Norwegian School of Economics
and Business Administration
Bergen, autumn 2010

The forward-spot spread in the natural gas market:

An empirical investigation of Henry Hub and NBP

By Jacob Christopher Falck and Jens Sørlie Kværner

Advisor: Professor Thore Johnsen

Master thesis within the main profile of Economic Analysis

Abstract

This study analyses forward-spot relationships at two of the world's largest hubs for natural gas (Henry Hub and NBP). We find that spot and forward prices are covariance-stationary. Testing the theory of storage shows that inventories are highly significant in explaining the basis. In particular, we find evidence of a positive cost-of-carry in both markets. Furthermore, in both markets, forward prices have on average exceeded subsequent spot prices. Under the assumption of rational expectations, this indicates a negative risk premium. In fact, the premium appears to be time-varying. Finally, expected inventories at a contract's maturity seem to be a more important determinant of the risk premium than the contractual length in the UK.

Preface

This dissertation completes our Master of Science in Economics and Business Administration at NHH. The choice of topic was initiated by Dag Willoch in Yara seeking students to perform an empirical study of the natural gas forward market. In particular, we were to test various well-established hypotheses regarding the spot-forward relationship.

Working on this thesis has been exciting and very rewarding. It has further enriched our understanding of the natural gas market and the statistical framework necessary to analyse it. In retrospect, we truly acknowledge our major in Economic Analysis for providing us with a solid basis for writing empirical analyses.

Throughout the work we also experimented with various stochastic models and different calibration techniques such as the Kalman filter. Some of these methods required a tremendous workload for the authors. We recognise that our lack of training in solving intractable problems using numerical methods led us to choose parsimony over complexity.

We would like to thank our advisor Professor Thore Johnsen for his good advice and economic insights. Dag Willoch in Yara deserves credit for coming up with an interesting topic and his help along the way. We are also immensely grateful to Bård Misund and Morten Færevåg in Statoil for providing us with the data. Finally, we wish to thank Jonas Andersson, Bård Støve and Andrzej Mikolaj Suslowicz for their helpful comments on different statistical issues.

Contents

1 Introduction

This paper investigates whether well-established theories concerning the relationship between spot and forward prices in the natural gas market hold up to empirical scrutiny. We consider two of the world's most central hubs for trading gas, namely NBP in the UK and Henry Hub in the US. We use prices dating back to 1999 up to the most recent data available. More importantly, as opposed to similar studies, we perform a comparative analysis using the same sample length for both markets.

We find that the theory of storage in its original form performs rather poorly. In particular, we observe that the role of the cost of capital is nonessential. However, the level of inventory is highly significant in explaining the basis. We detect a high forward-spot spread in the early autumn when inventories are plentiful, whereas the basis is deteriorating in the heating season when inventories are scarce. Moreover, a concave relationship between inventories and the basis is deemed appropriate in both markets. As opposed to other authors, we observe that not only shocks to the level of storage works to increase the basis, but rather that these add to the effect of absolute inventory levels.

Furthermore, we find the basis to be positive on average, implying a positive cost-of-carry in both markets. Much to our surprise, we find that the mean spread is larger in the UK than in the US. Intuitively, we would expect the opposite due to the higher variability in UK inventories, suggesting a higher convenience yield and hence a lower cost-of-carry. However, the large UK basis could be justified by the relatively higher concentration of suppliers in the UK, seeking economic rent by keeping forward prices "artificially" high. In addition, there could be a lack of arbitrageurs exploiting the price differential.

Through extensive testing of the unbiasedness hypothesis we find that forward prices are not unbiased predictors of subsequent spot prices in the UK. Instead they contain a significant time-varying bias. We argue that the bias inherent in forward prices reside heavily on the supply and demand for hedging. In particular, the UK market structure, with a pronounced hedging-demand for forward contracts and supply-side market power, dictates that forward prices are consistently above expected spot prices. Under the assumption of rational expectations, the bias may be interpreted as a negative forward risk premium.

The US market on the other hand, is more balanced in terms of hedging pressure. The supply-side is more competitive and large purchasers are subject to strict regulatory restrictions. Consequently, rejecting the unbiasedness hypothesis is somewhat harder in the US. Still, we are in fact able to detect a negative risk premium in the forward price. This is surprising, considering that most previous studies come to the opposite conclusion. We argue that this contradiction could be related to the emergence of index speculators in the recent decade, imposing an upward pressure on the demand for forward contracts. Moreover, the increased importance of shale gas could have altered the historical spot-forward relationship.

Finally, we discover a strong link between the market price of risk and the level of inventory at delivery in the UK. In fact, we find that inventories at a contract's maturity appear to be a more important determinant of the UK risk premium than the contractual length. This suggests that the risk premium may be interpreted from an insurance perspective; risk-averse agents' willingness to pay a premium depends on the level of anticipated price risk at maturity, which is mainly determined by the expected level of inventory. In the US however, it seems that the storage level at delivery is a less important factor in explaining the market price of risk. This is probably due to US inventories being less seasonal.

The paper is organised as follows. In the next section, we briefly introduce the most important market centres and the instruments traded. We then proceed by examining the data at hand and conducting unit-root tests for stationarity. Section 4 investigates the relationship between the forward-spot spread using the traditional theory of storage. This is followed by an extended analysis of the unbiasedness hypothesis. Section 6 presents a stochastic modelling approach in which we link the size and direction of the risk premium to the level inventory. Finally, the insights from previous sections are put into practise when examining the relationship between inventories at delivery and the risk premium.

2 The natural gas market

Trading points for natural gas are located at onshore terminals/hubs where gas is delivered to the pipeline network, with infrastructural capabilities such as storage and concentration of buyers and sellers. The most important market centre of natural gas in the US, with the highest daily trading volumes, is the Henry Hub located in Louisiana. Henry Hub is used as delivery point for the New York Mercantile Exchange's (NYMEX) natural gas futures contracts, and is a pricing reference point for virtually the entire North American natural gas market (Augustine et al., 2006). Equivalently, the UK's most important market centre, the National Balancing Point (NBP), is merely a notional hub, where all UK gas flows through, and is the pricing and delivery point for ICE structured natural gas futures. Futures are traded for a wide range of maturities, namely weeks, months, quarters and even years for both markets. Most of the trade in futures takes place in the over-the-counter market (OTC), however some exchanges also offer futures with physical delivery trough a hub. Furthermore a market for short-term delivery exists, commonly referred to as the spot market. Nevertheless, this is not an organised market in the sense of standardised spot contracts traded with publicly available prices (Benth et al., 2008). Hence, short-term trading is also mostly OTC. Still, several well-known objective day-ahead indexes are available, functioning as proxies for the actual spot price. The index for a given day is the volume-weighted average of transaction prices for gas to be delivered the following day.

3 Data

The study uses several sets of data. For the British market, the data covers daily spot and forward prices for the National Balancing Point (NBP) over the period from January 1999 through August 2010, although earlier data is available. This is because in 1998, a large interconnector pipeline between Bacton and Zeebrugge established a link between UK gas prices and the oil-indexed prices in continental Europe. This structural change means that pre-1998 data will no longer be relevant to the current market.

All the UK forward prices we analyse come from Heren Energy Ltd. The spot prices are day-ahead and weekend-ahead prices from the NBP ICIS Day Ahead Heren mid index. The equivalent US forward prices are obtained from NYMEX quotations for Henry Hub (henceforth referred to as HH). Furthermore, US spot prices come from the Louisiana Onshore South Henry Hub Platts mid index. All the spot and forward prices analysed are price assessments due to the lack of closing prices. Moreover, both UK and US forward contracts are delivered as a continuous flow over the delivery month. In the UK, gas is typically denominated in *British thermal units* (Btu) and contract prices are quoted in pence per therm, whereas in the US, gas transactions are denominated in USD per MMBtu.

The use of price assessments rather than actual close prices can potentially induce measurement errors. As long as these errors are unbiased, they will average out to zero. On the other hand, a bias in either direction will affect the validity of our estimates. Obviously this matter must be taken into account when interpreting the results.

We consider monthly forward contracts with delivery up to 6-months ahead. We omit longer contracts from the analysis due to their rather poor liquidity. Prices on forward contracts are chosen instead of futures for the same reason. The data series contain 127 and 94 missing values (constituting 4.7 percent and 3.1 percent of the total amount of data) for HH and NBP, respectively. The missing values where set equal to the average of the two closest observations. Furthermore, following the lead of the literature, the daily forward prices are made monthly by choosing the prevailing price from the weekday closest to the 20th each month. Hence, we always use a day close to the end of the month that exists for all months of the year. Spot prices are aggregated to monthly resolution by computing monthly arithmetic means in order to facilitate comparison between spot and forward prices. This is because monthly forward contracts have delivery over a month rather than at a particular date.

Some of the models require additional data such as interest rates and storage levels. The former include monthly (annualised) data on 1, 3 and 6 month LIBOR money market interest rates quoted in GBP and USD. Another pertinent question is which type of storage data ought to be applied. We use the sum of monthly base and working storage figures for the US market obtained from the Energy Information Administration (EIA, 2010a). This is consistent with the approach taken in previous literature (see e.g. Modjtahedi and Movassagh, 2005). We argue that due to interconnector pipelines, and hence strong link to the continental markets, storage data from continental Europe should also be included in the UK case. Moreover, Haff et al. (2008) find that using aggregate European storage numbers improves the results. Therefore, despite the relatively questionable quality of data for some countries, UK inventory data consist of monthly storage numbers from Europe provided by PIRA Energy Group.

3.1 Preliminary look at the data

In this section we analyse our data by considering numerous plots of the various time series. Although casual inspection does have its perils, visual patterns could indicate whether stationarity tests are needed to substantiate any first impressions. Furthermore, we look for key features such as trends, seasonality, large price spikes and tendencies towards volatility clustering. Fig. 3.1 plots the monthly spot price in addition to the 1-month and 6-months ahead forward prices on log-scale for NBP and HH. Note that the contracts have the same delivery date.

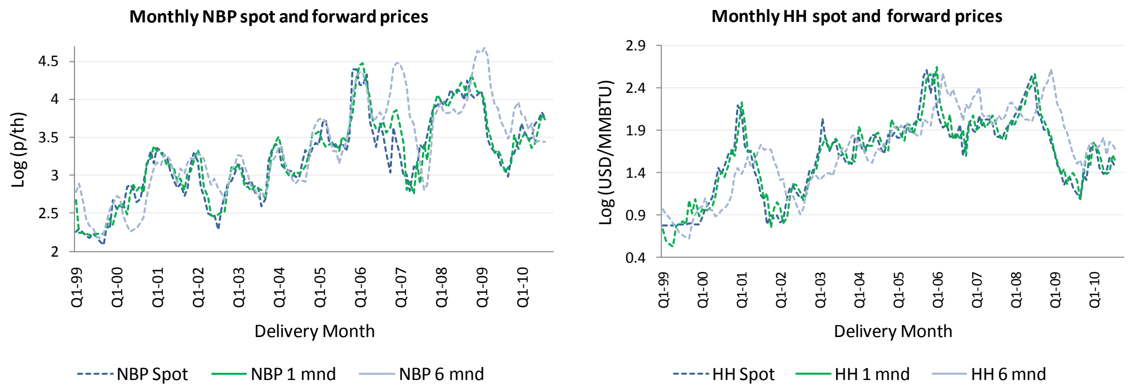

Fig. 3.1 Monthly spot and forward prices on log scale from January 1999 till August 2010

Several features are noteworthy. Firstly, there appears to be a positive trend in all the series. This immediately raises the question of whether the series are trend stationary or random walks with drifts. Obviously formal testing is warranted. Secondly, the 1-month forwards seem to trace the spot prices reasonably well and appear to be as volatile in both graphs. The 6-month forwards, however, miss the target quite often and, at least for the NBP case, appear to be somewhat more volatile. The former is naturally a characteristic we would expect to find in any forecast series. Following a relatively stable period from 2003 to 2007 with overall economic growth and less uncertainty, the volatility of the HH series has increased in recent years. The same pattern is evident in the UK where the volatility seems to be on the rise throughout the whole time span.

