)	-.1560	(-4.49)***	-.1555	(-3.55)***	-.0226	(-1.64)	.0262	(.27)	18.4%
3	M3-OCM	.0056	(.14)	-.2861	(-3.77)***	-.1886	(-3.55)***	-.0434	(-2.35)**	-.0372	(-.28)	17.7%
4	M4-OCM	.0034	(.08)	-.2257	(-2.79)***	-.1386	(-2.61)**	-.0699	(-3.61)***	-.0059	(-.04)	16.9%
5	M5-OCM	-.0028	(-.07)	-.0940	(-1.19)	-.0645	(-1.16)	-.0922	(-4.65)***	.0517	(.40)	17.0%
6	M6-OCM	-.0042	(-.10)	-.0043	(-.11)	.0224	(.39)	-.0912	(-4.95)***	.0785	(.69)	14.0%
7	M2-M1	.0025	(.23)	-.0714	(-3.35)***	-.0439	(-3.50)***	-.0162	(-3.14)***	-.0294	(-.81)	16.9%
8	M3-M1	.0075	(.42)	-.0979	(-2.57)**	-.0637	(-3.03)***	-.0358	(-4.58)***	-.0752	(-1.55)	28.8%
9	M4-M1	.0060	(.33)	-.0466	(-1.29)	-.0227	(-1.16)	-.0619	(-7.58)***	-.0682	(-1.23)	44.4%
10	M5-M1	.0000	(.00)	.0605	(1.69)*	.0384	(1.63)	-.0833	(-7.70)***	-.0232	(-.40)	43.7%
11	M6-M1	-.0024	(-.10)	.1404	(3.55)***	.1052	(3.69)***	-.0860	(-7.29)***	.0001	(.00)	41.0%

Table 5.4: Estimation results when controlling for crude oil price changes according to (5.10). Details on the notation are given in the explanation for Table 5.1 on p. 68.

annualized in order to match their maturities in all the regressions.[24] Only the coefficients in regressions 5, 6 and 11, in which the gas convenience yield has a long maturity, is in line with this expectation. Even here, the gain in the adjusted R^2 is in the order of 1% only. These results could be due to the fact that the gas market reacts to changes on the crude oil market with a precise time lag so we rerun both model specifications with 6-month and a 9-month lag for the oil price. The results remain basically unchanged to the tabulated ones, however. Apparently, any direct link of the gas convenience yield to the crude oil market does at best affect the longer maturity futures contracts. However, when projections for the very long term are required and market liquidity in the respective contracts is sufficient, it would be interesting to investigate this relationship again.

Based on the findings in this section, the basic linear model and the squared stock model seem to be the most promising specifications to continue with. Therefore, we analyze the robustness of these two models in the following.

[24] The formula for annualizing the data is based on discrete compounding (the tilde denotes annualized values): $\widetilde{y}_t := \widetilde{q}_{t,T} - \widetilde{c}_{t,T} - \widetilde{\rho}_{t,T} = \widetilde{r}_{t,T} - \frac{1}{T-t}(\frac{F_{t,T}-S_t}{S_t})$

No.	$\hat{\beta}_{cons}$		$\hat{\beta}_{stock}$		$\hat{\beta}_{stock^2}$		$\hat{\beta}_{temp}$		$\hat{\beta}_{CY_{oil}}$		R^2_{adj}
1	-.4785	(-1.29)	-1.6740	(-4.01)***	-.9712	(-2.54)**	-.0965	(-1.20)	1.1719	(.98)	18.6%
2	-.4001	(-1.90)*	-1.4972	(-5.54)***	-.8662	(-3.40)***	-.1705	(-3.31)***	.7727	(.95)	42.3%
3	-.4588	(-2.52)**	-1.1620	(-5.10)***	-.6850	(-3.20)***	-.2343	(-5.39)***	.6259	(.98)	48.2%
4	-.5846	(-3.93)***	-.6845	(-4.19)***	-.3715	(-2.39)**	-.2711	(-7.94)***	.5585	(1.18)	55.1%
5	-.7385	(-5.74)***	-.1992	(-1.84)*	-.0493	(-.47)	-.2785	(-10.12)***	.5815	(1.76)*	61.1%
6	-.8274	(-6.16)***	.1713	(1.79)*	.1686	(2.07)**	-.2458	(-10.49)***	.6084	(2.39)**	59.0%
7	-.3198	(-1.98)*	-1.0266	(-7.32)***	-.4658	(-3.46)***	-.1917	(-6.01)***	.1856	(.47)	52.7%
8	-.3912	(-3.01)***	-.7156	(-5.58)***	-.3559	(-3.07)***	-.2442	(-8.98)***	.2252	(.71)	63.1%
9	-.5362	(-5.48)***	-.3026	(-3.28)***	-.0979	(-1.14)	-.2728	(-13.02)***	.1705	(.75)	74.2%
10	-.7077	(-7.11)***	.0980	(1.36)	.1628	(2.22)**	-.2758	(-14.36)***	.2379	(1.36)	76.8%
11	-.7981	(-6 .57)***	.3918	(4.93)***	.3179	(4.31)***	-.2404	(-13.14)***	.3640	(2.09)**	71.9%

Table 5.5: Estimation results when controlling for the convenience yield of crude oil according to (5.11). Details on the notation are given in the explanation for Table 5.1 on p. 68.

5.6 Robustness tests

To prove that our linear specifications are robust, we first test the models' stability over time and their forecasting performance. Second, we show that the model keeps explanatory power when data is deseasonalized. For the first step we divide the time series used above in two subsamples, a calibration period and a prediction period, and evaluate the goodness-of-fit in the latter period compared to the earlier one. Since the model has autocorrelated error terms in some cases, a dynamic forecast must be used, and for the accuracy of estimation it is relevant to determine how many periods we would like to predict the convenience yield. Under the prevailing first order autocorrelation of the form $u_t = \rho u_{t-1} + \epsilon_t$ the k-period forecast is

$$\hat{y}_{t+k} = \hat{\beta}_{cons} + \hat{\beta}_{stock} x_{stock,t+k} + \hat{\beta}_{temp} x_{temp,t+k} + \rho^k \hat{u}_t \tag{5.12}$$

in the basic model (analogously in the squared stock model). The variance of the forecast error is a concave increasing function of the number of lags (cf. Appendix 4). We limit our robustness test to a one-period forecast. The estimation period is set to six full years plus one lag for the estimation of an autocorrelation coefficient where necessary.

To evaluate the goodness-of-fit, we use the root mean squared error (RMSE), the coefficient of determination and coverage counts. The RMSE is the estimated standard deviation of the residuals, which allows an absolute pairwise comparison between the same regressions in the two models. In turn, the R^2-statistic serves as a relative measure of fit. This time we use Feasible Generalized Least Squares (FGLS) to estimate the model in the presence of autocorrelation. This permits us to directly compare the model fit in the prediction and the corresponding calibration period based on the

R^2.[25] The coverage counts measure the number of observations which lie within the semi-standard deviation interval ($CC_{.5\sigma}$) and the one-standard deviation interval (CC_{σ}) around the actual value of the dependent variable. The standard deviation used is the RMSE from the calibration period. If the fraction of covered observations is close to normal, this will confirm the stability of the model. In those regressions in which the autocorrelation coefficient is not significant on a 5% confidence level the coefficient is set to zero and the model is reestimated in OLS form. In turn, all regression coefficients were included in the forecast independently of their significance in the calibration period. The reason is that the R^2-statistic in the prediction period are not meaningful if intercepts are dropped, as long as their point estimate is not close to zero in absolute terms. In addition, since there is no reasonable heteroscedasticity correction the confidence bounds and significance levels of the estimators should be regarded with caution. Apart from this reasoning, the decision about whether to drop insignificant coefficients should not produce a material economic difference in the results since these coefficients could equally harm or benefit the forecast. The in-sample results obtained for the basic and the squared stock model are presented in Tables 5.6 and 5.7 respectively.

Obs. denotes the number of observations in the calibration and prediction period. Superscript * (in the header) indicates the FGLS-transformation of the model and the corresponding coefficient of autocorrelation is ρ^*. Due to the fact that some of the regressions have a nonzero coefficient of autocorrelation, the possibility of comparison to the full sample results in Tables 5.1 and 5.2 is limited. In fact, the comparison only allows for a few evaluative comments on the stability of the model: The majority of results

[25] Note that the prediction R^2 can no longer be interpreted as the fraction of variance of the dependent variable explained by the model variance and that its lower bound is not zero, but -$\frac{n(\beta^*_{cons}-\bar{y}^*)^2}{\sum(y^*_t-\bar{y}^*)^2}$. However the R^2 can still be used as a relative measure of the goodness-of-fit in the above mentioned comparison.

		Calibration						Prediction				
No.	*Basis*	*Obs.*	$\hat{\beta}^*_{cons}$	$\hat{\beta}^*_{stock}$	$\hat{\beta}^*_{temp}$	$\hat{\rho}^*$	R^{2*}	*Obs.*	$RMSE^*$	R^{2*}	$CC_{.5\sigma}$	CC_{σ}
1	M1-OCM	73	-.0720***	-.0753***	-.0045	-.24	24.5%	20	.25	14.7%	7	13
2	M2-OCM	73	-.1402***	-.1580***	-.0201***	-	46.1%	20	.43	20.1%	8	12
3	M3-OCM	73	-.2089***	-.1925***	-.0460***	-	51.8%	20	.56	27.5%	8	14
4	M4-OCM	73	-.2505***	-.1574***	-.0763***	-	59.4%	20	.56	38.2%	9	15
5	M5-OCM	73	-.2725***	-.0703	-.1030***	-	63.8%	20	.49	49.7%	7	15
6	M6-OCM	73	-.2838***	.0364	-.1127***	-	57.5%	20	.46	52.5%	9	15
7	M2-M1	73	-.0440***	-.0617***	-.0139***	-	54.7%	39	.14	40.9%	10	22
8	M3-M1	73	-.0931***	-.0778***	-.0347***	.46	44.5%	39	.18	41.4%	17	28
9	M4-M1	73	-.1262***	-.0462***	-.0568***	.59	56.3%	39	.19	54.4%	22	25
10	M5-M1	73	-.1418***	.0203	-.0760***	.54	64.3%	39	.26	51.5%	15	25
11	M6-M1	73	-.1480***	.0927***	-.0794***	.62	58.9%	39	.33	38.9%	11	18

Table 5.6: Results of in-sample robustness test for basic model (5.6).

	Calibration						Prediction			
$No.$	$\hat{\beta}^*_{cons}$	$\hat{\beta}^*_{stock}$	$\hat{\beta}^*_{stock^2}$	$\hat{\beta}^*_{temp}$	$\hat{\rho}^*$	R^{2*}	$RMSE^*$	R^{2*}	$CC_{.5\sigma}$	CC_σ
1	-.0498*	-.0821***	-.0252	-.0053	-.23	25.7%	.24	21.3%	6	14
2	-.0608*	-.1838***	-.0890***	-.0227***	-	51.9%	.41	28.2%	6	14
3	-.1101**	-.2246***	-.1106**	-.0493***	-	55.8%	.53	34.6%	5	14
4	-.1852***	-.1787***	-.0732	-.0784***	-	60.6%	.54	41.9%	10	15
5	-.2776***	-.0687	.0057	-.1028***	-	63.8%	.49	49.5%	7	15
6	-.3543***	.0593	.0790	-.1104***	-	58.4%	.46	53.1%	10	16
7	-.0183	-.0679***	-.0274**	-.0150***	-	57.7%	.13	44.3%	14	21
8	-.0571	-.0885***	-.0394	-.0360***	.45	47.0%	.18	45.5%	15	29
9	-.1214**	-.0476	-.0052	-.0570***	.59	56.4%	.19	54.7%	22	25
10	-.1896***	.0345	.0526*	-.0744***	.55	65.8%	.26	51.0%	18	23
11	-.2375***	.1192***	.0989***	-.0765***	.63	64.2%	.32	40.9%	10	17

Table 5.7: Result of in-sample robustness test for squared stock model (5.8).

of the calibration period remains unchanged with respect to the significance of the parameters. In turn, some of the point estimates have changed even in the regressions which were not transformed, i.e. those which did not necessitate an autocorrelation correction. In regressions 2 to 5 of the basic model (Table 5.6) the stock coefficients have diminished by roughly 15%. The constant in regression 7 is reduced by more than 20%. Given that the most severe outliers observed in autumn 2006 are now contained in the prediction period, we conclude that the intercept and the stock coefficient are more sensitive to outliers than the temperature variable. Nearly the same observations can be made for the series of the squared-stock model which have not been transformed (Table 5.7). The only other change is that the squared stock coefficient has slightly lost significance. Nevertheless these changes are still fairly acceptable given the strong irregularities in the residuals' distribution.

Moving on to the goodness-of-fit measures, the obtained results look very satisfying in general in both Tables 5.6 and 5.7. Naturally, the regression series 1 to 6 have a somewhat weaker explanatory power than the remaining ones. Due to the shorter sample length of the forecast period and the more pronounced irregularities in the OCM price compared to the M1 series, a stable forecast is more difficult to achieve in these series. Not surprisingly, the regressions 1 to 3 are the weakest in terms of absolute R^2 and they also lose a higher fraction of the calibration R^2 in the forecast period. While regression 11 is not strongly robust either, roughly 40% of the convenience yield variance can still be explained. However, the squared stock model performs at least equally well as does the basic model and is especially ahead regarding the shorter maturities. More precisely, the forecast is improved by 6.6%, 8.1% and 7.1% in regressions 1 to 3 and by 3.4% and 4.1% in regressions 7 and 8 respectively when we compare the results in both tables. The RMSE also favours the latter model as this statistic is closely related to the R^2.

The coverage counts are not designed for a pairwise comparison since the squared stock model has lower RMSE values in the calibration period and, hence, tighter confidence bounds. Due to the limited sample size of the forecasts it is more meaningful to look at the counts in the aggregate: The sum of all observations projected within the smaller confidence bounds is 123 in both models. Applying the larger bounds the squared stock model achieves a total of 203, which is only one more than does the basic model. From these figures both models appear equally stable with regard to the calibration period. In addition, the simple average of the percentages covered in each of the 11 time series indicates that both models keep a roughly normal error distribution: 39.3% and 65.7% are the respective distribution values for the two intervals in the simple model, while 38.4% and 66.8% are the corresponding values in the extended model. They are sufficiently close to the corresponding normal values of 38.3% and 68.3%. To summarize, while the coverage counts state that both models are generally stable, the R^2-statistic and the RMSE show that the squared stock model is slightly more powerful and robust.

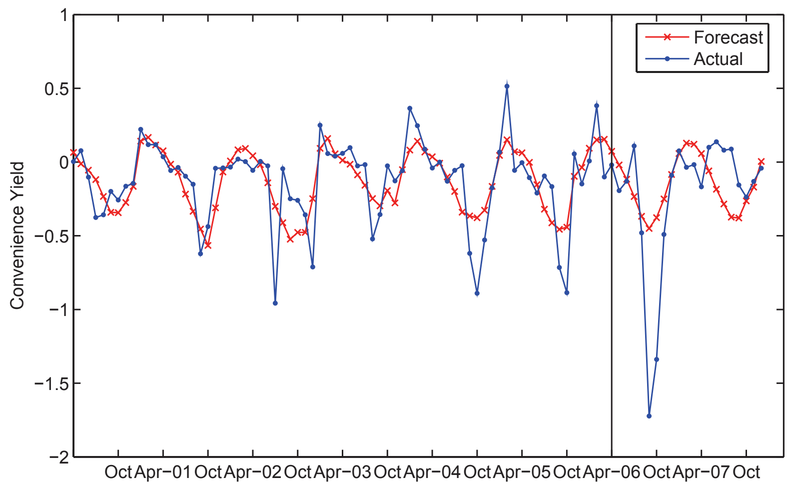
The in-sample test has already demonstrated the forecasting power of our favored model, most notably by the strong R^2's which appear even in the FGLS-transformed version. Next, we investigate the out-of-sample forecast of this model. We use a naive prediction for the two explanatory variables, monthly mean temperature and closing stock levels. We are well aware that our prediction ignores any fundamental insight into meteorology or the economics of storage, but this is not the primary focus of the study. With a sinusoidal seasonality function for both variables and an additional error correction term with one lag for the storage variable we can obtain already a very acceptable explanatory power. The R^2-values for temperature are 95.0% and 94.8% respectively for the two different 73-month calibration periods. The corresponding (transformed) values for the storage variable are 80.0% and 82.4%. We predict temperature and storage levels during the forecast period and compute convenience yield

No.	*Obs.*	$RMSE^*$	R^{2*}	$CC_{.5\sigma}$	CC_σ
1	20	.25	16.1%	6	13
2	20	.42	23.1%	6	11
3	20	.55	28.7%	5	13
4	20	.58	32.2%	10	11
5	20	.54	37.5%	8	12
6	20	.51	42.0%	9	14
7	39	.14	40.2%	10	20
8	39	.18	46.2%	16	28
9	39	.20	48.8%	12	25
10	39	.27	48.0%	14	23
11	39	.33	40.8%	10	21

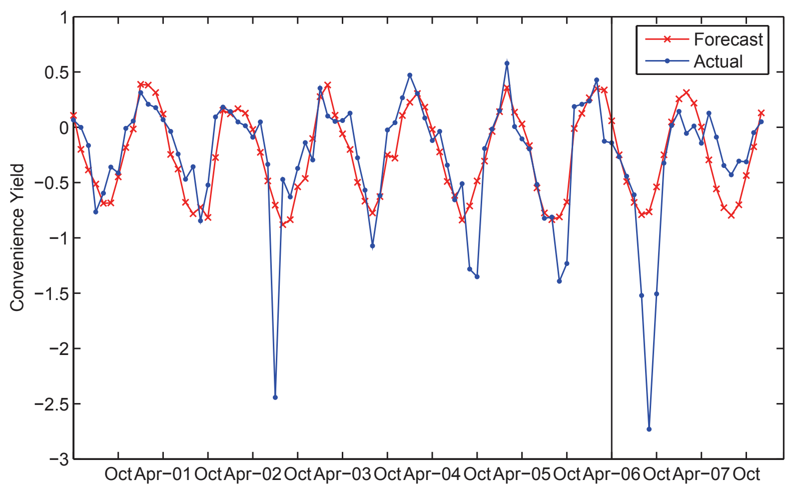
Table 5.8: Diagnostics for out-of-sample forecast of squared stock model (5.8). The notation is in accordance with Table 5.7.

estimates based on those predictions. The corresponding regression statistics are displayed in Table 5.8. Figure 5.6 graphs a set of the convenience yield forecasts and their actual values in comparison.

As expected, the out-of-sample forecast has somewhat lower R^2-values than the in-sample counterpart due to slight inaccuracies of the forecast of stocks and temperature. Yet, not many time series are strongly affected as is visible from comparing Tables 5.7 and 5.8. Only series 4, 5 and 6 have lost more than 6% of the R^2. Their RMSE has increased by .04 to .05 accordingly, quite in contrast to the very moderate increase of .01 in most other series. The coverage counts of series 1 to 6 are rather stable in the semi-standard deviation interval whereas the weakness of the former 3 series reappears, to some extend, in the one-standard deviation case.



Basis: M2 - OCM



Basis: M4 - OCM

Figure 5.6: Selected plots of convenience yield forecasts out of the sample.

The aggregate coverage counts decrease to 106 and 191 observations and the average rates of coverage drop to 34.5% and 60.9% respectively. This is below the desired level of a normally distributed variable. Having said that, an inclusion of fundamental predictors of storage (e.g. derived from a total market model) should ease this problem. Storage was deemed the more sensitive influential variable in the above in-sample analysis and our storage forecasts certainly leave some room for optimization. Apart from the coverage counts the out-of-sample test documents an acceptable level of explanatory power even under critical investigation which makes the model usable for forecasting purposes.

Finally, we investigate how the favored squared stock model performs with deseasonalized data in order to prove that it can well explain variation over and above seasonality.[26] This makes it necessary to use a higher data frequency. There are two practical benefits to this test regarding a possible application of our model with current reduced-form spot price models. Firstly, in spot price models we might want to work with daily or weekly data intervals as well. Secondly, seasonality is typically removed before estimating the parameters of the stochastic process so we can directly see if our model has additional explanatory power when combined with a reduced-form model. To obtain a consistent daily time series of convenience yields, we construct a complete futures term structure from the daily observations of our ICE month futures contracts and the OCM contract. This is done by fitting a cubic spline to the data based on time to maturity. From the daily term structure curves we then compute a set of interpolated futures prices with constant time to maturity for every daily observation.[27] The resulting convenience yield time series from these artificial futures prices accordingly has an equal term. Daily mean temperature

[26] We find very similar results for the basic linear model so they are not discussed.

[27] We choose to use the medium time to maturity of the six monthly contracts, i.e. 14, 45, 75, 106, 136 and 167 days respectively.

(Central England) from 01/00 to 12/08 comes from the HadCET database of the British Atmospheric Data Centre (BADC). The daily storage figures are closing levels from National Grid's operational data which are available since 10/05. This limits the number of observations to $N = 568$.

Storage figures are detrended as described in section 5.3. Annual seasonality is then removed from all data with a sinusoidal function of the type

$$f(t) = b_1 + b_2 sin(b_3 + 2\pi t/obs.\ p.a.). \tag{5.13}$$

The entire available length of each dataset is used for this to get the most accurate estimates possible. Other patterns of seasonality could not be detected. Prior to fitting the functions for the convenience yield series, we annualize these data in order to make the parameters comparable to existing estimation results for reduced-form models. For this purpose we also use continuous compounding this time as these studies do. That is, we compute the following annualized continuous time equivalent of the negative of (5.5)

$$\widetilde{q}_{t,T} - \widetilde{c}_{t,T} = \widetilde{r}_{t,T} - \frac{1}{T-t} \ln \left(\frac{F_{t,T}}{S_t} \right). \tag{5.14}$$

We also remove outliers with the procedure outlined in Benth et al. (2008) (p. 130)[28] and replace them by a linear interpolation between the neighboring data points. All seasonality coefficients turn out significant (at least at a 95% confidence level) except the b_1-coefficient of the storage data which have already been demeaned before. The residuals of the regressions (not shown) are highly autocorrelated (coefficient values $> 90\%$) despite the fact that the convenience yield is stationary. We identify an $AR(1)$-structure in the daily convenience yields and reestimate the model in first-difference

[28] A maximum of 3.5% of the observations in one series is affected which is in line with their removal rate.

form with the lagged dependent variable as an additional regressor. The latter one will tell us the (negative) rate of mean reversion on a daily basis. Furthermore, all convenience yield series have significant ARCH-effects of order 1 so we include such effects in the reestimated model. The final model equations are

$$\begin{aligned} y_t^s &= y_t - \left[b_1 + b_2 \sin\left(b_3 + \frac{2\pi t}{obs.\,p.a.}\right)\right] && (5.15a) \\ \Delta y_t^s &= \beta_{\Delta stock}\Delta x_{stock,t}^s + \beta_{\Delta(stock^2)}\Delta x_{stock^2,t}^s && (5.15b) \\ &\quad +\beta_{\Delta temp}\Delta x_{temp,t}^s + \beta_{y_{t-1}^s} y_{t-1}^s + \epsilon_t \\ \sigma_{\epsilon_t}^2 &= \beta_{cons} + \beta_{\epsilon_{t-1}^2}\epsilon_{t-1}^2 + u_t,\; u_t \overset{i.i.d.}{\sim} \mathcal{N}(0,\sigma^2). && (5.15c) \end{aligned}$$

(5.15a) represents the seasonal adjustment, (5.15b) is the main equation and (5.15c) the ARCH-equation. The results are given in Table 5.9. Since all constants on the left as well as the right hand side have been removed by the seasonality function, we estimate the model without a constant. This is helpful to assure a proper estimation of the ARCH-model. In turn, $\hat{b}_1$ reports the estimate of the intercept term in the seasonal function independent of the regression itself. Here, it can be interpreted as the average annualized spot convenience yield for the respective maturity. Since the model parameters have been obtained by maximum likelihood estimation, figures in parentheses denote z-statistics.

From the z-statistics for $\hat{\beta}_{y_{t-1}^s}$ one can see that the mean reversion parameter is significant which supports an Ornstein-Uhlenbeck convenience yield process as suggested in the reduced-form models by Gibson and Schwartz (1990) and Schwartz (1997). If we annualize (the negative of) these coefficients and compare them to those for crude oil in the aforementioned studies, the speed of adjustment seems to be in a similar range. Our annualized coefficients are between 8.6 (regression 10) and 27.3 (regression 1) whereas Gibson and Schwartz (1990), for instance, obtain 16.1. Annu-

		Main								ARCH	
No.	$\hat{b}_1$	$\hat{\beta}_{\Delta stock}$		$\hat{\beta}_{\Delta(stock^2)}$		$\hat{\beta}_{\Delta temp}$		$\hat{\beta}_{y_{t-1}}$		$\hat{\beta}_{\epsilon^2_{t-1}}$	$\hat{\beta}_{cons}$
1	-1.033	-10.377	(-5.93)***	.309	(.14)	-.181	(-8.76)***	-.108	(-8.25)***	.451***	1.017***
2	-.811	-5.582	(-5.99)***	1.047	(.78)	-.099	(-8.60)***	-.082	(-6.85)***	.428***	.294***
3	-.685	-3.529	(-5.70)***	.518	(.61)	-.068	(-8.79)***	-.075	(-6.33)***	.419***	.128***
4	-.569	-2.719	(-5.94)***	.625	(.95)	-.053	(-9.01)***	-.078	(-6.76)***	.436***	.076***
5	-.478	-1.915	(-4.99)***	.560	(.99)	-.039	(-8.25)***	-.074	(-7.06)***	.400***	.051***
6	-.417	-1.572	(-4.61)***	.252	(.49)	-.033	(-8.06)***	-.072	(-6.22)***	.376***	.039***
7	-.563	-1.207	(-3.29)***	5.836	(18.59)***	-.015	(-3.21)***	-.081	(-13.30)***	1.650***	.122***
8	-.480	.288	(.67)	1.551	(3.77)***	-.007	(-.92)	-.053	(-4.72)***	.243***	.067***
9	-.384	.197	(.72)	1.172	(4.44)***	-.005	(-1.22)	-.040	(-2.88)***	.246***	.033***
10	-.316	.112	(.48)	.724	(2.94)***	-.002	(-.62)	-.034	(-2.41)**	.282***	.023***
11	-.279	.072	(.44)	1.110	(5.92)***	-.002	(-.51)	-.069	(-6.67)***	.950***	.019***

Table 5.9: Regression results from squared stock model (5.15) for deseasonalized daily data.

alizing mean reversion speeds gives, of course, only an approximate idea of the process dynamics. The intercept of the seasonal function, in turn, points at a difference to the estimation results for crude oil. Obviously, all (net) convenience yield rates are strongly negative because the net storage cost and the storage risk premium exceed the value of the gross convenience yield. Gibson and Schwartz (1990) report a net convenience yield of 18.61%. The comparatively lower values for the short maturities are in line with the reasoning from section 5.5. That is, storage is also more costly for the short term than for the long term, provided that stock levels are often close to the capacity limit.

The temperature coefficient is strongly significant in all regressions in which the OCM price is used. The insignificance in regressions 8 to 11 is reasonable since the monthly contract M1 is less driven by spontaneous weather changes than the day-ahead price. Regression 7 could possibly be incorrectly estimated since the coefficients differ materially from those found for the following series and the ARCH coefficient signifies an exploding variance ($\hat{\beta}_{\epsilon^2_{t-1}} > 1$). It can thus be said that temperature generally contains significantly more than seasonality to explain the convenience yield. The storage variable does have slightly less explanatory power, but is also significant when the OCM spot price is used (1-6). As found in the first part of this section, storage levels might be slightly more sensitive to outliers in the data. Having said that, both variables clearly keep their explanatory power on a daily basis and with deseasonalized data.

In sum, this section has demonstrated that the basic model and the squared stock model (both linear in the parameters) qualify for forecasting purposes since in-sample as well as out-of-sample forecasts have shown a strong performance. In terms of explanatory power, the squared stock model was slightly ahead. Analyzing data with shorter time intervals is possible and the introduced additional explanatory variable, air temperature, proves to be highly significant even with deseasonalized data and is robust to a change in the data frequency. The traditional variable of interest, storage, also keeps its significance in this case, particularly for the short term. Therefore, this model enhances the explanatory power of reduced-form models when the fundamental convenience yield estimate is incorporated.

5.7 Conclusion

The analysis presented in this chapter has advanced the modelling of the convenience yield of natural gas. We have shown that, beyond the storage level of gas, temperature is an important additional explanatory variable. The reason is that a sizable fraction of demand reacts very sensitive to temperature changes and, thereby, impacts on the current scarcity of gas. This fraction consists, to a large extent, of heating demand from domestic and commercial customers. In a linear two-variable model up to 75% of convenience yield variability could be explained in-sample. Furthermore, we have presented two extensions of this basic model: A linear model with the squared stock level as an additional explanatory variable and a (smooth-transition) regime-switching model. We have also analyzed the influence of the crude oil market.

The regime-switching model has not yielded an improved explanatory power, but it has shown that the variability of the basis can also increase rather than decrease with rising stock levels. This finding is reasonable for the gas market when stock levels are much closer to the capacity limit than to complete depletion as was the case in our dataset. The squared stock model has supported Brennan's theory of an increase in the cost of physical storage implied in the net storage cost when stock levels are high and the storage horizon is sufficiently short. It has also slightly amended the explanatory power of the model, especially with respect to short-term convenience yields. A strong link of convenience yields to the crude oil market could not be established although long-term futures (and their bases) might be influenced by the crude oil convenience yield to a certain degree. This is in line with current literature where co-movements of gas and crude oil prices are documented, but no short-term relationship is found. We have shown that the slight superiority of the squared stock model holds in a forecasting exercise so this model qualifies for forecasting. An analysis of daily deseasonalized data has also demonstrated that integrating a fundamental convenience yield forecast into the pricing of reduced-form models could enhance the latter.

An implicit assumption of the analysis was that the established pipeline connections to the European mainland (e.g. the Bacton-Zeebrugge Interconnector), only affect the convenience yield in case of exogenous events. In fact, gas contracts in continental Europe are still largely oil-indexed and if prices on both sides of the canal (market based vs. oil price based) differ by more than the transportation cost, gas will flow to the market with the higher price. While the *level* of UK prices is certainly influenced by this arbitrage possibility, there is a priori no reason to assume a permanent distortion of the interest-adjusted basis. This means that the difference between futures and spot prices in UK is assumed to remain unchanged as long as the pipeline operates normally. In case of an extraordinary event the effect will be comparable to shocks to domestic production, which our model captures in the form of storage reactions. While this seems economically justified, the ultimate proof of this assumption is certainly complicated. It would hence be another interesting research question for a separate study, especially since Stronzik et al. (2008) state that the UK market for storage is not yet working efficiently.

Furthermore, we have mentioned that an important motivation for this study has been the liberalization of the continental European gas markets and the necessity to find accurate spot price models to value derivative contracts emerging in these markets. However, the transferability of our results to the European mainland also builds on an important assumption: The UK, just as the US, has been relying on its largely autonomous gas supply until recently. On the contrary, many other European countries heavily rely on imports, and in the future these imports will concentrate on an oligopoly of suppliers (Russia, Norway and Algeria to mention first).[29] Hence, if we want to transfer our results to these countries, we need to assume that this differing supply situation has no impact on the choice of variables determining the convenience value of gas in the mainland markets. While we believe that stock levels will rather gain importance in very import-dependent countries, it is conceivable nonetheless that the convenience yield process becomes more irregular. Possible reasons for this could be more frequent supply disruptions and the limited operational capacity of the storage system such as injection and withdrawal capacity. Therefore, a fundamental model taking account of operational capacity limits would similarly be an interesting idea for future research related to this topic.

[29] Cf. National Grid (2007), p. 45.

Chapter 6

Application to pricing

So far, we have shown that it is appropriate to model natural gas prices individually. Furthermore, it has been demonstrated that no spot price model specific to the gas market has been proposed in literature so far. Therefore, the motivation of this thesis is to derive and test such a model. Chapter 4 has pointed out that the convenience yield is a parameter which appears in both structural and reduced-form models, and that it can be used to include fundamental information into a reduced-form model framework. Consequently, chapter 5 has analyzed the convenience yield dynamics in the natural gas market in detail and has finally proposed a customized regression model to make fundamental predictions for this parameter.

The following chapter documents the development and testing of a hybrid spot price model, i.e. a reduced-form framework including fundamental information, with empirical data from the UK and the US gas markets. The first section is dedicated to the model conception. The basis is the stochastic convenience yield (SCY) model by Gibson and Schwartz (1990) and Schwartz (1997), whose closed-form solution for futures prices is derived below, before the model extension is explained. In the second section, the empirical data is analyzed, and the hybrid model is estimated and benchmarked against the SCY model. Different estimation methods are presented together with the corresponding results. The last section of the chapter concludes.

6.1 Model

In this section we show how one may accommodate the convenience yield projection from the squared stock model (5.8) in a reduced-form model of the spot price. In section 4.3, we have shown that among the models with a convenience yield variable, the SCY model by Gibson and Schwartz (1990) and Schwartz (1997) is uncomplex and flexible. The fact that the model is widely applied in the energy trading practice supports our assessment. Nevertheless, it does not take account of fundamental information regarding the determination of convenience yields in the market. Hence, we choose this model as the basis and extend it by including a convenience yield forecast. The SCY model has two risk factors, the spot price S and the convenience yield δ, and their dynamics are characterized by (4.3) and (4.5):

$$\begin{aligned}\frac{dS}{S} &= (\mu-\delta)dt+\sigma_1 dW_1\\ d\delta &= \kappa(\alpha-\delta)dt+\sigma_2 dW_2.\end{aligned}$$

Schwartz (1997) provides a closed-form solution for futures prices in the model together with a recursive procedure to estimate the value of the unobservable state variables from futures price data. We will adopt this way of estimation since gas futures are the most liquid contracts traded on exchange. To do so, one has to know the model's solution for the futures price upfront. The section starts with the derivation of the closed-form futures price before we derive our extended model which includes the fundamental convenience yield forecast.

6.1.1 Schwartz (1997) model

To find the futures price of the SCY model in closed form, one can apply the PDE method explained in section 3.2. Let us state the change in the futures price as a function of a change in the spot price and the convenience yield with the help of Ito's Lemma:

$$\begin{aligned} dF &= \eta(F)dt + \frac{\partial F}{\partial S}\sigma_1 S dW_1 + \frac{\partial F}{\partial S}\sigma_2 dW_2 \qquad (6.16) \\ \eta(F) &= \frac{\partial F}{\partial t} + \frac{\partial F}{\partial S}(\mu - \delta)S + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma_1^2 S^2 + \frac{\partial F}{\partial \delta}\kappa(\alpha - \delta) \\ &\quad + \frac{1}{2}\frac{\partial^2 F}{\partial \delta^2}\sigma_2^2 + \frac{\partial^2 F}{\partial S \partial \delta}\rho\sigma_1\sigma_2 S. \end{aligned}$$

The replicating portfolio must contain not only stakes in the commodity and the money market account (MMA), but also a position in a second future which is sensitive to changes in the convenience yield. The reason is that by owning the commodity the convenience yield risk cannot be hedged. At the time of investment the value of the second future position is zero so the replicating portfolio must have a value of zero as well:

$$P = \psi S + \chi B \stackrel{!}{=} 0.$$

We can eliminate one weight and set

$$\chi = -\frac{\psi S}{B}.$$

The changes in the replicating portfolio are therefore

$$dP = \psi dS - \frac{\psi S}{B} dB + \upsilon d\tilde{F},$$

where υ denotes the portfolio weight and $d\tilde{F}$ the change of the second future's price as given by (6.16). Plugging in this relationship gives

$$dP = \left[\psi(\mu - r)S + \upsilon\eta(\tilde{F})\right] dt + (\psi + \upsilon\frac{\partial\tilde{F}}{\partial S})\sigma_1 S dW_1 + \upsilon\frac{\partial\tilde{F}}{\partial\delta}\sigma_2 dW_2.$$

Equating the diffusion parts of the future and the replicating portfolio yields the following portfolio weights:

$$\begin{aligned}
\psi &= \frac{\partial F}{\partial S} - \frac{\partial F/\partial\delta}{\partial\tilde{F}/\partial\delta}\frac{\partial\tilde{F}}{\partial S} \\
\upsilon &= \frac{\partial F/\partial\delta}{\partial\tilde{F}/\partial\delta}.
\end{aligned}$$

By equating the drift terms and substituting these portfolio weights, one obtains

$$\eta(F) = \left(\frac{\partial F}{\partial S} - \frac{\partial F/\partial\delta}{\partial\tilde{F}/\partial\delta}\frac{\partial\tilde{F}}{\partial S}\right)(\mu - r)S + \frac{\partial F/\partial\delta}{\partial\tilde{F}/\partial\delta}\eta(\tilde{F}),$$

which can be rearranged to represent a constant ratio of excess returns to convenience yield sensitivity that holds for any two futures contracts:

$$\frac{\eta(F) - \partial F/\partial S(\mu - r)S}{\partial F/\partial\delta} = \frac{\eta(\tilde{F}) - \partial\tilde{F}/\partial S(\mu - r)S}{\partial\tilde{F}/\partial\delta} = \lambda_\delta.$$

This ratio is interpreted as the risk premium for convenience yield risk since it is the excess return for convenience yield risk divided by the "factor loading", i.e. the amount of convenience yield risk for one futures contract.[1] In consequence, the total expected return of the futures contract must be

$$\eta(F) = \frac{\partial F}{\partial S}\lambda_S S + \frac{\partial F}{\partial \delta}\lambda_\delta$$

with $\lambda_S = \mu - r$. Setting this equal to the drift in (6.16) and rearranging terms, one obtains the fundamental PDE of the stochastic convenience yield model

$$\begin{aligned}&\frac{\partial F}{\partial t} + \frac{\partial F}{\partial S}(\mu - \lambda_S - \delta)S + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}\sigma_1^2 S^2 \\ &+\frac{\partial F}{\partial \delta}\kappa[(\alpha - \delta) - \lambda_\delta] + \frac{1}{2}\frac{\partial^2 F}{\partial \delta^2}\sigma_2^2 + \frac{\partial^2 F}{\partial S \partial \delta}\rho\sigma_1\sigma_2 S = 0.\end{aligned}$$

The solution to this equation is rather involved. It is derived in Bjerksund (1991) and adopted by Schwartz (1997):

$$\ln F_{t,T} = \ln S_t - \delta\frac{1 - e^{-\kappa(T-t)}}{\kappa} + A(T-t) \tag{6.17}$$

$$\begin{aligned}A(T-t) = &\left(r - \alpha^* + \frac{1}{2}\frac{\sigma_2^2}{\kappa^2} + \frac{\sigma_1\sigma_2\rho}{\kappa}\right)(T-t) \\ &+\frac{1 - e^{-2\kappa T}}{4\kappa^3}\sigma_2^2 + \left(\hat{\alpha} + \frac{\sigma_1\sigma_2\rho}{\kappa} - \frac{\sigma_2^2}{\kappa^2}\right)\frac{1 - e^{-\kappa T}}{\kappa}\end{aligned} \tag{6.18}$$

The solution shows, among other properties, the implications of the model formulation for the term-structure of futures prices. The log-spot price serves as a level factor which shifts all futures prices up or down. In turn, the convenience yield impacts on the curvature of the term structure since

[1] Note that whenever it is referred to the term "convenience yield" in the model, it is to be regarded *net of storage cost* since a storage cost parameter does not occur in the model. The convenience yield variable may therefore become negative contrary to Brennan's definition.

its coefficient is non-linear in the time to maturity $T-t$. The model (6.17) serves as a performance benchmark for the extended version which we derive in the following subsection.

6.1.2 Extended model

As outlined in the introduction to this chapter, the SCY model does not completely meet our requirements. Firstly, it has originally been developed for the crude oil market instead of the gas market. Secondly, Gibson and Schwartz (1990) and Schwartz (1997) assume that variation in convenience yield can be purely explained by a mean-reverting stochastic process. Chapter 5 has shown that this is a very simplifying and debatable assumption. Authors such as Pirrong (1998), Clewlow and Strickland (2000) and Carlson et al. (2007) have also argued that the constant volatility parameter misspecifies the true price dynamics due to a negligence of fundamental factors which impact on volatility. For these reasons, we propose an extended model in which the convenience yield stems from two sources, a *fundamental component* y^s, driven primarily by national gas inventories and air temperature, as well as a *preference component* $\tilde{\delta}$, which stands for the speculation preferences of investors. Their attitude towards speculation is influenced by an interplay of numerous macroeconomic factors (e.g. the current price and price expectation of substitute goods) which cannot easily be modeled individually. Both components revert towards a long-run mean in line with a Neokeynesian model setup.[2] We allow for different mean-reversion speeds of the two variables. The total convenience

[2] Frictions in the spot market (e.g. pipeline capacity) lead to frequent temporary deviations from the long-run path. Bessembinder et al. (1995) show that this is a common phenomenon in commodity markets and that the market anticipates the readjustments to a considerable degree.

yield is defined as the sum of both components:

$$\delta = \tilde{\delta} + y^s. \tag{6.19}$$

In this way, we attempt to explain parts of the changing price volatility without introducing an additional stochastic variable. More precisely, our fundamental component is defined as the (seasonally adjusted) convenience yield from (5.15b)[3]

$$\begin{aligned} \triangle y_t^s &= \beta_{\Delta stock} \triangle x_{stock,t}^s + \beta_{\Delta(stock^2)} \triangle x_{stock^2,t}^s \\ &\quad + \beta_{\Delta temp} \triangle x_{temp,t}^s + \beta_{y_{t-1}^s} y_{t-1}^s + \epsilon_t. \end{aligned}$$

In the following, unless stated otherwise, y_t^s only stands for the instantaneous convenience yield proxied by the nearest maturity available in the market. We assume that a continuous-time limit of this equation exists since all explanatory variables are metrically scaled and can be measured at any moment in time.[4] In turn, the preference component $\tilde{\delta}_t$ follows the mean-reversion process defined by Gibson and Schwartz (1990) in (4.5)

$$d\tilde{\delta} = \kappa(\alpha - \tilde{\delta})dt + \sigma_2 dW_2. \tag{6.20}$$

Unfortunately, incorporating the forecast into the spot price model is slightly involved since y_t^s is not known as an analytic function of time. We recall that both storage levels and temperature are exogenous variables (or forecasts). Therefore, including y^s directly in the stochastic process of the SCY model is not possible and we cannot obtain the resulting futures price solution with standard Ito calculus. However, due to the deseasonalization

---

[3] The necessity of ARCH-coefficients depends on the measurement interval and remains to be explored.

[4] In case of a one-period forecast, the prediction in levels is given by $\hat{y}_{t+1}^s = y_t^s + \triangle \hat{y}_{t+1}^s$, where a hat denotes an estimate of the true variable.

explanatory variables in (5.15b) as well as the levels of the dependent variable y_t^s have mean zero. In addition, y_t^s has a normal distribution, provided that the explanatory variables are normal.[5] Since they are residuals of a seasonal function, this can be assumed. Hence, it is possible to maintain the distributional assumptions of the SCY model even if we apply it on a set of *synthetic futures prices* which we have "netted" from the influence of the fundamental component. In particular, creating such a price set maintains the normality of the stochastic increments. The expected benefit of the "netting" is that the fundamental convenience yield will explain some of the variability in convenience yields which cannot be explained through an Ornstein-Uhlenbeck process. After the netting, the reduced-form model and the implied spot prices are estimated with the synthetic futures. This procedure is expected to yield less noisy parameter estimates. Figure 6.1 shows the relationship of the expected spot price drift in the SCY model (called *standard model* in the following) and the *extended model* under the risk-neutral and the empirical measure. $\tilde{S}$ denotes the synthetic spot price which corresponds to S in the standard model. Since we only need two additional futures contracts in addition to the asset and the MMA to trade in the "synthetic" asset, we assume that the market is complete. The required assumption for the replication of the synthetic futures is, hence, fulfilled.

Next, we investigate the effect of the change of drift on the futures price. Commodity returns derive from two sources, an expected spot price drift which is given by $E_t^Q(\mu_{S,u}) = r - E_t^Q(\delta_u)$ in the SCY model, and an expected convenience yield $E^Q(\delta_u)$ for any time u until maturity of the future. If the convenience yield δ_u rises at time u , the futures contract becomes less attractive compared to the asset such that its value will immediately decrease and vice versa. Due to autocorrelation in the

[5] Any linear combination of independent normal distributions is normal itself.

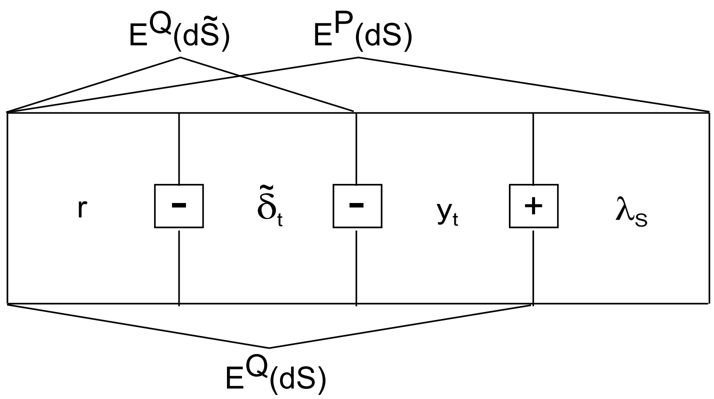


Figure 6.1: Components of spot price drifts: Standard vs. extended model.

(instantaneous) convenience yield, the change in the current convenience yield also changes its future expectations until maturity.