Furthermore, we observe a prominent seasonality in the prices. Because natural gas consumption is seasonal while production is not, we tend to find higher prices in the winter than in the summer. Inventories are built during the summer for use in the winter, putting an upward pressure on prices during periods of cold weather due to increased scarcity of gas. Although apparent in the US, this pattern seems to be more pronounced in the UK, probably a result of natural gas playing a more important role for heating there. Residential consumption (mainly used for heating) accounts for approximately 36 percent of total consumption in the UK in 2009 (Department of Energy and Climate Change, 2010). The US equivalent is approximately 21 percent (EIA, 2010b). The difference in seasonality is further emphasised in Fig. 3.2, showing monthly median spot prices from January 2000 to August 2010 for HH and NBP.

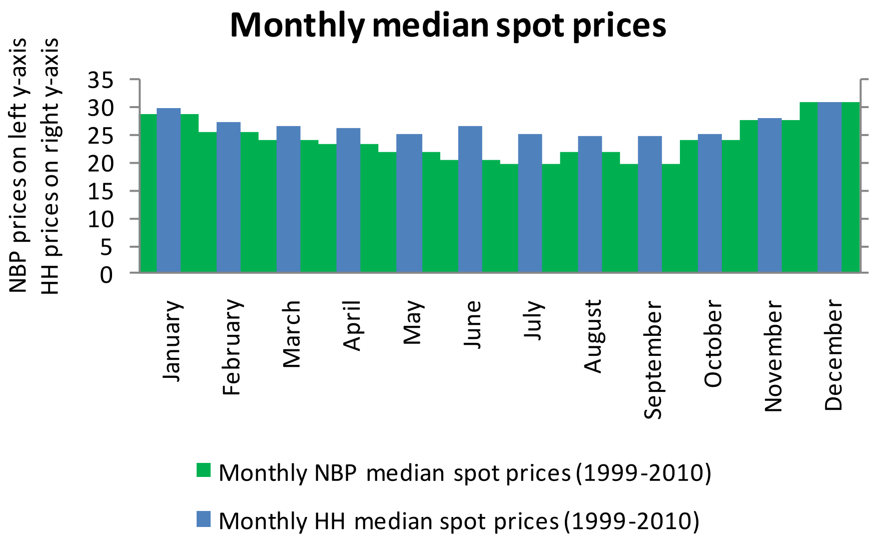

Fig. 3.2 Monthly median spot prices

Finally, we could possibly be dealing with a structural break in both our series due to the emergence of low-cost unconventional gas, especially shale gas. In 2008, shale gas production constituted approximately 8 percent of North American gas production and its share has been growing ever since due to advances in hydraulic fracturing and horizontal completions (Cohen, 2009). Not surprisingly, this has depressed natural gas prices, a feature that is evident in both markets. Moreover, large investments in LNG capacity have further increased the fall in the wholesale price of natural gas. Today, the US natural gas market is quite self sufficient. Therefore LNG, originally intended for North American consumers, is redirected to Europe, thereby establishing a closer link between the two markets. Reduced costs of transportation have strengthened the interconnection between the markets even further. Combined, these features could have a permanent influence on the prices of natural gas in both markets.

Fig. 3.3 plots the difference between forward and spot prices with the same delivery date from January 1999 through August 2010. From now on this difference is referred to as the *market forecast error*.

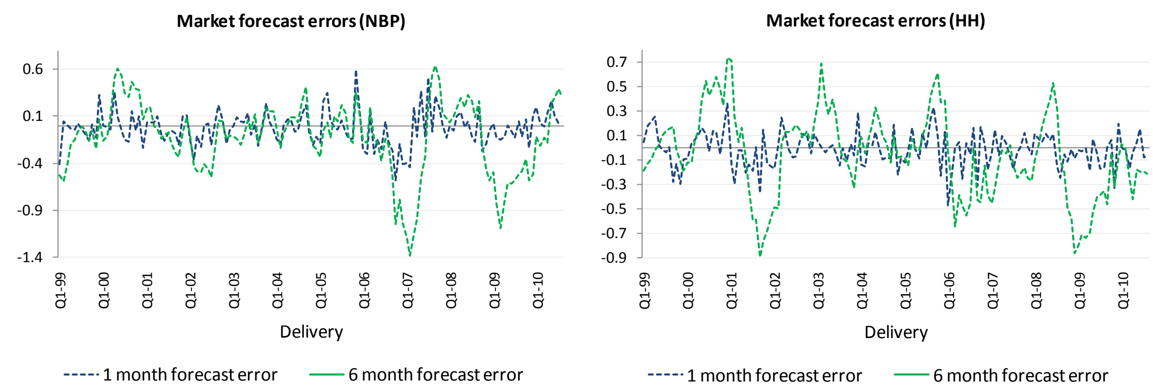

Fig. 3.3 Difference between 1- and 6-months ahead forward and spot price on log scale

As expected, the above figure shows that the 1-month ahead market forecast errors are closer to zero and less volatile than the 6-months ahead forecast errors. In addition, we find that the volatility of the forecast errors in the UK seem to have accelerated throughout the period. The same pattern is not evident in the US.

3.2 Unit-root test

Casual inspection of the time-series plots in Fig. 3.1, accompanied by a slowly decaying autocorrelation function (not reported here), call for formal unit-root testing[1]. Determining whether a series is trend-stationary rather than a random walk with drift by visual examination is difficult, if not to say impossible in some cases[2]. The seemingly innocuous difference between the two has profound consequences. For the former, shocks reflect only temporary departures from the trend, whereas for the latter, each shock will have a permanent effect on the mean.

Fig. 3.1 indicates that our time-series exhibit positive trends. Therefore our unit root tests are also carried out by including a drift, as well as a drift and a deterministic time trend. Moreover, we also test whether the forecast errors are stationary. Visual inspection of the forecast errors gives no indication of a trend and hence the tests do not include one. The Phillips-Perron test is a particularly relevant unit root test in the presence of moving-average error terms as it is robust

1 Unit root tests are used as a means of testing whether a series of data is stationary.
2 A formal definition of the term covariance-stationarity is provided in Appendix 1.

with respect to autocorrelation and heteroskedasticity (Modjtahedi and Movassagh, 2005). The augmented Dickey-Fuller test (ADF) handles autocorrelation by adding lags of the first difference of the dependent variable, but is unable to cope with a non-constant variance. Table 3.1 present the results from the unit-root tests on NBP data.

Unit-root tests for ${}_{t-k}F_t$				Unit-root tests for $({}_tF_t - {}_{t-k}F_t)$	
k	ADF p-value for t-statistics (No time trend)	ADF p-value for t-statistics (Drift)	ADF p-value for t-statistics (Time trend)	ADF p-value (No time trend)	Phillips-Perron p-value (No time trend)
0	0.840	0.014**	< 0.01***		
			Φ_3 = 5.93*		
1	0.801	0.013**	< 0.01***	< 0.01***	< 0.01***
			Φ_3 = 5.28		
2	0.961	< 0.01***	< 0.01***	< 0.01***	< 0.01***
			Φ_3 = 7.67**		
3	0.951	0.012**	< 0.01***	< 0.01***	< 0.01***
			Φ_3 = 7.28**		
4	0.807	< 0.01***	< 0.01***	< 0.01***	< 0.01***
			Φ_3 = 9.15***		
5	0.755	< 0.01***	< 0.01***	< 0.01***	0.014**
			Φ_3 = 7.91**		
6	0.706	< 0.01***	< 0.01***	< 0.01***	0.037**
			Φ_3 = 11.97***		
Critical values (Φ_3)					
1%			8.43		
5%			6.49		
10%			5.47		

The data used for testing are monthly spot and forward prices over the period from January 1999 to August 2010. One, two and three asterisks indicate rejection of the null hypothesis of a unit root at 10%, 5% and 1% significance levels, respectively. The number of lags of the first difference included in the ADF tests is based on the Schwartz-Bayesian Criterion. All prices are in natural logarithms. The Φ_3 test statistic tests the null hypothesis of a random walk with drift against the alternative that the data contain an intercept and/or a unit root and/or a deterministic time trend. We reject the null hypothesis for values of Φ_3 larger than the critical values. The lag truncations for the Phillips-Perron test for the forecast errors equal the order of the MA-process, k-1.

Table 3.1 Unit-root test for NBP

The ADF tests without a drift and deterministic time trend fail to reject the null hypothesis of non-stationarity for all lags k. However, adding a drift and a time trend dramatically alters this conclusion. We are now able to reject the null hypothesis of non-stationarity for all our series at the 1 percent significance level. Furthermore we find that both the trend and intercept terms are individually significant for all lags k (not reported here). To further investigate the issue of whether the series are trend-stationary or contain a unit root plus a drift term we apply the Φ_3 test statistic. We reject the null hypothesis of the latter model and conclude that our spot and forward prices are trend-stationary. This finding is consistent with the widely-held view that commodity prices should be mean-reverting, the "mean" being the real marginal cost of production. If this is indeed the case, we regard the observed trend as the real marginal cost related to the extraction and production of natural gas. An interesting suggestion by Modjtahedi and Movassagh (2005) is that the observed trend could also reflect Hotelling's theory of the evolution of the prices of exhaustible resources. However we will not dwell on this theory.

Table 3.2 present the results from our unit-root tests on HH data.

Unit-root tests for ${}_{t-k}F_t$				Unit-root tests for $({}_tF_t - {}_{t-k}F_t)$	
k	ADF p-value for t-statistics (No time trend)	ADF p-value for t-statistics (Drift)	ADF p-value for t-statistics (Time trend)	ADF p-value (No time trend)	Phillips-Perron p-value (No time trend)
0	0.771	0.016**	0.024**		
			$\Phi_3 = 3.05$		
1	0.849	< 0.01***	0.017**	< 0.01***	< 0.01***
			$\Phi_3 = 3.51$		
2	0.856	0.014**	0.017**	< 0.01***	< 0.01***
			$\Phi_3 = 3.34$		
3	0.837	0.030**	0.058*	< 0.01***	< 0.01***
			$\Phi_3 = 2.47$		
4	0.856	0.047**	0.093*	< 0.01***	< 0.01***
			$\Phi_3 = 2.07$		
5	0.909	0.064*	0.135	< 0.01***	< 0.01***
			$\Phi_3 = 1.79$		
6	0.995	0.058*	0.071*	< 0.01***	0.018**
			$\Phi_3 = 2.14$		
Critical values (Φ_3)					
1%			8.43		
5%			6.49		
10%			5.47		

The data used for testing are monthly spot and forward prices over the period from January 1999 to August 2010. One, two and three asterisks indicate rejection of the null hypothesis of a unit root at 10%, 5% and 1% significance levels, respectively. The number of lags of the first difference included in the ADF tests is based on the Schwartz-Bayesian Criterion. All prices are in natural logarithms. The Φ_3 test statistic tests the null hypothesis of a random walk with drift against the alternative that the data contain an intercept and/or a unit root and/or a deterministic time trend. We reject the null hypothesis for values of Φ_3 larger than the critical values. The lag truncations for the Phillips-Perron test for the forecast errors equal the order of the MA-process, k-1.

Table 3.1 Unit-root test for HH

The spot and forward price series from HH paint a somewhat different picture. Again, the initial ADF test fails to reject the null hypothesis of non-stationarity for all lags k. Following the above procedure, we add an intercept and a deterministic time trend to the test. For all lags k except for k = 5, we are now able to reject the null at the 10 percent significance level or less. Nevertheless we find that the time trend is not different from zero for any conventional significance level. Moreover, the Φ_3 test statistic fails to reject the null hypothesis of a random walk with drift. Accordingly, we now run the ADF test adding only a drift term. As shown in Table 3.2, we are now able to reject the null of a unit root also for lag 5.

These results are somewhat surprising to the authors. Corresponding results from previous research come to the opposite conclusion. However, our data span over a longer time-horizon than that of Haff et al. (2008). Furthermore, we cover a different period, although overlapping, than that used by Modjtahedi and Movassagh (2005). Our results are nonetheless consistent with the findings of Wei and Zhu (2006) in their study of the US market. As an obvious consequence of the above results, we find that all the forecast error series are stationary.