To quantify the effect on the futures price, we look at the discretized equations of the stochastic processes (4.3) and (4.5) under the equivalent martingale measure. Assume that the time steps are $\triangle t = \frac{T-t}{n}$ with $n \in \mathbb{N}$. The approximate system equations become

$$\begin{aligned}\triangle S_t &= (r - \tilde{\delta}_{t-1})S_{t-1}\triangle t + \sigma_1 S_{t-1}\triangle W_1^Q && (6.21a)\\ \triangle\tilde{\delta}_t &= \kappa(\alpha^* - \tilde{\delta}_{t-1})\triangle t + \sigma_2\triangle W_2^Q. && (6.21b)\end{aligned}$$

α^* represents the risk-neutral mean reversion level in the SCY model and is related to α in (4.5) by $\alpha^* = \alpha - \frac{\lambda_\delta}{\kappa}$. Evidently, if the convenience yield is at the mean level in t, its risk-neutral expectation is flat with

$$E_t^Q(\delta_u) = \alpha^* \ \forall \ u \in [t, T]. \qquad (6.22)$$

If either $\tilde{\delta}_t$ or y_t^s are not at the mean level, the market expectation for the future must be different due to serial correlation. By rearranging (6.21b) to

$$\tilde{\delta}_t = \kappa\alpha^*\triangle t + (1-\kappa\triangle t)\tilde{\delta}_{t-1} + \sigma_2\triangle W_2^Q$$

one sees that $CORR\left[\tilde{\delta}_t, \tilde{\delta}_{t-1}\right] = 1-\kappa\triangle t$. However, the autocorrelation in the fundamental component can be different. It is implicitly estimated in (5.15b) since

$$CORR\left[y_t^s, y_{t-1}^s\right] = 1-\beta_{y_{t-1}^s}.$$

Accordingly, the mean-reversion speed of y^s can be defined as

$$\bar{\kappa} \approx -\frac{\beta_{y_{t-1}^s}}{\triangle t}, \tag{6.23}$$

assuming that $\triangle t$ equals the step length in the regression. Since the convenience yield is a single risk factor, all uncertainty of the fundamental forecast is attributed to noise in the stochastic process such that the risk premium of y^s is zero. Given that the deseasonalized stock and temperature variables have an expected value of zero, we can, hence, write the risk-neutral conditional k-period forecast as

$$E_t^Q(y_{t+k\triangle t}^s) = E_t(y_{t+k\triangle t}^s) = (1-\bar{\kappa}\triangle t)^k y_t^s. \tag{6.24}$$

To derive the immediate response of convenience yield changes in the futures price (i.e. the level change), let us assume that the convenience yield has been at the steady state value α^*, but the fundamental component increases at t, expressed in a positive value of y_t^s. Consequently, the current period's value of the future compared to the asset changes by

$$\begin{aligned}\frac{\partial \ln F_{t,T}}{\partial y_t^s} \quad = \quad & -\Big(y_t^s + E_t^Q(y_{t+\triangle t}^s) + E_t^Q(y_{t+2\triangle t}^s) + ... + E_t^Q(y_{(n-2)\triangle t}^s) \\ & + E_t^Q(y_{(n-1)\triangle t}^s)\Big)\triangle t.\end{aligned}$$

This can be simplified to a geometric sum by inserting (6.24) such that

$$\frac{\partial F_{t,T}}{\partial y_t^s} = -\sum\nolimits_{k=0}^{n-1}(1-\bar{\kappa}\triangle t)^k y_t^s \triangle t = -\frac{1-(1-\bar{\kappa}\triangle t)^n}{1-(1-\bar{\kappa}\triangle t)}.$$

Then, by letting the number of steps in the time mesh $n \longrightarrow \infty$, we get the continuous time limit

$$\frac{\partial \log F_{t,T}}{\partial y_t^s} = -\frac{1-e^{-\bar{\kappa}(T-t)}}{\bar{\kappa}}.$$

Comparing this result to the convenience yield coefficient in the futures price solution (6.17) shows that in the extended model the sensitivity of $\log F_{t,T}$ w.r.t. to y_t^s is identical to that of $\tilde{\delta}_t$ under the stated assumptions.

In sum, the extended model can be characterized by the new additive convenience yield (6.19), the dynamics of the risk factors, i.e. the spot price and the preference component of the convenience yield

$$d\tilde{S} = (r-\tilde{\delta})\tilde{S}dt + \sigma_1 \tilde{S} dW_1^Q \tag{6.25a}$$
$$d\tilde{\delta} = \kappa(\alpha^* - \tilde{\delta})dt + \sigma_2 dW_2^Q, \tag{6.25b}$$

together with the continuous time equivalent of the fundamental convenience yield dynamics (6.26a), the relationship between observed and synthetic futures prices (6.26b) and the closed-form solution for synthetic futures prices (6.26c)

$$\begin{aligned} dy_t^s &= \beta_{dstock} dx_{stock,t} + \beta_{d(stock^2)} dx_{stock^2,t} \\ &\quad + \beta_{dtemp} dx_{temp,t} + \beta_{y_t^s} y_t^s \end{aligned} \tag{6.26a}$$
$$\ln \tilde{F}_{t,T} = \ln F_{t,T}^s + y_t^s \frac{1-e^{-\bar{\kappa}(T-t)}}{\bar{\kappa}} \tag{6.26b}$$
$$\ln \tilde{F}_{t,T} = \ln \tilde{S}_t - \tilde{\delta}_t \frac{1-e^{-\kappa(T-t)}}{\kappa} + A(T-t). \tag{6.26c}$$

The variable $\ln F^s_{t,T}$ denotes the input of futures prices which have been deseasonalized with the seasonal function (5.13), $\ln \tilde{F}_{t,T}$ denotes the resulting synthetic prices and $A(T-t)$ is given by the expression in (6.18). The following section estimates this model by following the order of the last three equations. While in (6.26a) and (6.26b) the aim is to estimate the left-hand side, in (6.26c) the synthetic spot prices will be estimated recursively, given (6.25a) and (6.25b). The details of the estimation procedure are also provided in the following section.

6.2 Empirical estimation

In the last section, we have shown the closed-form solution for futures prices in the two-factor SCY model by Gibson and Schwartz (1990) and Schwartz (1997). Furthermore, we have proposed an extended model with additional systematic convenience yield changes, which are driven by fundamental variables. In this section, we explain and document the estimation of the two models and compare the models' performance. The critical question is whether spot price forecasts are more precise and parameter estimates more robust for the extended model. We use US and UK gas market data for the analysis. Firstly, we calculate convenience yield projections for the US market, since the focus in chapter 5 was on the UK market only. We analyze the corresponding estimation results. Next, we compute the synthetic futures prices net of the fundamental convenience yield component and show that the resulting futures price innovations are in line with the distributional assumptions made in the last section. Thereafter, we cast (6.26c), (6.21a) and (6.21b) in state space form and explain how to apply the Kalman filter to estimate the spot price and the residual convenience yield $\tilde{\delta}$. We continue by running simulations of the Kalman filter algorithm to choose the most accurate estimation routine and apply this routine to the dataset at hand. Finally, standard and extended model are compared regarding the forecast errors produced.

6.2.1 Convenience yield projections

Fundamental convenience yields for the extended model are projected with the discretized version of (6.26a) given by (5.15).Since the next month's contract (M1) has proved to be a rather inaccurate proxy for changes in daily conditions, the convenience yield specification applied here uses the day-ahead price as a spot price proxy (i.e. the OCM contract for the UK dataset). For reduced-form models of the spot price, data frequency should be weekly or daily. The significance of the ARCH-coefficients in the convenience yield model does actually not matter for the point estimate of the forecast. However, due to the strong autocorrelation of the convenience yield, deciding on the desired forecast horizon is a relevant issue. The in-sample test of our model uses a one-period forecast since the same is used for the estimation of the reduced-form model (see the following subsection).

As for the UK dataset, little further treatment of the data is needed. We keep the daily frequency used in chapter 5 because of the short history of storage data available. The resulting time series of 568 observations is long enough such that the asymptotic properties of the estimation procedure will be fulfilled. Hence, the projection of the instantaneous UK convenience yield is already available from the estimation reported in Table 5.9 (line 1). Since we have only 6 liquid futures contracts available, an out-of-sample test is difficult to run with the existing dataset. Therefore, an update of the OCM and ICE month futures time series until 04/09 has been obtained. The number of additional observations is 339. Besides the one-period forecast, a longer-term forecast is used as a stress test. That is, we recursively generate a five-period dynamic forecast of the convenience yield differences. The following list sums up the steps to arrive at the convenience yield projections:

1. Compute days to maturity for futures contracts at each date
2. Apply spline to generate futures series with equal maturities
3. Calculate convenience yield time series according to (5.14)
4. Remove outliers and interpolate the respective observations
5. Deseasonalize convenience yields, stocks and temperature
6. Estimate convenience yield model in first differences
7. Generate (recursive) forecasts of first differences $\triangle_k \hat{\theta}_{t+k}$ and add forecast to the last observation ($\hat{\theta}_{t+k} = \theta_t + \triangle_k \hat{\theta}_{t+k}$)

The US dataset has a much longer history. Storage announcements, provided by the EIA, are available since 1994. In turn, the reporting is weekly (every Friday), not daily. Hence, we are obliged to use a weekly frequency of observations for the US market. Monthly US natural gas futures traded on the NYMEX are available since 1990. Maturities range from one month to several years. Market liquidity is very high at the short end, but declines significantly for longer terms. The official daily settlement quotes are obtained from Bloomberg. We choose to use all contracts with maturities up to 18 months. The limiting factor of the history of convenience yields is the spot market index. This is the counterpart of the OCM contract in UK. Both the ICE and Bloomberg provide such an index. The Bloomberg index is, in fact, available since 1991, but liquidity was insufficient in the first years. It is a volume-weighted price average of all (reported) day-ahead trades in the OTC market for delivery at Henry Hub during the last half hour before market closing at the NYMEX. The ICE index is much younger with a history dating back to 04/2001. The price differences between these indices are negligible so we use a composite index reported also by Bloomberg.

We choose a reasonable starting point for the sample by analyzing the historical price chart of the front month contract (cf. Figure 6.1). Most notably, it displays a significantly elevated volatility and a long-term upward trend starting in 2000, such that the beginning of that year constitutes a reasonable starting point. It roughly corresponds to the starting point of the UK time series (04/00) as well. The resulting dataset has a length of 513 observations and ranges from 01/00 to 09/09.

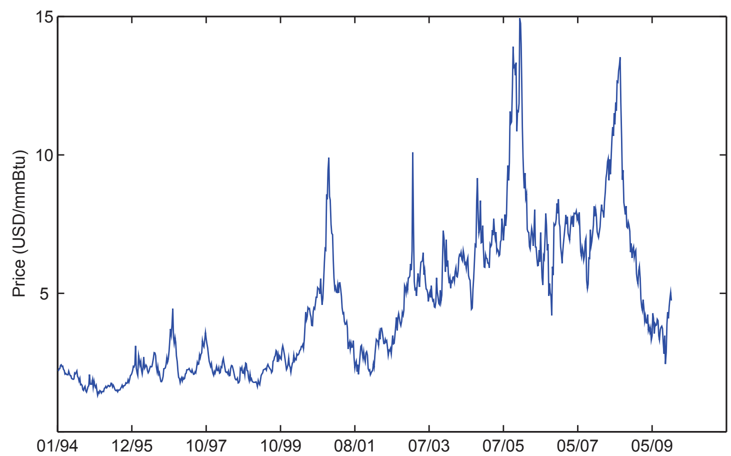


Figure 6.2: Historical weekly settlement prices of NYMEX front month future (1994-2009).

The maturity-matching USD interest rates are taken again from the BBA LIBOR database. However, the cumbersome part of the convenience yield model for the US is the computation of a "market" temperature. Two additional complications with regard to the UK counterpart occur. The minor one is the fact that not only heating, but also cooling activity in the economy will affect the gas demand in this market. Air conditioners are widely used, and air temperatures in many regions rise well above a comfortable room temperature level in the summer. As gas-fired power generation has a high importance in the US (cf. Appendix 5), gas demand will react to temperature increases above the threshold level. To account for this fact we use the sum of Heating and Cooling Degree Days (HDDs and CDDs) at each observation date to represent the temperature variable. HDDs are the amount by which the current temperature undershoots the comfortable room temperature level (set to $18°C$ for the US). CDDs are the amount by which it overshoots this level. Minimum and maximum daily temperatures instead of averages are used for the US to account for the fact that both heating and cooling activity can occur at the same day at different times or in different regions

$$\begin{aligned} HDD &= \max\{18°C - \text{min. daily temperature}; 0\} \\ CDD &= \max\{\text{max. daily temperature} - 18°C; 0\}. \end{aligned}$$

We keep denoting the sum of these two variables as the "temperature" variable in the following. It is immediately clear that the expected sign of the associated regression coefficient is positive, since higher "temperature" value in this case signifies higher demand.

The major problem is the geographical extension of the market. While the central England temperature is a viable proxy for the whole UK, the US hosts a variety of climes. Since the temperature variable is a demand proxy, the HDDs and CDDs of the entire US are approximated by a population-weighted average of the maximum and minimum daily temperature measured in the 15 most populated metropolitan areas of the US which account for one third of the total population. The weights to compute national HDDs and CDDs are also used by the National Climatic Data Center (NCDC). However, data from this source is available only on a monthly basis and until 1992. The population data is based on the year 2000 US census and is obtained from the National Bureau of Census. The daily maximum and minimum air temperature series are extracted from the MIDAS GL (Global Weather Observations) database of the BADC. The weather stations closest to the center of each metropolitan area are selected. In the rare case in which no observation at all was recorded for a particular day, a linear interpolation between the neighboring daily observations is used.

In the process of deseasonalization we follow the steps outlined in the list above. Between 12 and 21 (out of 513) observations per series are identified as outliers and replaced by a linear interpolation between neighboring observation dates. Summary statistics for the levels of the deseasonalized convenience yields (net of outliers) in comparison to the UK statistics are given in Table 6.2. To investigate whether the convenience yield model holds also in the US, we look at the whole term structure of convenience yields for months 1 to 18. Regarding the US data on a stand-alone basis, the table shows that with respect to the normality assumption made for the estimator $\hat{\theta}$, the main concern is the excess kurtosis, which is 7.25 in the nearest-to-maturity convenience yield. It diminishes gradually with rising maturity until it drops to an acceptable value of less than 1 for the ten-month yield. We therefore check for a distorting effect on the model performance before estimating the reduced-form model. Besides this phenomenon, the distributional assumptions are fulfilled to an acceptable extent. Particularly, the median does not lie far apart from the mean (of zero) and the effect of skewness and remaining outliers in levels (rather than in first differences) seems rather mild. We stress again that the extreme observations in levels are not spikes, but they mostly evolve over a number of periods. The observed pattern might be rather interpreted as a combination of temporarily slow mean-reversion and strong heteroscedasticity of the increments. Looking at the UK statistics (including now the update until 04/09) we see that the excess kurtosis is slightly smaller, but the standard deviation is higher which produces the absolutely higher minimum and maximum values. Apart from this, there is no remarkable difference in the distributional properties.

Table 6.3 shows the final regression results of the first-difference speci-

Basis	Med	Stdev	Skew	Kurt	Max	Min
US data						
M1-D1	.0692	1.77	-1.75	10.25	5.54	-9.79
M2-D1	.0078	.77	-1.59	8.29	2.15	-4.06
M3-D1	.0237	.57	-1.07	7.03	2.17	-2.79
M4-D1	-.0070	.49	-.04	7.36	2.42	-2.02
M5-D1	.0064	.41	.04	6.37	1.92	-1.58
M6-D1	.0116	.36	-.04	5.22	1.53	-1.23
M7-D1	.0138	.32	.00	4.79	1.30	-1.09
M8-D1	.0125	.30	-.13	4.46	1.15	-1.02
M9-D1	.0044	.28	-.13	4.21	1.07	-.98
M10-D1	.0027	.26	-.11	3.98	1.01	-.92
M11-D1	.0054	.24	-.10	3.85	.93	-.84
M12-D1	.0045	.23	-.09	3.69	.86	-.76
M13-D1	.0103	.22	-.08	3.80	.83	-.72
M14-D1	.0132	.21	-.06	4.02	.84	-.69
M15-D1	.0071	.20	.00	4.19	.83	-.67
M16-D1	.0068	.20	.02	3.98	.76	-.63
M17-D1	.0087	.19	-.05	3.77	.70	-.59
M18-D1	.0125	.18	-.08	3.56	.66	-.55
UK data						
M1-OCM	.0589	2.66	-1.01	8.37	12.50	-14.46
M2-OCM	.0381	1.57	-.63	6.45	7.50	-7.06
M3-OCM	.0572	1.08	-.55	5.59	4.89	-4.71
M4-OCM	.0124	.79	-.45	5.40	3.59	-3.52
M5-OCM	.0550	.67	.01	5.95	2.90	-2.80
M6-OCM	-.0003	.55	-.02	5.31	2.64	-2.46

Table 6.2: Summary statistics for annual deseasonalized spot convenience yields.

		Main				ARCH	
$Basis$	$\bar{y}$	$\hat{\beta}_{\Delta stock}$	$\hat{\beta}_{\Delta(stock^2)}$	$\hat{\beta}_{\Delta temp}$	$\hat{\beta}_{y_{t-1}}$	$\hat{\beta}_{\epsilon^2_{t-1}}$	$\hat{\beta}_{cons}$
M1-D1	-.358	-5.880 (-5.84)***	-3.274 (-1.63)	.664 (3.15)***	-.498 (-11.93)***	.318***	1.120***
M2-D1	-.253	-1.898 (-5.70)***	-.338 (-0.53)	.315 (5.29)***	-.284 (-13.71)***	.449***	.123***
M3-D1	-.202	-.401 (-1.74)*	-.123 (-0.26)	.181 (4.11)***	-.151 (-7.05)***	.286***	.058***
M4-D1	-.180	.087 (0.54)	.171 (0.56)	.117 (4.41)***	-.066 (-5.40)***	.417***	.029***
M5-D1	-.146	.058 (0.44)	.046 (0.19)	.089 (4.16)***	-.058 (-5.10)***	.378***	.019***
M6-D1	-.143	-.139 (-1.29)	-.131 (-0.61)	.082 (4.37)***	-.065 (-5.17)***	.323***	.014***
M7-D1	-.140	-.221 (-2.17)**	.032 (0.16)	.079 (4.35)***	-.068 (-5.12)***	.272***	.012***
M8-D1	-.130	-.243 (-2.57)**	.158 (1.02)	.077 (4.62)***	-.060 (-4.66)***	.217***	.010***
M9-D1	-.123	-.156 (-1.78)*	.078 (0.53)	.070 (4.38)***	-.062 (-4.85)***	.166***	.008***
M10-D1	-.107	-.071 (-0.86)	.110 (0.81)	.063 (4.11)***	-.058 (-4.57)***	.148**	.007***
M11-D1	-.093	-.015 (-0.21)	.040 (0.33)	.049 (3.68)***	-.046 (-3.78)***	.155**	.006***
M12-D1	-.072	-.061 (-0.91)	.088 (0.80)	.049 (3.95)***	-.051 (-4.04)***	.162**	.005***
M13-D1	-.068	-.125 (-2.06)**	.135 (1.39)	.048 (4.39)***	-.049 (-4.20)***	.212***	.004***
M14-D1	-.059	-.126 (-2.18)**	.088 (0.97)	.043 (4.49)***	-.046 (-4.25)***	.244***	.003***
M15-D1	-.051	-.054 (-1.01)	.118 (1.35)	.039 (4.25)***	-.042 (-4.00)***	.213***	.003***
M16-D1	-.045	-.040 (-0.75)	.137 (1.55)	.033 (3.51)***	-.045 (-4.13)***	.180***	.003***
M17-D1	-.039	-.006 (-0.13)	.123 (1.52)	.027 (3.25)***	-.033 (-3.19)***	.223***	.002***
M18-D1	-.034	-.049 (-1.12)	.085 (1.05)	.027 (3.21)***	-.044 (-4.02)***	.253***	.002***

Table 6.3: Regression results from convenience yield model (5.15) for the US.

fication (5.15b) for the US dataset (for the UK results see Table 5.9). The regression results are overall consistent with those for the UK. Since the convenience yield had to be slightly trend-corrected (stationarized) this time, we report $\bar{y}$, its median before deseasonalization, instead of $\hat{b}_1$ as an indicator for the average convenience yield. The temperature variable is strongly significant for all maturities. Mean-reversion speeds again tend to decrease with maturity, meaning that autocorrelation is strongest in long-term contracts. Here, the reversion speed computed with (6.23) lies between 26.8 for the shortest and 4.4 for the longest maturity. The reason for this phenomenon is that long-term contracts react less strongly to any temporary shocks in supply, demand or operations. A slight difference to the UK data exists regarding the importance of the storage variable. In the US market its significance is less pronounced than in the UK. One possible explanation is that supply is much more diversified and, therefore, potentially less volatile in the US.[6] In turn, the irrelevance of the squared storage term in this model specification is found in both datasets. Nevertheless, no misspecification is apparent from the results so we can use the fundamental component projected by the regression to compute the synthetic futures prices as described above. This is the topic of the next subsection.

6.2.2 Synthetic futures prices

In the following we describe the construction of the synthetic futures prices, the input for estimating the reduced-form model parameters afterwards. Let $\tilde{F}^e_{t,T}$ be the empirical counterpart of $\tilde{F}_{t,T}$, and $\bar{\kappa}^e$ the estimator for $\bar{\kappa}$,

[6] The UK's supply primarily comes from the North Sea whereas the major US supply regions currently are at the Gulf coast, Western Canada, Permian (Northern Texas), San Juan (New Mexico) and the Rocky Mountains to name only a few.

then following (6.26b) we have

$$\tilde{F}^e_{t,T} = F^s_{t,T} + \hat{y}^s_t \frac{1 - e^{-\bar{\kappa}^e (T-t)}}{\bar{\kappa}^e} \forall T \in [t, T_{\max}].$$

Variable $\hat{y}^s_t$ denotes the estimate for the instantaneous fundamental convenience yield component. To estimate the deseasonalized observed futures prices $F^s_{t,T}$, we use the interpolated futures prices from above, take logs and temporarily remove the time trend of each price series by linear regression. The seasonality function (5.13) is then estimated for the futures price series. We verify the significance of the seasonal parameters by looking at the confidence intervals and the functions' fit to the price series. Doing so unveils a difference between futures prices in the UK and the US, i.e. US prices show a less pronounced seasonality. Figure 6.3 shows the difference of seasonal influence for the futures contracts M1, M3 and M6 in both markets.

It is visible that the fit of the (smooth) seasonality function in the left charts is poorer. That is, a greater part of the dynamics of futures cannot be explained by pure seasonality. We test whether the influence of increased short-term power generation in the summer supports a half-yearly seasonal function, but the fit in this case even deteriorates. From an econometric perspective though, it does not matter whether or not we identify a pronounced seasonality as long as the netted time series are free of the deterministic effect.

To prove the validity of the approximate normality assumption for the futures price increments, we look at the histograms of the original (F) and the synthetic ($\tilde{F}$) futures price changes displayed in Figure 6.4.

It can be seen that the log-futures price changes are almost normally distributed and, hence, also symmetric. Therefore, the synthetic prices do not violate the model assumptions. In addition, the normal distribution

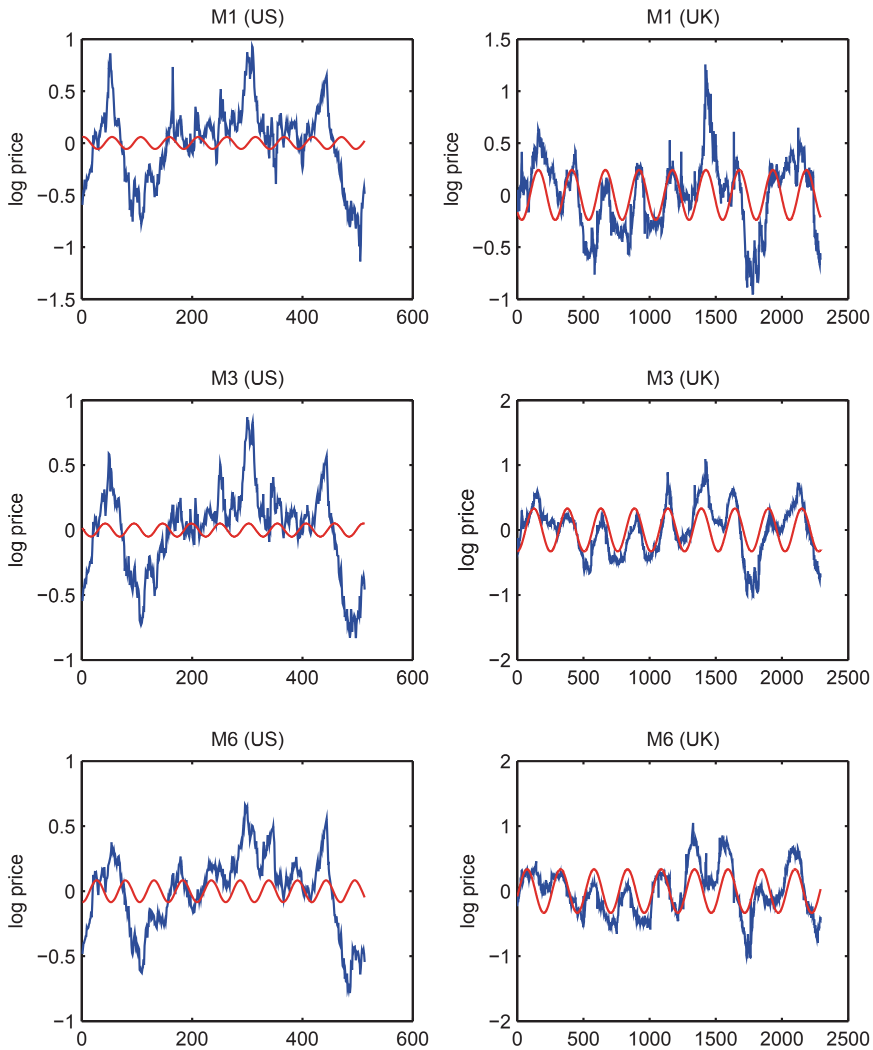


Figure 6.3: Comparison of seasonal fitting for US and UK futures.

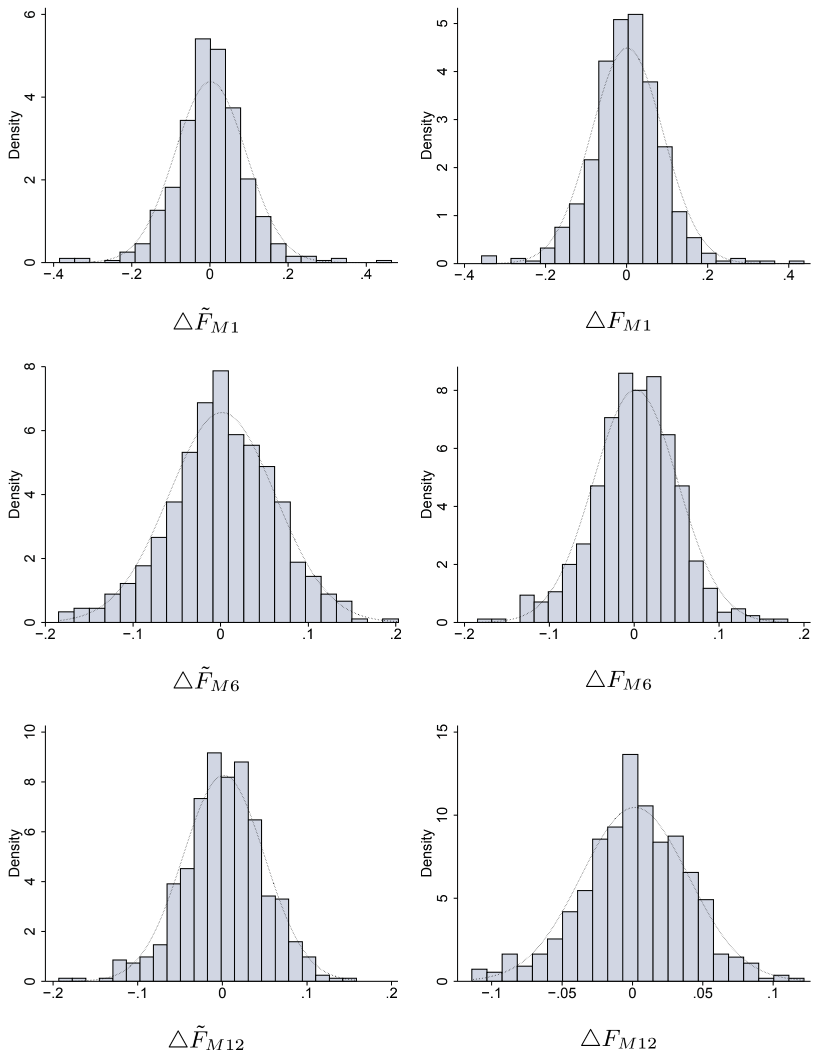


Figure 6.4: Histograms of US futures price changes (original and modified data).

of these increments is a prerequisite for the following reduced-form model estimation procedure. We describe the setup and the procedure in the following two subsections and test the reliability of the procedure thereafter. This is necessary to assure that estimates are not strongly influenced by a wrong statistical setup or numerical inaccuracies. Then, we continue by analyzing the empirical results.

6.2.3 State-space model

The empirical counterpart of the reduced-form model (6.26c) is a state-space model whose parameters can be estimated with the (discrete-time) Kalman filter as proposed in Schwartz (1997). We now shortly present the general idea of the Kalman filter before we elaborate on the main problems when estimating the state space model.

The Kalman filter is a recursive algorithm to estimate a vector of latent state variables of a time-series model. It iterates forward through the series. At each time step (i.e. observation date) it first corrects projections of the state variables (*"measurement update"*) made at the last step before it generates new projections for the next step (*"time update"*). The correction uses information from the observable (or measurement) variables which are linked to the state variables by the measurement equation of the model. The general form of this equation (Harvey (1989), ch. 3) is

$$y_t = Z_t \alpha_t + d_t + \epsilon_t, \tag{6.27}$$

in which y_t is an $N \times n$ vector of observations and α_t is an $m \times 1$ vector of state variables at time t. N corresponds to the length of the time series and n to the number of observable variables. In our case, the latter corresponds to the number of futures contracts used. The number of state variables m is 2, the synthetic spot price and the preference component of the convenience yield. It is important that the observables are linearly dependent on the state variables. This can be seen by the affine relationship created with design matrix Z_t and constant vector d_t. Finally, ϵ_t is a serially uncorrelated disturbance array with mean zero and covariance matrix H_t. In the SCY model $Z_t = Z$, $d_t = d$ and $H_t = H$, i.e. the elements of the measurement equation are time-invariant. In addition, H is diagonal.

The prediction uses a transition equation which creates a one-step forecast of the state variables conditional on their corrected values from the current time step. The transition equation of a state space model in general form is

$$\alpha_t = T_t \alpha_{t-1} + c_t + R_t \eta_t. \tag{6.28}$$

The variable T_t is an $m \times m$ time-dependent transition matrix which creates a linear relationship between the values of the state variables in successive periods, c_t is a constant and R_t is an $m \times g$ matrix linking the disturbance vector η_t to the state variables. Just as ϵ_t, η_t has mean zero and its elements are possibly correlated in the cross section (covariance matrix Q_t), but uncorrelated serially. In our case the transition system is time-independent, i.e. $T_t = T$, $c_t = c$, $R_t = R$ and $Q_t = Q$. We net the prices from the seasonal component before estimating the remaining parameters. This limits the number of parameters to be estimated with the rather involved Kalman filter algorithm. For this reason, we can generally adhere to the state space form of the standard SCY model. The following equations show the state space form used in Schwartz (1997) adapted to the notation of our extended model

Measurement equation system

$$
\begin{aligned}
y_{i,t} &= \ln F_{t,T_i} \quad \text{for } i = 1, ..., n \text{ and } t = 1, ..., N \\
Z_i &= \begin{pmatrix} 1 & -\frac{1-e^{-\kappa(T_i - t)}}{\kappa} \end{pmatrix} \quad \text{for } i = 1, ..., n \\
\alpha_t &= \begin{pmatrix} X_t \\ \tilde{\delta}_t \end{pmatrix}, \; X_t = \ln \tilde{S}_t \\
d_i &= A(T_i - t) \quad \text{for } i = 1, ..., n
\end{aligned}
$$

Transition equation system (first order approximation)[7]

$$
\begin{aligned}
T &= \begin{pmatrix} 1 & -\triangle t \\ 0 & 1 - \kappa \triangle t \end{pmatrix} \\
c &= \begin{pmatrix} (\mu - \frac{1}{2}\sigma_1^2)\triangle t \\ \kappa\alpha\triangle t \end{pmatrix} \\
R &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
Q &= \begin{pmatrix} \sigma_1^2 \triangle t & \rho\sigma_1\sigma_2\triangle t \\ \rho\sigma_1\sigma_2\triangle t & \sigma_2^2\triangle t \end{pmatrix}.
\end{aligned}
$$

The optimal parameter estimates are found by repeatedly running the Kalman filter over all observation dates N of the time series of futures contracts. At each run, a likelihood value is calculated. It is maximized by changing the parameter values by a superordinate optimization routine (*maximum likelihood estimation*). The parameter set of this state space formulation is $\Omega = \{r, \alpha, \kappa, \lambda, \mu, \rho, \sigma_1, \sigma_2, \boldsymbol{\xi}\}$ with $\boldsymbol{\xi}$ as the vector of standard deviations of measurement noise . The model as such is overdefined

---

[7] To set up the state space form, note that the SDE for the synthetic spot price under the empirical measure solves to $dX = (\mu - \frac{1}{2}\sigma_1^2 - \tilde{\delta})dt + \sigma_1 dW_1$ with $X = \ln S$.

since r and α are linearly interdependent. To solve this problem, Schwartz (1997) sets parameter r to the average of the 3-month T-bill rate observed during the estimation period. Here, μ is replaced by $r+\lambda_S$ and $r-\alpha$ is then replaced by trend variable v^*. The new parameter set contains one parameter less: $\tilde{\Omega} = \{\kappa, \lambda_S, \lambda_{\tilde{\delta}}, \rho, \sigma_1, \sigma_2, v^*, \boldsymbol{\xi}\}$ and the preference component $\tilde{\delta}$ now reverts to zero as in (4.9b). The optimization applies the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, a quasi-Newtonian routine which is frequently used for large-scale optimization problems.[8] The most delicate issue in estimating the model, however, is the choice of the precise Kalman filter routine. For this reason, the following subsection is dedicated to this problem.

6.2.4 Kalman filter algorithms

An important prerequisite of the standard Kalman filter algorithm is that the transition equation is stationary. This is not the case whenever matrix T has unit roots as is the case for the SCY model. A unit root induces a stochastic trend in one or more state variables. This leads to problems in the specification of the initial conditions for the covariance matrix of prediction errors. Generally speaking, this covariance matrix, labelled P, measures by how much, on average, the generated predictions deviate from the measurement. It is estimated in parallel with the state variables and its entries converge during the iteration over time steps. Its starting values are given by the unconditional variances and covariances of the involved state variables. However, for non-stationary variables the unconditional variance is infinite which applies, in our case, to the (log-)spot price.

---

[8] The applied Matlab routine is "csminwel" written by Christopher Sims. In several trial procedures it has proven to be most robust against cliffs in the likelihood function.

There are two solutions to this problem. A pragmatic solution given by Harvey (1989) is to initialize the prediction error covariance matrix, as a block diagonal matrix in which the non-stationary variables take large variances in the order of 10^7 and zero covariance with the stationary variables. This solution has several drawbacks. One general problem for any size of the system is that the covariance matrix can lose positive definiteness during the Kalman filter iterations such that the optimization breaks down. This problem is due to numerical inaccuracies caused by computer round-off and occurred regularly when the models for this study were tested. So-called square-root Kalman filters can be applied to circumvent this problem. They ensure positive definiteness of P throughout the iterations and diminish the rounding errors. The disadvantages of these methods are higher implementation efforts and time-consuming additional computation steps in the filter (most importantly matrix orthogonalizations). Most recommended regarding this trade-off is, for the majority of applications, the Bierman-Thornton UD Filtering.[9]

Another solution is to use one of the algorithms developed by De Jong (1991), Koopman (1997) and Koopman and Durbin (2003) for a truly diffuse initialization of P. In this case, severe rounding errors during the initial filtering steps can be avoided and P will most likely remain positive definite and converge as desired. The algorithms switch from an initial *diffuse mode* into a *non-diffuse mode* after a certain number of steps. This study uses the latter alternative, concretely the Koopman and Durbin (2003) algorithm, since it is computationally more efficient and directly addresses the cause of filter instability.

The computational steps are derived from the system of equations given by 6.27 and 6.28. The one-step forecasts are also called *prior* or *a priori* values whereas the corrected values are called *posterior* or *a posteriori*

[9] Cf. e.g. Grewal and Andrews (2001).

values. Details about the general setup of the filtering steps can be found in Harvey (1989), ch. 3.2.1 and the applied filter is described by Koopman and Durbin (2003).

6.2.5 Simulations

We previously stated that one important model requirement is robustness of the parameter values over time. However, robustness is not uniquely determined by the analytical model, but also by the estimation procedure. As the Kalman filter has high implementation requirements and the maximum likelihood principle necessitates a large sample size to yield unbiased estimates, we test the model in a simulation exercise with time series of 500 observations using a reasonable parameter configuration[10] and futures maturities matching those available in our gas market dataset. We benchmark our algorithm against an identical model implementation in Dynare, a preprocessor and a set of Matlab routines to solve Dynamic Stochastic General Equilibrium models (DSGEs). Dynare has the appealing advantage that it is able to compute quasi-analytic derivatives of the model equations with a set of first and second order Taylor approximations. Yet, maximum likelihood optimization in Dynare frequently terminates at points where the negative of the Hessian (the *information matrix*) is not positive definite. This means that even though the point estimates might be very close to the true optimal values, no reliable standard errors of the parameter estimates are obtained. We use this model as a benchmark in the simulation exercise. The simulation procedure is repeated 250 times and the mean and standard deviation of the resulting parameter estimates from each estimation run are computed. They are shown in Table 6.4.

[10] The parameter values are those estimated in Gibson and Schwartz (1990) for their oil futures data set.

		Numerical derivatives			Analytical derivatives		
	actual value	mean est.	rel. dev.	std. dev.	mean est.	rel. dev.	std. dev.
κ	16.00	14.84	-7.2%	2.89	17.76	11.0%	5.20
λ_S	.15	.16	4.7%	.12	.85	465.8%	2.25
λ_δ	-1.80	-1.71	5.1%	.96	-1.87	-3.8%	3.76
ρ	.32	.08	-76.5%	2.30	.29	-6.8%	1.43
σ_1	.35	.38	8.8%	.02	.82	133.5%	1.01
σ_2	1.12	.92	-17.7%	.69	1.07	-4.9%	1.00
υ^*	-.19	-.20	-4.9%	.06	-.16	15.6%	.17
ξ_1	.020	.024	18.5%	1.9e-3	.020	0.0%	1.0e-3
ξ_2	.025	.027	7.7%	1.4e-3	.025	0.0%	.9e-3
ξ_3	.030	.030	-1.7%	1.2e-3	.030	0.0%	1.1e-3
ξ_4	.030	.030	-1.7%	1.1e-3	.030	0.0%	1.2e-3
ξ_5	.030	.029	-1.9%	1.2e-3	.030	0.0%	1.1e-3

Table 6.4: Simulation results of the estimation routine for the reduced-form model.

The second and third column of each result block show the deviation of the mean point estimate relative to the actual parameter value and the standard deviation of the point estimate over all estimation runs, respectively. It is visible that both routines have identification problems with different parameters. Similar to the estimates obtained by Schwartz (1997) and Schwartz and Smith (2000), the risk-premia for the spot price and the convenience yield are estimated with considerable uncertainty when our own routine (numerical derivatives) is used. However, both estimates have a negligible mean bias. In turn, the coefficient of correlation ρ is strongly biased and has the highest standard deviation relative to its absolute value. However, overall accuracy of this routine seems to be much higher than that of the Dynare routine. Not only are the relative deviations smaller in sum, but also the standard deviations, except for the coefficient of correlation of the structural shocks. Compared to another simulation study for two-factor reduced-form models of interest rates by Bolder (2001), our estimation seems to have roughly the same level of accuracy. Hence, despite the bias visible from the table, it still seems reasonable to use this estimation routine in the following.

6.2.6 Reduced-form model estimation and results

The presentation and analysis of the estimation results for the two markets is done sequentially, i.e. we first show the in-sample and out-of-sample estimation results of both the standard and the extended model for the UK market, before we proceed to the analysis for the US market. The UK dataset is divided into an estimation and a forecast period as described in subsection 6.2.1. Given that for the US the time series is shorter, but a larger number of futures maturities is available, we use a different (quasi-)out-of-sample test in this case. That is, we use only a subset of the futures contracts to estimate the model and filter the state variables. The remaining contracts are then measured against the maturity-matching output prices of the estimated model. For this application, a (static) one-period convenience yield forecast is sufficient. Thereafter, we will test for parameter stability with the US dataset by changing the futures contracts used for estimation.

The parameter values generated by the Kalman filter routine for the original sample (until 12/07) are shown in Table 6.5. We realize that the risk premia are again estimated with high standard errors. A negative premium for convenience yield risk means that it pays for the futures investor

	Extended model		Standard model	
	point est.	std. err.	point est.	std. err.
κ	2.3070	(.289)	2.3279	(.297)
λ_S	-.3532	(.628)	-.7949	(.505)
λ_δ	-.7228	(1.817)	-2.3919	(1.759)
v^*	.8868	(.806)	1.2618	(.736)
ρ	.9004	(.016)	.8306	(.025)
σ_1	1.0855	(.048)	.8511	(.042)
σ_2	3.2461	(.266)	3.0926	(.259)
ξ_1	.1016	(.003)	.1444	(.005)
ξ_2	.0137	(.002)	.0075	(.002)
ξ_3	.0480	(.001)	.0594	(.002)
ξ_4	.0145	(.001)	.0125	(.001)
ξ_5	.0937	(.003)	.1100	(.003)
ξ_6	.1700	(.005)	.1980	(.006)
$logL$	4282.7		3989.4	

Table 6.5: Estimation results of the SCY-model for the UK using the Kalman filter.

to bear this risk because $\frac{\partial F}{\partial \delta}$ is negative. In turn, the spot price premium in both models is negative. The latter result is inconsistent with theory. Yet, in both model versions, the true parameter is not significantly different from zero on a 95% confidence level. A third parameter is estimated with considerable noise, as was the case in the simulation exercise. This time, it is the risk-neutral spot price trend. The estimate in the extended model is closer to our expectation. Since v^* is the difference of the riskless interest rate r and the mean convenience yield α, we can calculate the implied empirical mean convenience yield as $\alpha = r - v^*$. Assuming a mean short rate of roughly 5%[11] gives $\alpha \approx -.84$ and -1.21 in the extended and standard model, respectively. Both values are close to the estimate of $\hat{b}_1$for the short convenience yield (-1.03, cf. Table 5.9).

The parameters κ and ρ have very low standard errors, but their values deviate rather strongly from our expectation. As mentioned in section 5.6, the annualized mean-reversion coefficient calculated from the convenience yield regression is 27.3. The expectation for ρ was roughly 20%, given by the residuals from the seemingly unrelated regression model applied in Gibson and Schwartz (1990) (eqs. 10 and 11). As one can see, the estimate for κ is now much lower, 2.3, and the one for ρ much higher, .83 to .90. Interestingly, the same pronounced deviation occurs by comparing the estimates for κ and ρ in Gibson and Schwartz (1990) to those from the Kalman filter in Schwartz (1997), both generated from the same oil price dataset with an overlapping time window and the same frequency of observations. While the estimate for the mean-reversion speed is 16.1 in the former paper, the highest outcome in Schwartz (1997) is 1.9. The latter value is close to our results. The same holds for ρ which is found to be .32 in the former and at least .77 in the latter paper. While a pronounced deviation of ρ does not come at a surprise given our simulation results above, the deviation of κ cannot be explained only by numerical problems.

The standard deviation of the spot price shock σ_1 is very close to our a priori estimates of 1.11 and .90 for the extended and the standard model, respectively. σ_2, in turn, is much lower than expected. This is natural due to its immediate relationship to κ which also turned out lower than expected. Figure 6.5 shows an exemplary plot of the filtered spot price against the realized day-ahead price and the filtered values of the convenience yield against their a posteriori values. The values are taken from the standard

[11] The mean one-month GBP LIBOR during the period is 5.24%.

model. As can be seen, the filtered spot price matches the day-ahead price acceptably although the latter price was not contained in the calibration set. In addition, the (one-period) convenience yield projections match the posterior values quite accurately. As could already be inferred from the estimate for ρ, convenience yield moves quite closely together with the spot price. Figure 6.6 shows the fit of model futures prices to the synthetic futures. All of the prices appear to be closely matched while the best match is realized for the middle maturities.

The overall impression of the standard deviations of measurement errors is that the in-sample noise is slightly lower in the extended model. This is confirmed by looking at our goodness-of-fit evaluation. As shown in Table 6.6, we use the mean error (ME) the mean absolute error (MAE) and the root mean squared error (RMSE) to compare the two models. The ME shows the potential bias in the futures price estimates whereas the other two quantities show the variation contained in the forecast error. These measures are computed for the in-sample (one-period) forecast and an out-of-sample one-period and five-period forecast, respectively. The latter two projections use the given parameter estimates to fit the second period of the time series from 01/08 to 04/09 (339 additional observations). An measure of the in-sample fit is also provided by the log-likelihood value at the optimum in Table 6.5. Since the reduced-form model itself is identical for both estimations and the same number of observations is used, a direct comparison of the likelihood densities is possible.

Regarding the in-sample fit, we can see that there is no difference in the mean errors. However, taking into account the measures of variation we see that the extended model has a slight advantage over the standard model overall. It is less important if we move on to the one-period forecast. The likely reason for this is that the Kalman filter has two degrees of freedom at every observation date to fit the state variables to the one-period forecast. Therefore, the one-period forecast does not deteriorate as long as market

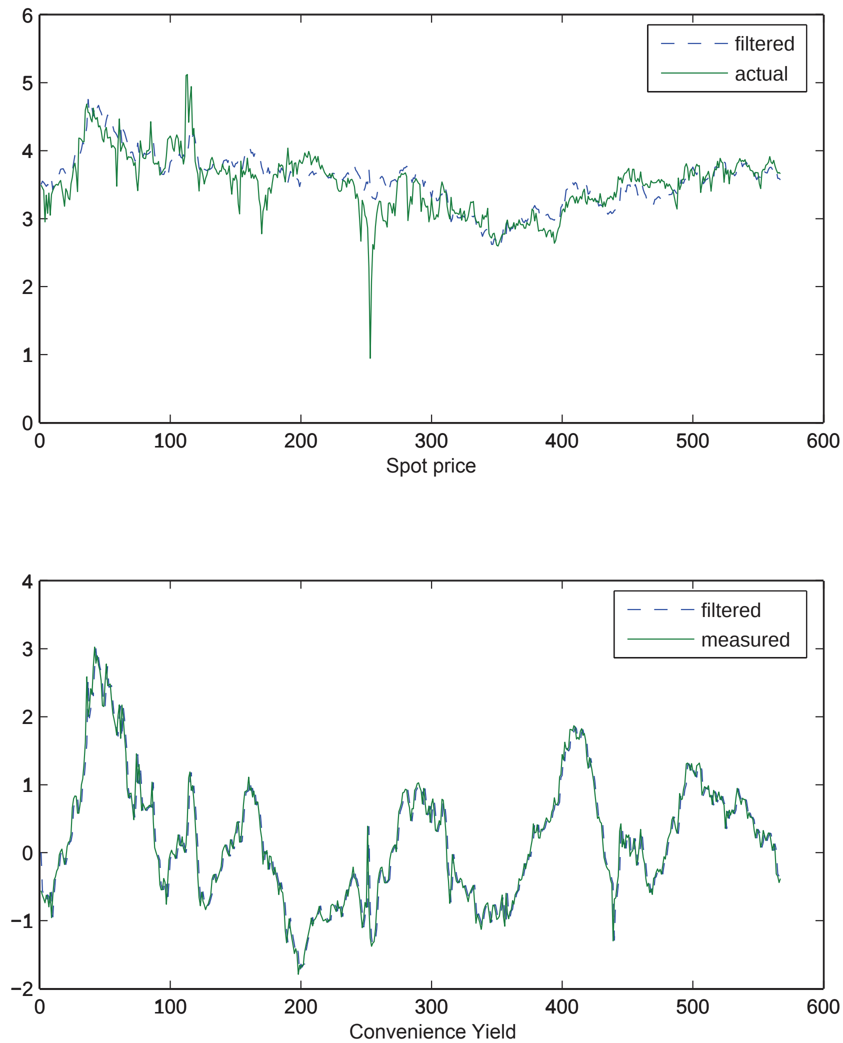


Figure 6.5: Filtered versus actual/measured values of state variables in UK (here: extended model).