Our findings suggest that there has been a positive trend in natural gas prices, probably a result of the period from 2003 to 2007 being one of high economic growth. Moreover, price volatility appears to be time-varying, a feature often found in financial markets. Particularly, the emergence of unconventional gas has served to increase volatility in recent years, possibly imposing a structural break in both markets. We also observe a pronounced seasonality in the prices of natural gas. Not surprisingly, this feature is more evident in the UK where natural gas, to a larger extent, is used for heating. Our unit-root tests show that all the series are covariance-stationary. This implies that no data transformations are needed to perform the subsequent regressions. In particular, we find that the UK series are trend-stationary, i.e. exhibiting reversion to an upward trend.

4 The spot-forward parity in the natural gas market

A forward contract is an agreement to buy or sell an asset at a certain future time at a predetermined price (Hull, 2008). It is mostly traded in the OTC market, and usually between two financial institutions, or between a financial institution and client. The one selling the forward is said to hold a *short position,* while the buying counterparty is said to hold a *long position.* Forward contracts on storable assets in efficient markets are priced according to the no-arbitrage argument. This is easy to verify. Letting r be the riskless k-period interest rate observed at time t, S_t the current spot price, and ${}_tF_{t+k}$ the price at time t of a k-period forward contract, we must have that

$${}_tF_{t+k} = S_t e^{rk} \qquad (1)$$

If Eq. (1) does not hold, say, for instance, that the left-hand side is bigger than the right-hand side, we could short the forward contract and simultaneously borrow an amount S_t in the bank to buy the asset. At maturity, after paying off the loan, we would receive a risk free profit ${}_tF_{t+k} - S_t e^{rk} > 0$. As more and more arbitrageurs short the forward contract, the price will converge to the no-arbitrage price in Eq. (1).

The no-arbitrage argument that underpins Eq. (1) requires that the underlying asset can be stored at no benefit nor cost. This is typically not true for a wide range of commodities. Hence, when pricing forward contracts, a clear distinction must be made between investment assets and consumption assets. Investment assets are held for investment purposes (e.g. a stock index), while consumption assets are primarily held for consumption (e.g. natural gas). The fact that the owner of a consumption asset might be reluctant to sell his commodity in the spot market, and secure future access by purchasing a forward on the same asset, implies that Eq. (1) is not necessarily applicable to pricing consumption assets. Hull (2008) exemplifies this with an oil refinery that might not be willing to sell oil and buy forward contracts on oil, since a shortage could potentially shut down production. Another example would be the convenience of having an inventory to meet unexpected demand. Such direct benefits from holding the physical asset are often referred to as a *convenience yield*. On the other hand, consumption assets are often subject to significant *storage costs*. Therefore the holder of the asset requires a compensation for storing the asset from one period to another. These costs are typically related to rent of storage space, insurance and shrinkage.

The theory of storage was initially put forth by Kaldor (1939), Working (1949), Brennan (1958) and Telser (1958). It attempts to explain the difference between contemporaneous spot and futures prices in terms of interest foregone, storage costs and the convenience yield (Fama and French, 1987). Despite the presence of the latter two factors in consumption asset markets, it is, under the assumption that the market is sufficiently competitive, and that investors and speculators can short the underlying asset, possible to obtain a no-arbitrage price for the forward contract. In such an informationally efficient market any attempt to deviate from the equilibrium price, will quickly be eliminated by arbitrageurs exploiting the price differential. Assuming that both the convenience yield and the storage cost can be expressed at a constant rate, the arbitrage-free price of a forward contract at time t with delivery at time T, can be written as

$$ {}_{t}F_{t+k} = S_t e^{(R-C+M)k} \tag{2} $$

Taking the logarithm and rewriting Eq. (2) we obtain

$$ \ln(\, {}_{t}F_{t+k}) - \ln(\, S_t) = {}_{t}R_{t+k} + {}_{t}M_{t+k} - {}_{t}C_{t+k} \tag{3} $$

The left-hand side of Eq. (3) is the logarithmic difference between the current price of a forward contract with delivery k-months ahead and the contemporary spot price S_t. This forward-spot spread is the so-called *basis*. Moreover, in Eq.(3) ${}_{t}R_{t+k}$ is the k-month nominal interest rate at time t, ${}_{t}M_{t+k}$ is the marginal cost of storing the commodity until delivery is due, and ${}_{t}C_{t+k}$ is the marginal convenience yield accrued during the holding period. All the variables on the right-hand side in Eq. (3) are continuously compounded. From Eq. (3) we see that the basis is increasing with the interest rate ${}_{t}R_{t+k}$. This should come as no surprise, as the gain from selling the asset in the spot market and placing the proceeds in a bank account is increasing with the rate earned on the deposit. Therefore the holder of the commodity requires a larger compensation to store the asset when interest rates are higher. In addition to the interest foregone, the holder of the asset also faces storage costs ${}_{t}M_{t+k}$ which drive up the basis even further. On the other hand, the benefits provided by ownership of the asset, i.e. the convenience yield ${}_{t}C_{t+k}$, dampens the upward pressure on the basis.

The convenience yield and storage costs are both relatively hard to measure explicitly. However, they should be related to the level of inventory. In particular, the theory of storage suggests a negative relation between the convenience yield and inventories, whereas the cost of storage is increasing in the level of inventory. In effect, the "net" storage cost $\left({}_{t}M_{t+k} - {}_{t}C_{t+k}\right)$ will be an increasing function of the current inventory level. Intuitively, the higher the level of storage, the lower the value of marginal storage will be. Because of the fear of shutting down production, producers will never allow the inventory to reach zero. Thus, the convenience yield becomes very large as the inventory is depleted. Moreover, the convenience yield, by definition a benefit, cannot turn negative. When inventories are plenty, the market does not expect any shortages in the near future, and hence the convenience yield will tend towards zero. Therefore, the convenience yield is a highly convex function of inventories as recognised by several authors, e.g. Pindyck (1990). This relationship is depicted in Fig. 4.1 below.

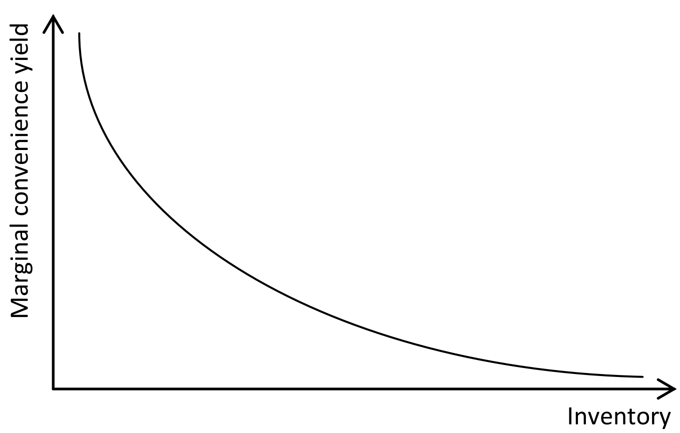

Fig. 4.1 The marginal convenience yield

In order to build storage facilities for natural gas, large investments are required. However, once such inventories are built, we assume, equivalent to Brennan (1958), that the marginal cost of an additional unit of storage is fairly constant. This is typically the case for commodities with continuous production (Gjølberg & Johnsen, 2004). Nevertheless, as soon as the inventory is full, producers of the commodity will face a runup in storage costs because the demand for a place to store the commodity greatly increases. This is because any excess supply not met by the demand of consumers will be lost in the sense that it must be handed away. For many goods, producers would, in such a situation, simply halt production until demand picks up or storage capacity is relieved. However, several economic factors are preventing gas producers to ease production. For instance, if production of natural gas from a well is halted, it may not be possible to restore the well's production due to reservoir and wellbore characteristics (Natural Gas Supply Association, 2010). In addition, natural gas is in many occasions a byproduct of oil production where extraction might be profitable despite very low gas prices. Consequently, the marginal storage cost per unit must equal the value of the lost proceeds from not being able to sell that unit, i.e. the prevailing spot price. Thus, there is a very high opportunity cost related to nearly full inventories. Therefore, when inventories are abundant, the net storage cost is highly positive, implying a large positive basis. On the other hand, when the aggregate inventory reaches a dangerously low level in relation to short-term expected demand, the value of holding an inventory greatly increases. As a result, the net storage cost drops and the basis should become negative[3]. This is illustrated in Fig. 4.2.

3 This will of course also depend on the cost of capital.

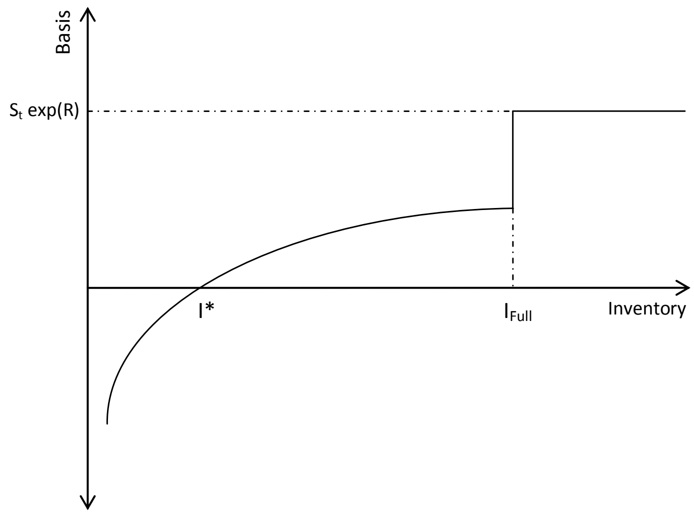

Fig. 4.2 The basis in relation to the level of inventory

As we can see from Fig. 4.2, the basis is increasing at a decreasing rate in the level of inventory. I* denotes the inventory level that equates the marginal convenience yield and the marginal cost of storage adjusted for the cost of capital. Moreover, when inventories are full, the basis reaches its maximum as the spot market, which balances real-time supply and demand, becomes a substitute for storage.

The above relationship between the convenience yield and inventories is easy and intuitive, but has led to some debate regarding the precise way in which inventory levels reveal information about the convenience yield (Volmer, 2009). Cartea and Williams (2007) entertain the idea of a dynamically updated optimal stocking policy and the facility owner's ability to smooth the marginal convenience yield over time. They stress that the convenience yield is influenced by deviations from average inventory levels rather than the actual levels which merely reflect seasonal changes in the stock. The rationale behind this train of thought is that such seasonal changes are already incorporated in the optimal stocking policy. While we are convinced that shocks in the level of storage will indeed impact the marginal convenience yield, we also believe that the absolute inventory level is important. We think that the value of holding the commodity will be larger in times of low inventories, regardless of shocks, because a completely drained inventory is likelier. Shocks will, however, further increase the value. We know that the total storage capacity is limited, especially in the UK where, until recently, inventories only covered a fortnight of average winter demand (Volmer, 2009). Therefore, even without shocks, there is a very real chance of dangerously low inventories, which according to theory would increase the convenience yield.

4.1 Empirical testing of the theory of storage

As proposed by Modjtahedi and Movassagh (2005), we begin by assuming a linear relationship between the net storage cost and the level of inventory:

$$\left({}_{t}M_{t+k} - {}_{t}C_{t+k} \right) = h_0 + h_1 I_t \quad (4)$$

By adding a coefficient to the interest rate we obtain the following estimable linear storage model

$$\ln({}_{t}F_{t+k}) - \ln(S_t) = h_0 + h_1 I_t + g\left({}_{t}R_{t+k} \right) + \varepsilon_t \quad (5)$$

If the theory of storage holds, we expect to find that the coefficient on inventory h_1 is greater than zero. Furthermore, as pointed out by Fama and French (1987), after controlling for the variation in the net storage cost, the k-period basis should vary one-for-one with the k-period interest rate. Thus we should find that the interest rate coefficient g equals unity. Also, if the model is fully specified, the intercept h_0 should be statistically insignificant.

In order to entertain the possibility of a concave relationship between the basis and the level of inventory, as illustrated in Fig. 4.2, we also estimate a concave storage model

$$\ln({}_{t}F_{t+k}) - \ln(S_t) = h_0 + h_1 \sqrt{I_t} + g\left({}_{t}R_{t+k} \right) + \varepsilon_t \quad (6)$$

From an econometric point of view, one could argue that a polynomial model would be preferable because it does not add any constraints with the respect to the curvature. However, when plotting the basis against the level of inventory, we find no pronounced turning point and hence we choose to estimate a model without one.

Finally, to address Cartea and Williams' (2007) view that the effect of storage on the basis is determined by the deviation from the expected seasonal storage level rather than the absolute level, we also propose a model using deseasonalised storage data. We deseasonalise our inventory data by *loess* as suggested by Cleveland et al. (1990)[4]. Loess is a nonparametric procedure that applies progressive smoothing and differencing to decompose data consisting of sums of curves. Denoting the actual inventory level for a particular month by I and the deseasonalised level by I^{+}, we compute a standardised measure of the "normal" inventory level by taking the logarithmic difference between the two. When substituting the actual storage levels in Eq. (5) by the normalised inventory series we obtain

$$\ln({}_{t}F_{t+k}) - \ln(S_t) = h_0 + h_1 \ln\left(I_t / I_t^{+}\right) + g\left({}_{t}R_{t+k}\right) + \varepsilon_t \qquad (7)$$

The theory of storage omits several variables possibly influencing the basis. Before continuing to the results, we briefly discuss two variables left to the error term ε_t. The first is the spot price volatility. Intuitively, demand for storage as a means of buffering fluctuations in production and consumption should depend positively on price volatility as higher uncertainty calls for a larger buffer. When price volatility is high, the value of insurance in the form of holding the underlying asset becomes greater. Thus, the basis is likely to be decreasing with the level of spot price volatility.

Another example of a variable potentially present in the error term is the price of crude oil. There are two explanations for why the price of crude oil may influence the basis. One straightforward explanation is that due to the interconnection in production between the two goods, demand for crude oil should affect natural gas prices. Empirical observations in the UK market confirm that the price of natural gas reacts to changes in the crude oil price with a lag of approximately 9 months (Volmer, 2009). Another possible explanation is that if prices on longer-term oil contracts drop, say for instance that a new cost-saving technology becomes available in the future, oil forwards will be relatively cheaper than natural gas forwards, reflecting lower expected production costs. Now, consumers with the ability to switch to oil will exploit the forward price

4 We used the "stl" function in R.

differential, resulting in decreased demand for natural gas forwards. The gas spot market on the other hand, is not likely to be influenced by such technological advances in the future. Hence the basis will decrease. Historically, several industries have switched between natural gas and residual fuel oil, using whichever energy source is available at the lowest price. It should be noted however, that over the last decade the number of facilities able to switch quickly between natural gas and refined petroleum products has declined (Brown & Yucel, 2007).

4.2 Results

We begin this section by briefly investigating the estimated basis for both markets. We restrict the analysis to the basis for 1-month ahead forward prices shown in Fig. 4.3.

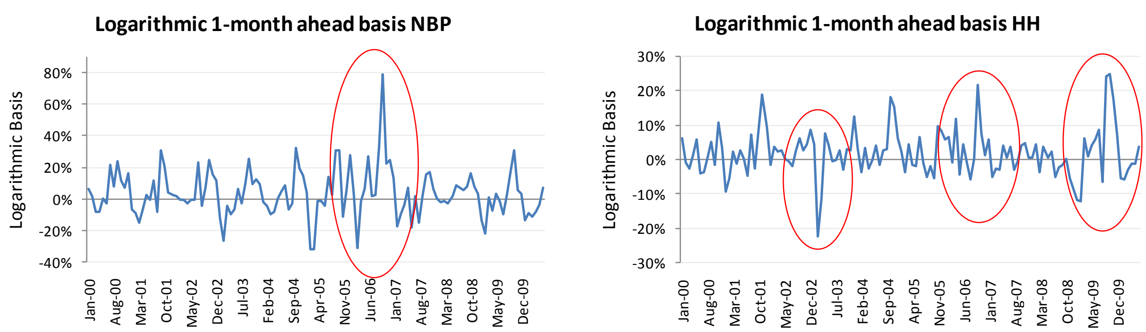

Fig. 4.3 One-month ahead logarithmic basis from January 2000 through June 2010

As emphasised by the red circles, there appears to have been some influential events the last decade resulting in large spikes in the basis. We begin by considering the historical basis in the UK. The cold winter of 2006 resulted in surging spot prices. In addition, interruptions to gas supplies from Russia transiting Ukraine further increased the pressure on UK spot prices since the European markets are highly interconnected. As a consequence, the basis deteriorated. When inventory levels are low, the convenience yield rises fast because inventories are used to meet the increase in demand. Therefore the inventory cannot properly act as a buffer, and a shock has a large impact on the current spot price. The change in forward prices is smaller because they incorporate the market's predictions of future supply responses to the increased demand.

Furthermore, we observe a large positive spike in the basis in the late autumn of 2006. We believe that seasonally mild weather conditions and an abundance of storage, combined with the fact that the Langeled pipeline from Norway started operating, reduced spot prices significantly below long-term contract prices (ICIS Heren, 2007).

We now proceed with an equivalent review of the US basis. In January 2003 the basis plunged. The reason for the sudden drop was the cold US winter season of 2002-2003. Hence, prices of natural gas skyrocketed due to a vast increase in demand. As a result, monthly volatility rose above 100 percent. Moreover, due to producers' limited ability to increase supply in the short run, storage levels fell to 44 percent below the 5-year average (EIA, 2007). All the above factors explain a convenience yield runup resulting in a large drop in the basis.

In late 2006, US inventory levels were significantly higher than the previous year and five-year averages. Simultaneously, forward prices rallied on expectations of a cold winter, even though the gas storage level was approaching an all-time high, and most analysts anticipated a downward pressure on the price (Pirog, 2006). Consequently, the basis in the US skyrocketed. In fact, these abnormal market conditions led to the downfall of the major hedge fund Amaranth in September 2006, who suffered huge losses on their long winter/short summer trading strategy. In particular, their poor bet on the March and April 2007 forward spread led to an immediate liquidity crisis due to higher margin calls to maintain other positions (Chincarini, 2007).

Finally, we find a large positive spike in the US basis in late 2009. Following the financial crisis, natural gas consumption contracted in most major global markets, falling by 4.7 percent in Europe and 1.8 percent in the US in 2009 (Economist Intelligence Unit, 2010). Spot prices dropped to their lowest level in 7 years. Contributing to the decline in prices were the reduced heating demand as well as higher than usual production. Moreover, at the end of November 2009, working natural gas in storage hit its highest monthly level on record and additions to storage continued past the official close of the injection season (EIA, 2010c). All the above factors contributed to an abnormally low convenience yield throughout the autumn of 2009 resulting in a surging basis.

In addition to the aforementioned influential events, a noteworthy feature stands out in the above graphs, namely that the mean basis appears to be positive in both markets. This would imply that on average, storage costs, adjusted for the cost of capital, exceed the value of the benefits of holding an inventory. In fact, the 1-month average basis amounts to approximately 3.6 percent (s.e. 0.013) and 1.8 percent (s.e. 0.006) in the UK and US, respectively.

We also observe that the UK basis reaches approximately 25 percent or more several times throughout the sample period. Due to the relatively high frequency of negative forward-spot spreads in our sample it might be objected that the positive average could be a result of outliers. Nevertheless, we find the sample median basis to be roughly 2 percent in the UK and 1 percent in the US. We know that price spikes occur more frequently in the UK. This feature should result in a relatively higher convenience yield in the UK, driving down the basis. Thus, our finding that the average UK basis is larger than its US counterpart is somewhat surprising. One explanation could be the relatively higher concentration of suppliers in the UK seeking economic rent by keeping forward prices "artificially" high. Moreover, there could be a lack of arbitrageurs exploiting the forward-spot price differential. Lastly, healthy scepticism is warranted as the findings could be influenced by the aforementioned issue regarding measurement errors.

Before moving on to the results from our regression models, we investigate how the basis varies throughout the year. Fig 4.4 shows the average monthly logarithmic basis and average monthly storage levels from January 2000 till June 2010.

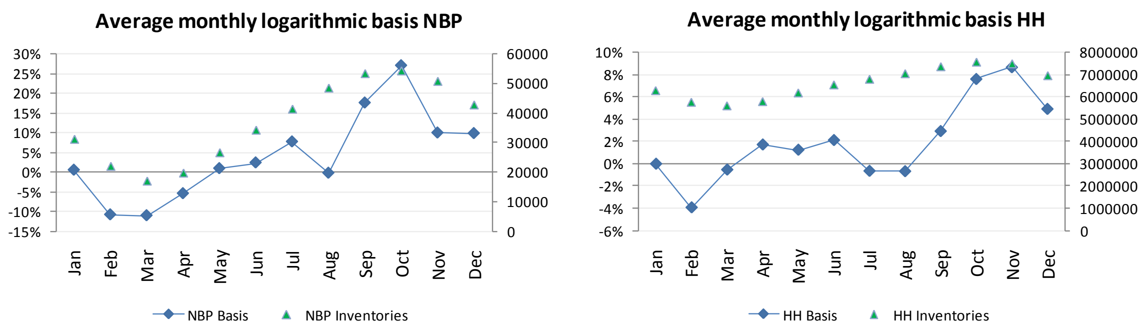

Fig. 4.4 One-month ahead average monthly logarithmic basis and average monthly inventories from January 2000 through June 2010. The left y-axis shows the basis and the right y-axis shows the inventory level

Fig. 4.4 shows that the basis is negative during periods of peak demand and increases as inventories are being filled for both markets. For the UK, the basis reaches its minimum level in March when inventories are scarce and peaks in October when inventories are at their highest level. Although less pronounced, the US basis seems to follow a similar pattern. Contrary to our expectations, for both markets, the basis declines in the late summer when a shortage of stock is unlikely. When replacing the sample average with the median, we come to the same conclusion, and hence this finding is not a result of outliers.

We now move on to analysing our regression models for NBP. We estimate the models for lag $k = 1$, 3 and 6. However, we believe that for longer-term contracts, the basis is mainly determined by the seasonality in the price levels rather than the inventory several months before delivery. For instance, consider estimating the net storage cost on a 6-months ahead forward contract in June. We know that prices are on average lower in the summer than in the winter. Therefore, the basis, as given by Eq. (3), will be very high, not because the net storage cost is particularly large, but rather because prices vary systematically throughout the year. We argue that the above reasoning is not sufficiently emphasised in the existing literature. Obviously, we therefore attach less importance to the results for lag k = 6. The results for NBP are reported in Table 4.1.

	Intercept		Storage Level		Root Storage Level		Interest Rate		
k	Estimate	P-value	Estimate	P-value	Estimate	P-value	Estimate	P-value	R^2
$\ln({}_tF_{t+k}) - \ln(S_t) = h_0 + h_1 I_t + g({}_tR_{t+k}) + \varepsilon_t$									
1	-0.21	< 0.01***	5.96E-06	< 0.01***			12.04	0.075*	0.34
3	-0.46	< 0.01***	1.58E-05	< 0.01***			0.21	0.956	0.54
6	-0.17	0.134	9.54E-06	< 0.01***			0.20	0.959	0.11
$\ln({}_tF_{t+k}) - \ln(S_t) = h_0 + h_1\sqrt{I_t} + g({}_tR_{t+k}) + \varepsilon_t$									
1	-0.40	< 0.01***			2.19E-03	< 0.01***	12.05	0.074*	0.34
3	-0.95	< 0.01***			5.72E-03	< 0.01***	0.27	0.945	0.53
6	-0.49	< 0.01***			3.60E-03	< 0.01***	0.19	0.960	0.11
$\ln({}_tF_{t+k}) - \ln(S_t) = h_0 + h_1 \ln(I_t / I_t^+) + g({}_tR_{t+k}) + \varepsilon_t$									
1	0.03	0.153	0.22	< 0.01***			11.22	0.086*	0.39
3	0.17	< 0.01***	0.53	< 0.01***			0.05	0.990	0.52
6	0.21	< 0.01***	0.32	< 0.01***			0.20	0.957	0.10

The sample period extends from January 2000 through June 2010. One, two and three asterisks indicate rejection of the null hypothesis on the 10%, 5% and 1% level, respectively.