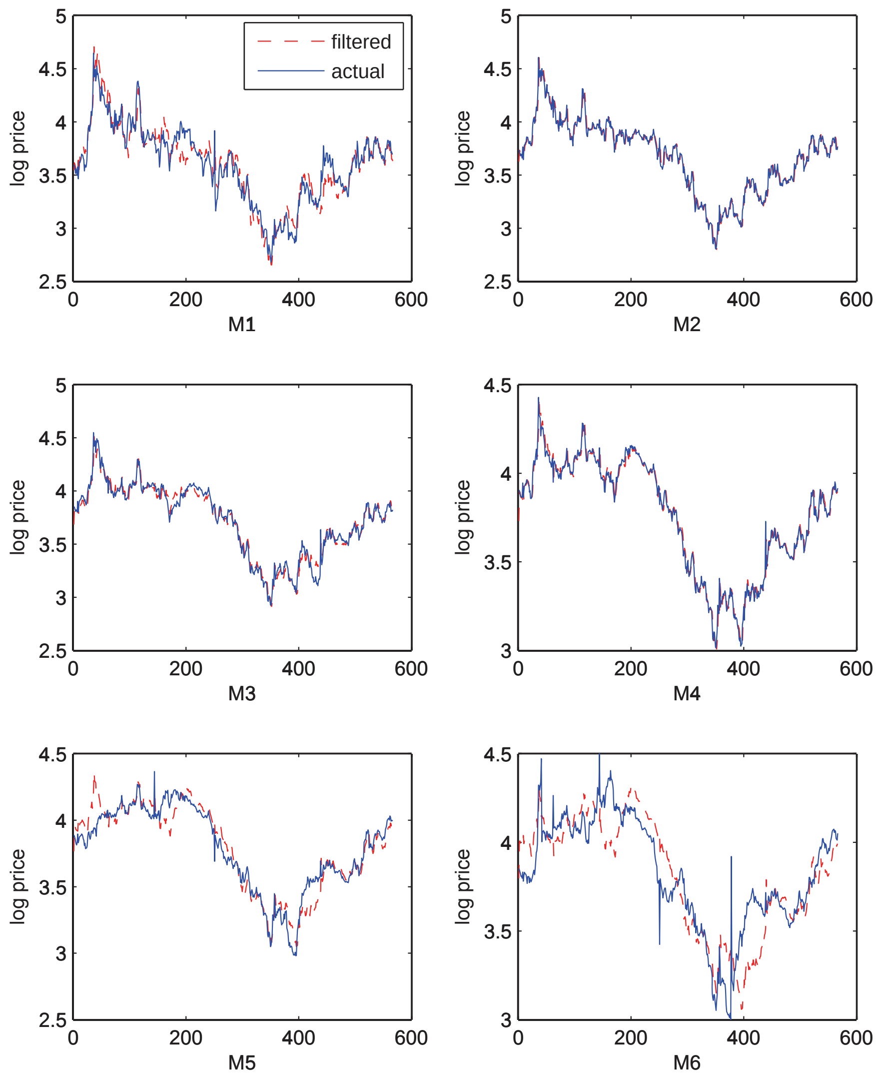


Figure 6.6: Filtered versus actual values of futures (obs. variables) in UK (here: extended model).

	Extended model			Standard model		
	ME	MAE	RMSE	ME	MAE	RMSE
In-sample						
M1	-0.0016	0.0914	0.1193	-0.0029	0.1194	0.1519
M2	-0.0008	0.0382	0.0529	-0.0013	0.0264	0.0404
M3	0.0021	0.0476	0.0634	0.0019	0.0503	0.0682
M4	-0.0016	0.0283	0.0415	-0.0019	0.0233	0.0368
M5	-0.0048	0.0764	0.1019	-0.0051	0.0872	0.1164
M6	0.0101	0.1335	0.1697	0.0096	0.1560	0.1960
OCM	-0.0619	0.1946	0.2827	-0.0638	0.2810	0.3874
Out-of-sample (1-period ahead)						
M1	-0.0141	0.0659	0.0812	-0.0157	0.0699	0.0865
M2	-0.0014	0.0238	0.0346	-0.0021	0.0242	0.0349
M3	0.0007	0.0332	0.0413	0.0003	0.0333	0.0414
M4	-0.0028	0.0223	0.0297	-0.0032	0.0217	0.0294
M5	-0.0176	0.0500	0.0631	-0.0182	0.0506	0.0637
M6	-0.0404	0.0866	0.1113	-0.0412	0.0884	0.1132
OCM	-0.0355	0.1383	0.1715	-0.0377	0.1430	0.1773
Out-of-sample (5-periods ahead)						
M1	0.0837	0.4137	0.4774	0.1243	0.4294	0.4950
M2	-0.0237	0.3053	0.3520	-0.0481	0.3278	0.3737
M3	-0.0475	0.2740	0.3087	-0.0746	0.2890	0.3242
M4	-0.0729	0.2352	0.2779	-0.1030	0.2427	0.2925
M5	-0.1072	0.2028	0.2745	-0.1407	0.2066	0.2944
M6	-0.1492	0.1853	0.2759	-0.1871	0.1966	0.3040
OCM	-0.1013	0.2064	0.2791	-0.1352	0.2130	0.3010

Table 6.6: Goodness-of-fit statistics for the UK market.

volatility remains the same and the parameters are stable over time. One can notice that the mean bias in the second set slightly increases for the M1 and M6 contract, otherwise the performance of both models remains acceptable.

The difference between the two specifications becomes most important in the five-period forecast. Quite obviously, the forecast deteriorates since the state variables are fit to the one-period forecast, and projections by four additional periods increase the forecast error quite significantly. Yet, the projections of the extended model are, on average, 32% closer to the true value than those of the standard model. The situation is similar for the measures of variation. Finally, we look at the fit to the true spot price, which was not included in the dataset for estimation. All statistics here seem to speak in favor of the extended model. Interestingly, the measures of variation are markedly smaller than for the benchmark, particularly in the in-sample comparison. In the forecast period, the lower mean bias becomes the more favorable statistic instead. In any case, the table documents that, regarding the UK market, the extended model yields more accurate point estimates both in-sample and out-of-the sample.

The estimation results for the US market are shown in Table 6.7. This time κ is estimated remarkably lower than for the UK. It can be assumed that the change in the frequency of observations (daily versus weekly) is responsible for this effect. That is, autocorrelation in the convenience yield increases with a lower frequency of observations since the innovations occurring between two observation dates even out. In addition, κ is pronouncedly smaller in the extended model, which goes along with a lower value of σ_2. A possible explanation is that a greater part of the convenience yield variation can be explained through past observations and, hence, supports the hypothesis that some noise in the process could be removed up-front by the fundamental model's projections. Decomposing v^* as above yields again a negative mean convenience yield α. Yet, it is

	Extended model		Standard model	
	point est.	std. err.	point est.	std. err.
κ	.7100	(.007)	1.3709	(.004)
λ_S	.0115	(.003)	-.0591	(.034)
λ_δ	-.2124	(.008)	-.3612	(.023)
v^*	.2420	(.006)	.3067	(.010)
ρ	.9492	(.008)	.9498	(.007)
σ_1	.5508	(.018)	.5447	(.012)
σ_2	.3266	(.014)	.6598	(.017)
ξ_1 (M1)	.0536	(.002)	.0791	(.003)
ξ_2 (M3)	.0193	(.002)	.0265	(.003)
ξ_3 (M6)	.0262	(.001)	.0499	(.002)
ξ_4 (M10)	.0267	(.001)	.0408	(.002)
ξ_5 (M14)	.0215	(.001)	.0312	(.002)
ξ_6 (M18)	.0326	(.002)	.0408	(.002)
$logL$	5612.7		4643.6	

Table 6.7: Estimation results of the SCY-model for the US using the Kalman filter.

larger (i.e. closer to zero) than in the UK market. The premium for convenience yield risk is again negative and, hence, consistent with theory. It is also lower in absolute terms than in the UK market. This makes sense because we have learned from the descriptive statistics in Table 6.2 that convenience yield variability is lower in the US so investors should require a smaller risk premium than UK investors.

The lower variability is also reflected in a slightly lower estimate of σ_2, although it is still roughly 90% and 80% of the corresponding values for UK, respectively, which seems slightly disproportionate. Again, the change in the observation interval could be the reason for this. The risk premium for the spot price is now slightly positive in the extended model and, in turn, slightly negative in the standard model. However, the latter value is not significantly different from zero on a 95% confidence level such that the former estimate is more reliable. ρ and σ_1, respectively, are very similar in both extended and standard model, and the fact that ρ again turns out much higher than the initial guess does not come at a surprise. Finally, the lower standard errors of the measurement innovations as well as the higher function value of the log-likelihood suggest that the in-sample fit again favors the extended model. We will investigate this further by discussing the following goodness-of-fit measures displayed in Table 6.8.

Looking at the in-sample statistics, it appears that the mean error is, on average, slightly larger in the extended version. In turn, we can observe a marginal advantage in the measures of variation. This is somewhat surprising given the clearly smaller measurement and state noise variation. The matching accuracy with respect to the spot price (proxied by the day-ahead price) shows a similar pattern. That is, the mean error is larger, but both MAE and RMSE are smaller in the extended model.

In the (quasi-)out-of-sample test, parameters and state variables from the estimation are used to generate model prices for the remaining maturities of futures and to record the corresponding prediction errors. This time, there is barely any difference in the out-of-sample performance of the two models. Since the futures prices are correlated along the term structure, it is not surprising that the results do not deviate materially from the in-sample results. All in all, we can constitute that the fundamental convenience yield projections seem to permit for a mentionable improvement when the prices of the estimation set of contracts are forecasted. In turn, no improvement occurs when we price the cross-section, i.e. when the goal is to infer the entire term structure of futures prices from a small set of contracts.

Having compared the two model versions, what general conclusions can we draw from the empirical analysis so far? First of all, the (net) mean convenience yield is dominated by the enormous physical cost of storage in the gas market, since it is negative in all cases. Therefore, it is likely that convenience yield variation is influenced by price variations in the market for short-term storage (for which data is not publicly available so far). Besides this, assuming that the above estimates are close to the true values, autocorrelation in the preference component of the convenience yield is extremely high in both markets. Day-on-day correlation in the UK market is in the order of 99% and the weekly figure for the US is roughly 98%. This would mean that speculation preferences change very slowly. In addition, this calls the necessity of the second stochastic factor into question. Yet, the reliability of the estimates is debatable given the strong deviation from the first-guess values for ρ that have been mentioned above.

To further investigate this reliability we test the robustness of both models to a change in the dataset of futures contracts used for the estimation. Since we need a certain number of contracts as a precondition to obtain stable state variable estimates and parameters, we can only use the US market dataset for this purpose. In a first estimation run we only use contracts at the *short end of available maturities* (i.e. contracts M1 to M6) for calibration before, in a second run, we use only the *long-dated maturities.*(M13 to M18). The results are shown in Table 6.9.

Naturally, the log-likelihood density at the optimum is higher for both models when only the short or long maturities are used. The reason is that neighboring maturities have a higher correlation and, hence, are easier to fit jointly. Looking at the parameters, the first problem that we find is the instability of mean-reversion speed κ when altering the contracts for esti-

	Extended model			Standard model		
	ME	MAE	RMSE	ME	MAE	RMSE
In-sample						
M1	-.0205	.0680	.0911	-.0155	.0751	.1019
M3	.0017	.0553	.0720	.0012	.0534	.0693
M6	.0035	.0498	.0636	.0017	.0538	.0683
M10	-.0013	.0457	.0579	-.0011	.0466	.0589
M14	-.0078	.0391	.0502	-.0074	.0362	.0465
M18	.0061	.0384	.0505	.0025	.0414	.0547
Out-of-sample						
D1	-.0430	.0747	.0993	-.0361	.0938	.1281
M2	-.0052	.0590	.0762	-.0037	.0580	.0758
M4	.0027	.0542	.0700	.0010	.0556	.0710
M5	.0033	.0529	.0680	.0013	.0572	.0723
M7	.0026	.0482	.0617	.0012	.0493	.0631
M8	.0014	.0474	.0609	.0005	.0460	.0603
M9	.0009	.0466	.0598	.0006	.0458	.0594
M11	-.0048	.0441	.0558	-.0042	.0469	.0583
M12	-.0073	.0412	.0519	-.0066	.0432	.0534
M13	-.0086	.0390	.0502	-.0080	.0376	.0477
M15	-.0055	.0398	.0514	-.0056	.0377	.0492
M16	-.0022	.0393	.0508	-.0032	.0397	.0512
M17	.0012	.0397	.0510	-.0010	.0409	.0522

Table 6.8: Goodness-of-fit statistics for the US market.

	Extended model				Standard model			
	short maturities		long maturities		short maturities		long maturities	
κ	.0048	(.006)	1.3929	(.119)	1.9966	(.192)	1.1183	(.290)
λ_S	.1398	(.137)	.1573	(.107)	-.1736	(.204)	.2753	(.198)
λ_δ	-.0316	(.005)	.0211	(.069)	-.8754	(.516)	.1973	(.300)
v^*	.2462	(.005)	.0050	(.049)	.4836	(.249)	.7581	(.162)
ρ	.5908	(.044)	.5619	(.047)	.8483	(.017)	.9871	(.006)
σ_1	.5406	(.018)	.3551	(.011)	.6572	(.025)	1.3675	(.280)
σ_2	.1167	(.007)	.2388	(.021)	1.6300	(.111)	1.9160	(.685)
ξ_1	.0527	(.002)	.0237	(.001)	.0594	(.002)	.0344	(.002)
ξ_2	.0287	(.001)	.0025	(.002)	.0136	(.002)	.0108	(.001)
ξ_3	.0248	(.001)	.0135	(.001)	.0287	(.001)	.0223	(.001)
ξ_4	.0046	(.001)	.0151	(.001)	.0259	(.001)	.0208	(.001)
ξ_5	.0024	(.001)	.0249	(.001)	.0122	(.001)	.0223	(.001)
ξ_6	.0187	(.001)	.0383	(.001)	.0386	(.001)	.0467	(.002)
$logL$	6695.5		6816.5		5474.0		6172.7	

Table 6.9: Estimation results with changing sets of maturities.

mation. In the original specification, κ is highest in the long-end dataset while it declines for the mixed set used above (*spanning dataset*) and, finally, for the collection of short maturities. The first value is almost twice as much as the last one. In the extended specification, the order is reversed with the additional problem that the parameter cannot be identified when only the short maturities are used for estimation. Regarding the empirical evidence from convenience yield regressions, we know that the latter pattern, i.e. higher mean reversion with rising time to maturity, is not supported by the data. This can be a consequence of the a priori modification of the dataset with convenience yield projections. In any case though, it seems that the stochastic process of the instantaneous convenience yield cannot be reliably estimated from any of the sets of contracts.

Corresponding to the instability in κ over datasets, we can also observe severe changes in σ_2. The respective estimates in the extended model are much lower than in the standard model. Yet, the estimates of σ_2 in the standard model again vary strongly over datasets. More precisely, the spanning set yields an estimate of .66, whereas the short-term and long-term set yield estimates of 1.63 and 1.92 respectively, i.e. between 2 and 3 times as much. We point out that although the relative difference between the estimates for σ_2 in the extended model are slightly smaller, the general problem is the same.

Unfortunately, these changes are not directly proportional to those in κ, in which case the situation would reduce to a single identification problem. Reviewing the results from the simulation exercise, however, it is likely that the variation with respect to σ_2 has also numerical reasons and does not need to be due to a general misspecification. In this case, an improved version of the optimization and filtering algorithm would eliminate the instability. The risk premia are rather stable in the extended model (at least in absolute terms), while they vary strongly in the standard model. v^* changes in both models, however the two significant estimates in the extended model are nearly identical. They correspond to a mean convenience yield of -.21 approximately (given the decomposition $v^* = r - \alpha$ from above and a mean short rate of 3% in the sample period). The standard model, in turn yields estimates between -.28 (spanning set) and -.73 (long-end set). The median value of the sample is -.35 for the short-end convenience yield, and it decreases in absolute terms for longer maturities (cf. Table 6.3).

Therefore, the estimates generated by the standard model are not consistent. In turn, it is comprehensible that the extended model is not able to identify the instantaneous convenience yield only from long-term contracts and estimates in the short-end and the spanning dataset are closer to our expectation. Notwithstanding, the empirical statistics hint at the problem that a single (mean) convenience yield applying to all maturities of futures contracts seems to be a problematic assumption. Further discussion on this issue will follow below. Looking at the volatility of spot price innovations, σ_1, both models produce different results for short and long maturities, respectively. This effect is somewhat smaller in the extended model where they range between .36 and .55 compared to .54 and 1.36 in the standard model. Yet, it is still not immediately clear why such a difference in the estimate for σ_1 occurs across datasets. The stability problem with respect to the coefficient of correlation ρ is obviously due to numerical problems as shown by the simulation study. Thus, the differing values of this parameter across datasets are not surprising.

All in all, one can see that the choice of data used to estimate the model has a crucial impact on the outcome. For this reason, it is always desirable to use a dataset covering both ends of the term structure as long as market liquidity in the contracts is sufficient. The proposed extension of the SCY model with a fundamental convenience yield forecast is favored over the benchmark. An advantage seems to materialize when forecasts of more than one time period are made. A necessary condition is that forecasts of temperature and storage levels are sufficiently reliable for the forecast horizon considered. However, the economic importance of this advantage becomes rather small in light of the problems regarding parameter stability discovered above.

Our strongest conceptual concern is the apparent instability in the mean-reversion speed. A straightforward reason for this phenomenon would be that a single convenience yield is an oversimplified picture of reality. Analogously to interest rates, convenience yields with different terms might not be equal. This is supported by the regression results across convenience yield maturities found in Tables 5.9 and 6.3, including the estimates for the mean-reversion parameter. Another support for this conjecture comes from an analysis of the measurement errors of the Kalman filter. An exemplary plot of these residuals is shown in Figure 6.7 using the standard model applied to the US dataset as an example. It is apparent that the errors for the longer term futures (M6 to M18) are not free of autocorrelation. While it was postulated by Schwartz (1997) (p. 931) that "*[...] the serial correlation and cross correlation in the log prices is attributed to the variation of the unobservable state variables*", this does not seem to work particularly well in reality. Again, a plausible reason for this would be that mean reversion speeds (and thus autocorrelation) in the convenience yield differ across maturities. In this spirit, it might be worthwhile to test whether a deterministic function of the time to maturity could be imposed on convenience yields, thereby allowing for a term structure with different mean reversion speeds. This is an interesting research question emerging from our analysis. Further insight into the main drivers of the (in)accuracy of forecast could be obtained with a sensitivity analysis, which investigates the impact of one parameter's change on the goodness-of-fit. This would also help prioritizing issues on the model design front.

In turn, some of the parameter variation across datasets is likely to have a numerical origin. While many studies including Schwartz (1997) have not

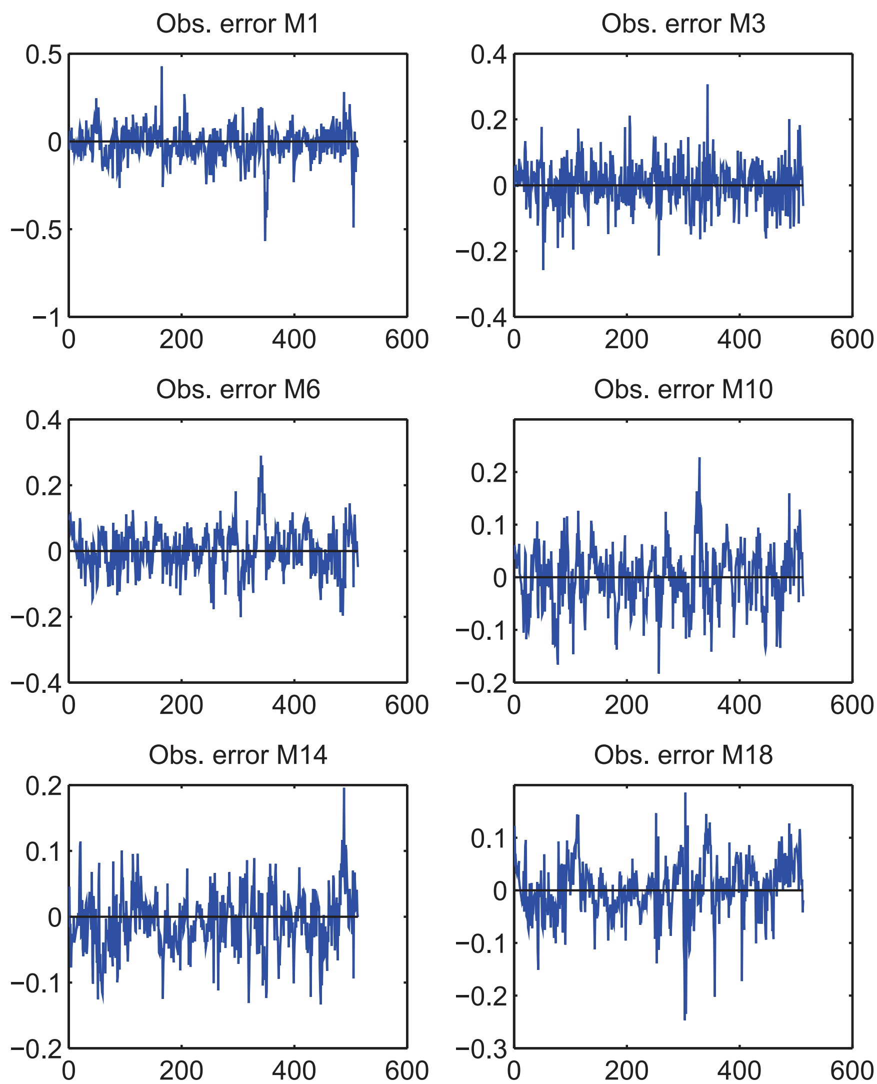


Figure 6.7: Measurement errors of the standard SCY model (US, spanning dataset).

made explicit the unbiasedness of the applied optimization algorithm, we have demonstrated in our simulation study that ρ is estimated with a severe lack of precision. The accuracy of this estimate is not unimportant though, since it transmits convenience yield shocks to the spot price. Given that fine-tuning the maximum likelihood routine can become complicated and the overall improvement potential is uncertain, we try instead two different approaches. One of them still relies on the Kalman filter, but takes explicit account of parameter uncertainty. That is, we change from a frequentist to a Bayesian estimation. The other approach used is the "implied method" proposed and applied in Cortazar and Schwartz (2003) and Lucia and Schwartz (2002). It is designed much simpler than the Kalman filter routine, with the downside of not providing confidence bounds for the parameter estimates. Whether one of these methods can improve parameter stability will be investigated in the following. The numerical results will also serve as a cross check to validate or reject individual estimates obtained with the ML-approach above.

6.2.7 Alternative estimation methods

In this subsection, we deepen the investigation of the robustness of the extended model from a numerical point of view by exploring the performance of alternative algorithms. First, we apply a Bayesian estimation and, second, the "implied method". As stated above, one appealing advantage of bayesian estimation is that model parameters are assumed to be uncertain. In other words, one explicitly takes account of model uncertainty. Using such an estimation method could thus unveil an instability of the parameter estimates found above. Another advantage is that non-normal distributions of the parameters are allowed, which could help increase the robustness of the parameters' point estimates.[12] The ultimate goal of bayesian estimation is to find the joint distribution of the parameters given the observed data, the *posterior distribution*, and its moments. If we denote the true parameter vector and the random realization by Ω and ω and the population and sample of the data by Y and y, respectively, then this density function can be formally stated as $f_{\Omega|Y}(\omega|y)$. Following from the application of Bayes' rule, the posterior density is computed as

$$f_{\Omega|Y}(\omega|y) = \frac{L(\omega;y)f_{\Omega}(\omega)}{f_Y(y)} \tag{6.29}$$

in which $L(\omega;y)$ is a likelihood function of the parameters given the data and $f_{\Omega}(\omega)$ is the so-called *prior distribution* of the parameters that is specified by the statistician. $f_Y(y)$ is the marginal (unconditional) density of the data. Since $f_Y(y)$ is independent of the parameters, $c = f_Y(y)^{-1}$ is a constant (called *normalization constant*) and does not influence the density order, but only its level. For this reason, dropping c and taking logs in the equation above shows that the log-posterior kernel (*un-normalized posterior density*) is still proportional to the sum of the log-likelihood and the log-prior density:

$$\log \mathcal{K}_{\Omega|Y}(\omega|y) \propto \log L(\omega;y) + \log f_{\Omega}(\omega). \tag{6.30}$$

The equation also shows that $\log \mathcal{K}$ results from a combination of the predetermined parameter density given the data and the subjective prior.[13] Usually, the prior densities are borrowed from well known analytical distributions whose basic properties match the expectation of the statistician

[12] Note that population parameters in the frequentist approach (ML) do not have a distribution. The standard errors computed only say *how frequently* the true parameter is contained in a certain interval around the point estimate.

[13] With simple analytical examples, one can show, in fact, that this combination is convex with weights depending on the particular distributions used.

regarding the parameters. The sensible issue of bayesian estimation is that the specification of these prior distributions should not take into account information from the current dataset. The idea is that rather expertise, economic intuition and past experience should play a role. In addition, the prior distribution should not dominate the posterior distribution as long as the parameter value is not known for certain.

Finding the normalized posterior distribution starts by finding the distribution of $\log \mathcal{K}$. This is often not possible analytically. Since the distribution of the likelihood in our case depends on the Kalman filter outcomes in a complicated way, no analytical distribution of the likelihood and, hence, of the posterior kernel is available. In this situation, one has to revert to simulations of the latter one. The Metropolis-Hastings algorithm is an efficient and widely used sampling method for this purpose, which is implemented in Dynare. The outcomes are the single parameters' conditional distributions.[14] The algorithm reverts to the central limit theorem and assumes that the conditional distribution of the parameters is asymptotically normal near the posterior mode. It uses an elaborated acceptance rule to generate a sequence of random sample values, which constitute a Markov chain, from a proposal distribution . For a sufficiently large number of draws, the density of these realizations converges to the continuous posterior distribution of the parameters. The technical details of the implemented algorithm (here: Random Walk Metropolis-Hastings) are found in Griffoli (2009).

To find the mode of the posterior kernel, the starting point of the sampling algorithm, the density is maximized, which is essentially the same as maximizing the likelihood, but with some positive weight imposed on

[14] They completely characterize, according to the *Clifford-Hammersley-Theorem*, the joint posterior distribution (cf. Johannes and Polson (2003), p. 12ff.).

the a priori distribution.[15] We specify an inverse gamma prior distribution for all parameters which are not allowed to become negative (i.e. $\kappa, \sigma_1, \sigma_2$ and ξ_1 to ξ_6), a normal prior for λ_S ,λ_δ and v^* and a uniform prior for ρ. In the following, two Markov chains of 20,000 draws each are generated, of which the second 10,000 are kept. The resulting posterior distributions are shown in Figure 6.8 and 6.9 (black lines), together with the specified prior distributions (grey lines) and the estimated posterior mode (dashed horizontal line).

It can be seen that the posterior densities of the first four parameters, i.e. $\kappa, \lambda_S, \lambda_\delta$ and v^*, are symmetric and not dominated by their respective prior. In addition, the mode computed by maximization of the posterior kernel approximately coincides with the mode of the sampled distribution. However, the plot for ρ shows that the identification problem discovered in the simulation exercise with the ML-approach reappears in the simulated distribution. In fact, the posterior density extends over the whole range of possible values from -1 to 1. This confirms that the estimate for ρ will always be very sensitive to noise in the data or to slight biases in the other parameter's estimates. It also shows how misleading the standard error estimates from the information matrix can be when the Hessian is computed in a very small region around the mode. The empirical confidence interval can be much larger than indicated by a single estimation run. In turn, the sampled distribution for σ_1 shows another problem: since the mode estimate exactly coincides with the hump of the prior distribution *and* the simulated mode, it is very likely that the prior has dominated the estimation. In addition, a second hump exists near the previously found point estimate of .55 and it cannot be told from the plot whether the density of the likelihood kernel alone would have been higher at this point than at the estimated mode. The posterior for σ_2 lies far apart from the

---

[15] Recall that by the asymptotic normality assumption the "posterior density" of the ML-estimator is symmetric, i.e. mode and mean coincide.

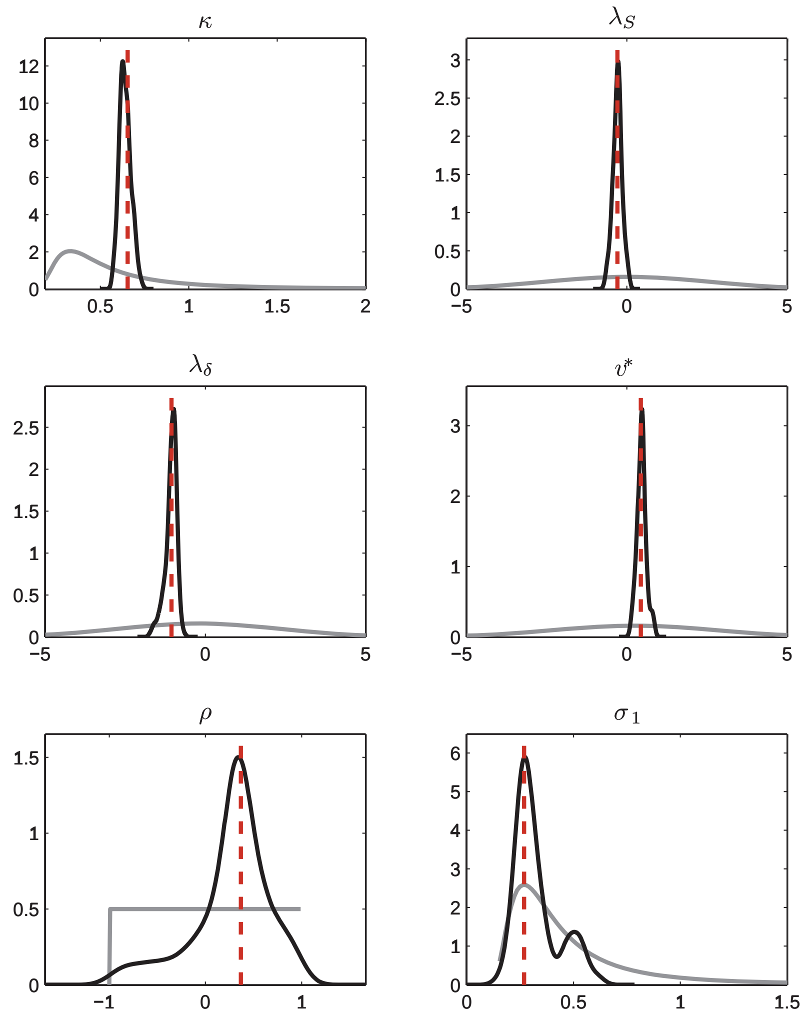


Figure 6.8: Prior and posterior distribution of parameters with US data (spanning data, part 1).

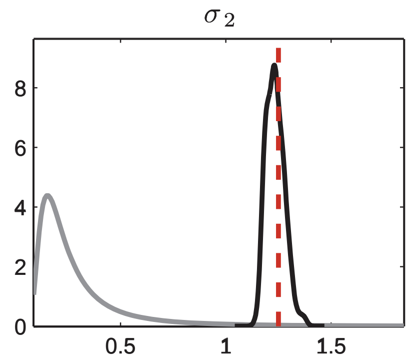


Figure 6.9: Prior and posterior distribution of parameters with US data (spanning data, part 2).

expected range, but does not appear to be biased by the prior. Briefly, the diagnostic checks show that except for σ_1 none of the parameter estimates seems to be biased by our priors.

Table 6.10 provides information on the prior specification and the corresponding posterior estimates. The displayed log-likelihood values correspond to the density at the posterior mode and the modified harmonic mean estimator respectively. The prior mean corresponds to the point estimate obtained with the ML estimation. The prior standard deviation was chosen with respect to the variation of the estimates across datasets (comparing Tables 6.7 and 6.9). An exception is ρ, for which only the bounds of the uniform distribution (-1 and 1) were fixed. The posterior mean is computed from the sampled densities and is displayed together with the 95% confidence interval. Comparing posterior modes and means shows only marginal differences due to the symmetry of most of the distributions and to the fact that the analytic mode was calculated with sufficient precision. Yet, the comparison with the ML results (or the prior means) requires some

	prior mean	post. mode (std. error)		posterior mean (conf. interval)		prior distr.	prior std. dev.
κ	.71	.6548	(.036)	.6399	(.592, .697)	invg	5.00
λ_S	.01	-.2934	(.159)	-.2949	(-.545, -.048)	norm	2.50
λ_δ	-.21	-1.0604	(2.502)	-1.0728	(-1.371, -.809)	norm	2.50
v^*	.24	.4350	(.140)	.4583	(.184, .676)	norm	2.50
ρ	.00	.3688	(7.442)	.2626	(-.318, .998)	unif	.58
σ_1	.55	.2692	(.108)	.3243	(.219, .537)	invg	1.25
σ_2	.33	1.2506	(.049)	1.2350	(1.169, 1.300)	invg	2.50
ξ_1	.05	.0483	(.002)	.0485	(.045, .051)	invg	.10
ξ_2	.02	.0199	(.002)	.0203	(.018, .023)	invg	.10
ξ_3	.03	.0254	(.001)	.0256	(.024, .027)	invg	.10
ξ_4	.03	.0273	(.001)	.0275	(.026, .029)	invg	.10
ξ_5	.02	.0192	(.001)	.0194	(.017, .022)	invg	.10
ξ_6	.03	.0318	(.002)	.0313	(.029, .034)	invg	.10
$logL$		5708		5658			

Table 6.10: Bayesian estimation results for the extended model (US data).

discussion. The estimates for the risk premia deviate remarkably from the former ones. λ_S is now negative and the confidence interval does not include zero. λ_δ remains negative, but the absolute value is now five times as high as that of the prior mean. The parameters σ_1 and σ_2 also have a posterior distribution which does not include the previous point estimate. The only value which remains close to the expectation is κ (and the measurement errors which are, however, only by-products of the estimation). In other words, all parameters but κ are obviously plagued by numerical problems since two estimation routines should return the same estimates when using a given dataset.

We now take, again, a look at the change in parameters when the dataset is varied as a check of numerical robustness of the bayesian approach. In the results from the short term dataset, the posterior distributions become somewhat less symmetric as shown by the diagnostic plots (Figures 6.10 and 6.11). The mass of κ seems to have shifted away from the mass of the prior. The distributions of κ and λ_δ have apparently widened in addition. The posteriors of both σ_1 and σ_2 now seem to be dominated by the respective priors and both have an additional local mode. The distribution of σ_2 has also widened significantly. Looking at the tabulated results (Table 6.11) the striking difference to the ML result is that κ is now significantly positive and even twice as high as in the spanning dataset. Its confidence interval appears, indeed, to have doubled as well, but it remains proportionate from an absolute viewpoint. The risk premia λ_S and λ_δ deviate strongly from both the ML estimate with the same dataset and the one from the spanning dataset above. Furthermore, the confidence interval from the latter parameter has doubled. v^* is now more than two times as high as in the ML routine, and ρ appears to be almost a slack variable again. Finally, σ_1 is strongly right-skewed and σ_2 has decreased by 50% and is not close to its ML counterpart. Briefly, the stability of parameters across datasets is not amended compared to the ML estimation. The only helpful fact is that skewed parameter distributions are identified and the posterior mean can be computed. Using the latter one as a point estimate is likely to increase robustness. Unfortunately though, this advantage does not seem to heal the general instability of the estimates.

Looking at the diagnostic plots for the long-term dataset (Figures 6.12 and 6.13), multimodality problems reappear. Particularly in the case of λ_S one would definitely prefer using the posterior mean estimate to using any of the two modes with almost equal density. As for λ_δ the mode was not computed very exactly and a second local mode from the simulated

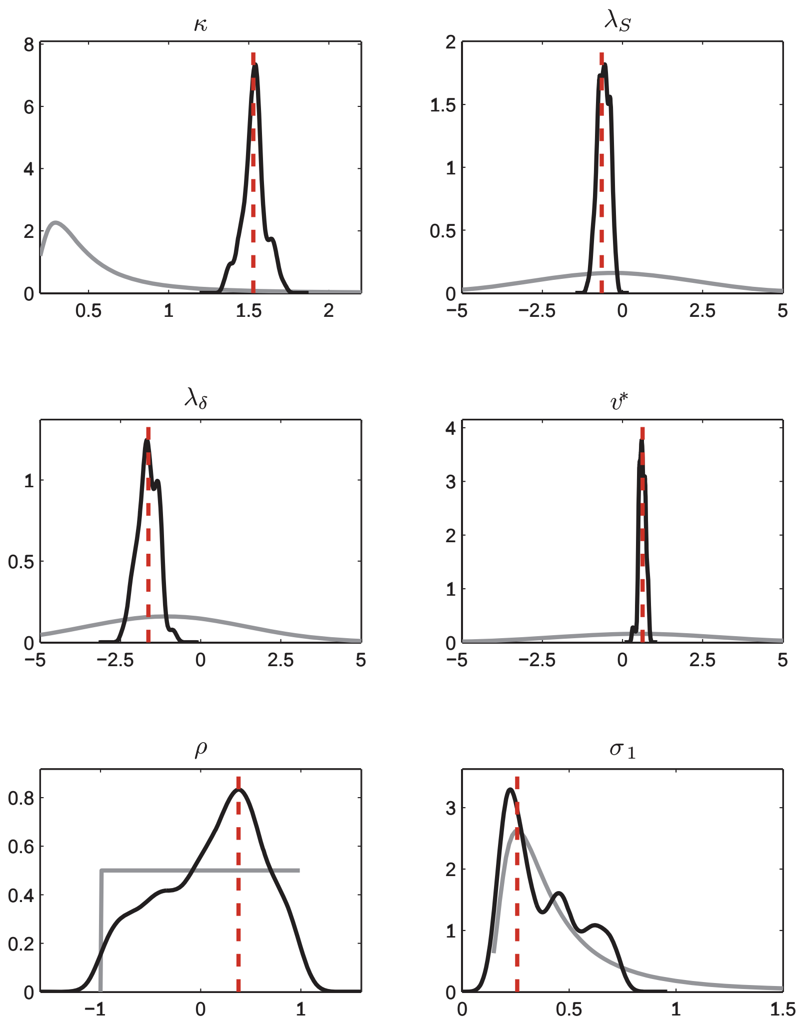


Figure 6.10: Prior and posterior distribution of parameters with US data (short-term data, part 1).

	short maturities				long maturities			
	post. mode	(std. error)	post. mean	(conf. interval)	post. mode	(std. error)	post. mean	(conf. interval)
κ	1.528	(.072)	1.533	(1.426, 1.676)	1.302	(.074)	1.300	(1.197, 1.434)
λ_S	-.647	(.147)	-.589	(-.881, -.264)	.202	(.291)	.371	(.001, .784)
λ_δ	-1.630	(2.522)	-1.649	(-2.176, -1.169)	-.112	(2.560)	.037	(-.882, .608)
v^*	.627	(.089)	.616	(.487, .800)	-.067	(.267)	-.207	(-.530, .212)
ρ	.378	(15.146)	.098	(-.748, .921)	-.333	(15.962)	-.120	(-1.000, .642)
σ_1	.258	(.105)	.373	(.144, .642)	.258	(.106)	.317	(.167, .501)
σ_2	.649	(.291)	.681	(.349, 1.019)	.624	(.256)	.742	(.389, 1.201)
ξ_1	.044	(.002)	.045	(.042, .048)	.022	(.001)	.022	(.020, .023)
ξ_2	.026	(.001)	.027	(.025, .028)	.007	(.001)	.007	(.005, .008)
ξ_3	.025	(.001)	.025	(.024, .026)	.014	(.001)	.014	(.014, .015)
ξ_4	.010	(.001)	.010	(.009, .011)	.017	(.001)	.016	(.015, .017)
ξ_5	.008	(.001)	.008	(.007, .009)	.026	(.001)	.026	(.024, .027)
ξ_6	.023	(.001)	.023	(.021, .024)	.036	(.001)	.036	(.034, .038)
$logL$	6310		6259		6678		6628	

Table 6.11: Bayesian estimation results for short- and long-maturity futures (US data).

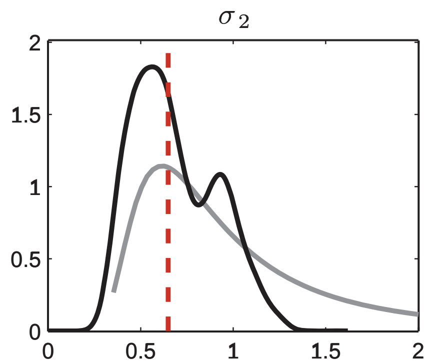


Figure 6.11: Prior and posterior distribution of parameters with US data (short-term data, part 2).

density appears and widens the distribution. The latter phenomenon might be induced by the prior, whose mode coincides with this local hump. The distribution for v^* is also asymmetric, and the mode from the maximization is slightly inaccurate. The distribution of σ_1 is now unimodal. In such a case, the prior's influence might sometimes be beneficial in that it induces just enough curvature in the posterior density to allow the maximization routine to identify an optimum at all. Yet, given the coincidence of prior and posterior mode, we do not get to know the value which the data alone would produce. Finally, also σ_2 is influenced by its prior again, and it could be that the mode of the data would otherwise be greater than 1, where the second hump of the density is located.

The results for the long-term dataset in Table 6.11 also show a mixed picture with respect to both accuracy of the point estimates as well as the similarity to the ML results. κ is sufficiently close to the estimate in Table 6.9. λ_S is now positive and pronouncedly larger than in the ML estimation. λ_δ, in turn, is close to the earlier estimate in Table 6.9, but the confidence

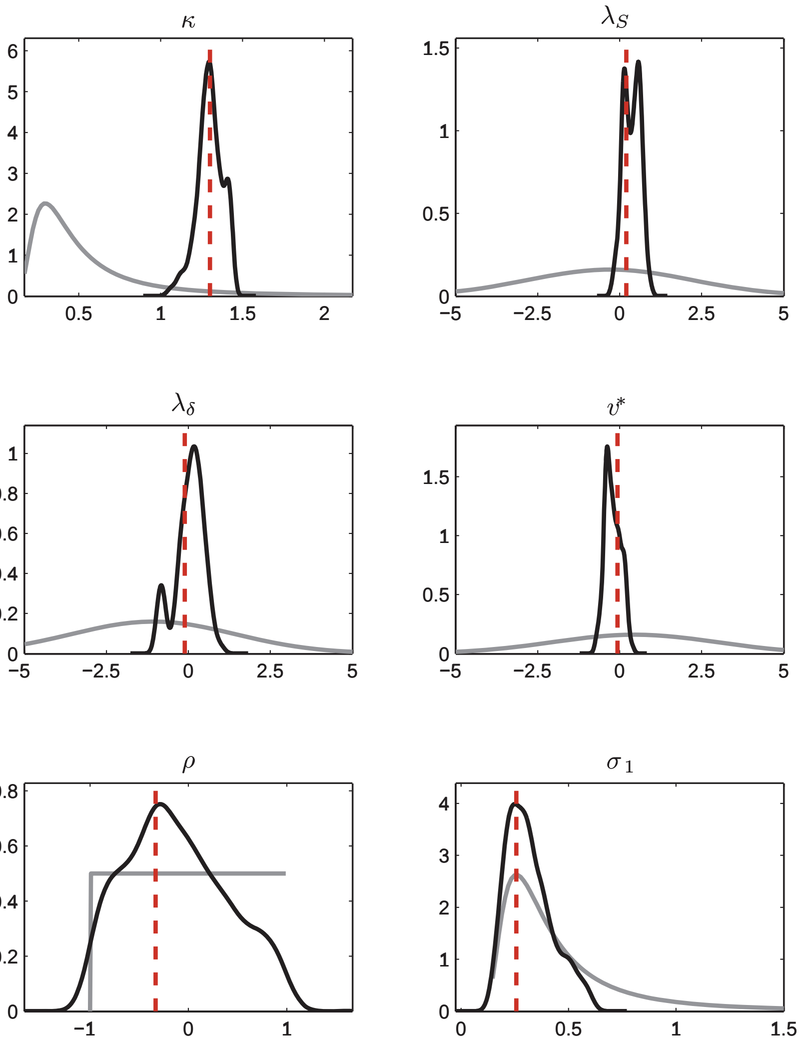


Figure 6.12: Prior and posterior distribution of parameters with US data (long-term data, part 1).

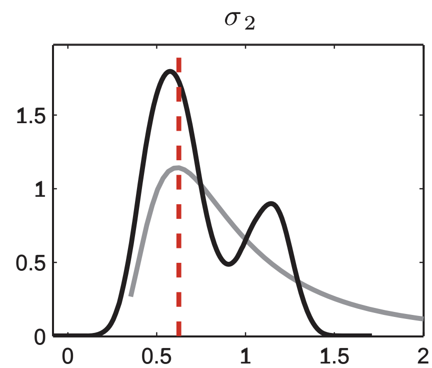


Figure 6.13: Prior and posterior distribution of parameters with US data (long-term data, part 2).

interval has widened even more compared to the short-maturity dataset. The risk-neutral drift v^* turns out negative which corresponds to a positive convenience yield. This is not supported by our preliminary analysis of the computed convenience yields. Yet again, the confidence region is large so the posterior density extends into the positive region. The latter finding shows that the identification problems of the convenience yield with only the long-term contracts are not specific to the ML routine but a more general problem. Correlation parameter ρ can again be considered indeterminate given the large confidence interval. In addition, σ_1 and σ_2 seem to be influenced by their prior distribution and are not similar in value to their counterparts in the ML estimation. The confidence interval for both of them is large with at least $\pm 50\%$ deviation around the mean estimate.