Table 4.1 NBP regression models

Let us begin by investigating the results from the linear storage model. Firstly, we observe that the storage level is highly significant for all lags. Consistent with theory, the estimates for the coefficient on inventory, h_1, are all positive. Hence, the basis is increasing with the net storage costs, albeit the magnitudes are very small. The latter is due to the lack of scaling of the inventory data. Furthermore, because inventories may be replenished in the future, current inventory levels should have a stronger effect on the shorter-term forward prices. Consequently, the t-ratios should be decreasing with time to maturity. Although not provided in the table, we find that the t-ratio is highest for lag k = 3 and lowest for lag k = 6. One possible explanation could be that 3 months is not enough time to resupply a severely depleted inventory, whereas 6 months is. Therefore, the 3-months ahead forward price could be as sensitive to the level of storage as the one-month ahead contract. Secondly, the interest rate is only significant for the first lag, suggesting that, contrary to theory, the cost of capital does not affect the basis for longer-term contracts. Nevertheless, although not statistically significant, the estimates for the remaining lags are not terribly far from unity. Moreover they are all positive. In fact, testing the hypothesis of the interest rate coefficient g being equal to unity shows that the estimates are not significantly different from one. Quite surprisingly, the same holds for k = 1 on the 10 percent significance level. Note that these tests are not given in the above table. Finally, the intercept h_0 appears to be significantly different from zero. Our results are generally consistent with the findings of Haff et al. (2008) for the UK market.

Moving on to the concave storage model, the estimates of the interest rate are equivalent to those of the linear storage model. Moreover, we note that the intercept h_0 is still significantly different zero. Again, we find that the inventory is highly significant in determining the basis. This indicates that a concave relationship between inventories and the basis could be appropriate.

Finally, we consider the results from the deseasonalised storage model. Our first discovery, although not very surprising, is that using a standardised inventory series gives estimates that are more sensible with respect to the magnitude. Furthermore, since this model is a log-log regression with respect to the inventory variable, the interpretation of its coefficient is somewhat different than the other models. Now, a 1 percent increase in the normalised inventory level yields, ceteris paribus, an increase in the basis of 0.2 percent for lag k = 1. Recall that the theory of storage posits that the intercept term h_0 should be insignificant, the coefficient on inventory h_1 positive and the interest rate coefficient g not significantly different from unity. For most lags, the deseasonalised model does not appear to yield more appropriate results than the previous models. However, for k = 1, the deseasonalised model gives the best overall fit according to theory. Fig. 4.5 attempts to graphically illustrate the relationship between the basis, the inventory level and the interest rate using the latter model.

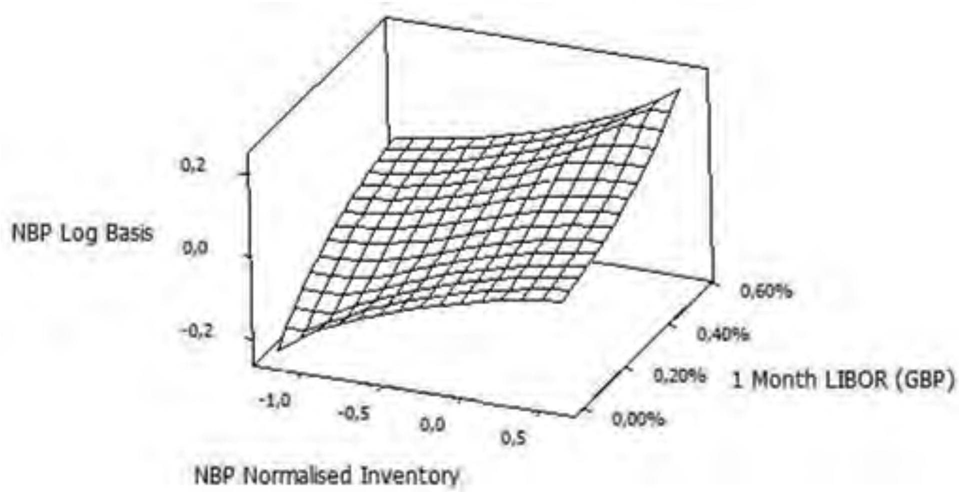

Fig. 4.5 Graphical presentation of the theory of storage on NBP data

Let us define the variable B as the logarithmic basis and NI as the normalised inventory. Fig. 4.5 confirms that $\partial \mathrm{B}/\partial \mathrm{r} > 0$, i.e. the basis is increasing in the cost of capital r. We also observe that $\partial \mathrm{B}/\partial \mathrm{NI} > 0$ and $\partial^2 \mathrm{B}/\partial \mathrm{NI}^2 < 0$ (concavity). Therefore, the curvature of the basis with respect to inventories is concave regardless of whether we control for a trend and seasonality. Consequently, we must have that $|\partial \mathrm{S}/\partial \mathrm{NI}| > |\partial \mathrm{F}/\partial \mathrm{NI}|$. This supports our suggestion that spot markets work as a substitute for storage.

Having analysed the results from the UK, we now turn to the equivalent analysis for the US market. The results from running the above regressions on our HH data are shown in Table 4.2 on the next page. Note that for the first two models the estimated storage coefficients are not directly comparable to those from the UK since the US storage figures are measured in million cubic feet rather than million cubic metres.

k	Intercept		Storage Level		Root Storage Level		Interest Rate		R^2
	Estimate	P-value	Estimate	P-value	Estimate	P-value	Estimate	P-value	
$\ln({}_tF_{t+k}) - \ln(S_t) = h_0 + h_1 I_t + g({}_tR_{t+k}) + \varepsilon_t$									
1	-0.22	< 0.01***	3.80E-08	< 0.01***			-3.24	0.450	0.18
3	-0.56	< 0.01***	1.00E-07	< 0.01***			-3.12	0.153	0.41
6	-0.40	< 0.01***	9.14E-08	< 0.01***			-4.85	< 0.01***	0.22
$\ln({}_tF_{t+k}) - \ln(S_t) = h_0 + h_1 \sqrt{I_t} + g({}_tR_{t+k}) + \varepsilon_t$									
1	-0.46	< 0.01***			1.91E-04	< 0.01***	-3.30	0.443	0.17
3	-1.19	< 0.01***			5.05E-04	< 0.01***	-3.20	0.147	0.40
6	-0.99	< 0.01***			4.65E-04	< 0.01***	-4.87	< 0.01***	0.22
$\ln({}_tF_{t+k}) - \ln(S_t) = h_0 + h_1 \ln(I_t / I_t^+) + g({}_tR_{t+k}) + \varepsilon_t$									
1	0.05	< 0.01***	0.24	< 0.01***			-7.90	0.070*	0.15
3	0.15	< 0.01***	0.58	< 0.01***			-7.08	< 0.01***	0.29
6	0.24	< 0.01***	0.32	0.026**			-6.43	< 0.01***	0.11

The sample period extends from January 2000 through June 2010. One, two and three asterisks indicate rejection of the null hypothesis on the 10%, 5% and 1% level, respectively.

Table 4.2 HH regression models

We will restrict the following paragraph to similarities and contradictions between the two markets. Equivalent to our results for NBP, we find the storage level to be highly significant for all the estimated models. Moreover, the intercepts are significantly different from zero. Surprisingly, in stark contrast to what theory dictates, all the estimated interest rate coefficients are negative and generally insignificant for the first two models. One possible explanation to the odd interest rate coefficients could be the recent financial turmoil resulting in non-functioning money markets. Thus, the extent to which LIBOR rates reflect the actual cost of capital after 2008 is questionable. In contrast to what we observed for NBP, we now also find that the interest rate is significantly different from unity for several lags for all three models (not reported here). In fact, all the three models seem to suffer from some degree of functional form misspecification.

Our findings suggest that there have been several historical events resulting in abnormal values of the basis. These are typically a result of extreme weather conditions and interruptions to supply. Moreover, the basis has been positive on average in the period analysed, reflecting a positive cost-of-carry. However, in both markets a negative basis occurs rather frequently, indicating a relatively high convenience yield. Surprisingly, we find the average basis to be higher in the UK than in the US, possibly a result of a larger concentration of suppliers. The theory of storage suggests that the cost of capital should serve to increase the forward-spot spread. However, we find that it is not particularly relevant in predicting the basis. The level of storage is, on the other hand, very important. Indeed, it seems that a concave relationship between the basis and the level of inventory is appropriate. Finally, we confirm that not only shocks to inventories, as posited by Cartea and Williams (2007), but also absolute storage levels are important determinants of the basis.

Contrary to our beliefs, we find that the monthly average basis tends to drop mid-summer when scarcity of gas should be unlikely. In addition, the UK basis appears to be larger than any reasonable cost of carry would predict. One explanation for this is that agents in the natural gas market do not behave rationally. We have mentioned that the arbitrage-free forward price given by Eq. (2) only holds under the assumption that the market is sufficiently competitive. Hence, our finding may be a result of market participants not exploiting potential mispricing. Alternatively, we may argue that our theory does not hold because only a few agents are in fact able to take advantage of the arbitrage opportunities available to them. Finally, forward prices could exceed spot prices as a result of the interaction between buyers and suppliers in terms of hedging pressure and the role of speculators in providing insurance to these market players. Therefore, we now turn our focus to forward contracts' ability to predict future prevailing spot prices.

5 The unbiasedness hypothesis

The theory of storage attempts to explain the difference between spot and forward prices quoted on the same date, i.e. the contemporaneous basis. We now consider the difference between forward prices and the prevailing spot prices by investigating the well-known, yet heavily debated, unbiasedness hypothesis. The theory of unbiasedness postulates that forward prices are unbiased predictors of future spot prices (Haff et al., 2008). This raises the question about how forward prices are determined in the market. If the forward prices are set to match the expected future spot price, then on average, forward prices are unbiased estimates of future spot prices. Consequently, the payoff from a forward contract is a result of unexpected deviations in the spot market. These are by definition unpredictable, and hence will be zero on average.

As recognised by several authors, there are however arguments for either an upward or a downward bias distorting this relationship. Under the assumption of rational expectations such a bias can be interpreted as a *risk premium*. Here, we denote a situation where the forward price is below the expected future spot price as a state of normal backwardation. On the other hand, we refer to a situation where the forward is above the expected future spot price as contango (Hull, 2008). However, it should be noted that these terms are sometimes used in relation to the basis. Fig. 5.1 is meant to illustrate a situation of normal backwardation.

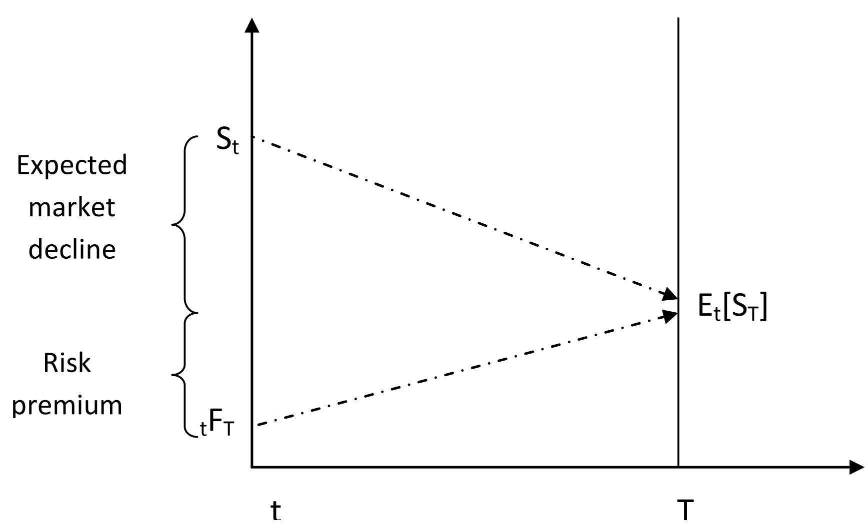

Fig. 5.1 Normal backwardation in the forward market

Here we display a situation where the forward price is initially set below the expected spot price. Thus, the long position is rewarded a risk premium in the sense that the predetermined price is lower than what a buyer is expected to pay in the future spot market. Obviously, the opposite is true when markets are in contango. The existence of a bias can have important practical implications. A large body of economic agents, not necessarily players in the forward market, base their production and consumption decisions on forward prices being estimates of future spot prices. Furthermore, a bias, in either direction, will greatly influence the costs and benefits of hedging.

A well-known hypothesis concerning the bias in futures prices as predictors of future spot prices is the traditional risk-management view. Keynes (1930) is perhaps its most famous advocate, stating that the market works as an intermediator facilitating the transfer of price risk from hedgers to those who are willing to bear it for compensation. However, Keynes only argued in favour of normal backwardation. As a response to the lack of empirical support for the forward market being backwardated at all times, a hedging-pressure view emerged. Intuitively, producing agents should hold a short position in forwards, whereas consuming players should hold a long position, both attempting to avoid future price uncertainty. In a balanced market, i.e. when the hedging demand of the two parties match, the unbiasedness hypothesis should hold. In an unbalanced market however, the party least compelled to hedge will receive a compensation in the form of a risk premium.

The hedging-pressure hypothesis implies that the market structure is highly relevant in predicting the size and direction of a potential risk premium. The UK gas market is characterised by a few large producers seeking exposure towards spot market fluctuations. This is, in part, due to the fact that their shareholders encourage such exposure in their presumably otherwise well-diversified portfolio. In addition, the produced volumes are far greater than what is being traded in the forward market, limiting the hedging possibilities available. The supply side has seen further consentration in recent years through consolidation, resulting in increased market power among the suppliers (UK Parliament, 2008). Therefore, it could be suggested that UK producers keep forward prices higher than the price that would prevail in a more competitive market. On the demand side, numerous purchasing agents are unable to pass on their potentially high energy costs related to spot price exposure to the end-user (Haff et al., 2008). In effect there is a great demand for forward contracts. However, the fact that forward prices are too high for many buyers to participate in this market somewhat mitigates the demand for these contracts. In contrast to a more competitive market where we would expect this to depress prices, we believe that the supply side market power is sufficiently large to maintain forward prices above the expected spot price (contango). The hedging demand from buyers able to partake in the forward market further adds to this expectation.

In the US market on the other hand, buyers are mostly regulated local distribution companies (LDCs) able to pass on any cost increases to the end-user, whereas production is spread on several small companies (Haff et al., 2008). The LDCs are regulated by state legislation and thus unable to exploit any market power. Their main objectives are to ensure reliable delivery of gas and to minimise costs for end-users. Hence, their participation in the forward market is limited to times when they cannot solely rely on the spot market to ensure availability of gas. Furthermore, the LDCs are only allowed to partake in the shorter-term bidweek forward market (Borenstein et al., 2007)[5]. As a result, in contrast to the UK, we do not expect any pronounced hedging pressure from the largest purchasers of natural gas. Finally, the US market is more transparent than the UK market in the sense that more information concerning prices is made publicly available (The National Petroleum Council, 2007). Consequently, it is more likely that arbitrageurs will exploit any systematic mispricing. We therefore believe that rejecting the unbiasedness hypothesis is unlikelier in the US compared to the UK.

5.1 Statistical issues

There are many statistical issues present when testing the unbiasedness hypothesis that may severely affect the outcome. Firstly, if futures prices are indeed unbiased predictors of future spot prices, the expected value of the forecast errors is zero ($E\left[S_t - {}_{t-k}F_t\right] = 0$). Hence, any deviation will translate into a risk premium. Note that for lag $\mathrm{k} > 1$, time to maturity exceeds the sampling interval, and consequently the forecast error follows a moving-average process of order $\mathrm{k} - 1$ (see Appendix 2 for a general proof of this statement). We verify this by examining the

5 "Bidweek" refers to the last five trading days of each month and the contracts traded there have delivery in the upcoming month.

empirical ACF and PACF plots of the market forecast errors, the former showing significant spikes up to lag $k-1$ and the latter displaying an oscillatory decaying pattern[6]. Indeed, this was the case for both markets.

Secondly, an interesting observation by Williams and Wright (1991) on previous testing of unbiasedness is that the very nature of storable commodities imposes a statistical difficulty. Wooldridge (2009) defines the time series analogue of the zero conditional mean assumption as follows

$$E(u_t|\boldsymbol{X}) = 0, t = 1,2, \dots , n. \tag{8}$$

where u_t is the error term and $\mathbf{X}$ denotes the explanatory variables for all time periods t. This assumption requires more than just contemporaneous exogeneity; u_t must be uncorrelated with both past and future values of the explanatory variables (and obviously present values). To return to the problem of storable goods, consider a regression of the prevailing spot price on lagged forward prices. Suppose now that current production of the good is unusually high. This abundance of supply will depress the current spot price and the error term. The possibility of storage enables producers to store some of their production and rather sell it in the forward market. Now the forward prices will decline as well. Consequently the error term in our initial regression will be correlated with future values of the explanatory variable, thus violating the above assumption and invalidating inference. We confirm that this is the case for both markets by investigating empirical cross-correlations between the error terms and future forward prices. As expected, these are all positive.

We will perform our tests of the unbiasedness hypothesis on log-scale. There are several reasons for doing so. To begin with, it is common practice in the existing literature (e.g. Modjtahedi and Movassagh (2005), Cartea and Williams (2007) and Haff et al. (2008)). Moreover, we wish to compare our results to those of other authors. Finally, a log transformation palliates the effect of the price spikes often found in natural gas prices. However, because the logarithm is a concave

6 ACF and PACF are short-hand notations for the autocorrelation function and the partial autocorrelation function, respectively.

function, our results cannot be directly translated into linear space due to Jensen's inequality (Haff et al., 2008).

5.2 Forward prices as unbiased predictors

Obviously expected spot prices cannot be directly observed. The approach in numerous studies is therefore to compare the forward price at time t (with delivery at time t + k) to the subsequent spot price. The challenge with this approach is to distinguish between differences due to non-rational expectations or inefficiencies in the pricing of forward contracts and actual risk premia (Gjølberg & Brattested, 2010). Examining the unbiasedness hypothesis is usually carried out by testing the null hypothesis that the mean forecast error is zero, where a rejection translates into a risk premium. Nevertheless, the null's assumption of a zero risk premium is problematic and several studies have disputed the unbiasedness hypothesis. In addition to the possibility of non-rational behaviour, the forward price may also be a biased forecast of prevailing spot prices due to changes in the market's perception of risk, including changes in the net hedging demand (Gjølberg & Brattested, 2010).

We now proceed to describing our first test of unbiasedness. From the section on statistical issues, we know that our forecast errors follow a moving-average process of order $\mathrm{k}-1$ when time to maturity exceeds the sampling interval. To overcome the overlapping observation problem, we estimate the following equation

$$\left(S_t - {}_{t-k}F_t\right) = \alpha + \varepsilon_t + \sum_{i=1}^{k-1} \beta_i \varepsilon_{t-i} \tag{9}$$

where α is the intercept and ε_t is a white noise process. Formally, a sequence $\{\ {}_{\mathrm{t}}\}_{\mathrm{t=0}}$ is a white noise process if each value in the sequence has a mean of zero, a constant variance, and is uncorrelated with all other realisations (Enders, 2009). Thus the expected forecast error has an unconditional mean equal to   regardless of the presence of moving-average errors. However, due to the possibility of a seasonal component in the forecast error variance, we could have unconditional heteroskedastic errors. For example, the variability in gas prices typically increases in the winter when inventories are scarce. Thus, it is likely that the forecast errors will display a similar pattern. Whenever the standard errors are heteroskedastic, inference will be invalid. By estimating the above equation as an appropriate moving-average process we account for the autocorrelation in the error term. Hence, having rendered the issue of autocorrelation obsolete, we may then solely direct our focus towards the unconditional heteroskedasticity problem. We will therefore estimate Eq. (9) by using heteroskedasticity-robust standard errors[7]. As usual, we start out by analysing the results from NBP given in Table 5.1 below.

		Intercept			Median
Market forecast error	NOB	Mean error	Standard error	P-value	error
Spot price forecasts					
$S_{t} - {}_{t-1}F_{t}$	128	-0.033	0.0161	0.039**	-0.043
$S_{t} - {}_{t-2}F_{t}$	128	-0.077	0.0271	< 0.01***	-0.047
$S_{t} - {}_{t-3}F_{t}$	128	-0.107	0.0406	< 0.01***	-0.074
$S_{t} - {}_{t-4}F_{t}$	128	-0.123	0.0482	< 0.01***	-0.085
$S_{t} - {}_{t-5}F_{t}$	128	-0.134	0.0585	0.022**	-0.067
$S_{t} - {}_{t-6}F_{t}$	128	-0.139	0.0672	0.039**	-0.087
1-month-ahead price forecasts					
${}_{t-1}F_{t} - {}_{t-2}F_{t}$	128	-0.042	0.0130	< 0.01***	-0.037
${}_{t-1}F_{t} - {}_{t-3}F_{t}$	128	-0.070	0.0231	< 0.01***	-0.051
${}_{t-1}F_{t} - {}_{t-4}F_{t}$	128	-0.086	0.0356	0.016**	-0.048
${}_{t-1}F_{t} - {}_{t-5}F_{t}$	128	-0.094	0.0435	0.031**	-0.046
${}_{t-1}F_{t} - {}_{t-6}F_{t}$	128	-0.099	0.0538	0.066*	-0.043

The sample period extends from January 1999 through August 2009. One, two and three asterisks indicate rejection of the null hypothesis of zero unconditional means at 10%, 5% and 1% significance levels, respectively. The reported standard errors are made robust to heteroskedasticity by applying Stata's "robust" command.

Table 5.1 NBP market forecast errors

From Table 5.1, we observe that the standard errors increase with the lag length. The mean market forecast errors follow a similar pattern. This is, of course, to be expected since forecasts should improve as relevant information arrives. The closer we are to maturity, the more information is available. Interestingly, the market forecast errors are all negative and statistically

7 We used the "robust" command in Stata.

significant – a clear sign of contango, also on the original scale. Therefore we reject the null hypothesis of unbiasedness, in favour of the alternative, namely the existence of a risk premium. Since the prevailing spot prices are consistently below the price "predicted" by the forward contracts, producers of natural gas seem to be at the receiving end. This does not come as a surprise in a market where producers have market power and there is a vast demand for forward contracts. Because forward prices are relatively expensive, anyone interested in the physical commodity will on average purchase gas at a lower price by solely relying on the spot market. These results are consistent with the findings of previous authors such as Haff et al. (2008) and Cartea and Williams (2007). We arrive at the same conclusion for forecasts of the 1-month ahead forward price, strengthening the evidence of contango. We observe that the forecast errors appear to be larger for predictions of the spot price than for the 1-month ahead price. The same seems to hold for the standard errors. This is probably a result of the spot market being more volatile than the forward market.

We may suspect that the above conclusion could be due to outliers. When considering a time series with several outliers the median error forecasts could be more representative of "typical" forecast errors. For the spot forecast errors in Table 5.1, the medians seem to follow the same pattern as the means. However, as shown explicitly in Fig. 5.2, they are generally smaller than the means, which could mitigate our rejection of the null hypothesis.