All in all, the bayesian estimation does not appear to be remarkably more stable than the ML-approach as long as the prior information is mainly taken from a preliminary analysis of the data itself. In our application, even cautiously adding such information and still allowing for pronounced variation might already have distorted some parameters' estimates. Yet, it could be seen that the assumption of asymptotic normality made in the ML-approach does not hold for the entire density space of the parameters. Most of the distributions can, in fact, be asymmetric or multimodal at some point. In this case, at least theoretically, the "maximum likelihood" estimator should not be the best pick since it is very sensitive to estimation uncertainty. This problem will particularly exist when a complicated recursive procedure such as the Kalman filter is involved and no analytical conditional posterior distributions can be computed. We have seen, for instance, that ρ remained basically unidentified in all bayesian estimations. Put differently, our standard errors from the ML estimation are very misleading since they only seem to hold in a very narrow space around the maximum likelihood estimator. This observation is of particular interest since it casts some doubt on the reliability of results for ρ in earlier studies such as Schwartz (1997), which are obtained with the maximum likelihood procedure. If the point estimate for ρ in his reduced-form model differs strongly from the one found by a simple regression analysis as in Gibson and Schwartz (1990), it is probably not very robust to changes in the dataset.

This leaves the question whether a simpler procedure does better in identifying the model parameters. With this motivation, we test, as a final alternative, the algorithm termed "implied state variable and parameter estimates" by Schwartz and Smith (2000). Applications of this technique can be found, for instance, in Lucia and Schwartz (2002), Cortazar and Schwartz (2003), Wilkens and Wimschulte (2007) and Yan and Li (2008). It consists of a 3-step loop. Since the futures price equation (6.26c) is linear in the state variables, the first step is to perform OLS regressions over the cross section of futures contracts at each observation date to obtain the state variables. The regression equation is equivalent to (6.27) with d_t being subtracted from the left-hand side. For this step, an initial estimate of the parameters has to be supplied. Next, the time series of state variable estimates is used to calculate σ_1, σ_2 and ρ from the state equations. In the third step, the latter parameters and the state variable estimates are assumed as given. The full futures dataset is then used to minimize the squared residual between model prices and observed prices using a non-linear optimization over κ, λ_δ and v^*. The loop is run until convergence in both parameter values and the squared residuals from the last step has occurred.[16] Risk premium λ_S is not needed in the loop, but can be recovered from (the mean of) the deterministic trend in the log-spot price equation, given v^* and σ_1.

While the procedure itself is easier to implement and computationally more efficient than the Kalman filter, the conditions under which convergence occurs are somewhat restrictive. Firstly, κ should not be too high. This is because otherwise the convenience yield shock will even out very quickly which means that the coefficient of the convenience yield in the futures price equation is nearly identical for the majority of (medium to long term) contracts. This means that the data matrix for the OLS regressions does almost lose its full rank and the regression results become inaccurate. This can lead to slow convergence or even divergence of the procedure. Secondly, the "measurement errors", i.e. the residuals of the futures price, are not supposed to be too important compared to the state noise (cf. Schwartz and Smith (2000), fn. 8). Since the model does not explicitly penalize noise in the state variables, some measurement noise will always be attributed to these variables. This effect leads to a bias in other parameters as well. Yet, it is less severe when the mean-reversion speed is rather low, i.e. the first restriction is respected. Finally, the initial values play a crucial role for both the speed and the success of convergence.

In turn, an advantage of this method is that it does not need a large sample since the estimator of both the OLS-regression and the nonlin-

[16] Cf. Cortazar and Schwartz (2003), p. 225.

ear optimization step produce efficient estimators non only in the limit. In addition, it is easy to retrieve unbiased parameter estimates when the measurement error is close to zero. Therefore, the question is how much

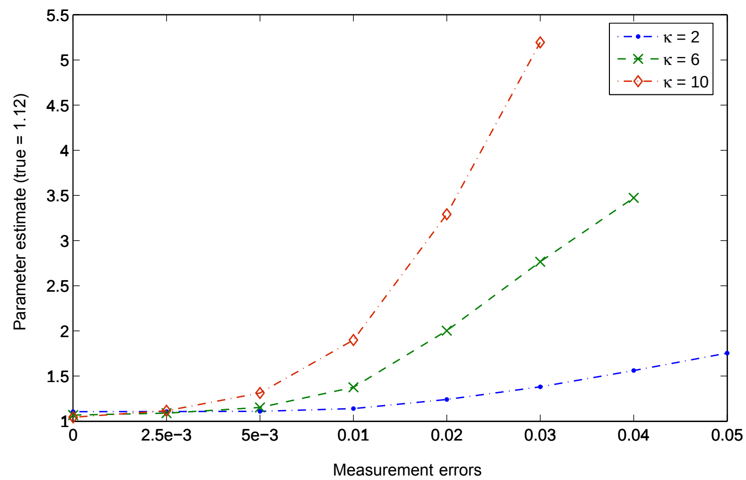


Figure 6.14: Analysis of estimation bias for σ_2 using implied parameter estimates.

residual noise this method can bear without becoming unacceptably inaccurate. Figure 6.14 illustrates the dimension of the first two restrictions with a simulated example for the bias in σ_2. 50 estimations from simulated datasets have been run for 8 different measurement error standard deviations and 3 different values of κ. The figure shows that for $\kappa \geqslant 6$ the bias becomes unacceptably high if the size of the measurement errors is in the order estimated for our natural gas futures data. Yet, for $\kappa = 2$, which is roughly the highest value that we have found in the US futures dataset, the error in σ_2 appears to be tolerable. Table 6.12 shows the mean (and standard deviation) of all parameters over the 50 iterations for each different size of the measurement noise when the true κ equals 2.

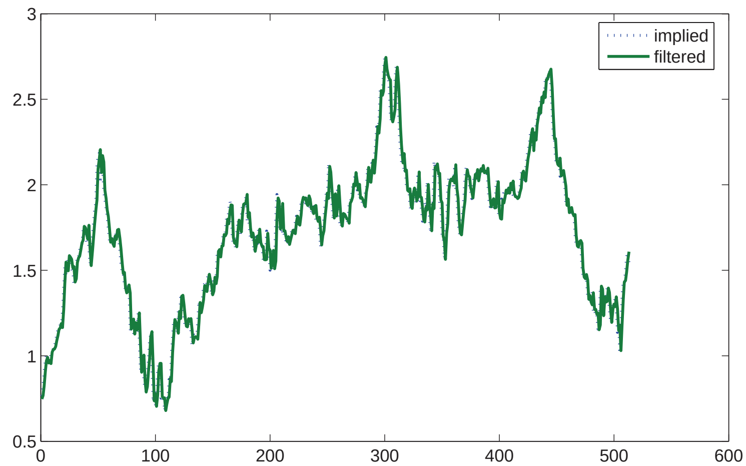
ξ_i	0	.0025	.005	.01	.02	.03	.04	.05
κ	2.000	2.000	2.001	1.998	1.998	1.985	1.972	1.962
(2.0)	(.001)	(.002)	(.007)	(.013)	(.027)	(.047)	(.051)	(.071)
λ_S	.100	.109	.114	.105	.120	.126	.115	.142
(.12)	(.095)	(.116)	(.122)	(.112)	(.127)	(.107)	(.104)	(.109)
λ_δ	-1.805	-1.803	-1.803	-1.800	-1.807	-1.801	-1.794	-1.808
(-1.8)	(.021)	(.017)	(.022)	(.023)	(.035)	(.051)	(.058)	(.082)
v^*	-.161	-.161	-.161	-.159	-.150	-.140	-.124	-.102
(-.16)	(.003)	(.002)	(.002)	(.003)	(.005)	(.009)	(.011)	(.014)
ρ	.292	.299	.302	.335	.414	.503	.582	.638
(.32)	(.043)	(.031)	(.043)	(.040)	(.037)	(.032)	(.041)	(.026)
σ_1	.344	.346	.347	.361	.390	.437	.498	.563
(.35)	(.012)	(.010)	(.010)	(.012)	(.013)	(.012)	(.017)	(.018)
σ_2	1.106	1.107	1.110	1.140	1.240	1.380	1.561	1.754
(1.12)	(.035)	(.032)	(.028)	(.038)	(.035)	(.058)	(.072)	(.061)
ξ_1	.000	.001	.003	.006	.012	.018	.024	.029
ξ_2	.000	.002	.004	.008	.017	.025	.034	.042
ξ_3	.000	.002	.004	.009	.018	.027	.036	.044
ξ_4	.000	.002	.004	.008	.016	.024	.032	.040
ξ_5	.000	.002	.004	.007	.014	.022	.029	.036

Table 6.12: Simulation results using implied parameter estimates.

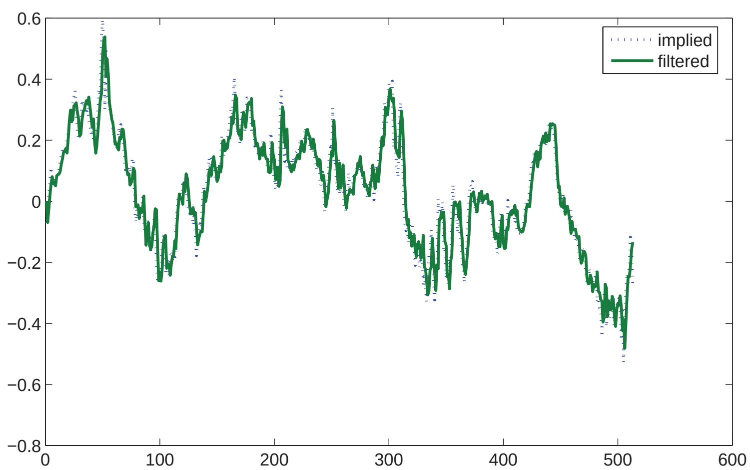
The actual parameter values are given below the respective parameter name, i.e. in the left-most column. The other columns show, from left to right, the estimation results with ascending standard deviation of the residuals. The overall impression is that the standard deviation of the parameters is remarkably lower than in the Kalman filter simulation. The only parameter which still carries economically relevant estimation uncertainty is λ_S. While mean estimates of ρ, σ_1, σ_2 remain slightly below the true values until the ξ's are increased to .005, they deteriorate for values higher than .01. v^* also becomes less accurate for higher measurement errors. However, the remaining parameters are not strongly affected and remain rather stable. Hence, we can conclude that, given $\kappa = 2$, the model remains a testable alternative for estimation. Yet, one might need to slightly downscale the state noise variability parameters according to the estimated amount of measurement noise.

Table 6.13 shows the estimation results with this method using the 3 different US synthetic futures datasets of the extended model (short, mixed and long maturities). Both the unweighted and weighted least squares (WLS) results are shown. The weights for the latter have been taken from the ML results of the Kalman filter estimation. We see that the results with weighted residuals get close to the ML results generated with the Kalman filter (Table 6.7) for the spanning and the short-term dataset.[17] An exception is σ_2 in the latter one, which is higher now. Figure 6.15 shows graphically that this also applies to the state variable estimates (here: spanning dataset), which are graphed together with the Kalman filter estimates. The precondition is that the "measurement errors" in the WLS version are weighted accordingly. With respect to the long-term dataset, apparently the same problem which has been outlined above for high values of κ has now occurred for $T-t$, the time to maturity. That is, if we only pick long-term futures, the regressor $-\frac{1-e^{-\kappa(T-t)}}{\kappa}$ of the convenience yield variable in the futures price equation is too similar in the cross-section and parameters can no longer be correctly identified. For evidence, note that the time for convergence becomes excessively long if we start with an initial value for κ in the order of the ML result (1.39).

[17] This is in line with the theoretical reasoning given in Schwartz and Smith (2000), fn. 8.



Spot price variable



Convenience yield variable

Figure 6.15: Plots of implied against filtered state variables (US, spanning dataset, WLS).

	span	span (WLS)	short	short (WLS)	long	long (WLS)
κ	.7260	.6588	3.8867	3.2E-5	.0003	-.0004
λ_S	-.0003	.0300	-.2055	.0620	.2146	.2455
λ_δ	.2722	-.2002	-1.4783	-.0001	.0356	.0685
v^*	-.4314	.2441	.5018	.2689	-.8998	.1443
ρ	.8608	.8722	.7427	.6242	.8896	.9165
σ_1	.5969	.5693	.6794	.5900	.6836	.7821
σ_2	.4882	.4041	2.1405	.6684	.4437	.5356
ξ_1	.0258	.0465	.0175	.0492	.0163	.0240
ξ_2	.0260	.0111	.0174	.0259	.0104	.0005
ξ_3	.0230	.0236	.0219	.0223	.0164	.0131
ξ_4	.0254	.0250	.0140	.0031	.0191	.0135
ξ_5	.0226	.0169	.0107	.0006	.0219	.0220
ξ_6	.0210	.0256	.0186	.0190	.0220	.0390

Table 6.13: Estimation results using implied parameter estimates. "span" denotes the spanning dataset (mixed maturities) as opposed to "short" and "long" maturities.

Obviously, even to obtain the results of the other two datasets, the residual weights must be known up-front. Therefore, a comparison with the unweighted results is in order. Overall, especially the estimates for λ_S, λ_δ and v^* appear to be sensitive to the change in the residual weighting. The estimate of κ also changes its value significantly in the short-maturity sample. Yet, the problem here is that the weight of the M1 contract is particularly low compared to the weights of the longer-term contracts M4 and M5 in the same sample. In fact, the latter contracts have measurement error standard deviations of .0031 and .0006 respectively in the weighted least squares version. This compares to a standard deviation of .0492 for the M1 contract, which practically means that the fit to the latter has almost no influence on the optimal parameter set. This shows that this is not simply a numerical problem, but rather a problem of the disparity in (maturity-matching) mean reversion speeds between the different contracts (or, more precisely, its implied convenience yields). From our regressions for different terms of convenience yields in 5.9 and 6.3), we have seen that the yield with the shortest maturity has the highest mean reversion speed.

To summarize, our analysis suggests that the implied method is a feasible alternative, but is only applicable under somewhat restrictive circumstances. That is, longer maturities in the dataset should not be too similar, and neither the mean-reversion speed nor the measurement noise should be very high. The method can produce results which are similar to the Kalman filter, but the risk premia and also v^* show at least the same sensitivity as in the latter estimation procedure. In fact, sensitivity can even become more severe if the mentioned restrictions are not sufficiently met. We conclude this subsection with the confidence that the ML estimation of our model is not "outperformed", in terms of robustness, by any of the two presented alternatives. In turn, the results above have provided support for our conjecture that a further extension of the SCY model for flexibility in mean-reversion speeds could be beneficial - at least for commodities with pronounced seasonality and high convenience yields such as natural gas.

6.3 Conclusion

In this chapter, we have derived an extended model for the gas spot price based on the SCY model by Gibson and Schwartz (1990) and Schwartz (1997). The extension consists of a second component of the convenience yield, which is not taken into account by these authors. This so-called fundamental component is driven by gas storage levels and air temperature.

It has to be distinguished from the preference component, the stochastic mean-reverting convenience yield variable known from the SCY model. We claim that the preference component is primarily driven by the investors' risk attitude and not by fundamental factors. We have estimated the extended model with the convenience yield model from chapter 5 together with a Kalman filter routine. Data from both the UK and the US market have been used. The resulting parameters as well as the predictions of spot and futures prices have been compared to those of the standard SCY model. Parameter estimates do not seem to be systematically distorted through our extension. Quite opposingly, some of them are more plausible in the extended specification.

In both versions, mean reversion of the stochastic convenience yield component (i.e. the preference component) turns out to be very slow, i.e. autocorrelation is in the order of 98% even with weekly instead of daily data intervals. This result is similar in size to the corresponding estimates in Schwartz (1997) for crude oil (cf. Table VI there), but it does not conform to the autocorrelation in the instantaneous convenience yield time series from the cost-of-carry equation. Week-on-week autocorrelation in the latter case is only 0.5 in the US market.[18] This means that the standard model fails to explain the mean reversion of the total convenience yield. Yet, testing the parameter robustness with different sets of futures maturities shows that even the extended model is not able to output an estimate for κ which is independent of the maturities of futures contracts in the sample. A possible explanation is that the assumption of a flat convenience yield is an oversimplified picture of reality.

The risk premia for spot price and convenience yield risk could not be identified with precision, but at least the second one has the correct sign in

[18] Cf. Table 6.3, line 1. Autocorrelation is given by $1 - \hat{\beta}_{y_{t-1}}$.

most cases. That is, it is negative and, hence, it pays to bear convenience yield risk. Also the risk-neutral spot price trend v^* (net of convenience yield) is in line with our expectation. The implied net mean convenience yield rate α is negative in all cases and shows once again that the cost of physical storage which is included in the net convenience yield, is much higher in the gas market than in the crude oil market.[19] This is in line with the empirical evidence found in chapter 5.

The remaining model parameters are estimated with a lack of precision, especially the correlation parameter for the stochastic increments ρ. Both a simulation with the Kalman filter routine and a bayesian estimation (Markov Chain Monte Carlo method) have shown that the point estimates of this parameter are not reliable. Therefore, the assertion that the high correlation of the convenience yield and the spot price makes a two-factor model with convenience yield generally inefficient, is flawed. The real problem seems to be that the coupling of the stochastic processes in Schwartz (1997) leads to a numerical identification problem. A way to circumvent this problem could be to redefine the spot price process in terms of the spot price trend net of convenience yield and find a closed-form solution. It is likely that in this case the spot price and convenience yield disturbances σ_1 and σ_2 would also become easier to identify and less sensitive to changes in the dataset.

While these problems apply to both the standard and the extended model, the goodness-of-fit comparison favors the latter one. For instance, in the five-period out-of-sample forecast for the UK market, the mean bias is reduced by 32% on average. This holds particularly for the spot price, but also for the fit to observed futures prices. The (quasi-)out-of-sample test for the US data applies an estimated model and the filtered state variables to price futures of other maturities in the same time period. In

[19] Cf. Schwartz (1997), Table VI.

this cross-sectional test both models seem to perform equally well. Hence, we may conclude that the hybrid model developed in this study provides slightly more accurate forecasts.

Yet, this effect is rather small compared to the change of parameter estimates in response to an alteration of the futures maturities in the dataset. We have investigated the two extreme cases with only short-term and only long-term maturities. In fact, the estimates vary strongly, although this effect is marginally smaller in the extended model. The two alternative estimation methods explored, the bayesian estimation and the method of implied state variable and parameter estimates, have not been able to improve numerical stability considerably. Since several parameters' posterior distributions in the bayesian approach have multiple modes and some are dominated by the prior, it is not clear in how far the results are more trustable. Estimation outcomes from the implied method are very similar to the Kalman filter when the residuals are weighted according to the measurement errors in the latter approach. If the residuals are unweighted, which is the best guess if their variability in unknown, outcomes differ materially.

In sum, the extended model for the gas spot price has been able to slightly improve the accuracy in forecasting. Nevertheless, the conceptual and numerical problems necessitate further investigation. It is widely known that the robustness of multidimensional models is a delicate issue. As a first important step, this chapter has pointed to the most relevant issues, and it has put in perspective the criticism of the convenience yield models for derivatives pricing.

Chapter 7

Conclusion

This thesis is motivated by the ongoing growth of exchange-based natural gas trading and, especially, the liberalization steps recently undertaken in the gas markets of the European Union. It has been argued that these steps, together with increasing global interest in CO_2-efficient power generation, have lead to a rising demand for short-term financial risk management for this commodity. In consequence, the question is what the fair price for the corresponding financial derivatives would be. It has been pointed out that gas has very specific physical and economic properties which distinguish it from other (energy) commodities. This is the reason why the aforementioned question should be studied individually in the natural gas market.

The basis of pricing a financial derivative is, almost always, an appropriate estimate of the expected future spot price, which was shown in chapter 3. Therefore, the thesis has looked at both reduced-form and structural commodity spot price models found in literature. It was found that no specific model for natural gas exists to date. Besides this, the choice of a model has been described as a trade-off between a number of model selection criteria. The most severe problem of structural models is the difficulty of empirical implementation while reduced-form models regularly neglect some of the economic properties of commodity markets. The aim of the thesis is to solve the trade-off by combining the advantages of structural and reduced-form models.

Since the two-factor reduced-form models provide an acceptable degree of in-sample fit and particular concerns have been raised about the convenience yield specification, the stochastic convenience yield model by Schwartz (1997) has been selected as the basis to construct a hybrid model. In chapter 5, the thesis tested the influence of several fundamental variables on convenience yield dynamics. It has turned out that gas storage levels and air temperature seem to be important and measurable determinants. With the help of these quantities, fundamental convenience yield forecasts have been generated for the UK and, in chapter 6, as well for the US gas market.

In chapter 6, an extended stochastic convenience yield model was derived. It redefines the convenience yield as the sum of a preference component of investors and a fundamental component. By constructing synthetic futures prices net of the fundamental component, the same estimation procedures as for pure reduced-form models can be applied to this model. While parameter estimates do not become more robust, the forecasting accuracy and the in-sample fit to both spot and futures prices is improved. However, we recommend further modifications to the model setup prior to an application to derivative pricing since parameter stability seems to have amended only at the margin.

Besides the remaining numerical and conceptual problems, the most important limits of this study are the somewhat restrictive assumptions imposed on the convenience yield in the extended model. One of these assumptions is that the fundamental and the preference component are uncorrelated. This requires that the risk attitude of investors is not influenced by storage levels or climate conditions in the market which can be debated. In addition, we assume that the fundamental component has no separate market price of risk. This neglects the risk inherent in the forecast of the fundamental component which might become considerable in size, depending on the frequency of observations and the forecast window.

Despite these limits, some implications for the derivative pricing practice should be noted. First of all, it is worthwhile to retain the concept of convenience yield for usage in spot price models. It it rooted in macroeconomic theory and can be linked to fundamental variables, in contrast, for instance, to the short-term component in the Schwartz and Smith (2000) two-factor model. According to our analysis, the high correlation of the stochastic increments of the risk factors primarily results from an econometric identification problem. In fact, a restated model with decoupled SDEs might solve this problem if closed-form solutions exist. Even if this problem remains unsolved, our model will most likely remain favorable compared to a pure reduced-form model with two rather independent factors. Clearly, the ability of a model without embedded fundamental relationships to fit a particular price dataset goes along with the considerable risk of misspecification.

Another practical implication of this thesis is that practitioners should be very cautious when estimating any of the multi-factor reduced-form models with the "implied method" suggested by Schwartz and Smith (2000). As has been shown in subsection 6.2.7, the results can be severely biased depending on the size of the idiosyncratic variation in futures prices and the degree of mean reversion in the second stochastic factor (the convenience yield or the short-term price component). Finally, an important insight for practitioners is that the mean-reversion estimate of the instantaneous convenience yield seems to understate the true mean reversion parameter and varies depending on the maturities of futures in the dataset. Therefore, the choice of contracts to estimate the model parameters should depend on the time-to-maturity of the derivatives which are to be valued.