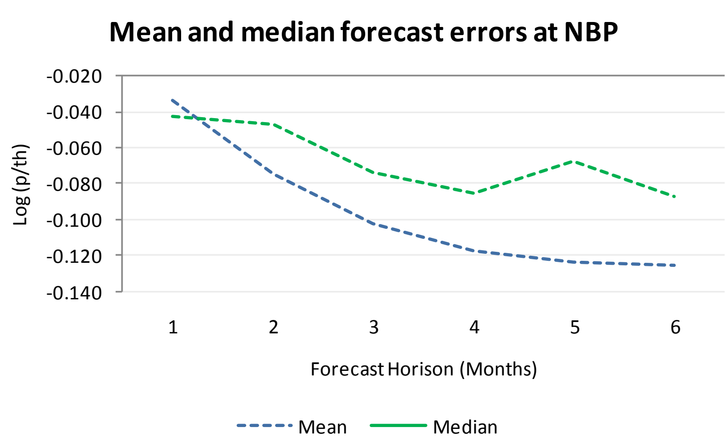

Fig. 5.2 NBP mean and median forecast errors

Having analysed the results from the UK, we now turn to the corresponding results from the US, which are found in Table 5.2 below.

Market forecast error	NOB	Intercept: Mean error	Intercept: Standard error	Intercept: P-value	Median error
Spot price forecasts					
$S_t - {}_{t-1}F_t$	128	-0.013	0.0129	0.302	-0.012
$S_t - {}_{t-2}F_t$	128	-0.034	0.0252	0.184	-0.028
$S_t - {}_{t-3}F_t$	128	-0.050	0.0383	0.193	-0.032
$S_t - {}_{t-4}F_t$	128	-0.054	0.0439	0.215	-0.051
$S_t - {}_{t-5}F_t$	128	-0.060	0.0598	0.320	-0.041
$S_t - {}_{t-6}F_t$	128	-0.056	0.0612	0.362	-0.023
1-month-ahead price forecasts					
${}_{t-1}F_t - {}_{t-2}F_t$	128	-0.020	0.0134	0.141	-0.023
${}_{t-1}F_t - {}_{t-3}F_t$	128	-0.034	0.0252	0.177	-0.036
${}_{t-1}F_t - {}_{t-4}F_t$	128	-0.039	0.0355	0.268	-0.017
${}_{t-1}F_t - {}_{t-5}F_t$	128	-0.039	0.0440	0.372	-0.022
${}_{t-1}F_t - {}_{t-6}F_t$	128	-0.039	0.0585	0.505	-0.004

The sample period extends from January 1999 through August 2009. One, two and three asterisks indicate rejection of the null hypothesis of zero unconditional means at 10%, 5% and 1% significance levels, respectively. The reported standard errors are made robust to heteroskedasticity by applying Stata's "robust" command.

Table 5.2 HH market forecast errors

Again we observe that both the standard errors and the mean forecast errors increase with the lag length. Furthermore, equivalent to our findings from NBP, the mean forecast errors are all negative. The standard errors are roughly the same for both markets. However, the forecast errors are much smaller, in fact only about half the size. As a result, none of the HH forecast errors are significantly different from zero. Thus, we fail to reject the unbiasedness hypothesis. This is indeed what we expected due to the US market being more efficient in the sense that arbitrageurs to a larger extent exploit mispricing. In addition, we believe that the US market is more balanced with respect to hedging pressure.

Our results are largely consistent with earlier research. Modjtahedi and Movassagh (2005) also fail to reject the null, withal there is a dramatic difference. They find that the mean forecast errors are positive. It should however be pointed out that Modjtahedi and Movassagh considered market data from 1993 through 2004, while we consider more recent data. This could indicate that the US market has undergone significant changes during the last decade. For instance, the number of participants in the natural gas forward market has increased dramatically following the entrance of index speculators around 2002 (Wray, 2008). They are typically pension funds, university endowments, life insurance companies, sovereign wealth funds, and banks. Most importantly, index speculators only take long positions. They turned their attention to commodity markets after discovering that the performance of commodities was not correlated with the return on equities and thus could be used as means of diversification. The volume held by these speculators has surged. Indeed, the value of open interest was approximately 23.6 billion USD in 2002, whereas the value had increased to roughly 87.3 billion USD in 2008 (Wray, 2008)[8]. As a result, the demand for forward contracts has risen quite significantly during the last decade. Another explanation as to why our results deviate from those of Modjtahedi and Movassagh could be that the US price series seem to exhibit a structural break following the emergence of shale gas. As the cost of extracting shale gas fell, the supply curve shifted to the right, resulting in greater quantities available at lower prices. It is however, unclear whether the introduction of shale gas has affected spot and forward prices differently, i.e. if it has influenced the forecast errors.

Yet again, the median spot forecast errors seem to be increasing with time to maturity. Nevertheless, another difference between the two markets is noteworthy. As opposed to our results from the UK, where the medians were generally smaller than the means, the medians are now largely equal to the means. Thus, on the contrary, our conclusion that we fail to reject the unbiasedness hypothesis does not risk being invalidated by outliers.

8 The open interest is the number of long positions or, equivalently the number of short positions.

5.3 Cointegration tests

The cointegration tests of this section will allow us to study features of the spot-forward relationship beyond unbiasedness, namely whether a potential bias is varying over time. Consider now the following linear regression

$$S_t = \alpha_k + \beta_k ({}_{t-k}F_t) + e_{k,t} \tag{10}$$

Under the null hypothesis of unbiasedness we must have that ${}_k = {}_k - 1 = 0$. Suppose that the beta coefficient does not equal unity, in particular that beta is lower than one. Every time a change in the forward price occurs, the spot price will change by a smaller amount according to Eq. (10). The market forecast error for that particular contract will also change by 1 - ${}_k$ for every one-unit change in the forward price. Thus, rejecting the null of unbiasedness will not only imply a risk premium, but one that is time-varying.

Earlier we stressed that the error term in Eq. (10) will be correlated with future values of ${}_{t-k}F_t$. We deal with this by defining a variable $v_{k,t}$ as suggested by Stock and Watson (1993)

$$v_{k,t} = {}_tF_{t+k} - {}_{t-1}F_{t+k-1} \tag{11}$$

Now, $e_{k,t}$ will be correlated with both current and future values of $v_{k,t}$, allowing us to consider a linear projection of the form

$$e_{k,t} = \sum_{j=0}^{p} \gamma_j v_{k,t+j} + z_{k,t} \tag{12}$$

where, by construction, $z_{t,k}$ is uncorrelated with $v_{t,k+j}$ for all $j = 0, \dots, p$. Substituting Eq. (12) in (10) gives

$$S_t = \alpha_k + \beta_k ({}_{t-k}F_t) + \sum_{j=0}^{p} \gamma_j v_{k,t+j} + z_{k,t} \tag{13}$$

For regressions on the 1-month ahead forward we simply substitute S_t by ${}_{t-1}F_t$. After having transformed our initial model, $z_{t,k}$ is uncorrelated with the explanatory variables. Hence our model no longer violates the zero conditional mean assumption, thereby allowing us to find consistent estimates of the regression parameters. Nevertheless, the issue regarding autocorrelated error terms remains; because $e_{k,t}$ follows a moving average process of order $k-1$, by implication, $z_{k,t}$ behaves in the same manner. Therefore we estimate Eq. (13) by OLS and adjust the t and F-statistics with a Newey-West correction. This procedure works by computing HAC standard errors which attempt to overcome autocorrelation and heteroskedasticity in the error term. The lag length used in the correction equals the order of the moving average process. Stock and Watson (1993) term the approach dynamic OLS (henceforth known as DOLS).

In order to properly estimate Eq. (13), we must also choose an appropriate lag length for our $v_{k,t}$ terms. One possible procedure is to start out using $p = k$ lags in our regression, continuing to add lags until the last added coefficient γ_j is insignificant. As it turns out, no additional lags are needed. Furthermore, from Eq. (13), a discovery that $z_{k,t}$ is stationary will imply that the error term, $e_{k,t}$, is also stationary. If this holds, it follows that the spot and forward prices are cointegrated with cointegrating vector (,). Therefore, a simple unit-root test on $z_{k,t}$ will suffice. From our preliminary studies, we already know that $e_{k,t}$ is stationary for all lags k. However, for the sake of argument, the results from our unit-root tests of $z_{k,t}$ will still be reported. Yet again, we proceed by first considering the results from NBP, given in Table 5.3.

				P-values				
Lead Price	Lagged price	Estimate of α	Estimate of β	α=0	β=1	α=0 & β=1	PP-statistic on residuals	ADF-statistic on residuals
S_t	${}_{t-1}F_t$	0.068	0.968	0.301	0.137	0.012**	< 0.01***	< 0.01***
	${}_{t-2}F_t$	0.160	0.927	0.166	0.045**	< 0.01***	< 0.01***	< 0.01***
	${}_{t-3}F_t$	0.254	0.889	0.123	0.030**	< 0.01***	< 0.01***	< 0.01***
	${}_{t-4}F_t$	0.385	0.845	0.079*	0.025**	< 0.01***	< 0.01***	< 0.01***
	${}_{t-5}F_t$	0.607	0.778	0.041**	0.019**	0.013**	0.014**	< 0.01***
	${}_{t-6}F_t$	0.774	0.728	0.040**	0.022**	0.030**	0.072*	< 0.01***
${}_{t-1}F_t$	${}_{t-2}F_t$	0.145	0.943	0.04**	< 0.01***	< 0.01***	< 0.01***	< 0.01***
	${}_{t-3}F_t$	0.227	0.909	0.069*	0.019**	< 0.01***	< 0.01***	< 0.01***
	${}_{t-4}F_t$	0.402	0.853	0.036**	0.014**	< 0.01***	< 0.01***	< 0.01***
	${}_{t-5}F_t$	0.542	0.810	0.033**	0.018**	0.022**	0.044**	< 0.01***
	${}_{t-6}F_t$	0.682	0.768	0.034**	0.022**	0.044**	0.080*	< 0.01***

The sample period extends from January 1999 through August 2009. One, two and three asterisks indicate rejection of the null hypothesis at the 10%, 5% and 1% significance levels, respectively. The number of lags of the first differences included in the ADF tests is based on the Schwartz Bayesian Criterion. The lag truncations for the Phillips-Perron test for the forecast errors equal the order of the MA-process, k-1. P-values for the regression results are computed using the Newey-West adjusted t and F-statistics.

Table 5.3 NBP cointegration test

We find that both and start with values close to zero and one for lag $k = 1$, respectively. The deviations from these values become larger as the lag length increases. The joint hypothesis $_k = {}_k - 1 = 0$ is rejected for all lags for the regressions on both the spot price and the 1-month ahead forward prices, giving rise to the rejection of our null hypothesis of unbiasedness. The conscious reader will now recall that these results are equivalent to what we found in section 5.2. Furthermore, we find that the p-values for the joint hypothesis $_k = {}_k - 1 = 0$ are increasing with the lag length k, although we reject the hypothesis for all lags. Thus, it seems that the risk premium inherent in forward contracts mainly originates from the final months before delivery. One possible explanation to this could be that the risk premium is tied to the prevailing inventory level. Intuitively, as more information about future inventories becomes available, market participants can better assess the risk of price spikes at maturity. Finally, the cointegration test's rejection of unbiasedness does not only imply a risk premium, but a time-varying one.

Having examined the results from NBP, we now investigate our results from the US found in Table 5.4 below.