We conclude this study with a short overview of interesting topics for future research. One such topic is the further extension of our model to accommodate a term structure of convenience yields. This seems reasonable given a number of hints which occurred during the empirical analysis. One possibility might be a log-linear relationship between the standard deviations of the convenience yield innovations which would attribute smaller shocks to long-term convenience yields. Mean-reversion speeds could be related by a hyperbolic function of maturity such that they decline with rising maturity and approach a certain minimum level (greater than zero). The idea of a term structure of convenience yields is already incorporated in a very general theoretical model by Miltersen and Schwartz (1998) which nests many of the models described in this study. Empirically, it may be preferable to estimate these term structure relationships via regression analysis and then run a Kalman filter estimation to determine the remaining model parameters.

Further research could also test how sizable the advantage of our extended model becomes for natural gas derivatives whose prices directly depend on spot price volatility (e.g. long-term futures options). In addition, the impact of the prices of substitutes on the convenience yield of natural gas could be investigated in more detail to validate the rather general specification of the preference component in our model. Finally, our assumption that the fundamental convenience yield component has no market price of risk could be relaxed, and forecast uncertainty could be directly attributed to the fundamental component. It would be interesting to see the modifications in the model for this case as well as the impact on the estimated size of spot price noise.

Appendix

Appendix 1
Importance of domestic and commercial gas consumption in UK

The following statistic from the UK Dept. of Business Enterprise and Regulatory Reform (BERR) demonstrates the importance of domestic and commercial gas demand based on their annual average share in total consumption. Commercial consumption includes public administration customers.

(in bcm)	2003	2004	2005	2006	2007
Total Consumption	103.14	105.39	102.42	97.48	98.36
Domestic Consumption	35.94	36.87	35.71	33.93	32.54
Commercial Consumption	7.80	8.33	7.94	7.73	7.49
Share of Total Cons.	42.42%	42.88%	42.63%	42.74%	40.70%

Appendix 2
Underground working gas capacity in relation to net import requirement in selected Western European Countries (12/2006)

The table below shows aggregate consumption and production as well as the underground working gas capacity of the United Kingdom (UK), the Netherlands (NL), Germany (GER), France (FR) and Italy (IT) at the end of 2006. The former two countries still have a largely autonomous supply, while GER, FR and IT are very dependent on imports already. This is demonstrated by the difference between consumption and production, the net import requirement. It is apparent that the amount of underground storage capacity is directly related to the net import requirement. Consumption and production data comes from the IEA's annual gas statistics whereas the working gas volumes come from the International Gas Union.

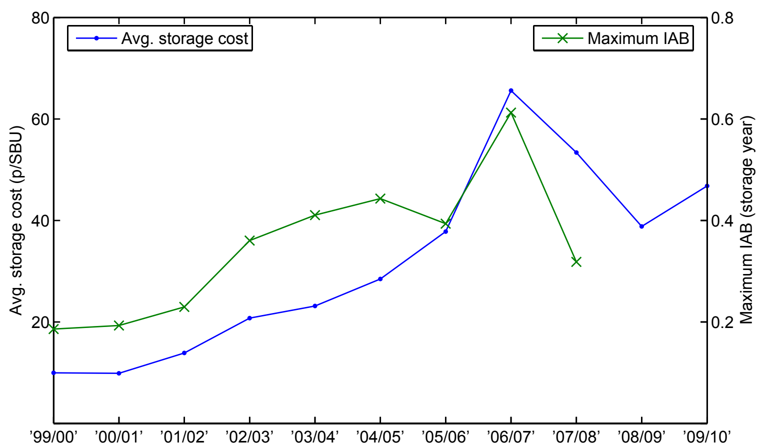
(in mcm)	UK	NL	GER	FR	IT
Consumption	96,974	41,487	95,248	47,292	83,126
Production	87,095	67,001	16,996	1,280	10,876
Net import requirement	9,879	0	78,252	46,012	72,250
Working gas	3,267	5,000	19,180	11,643	17,415

Appendix 3
Historical development of season-ahead storage cost

The most important part of the storage cost which influences the basis variability of gas is the capacity charges. Season-ahead storage capacity is sold in packages, so-called Standard Bundled Units (SBUs). They comprise a standard fraction of injection capacity, space capacity and withdrawal capacity rights for a limited time period (e.g. 67 days at Centrica). Storage facilities regularly offer numerous products with differing priority of service (firm versus interruptible) and renomination flexibility for the shipper of gas. Yet, price differences between these packages are negligible compared to the variability of these capacity charges throughout the year. The charges for the physical operation of storing gas are separate costs. These costs are based on actual throughput and also comprise an individual rate for injection, space and withdrawal respectively. Since their amount is each in the order of roughly 1/50 of the capacity charges only, they are economically negligible for us as well and, hence, not analyzed here.

Capacity already starts to be sold well ahead of maturity. After the storage operator has ended the auction for the next storage year, it continues to be traded on secondary markets. While depleted reservoirs accommodate by far the most important fraction of stored gas both in the UK and in the US, the UK market has a quasi-monopolistic operator, Centrica Storage Ltd. The company operates the Rough field in the North Sea. The prices for Rough storage capacity could be taken as market prices, but the company only publishes an average price of all SBUs sold for the next storage year (starting 1 May) on its homepage following legal obligations. These SBU prices are either fixed or indexed to the winter-summer spread of quarterly futures prices during an index period.

The graph below plots this average season-ahead price together with the maximum interest adjusted basis (IAB) one month ahead (M2 - M1) for the same storage year. The market seems to have anticipated the conditions in the following winter very well as shows the parallel development of the two lines. The upward trend can be attributed to the fact that capacity step-ups have not been keeping pace with the increasing dependence on imported gas.



Appendix 4
Variance of the forecast error in a linear regression with an AR(1) error process

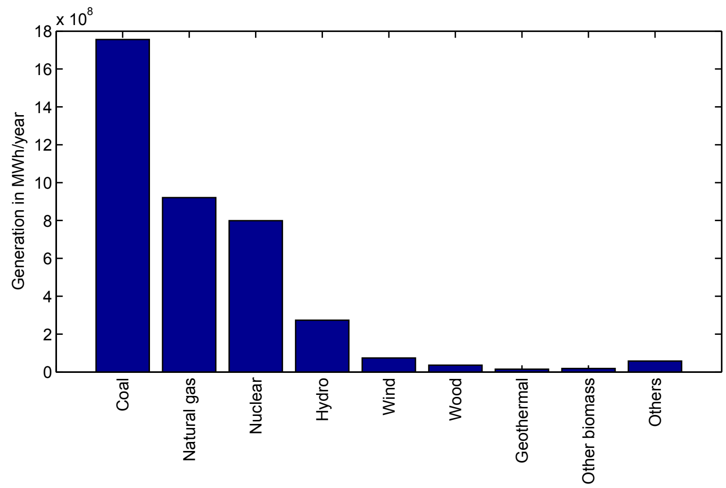
The variance of the conditional forecast error $v_{t+k}|u_t$ is obtained by recursively inserting u_{t+k}, $u_{t+k-1}, ...,$ u_{t+1}, u_t in the error generating process $u_t = \rho u_{t-1} + \epsilon_t$. The uncorrelated residuals remaining from each insertion step will then compose an MA$(k-1)$-process of the form

$$\begin{aligned}\mathrm{Var}(v_{t+k}|u_t) &= \rho^0\mathrm{Var}(\epsilon_{t+k}) + \rho^2\mathrm{Var}(\epsilon_{t+k-1}) + \rho^4\mathrm{Var}(\epsilon_{t+k-2}) + ... + \\ &\quad \rho^{2(k-1)}\mathrm{Var}(\epsilon_{t+1}) = \sum\nolimits_{i=0}^{k-1} \rho^{2i}\mathrm{Var}(\epsilon_{t+k-i})\end{aligned}$$

meaning that the forecast error is concave increasing in the forecast horizon k. If ϵ_t is assumed homoscedastic with $\mathrm{Var}(\epsilon_t) = \sigma^2$ we have that $\mathrm{Var}(v_{t+k}|u_t) = \sigma^2 \sum_{i=0}^{k-1} \rho^{2i}$, i.e. the relative advantage of one model does not change if the number of lags increases.

Appendix 5
US electric generation by energy source (2009)

The following chart shows electric generation data for the United States provided by the Energy Information Administration (EIA) for the year 2009.



References

APX-ENDEX (2011). Annual Report 2010.

Avellaneda, M., Buff, R., Friedman, C., Grandechamp, N., Kruk, L., and Newman, J. (2001). Weighted Monte Carlo: A New Technique for Calibrating Asset Pricing Models. *International Journal of Theoretical and Applied Finance*, 4, 91–119.

Bailey, W., and Chan, K.C. (1993). Macroeconomic Influence and the Variability of the Commodity Futures Basis. *The Journal of Finance*, 48, 555–573.

Benth, F.E., Benth, J.S., and Koekebakker, S. (2008). *Stochastic Modelling of Electricity and Related Markets*. Singapoore: World Scientific.

Bessembinder, H., Coughenour, J.F., Seguin, P.J., and Smoller, M. Monroe. (1995). Mean Reversion in Equilibrium Asset Prices: Evidence from the Futures Term Structure. *The Journal of Finance*, 50, 361–375.

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

Björk, T., Kabanov, Y., and Runggaldier, W. (1997). Bond Market Structure in the Presence of Market Point Processes. *Mathematical Finance*, 7, 211–239.

Black, F. (1976). The Pricing of Commmoditiy Contracts. *Journal of Financial Economics*, 3, 167–179.

Black, F., and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81, 637–654.

Bolder, J. (2001). Affine Term-Structure Models: Theory and Implementation. Working Paper. Bank of Canada. October 2001.

Boyle, P.P. (1977). Options: A Monte Carlo Approach. *Journal of Financial Economics*, 4, 323–338.

Boyle, P.P. (1986). Option Valuation Using a Three Jump Process. *International Options Journal*, 3, 7–12.

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

Brennan, M.J., and Schwartz, E.S. (1978). Finite-Difference Methods and Jump Processes Arising in the Pricing of Contingent Claims. *Journal of Financial and Quantitative Analysis*, 13, 461–474.

Brennan, M.J., and Schwartz, E.S. (1985). Evaluating Natural Resource Investments. *The Journal of Business*, 58, 135–157.

Carlson, M., Khokher, Z., and Titman, S. (2007). Equilibrium Exhaustible Resource Price Dynamics. *The Journal of Finance*, 62, 1663–1703.

Cartea, A., and Williams, T. (2008). UK gas markets: The Market Price of Risk and Applications to Multiple Interruptible Supply Contracts. *Energy Economics*, 30, 829–846.

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

Cassasus, J., Collin-Dufresne, P., and Routledge, B.R. (2005). Equilibrium Commodity Prices with Irreversible Investment and Non-Linear Technologies. Working Paper. Pontificia Universidad Catolica de Chile. November 2005.

Cassasus, J., Liu, P., and Tang, K. (2009). Commodity Prices in the Presence of Inter-commodity Equilibrium Relationships. Working Paper. Pontificia Universidad Catolica de Chile. May 2009.

Chambers, M.J., and Bailey, R.E. (1996). A Theory of Commodity Price Fluctuations. *Journal of Political Economy*, 104, 924–957.

Chen, Z., and Forsyth, P.A. (2006). Stochastic Models of Natural Gas Prices and Applications to Natural Gas Storage Valuation. Working Paper. University of Waterloo (Ontario, CA). November 2006.

Clewlow, L., and Strickland, C. (2000). *Energy Derivatives: Pricing and Risk Management*. London: Lacima Publications.

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

Courtadon, G. (1982). A More Accurate Finite Difference Approximation for the Valuation of Options. *Journal of Financial and Quantitative Analysis*, 17, 75–100.

Cox, J.C., and Ross, S.A. (1976). The Valuation of Options for Alternative Stochastic Processes. *Journal of Financial Economics*, 3, 145–166.

Cox, J.C., Ross, S.A., and Rubinstein, M. (1979). Option Pricing: A Simplified Approach. *Journal of Financial Economics*, 7, 229–263.

De Jong, P. (1991). The Diffuse Kalman Filter. *The Annals of Statistics*, 19, 1073–1083.

Deaton, A., and Laroque, G. (1992). On the behaviour of Commodity Prices. *Review of Economic Studies*, 59, 1–23.

Deaton, A., and Laroque, G. (1996). Competitive Storage and Commodity Price Dynamics. *Journal of Political Economy*, 104, 896–923.

Dietert, J.A., and Pursell, D.A. (2000). *Underground Natural Gas Storage*. `http://www.simmonsco-intl.com/files/63.pdf`. Online. [Accessed 14 July 2010].

Dincerler, C., Khokher, Z., and Simmin, T. (2005). An Emprirical Analysis of Commodity Convenience Yields. Working Paper. Pennsylvania State University. June 2005.

Douglas, S., and Popova, J. (2008). Storage and the Electricity Forward Premium. *Energy Economics*, 30, 1712–1727.

Duffie, D., and Glynn, P. (1996). Efficient Monte Carlo Estimation of Security Prices. *The Annals of Applied Probability*, 5, 897–905.

Duffie, D., Pan, J., and Singleton, K. (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica*, 68, 1343–1376.

Epps, T.W. (2007). *Pricing Derivative Securities*. 2 edn. Singapoore: World Scientific.

Evans, L.C. (1998). *Partial Differential Equations*. Oxford: Oxford University Press.

Eydeland, A., and Geman, H. (1999). *Energy Modelling and the Management of Uncertainty*. London: Risk Books.

Fama, E.F., and French, K. (1987). Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage. *The Journal of Business*, 60, 55–73.

Fama, E.F., and French, K. (1988). Business Cycles and The Behavior of Metals Prices. *The Journal of Finance*, 43, 1075–1093.

Gao, A.H., and Wang, G.H.K. (2005). Asymmetric Volatility of Basis and the Theory of Storage. *Journal of Futures Markets*, 25, 399–418.

Geman, H. (2006). *Commodities and Commodity Derivatives.* Chichester: Wiley.

Geman, H., and Roncoroni, A. (2006). Understanding the Fine Structure of Electricity Prices. *Journal of Business*, 79, 1225–1261.

Gibson, R., and Schwartz, E.S. (1990). Stochastic Convenience Yield and the Pricing of Oil Contingent Claims. *The Journal of Finance*, 45, 959–976.

Glasserman, P., and Yu, B. (2005). Large Sample Properties of Weighted Monte Carlo Estimators. *Operations Research*, 53, 298–312.

Goldfeld, S.M., and Quandt, R.E. (1972). *Nonlinear Methods in Econometrics.* Amsterdam: North-Holland.

Grewal, M.S., and Andrews, A.P. (2001). *Kalman Filtering: Theory and Practice Using Matlab.* Chichester: Wiley.

Griffoli, T.M. (2009) (November). *Dynare v4 - User Guide.*

Haase, N. (2008). *European Gas Market Liberalisation: Are Regulatory Regimes Moving Towards Convergence?* Tech. rept. Oxford Institute for Energy Studies.

Haff, I.H., Lindqvist, O., and Loland, A. (2008). Risk Premium in the UK Natural Gas Forward Market. *Energy Economics*, 30, 2420–2440.

Harrison, M.J., and Kreps, D. (1979). Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory*, 20, 381–408.

Harrison, M.J., and Pliska, S.R. (1981). Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochatic Processes and Their Applications*, 11, 215–260.

Harvey, A.C. (1989). *Forecasting structural time series models, and the Kalman filter*. Cambridge: Cambridge University Press.

Heston, S. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 6, 327–343.

Hotelling, H. (1931). The Economics of Exhaustible Resources. *Journal of Political Economy*, 39, 137–175.

Huisman, R., and Mahieu, R. (2003). Regime jumps in electricity prices. *Enegery Economics*, 25, 425–434.

Hull, J.C. (2009). *Options, Futures and Other Derivatives.* New Jersey: Pearson.

Hull, J.C., and White, A. (1987). The Pricing of Options on Assets with Stochastic Volatilities. *The Journal of Finance*, 42, 281–300.

Hull, J.C., and White, A. (1994a). Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models. *The Journal of Derivatives*, 2 (Fall), 7–16.

Hull, J.C., and White, A. (1994b). Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models. *The Journal of Derivatives*, 2 (Winter), 37–48.

IEA (2011). *World Energy Outlook 2011 - Special Report (Are we entering a golden age of gas?).* `http://www.worldenergyoutlook.org/golden_age_gas.asp`. Online. [Accessed 17 March 2011].

ILEX (2004). *Gas Prices in the UK.* `http://www.oilandgas.org.uk/issues/gas/ilexreport.pdf`. Online. [Accessed 22 March 2009].

Jaillet, P., Ronn, E.I., and Tompaidis, S. (2004). Valuation of Commodity-Based Swing Options. *Management Science*, 50, 909–921.

Johannes, M., and Polson, N. (2003). MCMC Methods for Continuous-Time Financial Econometrics. Working Paper. Columbia University. December 2003.

Johnson, H., and Shanno, D. (1987). Option Pricing When the Variance is Changing. *Journal of Financial and Quantitative Analysis*, 22, 143–151.

Kaldor, N. (1939). Speculation and Economic Stability. *Review of Economic Studies*, 7, 1–27.

Kogan, L., Livdan, D., and Yaron, A. (2009). Oil Futures Prices in a Production Economy with Investment Constraints. *The Journal of Finance*, 64, 1345–1375.

Koopman, S.J. (1997). Exact Initial Kalman Filtering and Smoothing for Nonstationary Time Series Models. *Journal of the American Statistical Associtation*, 92, 1630–1638.

Koopman, S.J., and Durbin, J. (2003). Filtering and Smoothing of State Vector for Diffuse State-Space Models. *Journal of Time Series Analysis*, 24, 85–98.

Litzenberger, R.H., and Rabinowitz, N. (1995). Backwardation in Oil Futures Markets: Theory and Empirical Evidence. *The Journal of Finance*, 50, 1517–1545.

Lucia, J.J., and Schwartz, E.S. (2002). Electricity Prices and Power Derivatives: Evidence from the Nordic Power Exchange. *Review of Derivatives Research*, 5, 5–50.

Marckhoff, J., and Muck, M. (2008). Die Bewertung von Stromderivaten mit Hilfe von Reduced-Form-Modellen. *Pages 295–320 of:* Oehler, A., and Terstege, U. (eds), *Finanzierung, Investition und Entscheidung: Einzelwirtschaftliche Analysen zur Bank- und Finanzwirtschaft*. Berlin: Springer.

Marckhoff, J., and Muck, M. (2009). Jump Risk Premia in Short-Term Spread Options: Evidence from the German Electricity Market. Working Paper. University of Bamberg. April 2009.

Merton, R.C. (1973). Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science*, 4, 141–183.

Merton, R.C. (1976). Option Pricing When Underlying Stock Returns are Discontinous. *Journal of Financial Economics*, 3, 125–144.

Miltersen, K.R., and Schwartz, E.S. (1998). Pricing Options on Commodity Futures with Stochastic Term Structures of Convenience Yields and Interest Rates. *The Journal of Financial and Quantitative Analysis*, 33, 33–59.

Modjtahedi, M., and Movassagh, N. (2005). Natural Gas Futures: Bias, Predictive Performance and the Theory of Storage. *Energy Economics*, 27, 617–637.

Muck, M. (2006). *Portfoliomanagement and Derivatives*. Professional dissertation, WHU - Otto Beisheim School of Management.

Muck, M., and Rudolf, Markus. (2005). Improving Discrete Implementation of the Hull and White Two-Factor Model. *Journal of Fixed Income*, 14, 67–75.

Nagelkerke, N.J.D. (1991). A Note on a General Definition of the Coefficient of Determination. *Biometrika*, 78, 691–692.

National Grid (2007). *Gas Transportation Ten Year Statement 2007*. `http://www.nationalgrid.com/uk/Gas/TYS/`. Online. [Accessed 23 March 2009].

Neumann, A., and von Hirschhausen, C. (2004). Less Long-Term Gas to Europe? A Quantitative Analysis of European Long-Term Gas Supply Contracts. *Zeitschrift für Energiewirtschaft*, 28, 175–182.

Panagiotidis, T., and Rutledge, E. (2007). Oil and Gas Markets in the UK: Evidence from a Cointegrating Approach. *Energy Economics*, 29, 329–347.

Pirrong, S.C. (1998). High Frequency Price Dynamics and Derivatives Prices for Continuously Produced, Storable Commodities. Working Paper. Washington University. September 1998.

POST (2004). *The Future of UK Gas Supplies*. `http://www.parliament.uk/documents/upload/POSTpn230.pdf`. Parliamentary Office of Science and Technology. Postnote No. 230 (October 2004). Online. [Accessed 23 March 2009].

Price, K.W. (1997). Competition and Regulation in the UK Gas Industry. *Oxford Review of Economic Policy*, 13, 47–63.

Ribeiro, D.R., and Hodges, S.D. (2004). Equilibrium Model for Commodity Prices: Competitive and Monopolistic Markets. Working Paper. University of Warwick. August 2004.

Ross, S.A. (1978). A Simple Approach to the Valuation of Risky Streams. *The Journal of Business*, 51, 453–475.

Routledge, B.R., Seppi, D.J., and Spatt, C.S. (2000). Equilibrium Forward Curves for Commodities. *The Journal of Finance*, 55, 1297–1338.

Routledge, B.R., Seppi, D.J., and Spatt, C.S. (2001). The Spark Spread: An Equilibrium Model of Cross-Commodity Price Relationships in Electricity. Working Paper. Carnegie Mellon University. May 2001.

Schwartz, E.S. (1977). The Valuation of Warrants: Implementing a New Approach. *Journal of Financial Economics*, 4, 79–93.

Schwartz, E.S. (1997). The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. *The Journal of Finance*, 52, 923–973.

Schwartz, E.S., and Smith, J.E. (2000). Short-Term Variations and Long-Term Dynamics in Commodity Prices. *Management Science*, 46, 893–911.

Seifert, J., and Uhrig-Homburg, M. (2007). Modelling Jumps in Electricity Prices: Theory and Empirical Evidence. *Review of Derivatives Research*, 10, 59–85.

Seppi, D.J. (2002). Risk-Neutral Stochastic Processes for Commodity Derivative Pricing: An Introduction and Survey. *Chap. 1.1 of:* Ronn, E.I. (ed), *Real Options and Energy Management*. London: Risk Books.

Shively, B., and Ferrare, J. (2007). *Understanding Today's Natural Gas Business*. San Francisco: Enerdynamics.

Stronzik, M., Rammersdorfer, M., and Neumann, A. (2008). Theory of Storage - An Empirical Assessment of the European Natural Gas Market. *5th International Conference on the European Electricity Market*, 1–6.

Sturm, Fletcher J. (1997). *Trading Natural Gas - A Nontechnical Guide*. Tulsa: Pennwell.

Telser, L.G. (1958). Futures Trading and the Storage of Cotton and Wheat. *Journal of Political Economy*, 66, 233–255.

Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics*, 5, 177–188.

Volmer, T. (2011). A Robust Model of the Convenience Yield in the Natural Gas Market. *Journal of Futures Markets*, 31, 1011–1051.

Wei, S.Z.C., and Zhu, Z. (2006). Commodity Convenience Yield and Risk Premium Determination: The Case of the U.S. Natural Gas Market. *Energy Economics*, 28, 523–534.

Wiersema, U.F. (2008). *Brownian Motion Calculus.* Chichester: Wiley.

Wilkens, S., and Wimschulte, J. (2007). The Pricing of Electricity Futures: Evidence from the European Energy Exchange. *Journal of Futures Markets*, 27, 387–410.

Williams, J.C., and Wright, B.D. (1991). *Storage and Commodity Markets.* Cambridge: Cambridge University Press.

Yan, W., and Li, S. (2008). A class of porfolio selection with a four-factor futures price model. *Annals of Operations Research*, 164, 139–165.