				P-values				
Lead Price	Lagged price	Estimate of α	Estimate of β	α=0	β=1	α=0 & β=1	PP-statistic on residuals	ADF-statistic on residuals
S_t	$_{t-1}F_t$	-0.019	1.002	0.611	0.924	0.029**	< 0.01***	< 0.01***
	$_{t-2}F_t$	0.073	0.934	0.288	0.104	0.033**	< 0.01***	< 0.01***
	$_{t-3}F_t$	0.127	0.892	0.227	0.066*	0.025**	< 0.01***	< 0.01***
	$_{t-4}F_t$	0.218	0.836	0.08*	0.014**	< 0.01***	< 0.01***	< 0.01***
	$_{t-5}F_t$	0.349	0.760	0.016**	< 0.01***	< 0.01***	0.020**	< 0.01***
	$_{t-6}F_t$	0.488	0.681	< 0.01***	< 0.01***	< 0.01***	0.069*	< 0.01***
$_{t-1}F_t$	$_{t-2}F_t$	0.096	0.930	0.070*	0.027**	0.037**	< 0.01***	< 0.01***
	$_{t-3}F_t$	0.152	0.888	0.094*	0.028**	0.023**	< 0.01***	< 0.01***
	$_{t-4}F_t$	0.259	0.824	0.02**	< 0.01***	< 0.01***	< 0.01***	< 0.01***
	$_{t-5}F_t$	0.389	0.749	< 0.01***	< 0.01***	< 0.01***	0.014**	< 0.01***
	$_{t-6}F_t$	0.527	0.670	< 0.01***	< 0.01***	< 0.01***	0.025**	< 0.01***

The sample period extends from January 1999 through August 2009. One, two and three asterisks indicate rejection of the null hypothesis at the 10%, 5% and 1% significance levels, respectively. The number of lags of the first differences included in the ADF tests is based on the Schwartz Bayesian Criterion. The lag truncations for the Phillips-Perron test for the forecast errors equal the order of the MA-process, k-1. P-values for the regression results are computed using the Newey-West adjusted t and F-statistics.

Table 5.4 HH cointegration test

Equivalent to what we found for NBP, both and start with values close to zero and one for lag k = 1. Once more the deviations from these values become larger as k increases. Again, we reject the unbiasedness hypothesis for all leads and lags, strongly indicating a risk premium that varies over time. We recall that we failed to reject our initial test of unbiasedness for HH in section 5.2, but that all the market forecast errors were negative. Hence, there seem to be some similarities between our test results. A rejection of the null is likelier for the cointegration test because we have added a new restriction, namely that the forward price coefficient must equal unity to ensure unbiasedness. Overall, the results from the US market tell the same story as those from the UK. Nevertheless, our results are somewhat different than those of Modjtahedi and Movassagh (2005) who fail to reject the unbiasedness hypothesis for equivalent lags using DOLS. However, when estimating Eq. (13) by dynamic GLS adding GARCH to the error processes (controlling for conditional heteroskedasticity) their results are more in line with our findings.

5.4 Martingale tests

Another alternative, rendering the moving-average error problem obsolete, is to consider martingale tests

$$E_{t-k-1}\left[{}_{t-k}F_t\right] = {}_{t-k-1}F_t \tag{14}$$

The forward price series are said to have a martingale property if the forward price at time $t-k-1$ is an unbiased predictor of the prevailing forward price at time $t-k$, when both have the same end date (Haff et al., 2008). Note the subtle difference between this expression and the above test of unbiasedness in section 5.3. Since the price increments now only cover one time period, time to maturity no longer exceeds the sampling interval. Therefore, the price increments in the regression below will be white noise processes if Eq. (14) is fulfilled.

$${}_{t-k}F_t = a_k + b_k({}_{t-k-1}F_t) + e_{k,t} \tag{15}$$

Thus, martingale tests do not require HAC standard errors to ensure valid inference. However, the Stock-Watson correction of $e_{k,t}$, to correct for the aforementioned violation of the zero conditional mean assumption, is still warranted. Moreover we can now possibly detect a Samuelson effect in our series. Samuelson (1965) demonstrates that the variance of a forward price may not be constant over the contractual life, specifically that it should increase as the contract approaches delivery. In economical terms, the rationale behind the effect is that as a contract gets closer to maturity it is further influenced by the arrival of information. This is because the market lacks time to adjust before delivery takes place. Table 5.5 confirms that the Samuelson effect is indeed present in the natural gas market.

		Standard deviation	
Price change	NOB	NBP	HH
S_t - ${}_{t-1}F_t$	128	0.182	0.146
${}_{t-1}F_t$ - ${}_{t-2}F_t$	128	0.147	0.152
${}_{t-2}F_t$ - ${}_{t-3}F_t$	128	0.129	0.146
${}_{t-3}F_t$ - ${}_{t-4}F_t$	128	0.114	0.132
${}_{t-4}F_t$ - ${}_{t-5}F_t$	128	0.106	0.116
${}_{t-5}F_t$ - ${}_{t-6}F_t$	128	0.096	0.106

The sample period extends from January 1999 through August 2009.

Table 5.5 Standard deviations of price changes

We observe that the standard deviation of the price changes is growing as a contract approaches delivery. Not surprisingly, the Samuelson effect appears to be somewhat less pronounced in the US, where the price volatility in both the spot and forward market is generally lower.

Before we continue to the results, an additional note on the martingale regression is necessary. As posited by Movassagh and Modjtahedia (2005), due to the possibility of a time-varying risk premium, the martingale property is not a sufficient condition for unbiasedness. Imagine a time-varying risk premium $\mathrm{rp_t}$ so that the forward price at time t – k – 1 is given by

$$ {}_{t-k-1}F_t = E_{t-k-1}[S_t] + rp_{t-k-1} \tag{16} $$

Hence, the expected forward price materialising one period from now (i.e. t – k) is

$$ E_{t-k-1}\left[{}_{t-k}F\right]_t = E_{t-k-1}[S_t] \; + \; E_{t-k-1}[rp_{t-k}] \tag{17} $$

Now, substituting Eq. (16) in (17) yields

$$ E_{t-k-1}\left[{}_{t-k}F_t\right] - {}_{t-k-1}F_t = E_{t-k-1}[rp_{t-k} - rp_{t-k-1}] \tag{18} $$

which is identical to Eq. (14) when a time-varying risk premium is present. Therefore, the change in the forward price from one period to the next is given by the expected change in the risk premium. Thus, if a risk premium exists but is constant or slow-moving, Eq. (15) will fail to reject the null hypothesis of unbiasedness. Nonetheless, if we are in fact able to reject the null, we will have direct evidence of a bias that varies over time (both on log-scale and in linear space). Table 5.6 shows the results from the martingale regressions given by Eq. (15) for NBP.

				P-values			
Lead Price	Lagged price	Estimate of *a*	Estimate of *b*	*a* =0	*b* =1	*a* =0 & *b* =1	R^2
S_t	${}_{t-1}F_t$	0.068	0.968	0.333	0.150	< 0.01***	0.942
${}_{t-1}F_t$	${}_{t-2}F_t$	0.145	0.943	0.101	0.015**	< 0.01***	0.942
${}_{t-2}F_t$	${}_{t-3}F_t$	0.102	0.960	0.203	0.054*	< 0.01***	0.958
${}_{t-3}F_t$	${}_{t-4}F_t$	0.112	0.962	0.151	0.032**	0.041**	0.967
${}_{t-4}F_t$	${}_{t-5}F_t$	0.099	0.969	0.170	0.053*	0.135	0.972
${}_{t-5}F_t$	${}_{t-6}F_t$	0.077	0.977	0.191	0.108	0.272	0.977

The sample period extends from January 1999 through August 2009. One, two and three asterisks indicate rejection of the null hypothesis on the 10%, 5% and 1% level, respectively.

Table 5.6 NBP martingale regression

First we note that all the estimates of the intercept coefficient *a* are positive and insignificant. However, except for the first and last series, the estimates of the forward price coefficient *b* are significantly different from one. Because the martingale test is not particularly good at distinguishing between unbiasedness and the case of a constant or slow-moving risk premium, it has a rather low *power*. Formally, the power of a test is equal to the probability of rejecting a false null hypothesis. Even so, looking at the p-values for the joint test $a_k = (b_k - 1) = 0$ we are in fact able to strongly reject the null for the first four series, meaning that we have direct evidence of a bias that varies over time. This is consistent with our findings in section 5.3. Moreover, it is interesting to observe that we are unable to reject the null of the 6 and 5-months ahead forwards being unbiased predictors of the 5 and 4-months ahead forwards, respectively. This could be due to the relatively low liquidity of these longer-term contracts.

Our results are not particularly consistent with those of Haff et al. (2008) who fail to reject the null hypothesis for most of their series. However, due to their smaller data set yielding a lower power of the test, they also perform the analysis by regressing all price change series at the same time, treating them as a single data set. Their rejections of the null hypothesis were then stronger for that set, but still not in accordance with the above results as they were only able to reject the null for two of their five forward price series. The results from HH are reported in Table 5.7

				P-values			
Lead Price	Lagged price	Estimate of *a*	Estimate of *b*	*a* =0	*b* =1	*a* =0 & *b* =1	R^2
S_t	${}_{t-1}F_t$	-0.019	1.002	0.790	0.889	0.055*	0.977
${}_{t-1}F_t$	${}_{t-2}F_t$	0.096	0.930	0.108	0.021**	0.025**	0.903
${}_{t-2}F_t$	${}_{t-3}F_t$	0.080	0.943	0.112	0.054*	0.082*	0.915
${}_{t-3}F_t$	${}_{t-4}F_t$	0.090	0.945	0.058*	0.040**	0.117	0.933
${}_{t-4}F_t$	${}_{t-5}F_t$	0.098	0.943	0.018**	< 0.01***	0.035**	0.951
${}_{t-5}F_t$	${}_{t-6}F_t$	0.098	0.945	0.018**	< 0.01***	0.021**	0.958

The sample period extends from January 1999 through August 2009. One, two and three asterisks indicate rejection of the null hypothesis on the 10%, 5% and 1% level, respectively.

Table 5.7 HH martingale regression

For all but one of the series, we strongly reject the null hypothesis of unbiasedness in favour of a time-dependent risk premium, confirming the results from section 5.3. We are now also able to reject the unbiasedness hypothesis for 6 and 5-months ahead contracts, as opposed to NBP. One possible explanation is that US long-term contracts are more liquid than their UK counterparts. Modjtahedi and Movassagh (2005), on the other hand, fail to reject the null hypothesis for all the lags we are considering. By modelling the regression residuals as GARCH-processes they are, however, able to reject the null for all the series reported in Table 5.7, with the exception of $k = 1$.

Our findings suggest that forward prices are not unbiased predictors of subsequent spot prices in the UK, but rather that they contain a significant risk premium. In particular, the UK market structure with a pronounced hedging-demand for forward contracts and supply-side market power dictates that forward prices are consistently above expected spot prices. Moreover, we find that rejecting the unbiasedness hypothesis is somewhat harder in the US. We believe that this is a result of the US market being more competitive on the supply side as well as strict regulatory restrictions on large purchasers. Nevertheless, although not statistically significant, the US market forecast errors are all negative, indicating the notion of contango. Furthermore, we argue that due to the emergence of index speculators, the demand for forward contracts has increased significantly in the recent decade, possibly changing the direction of US risk premia. Our cointegration tests provide strong evidence of a time-varying risk premium in both markets. The martingale tests in the final section confirm that this is indeed the case. Finally, we find evidence of the Samuelson effect in both markets.

6 Modelling the gas forward price

So far we have found evidence favouring the existence of contango in both markets. However, the very existence of markets for derivative securities indicates that future gas prices are uncertain, and thus should be modelled as a stochastic process. The purpose of this section is therefore to apply stochastic modelling as an alternative approach to estimating the risk premia. Moreover, by properly modelling the true price dynamics as a stochastic process we are able to incorporate the distinct characteristics of the natural gas market such as seasonality and sudden price spikes. These characteristics provide useful insights in exploring the size and direction of the risk premia. Particularly, we investigate the effect of the prevailing inventory level at a forward contract's delivery date.

A widely-held view regarding commodity prices is that they should be mean-reverting, the "mean" being the real marginal cost of production. As such, an Ornstein-Uhlenbeck process, i.e. a stochastic process which exhibits reversion to mean, could appropriately imitate the price dynamics of natural gas. This model has been used extensively by other authors (e.g. Vasicek (1977) and Schwartz (1997)) to model both interest rates and commodity prices such as oil and copper. Schwartz' one-factor model is an extension of a geometric Brownian motion, where the logarithmic spot price is assumed to follow an Ornstein-Uhlenbeck process. In this paper we have seen that the natural gas market is highly seasonal, particularly in the UK. We therefore believe that the Schwartz model in its simplest form will come out short as prices are systematically changing throughout the year. In addition, Schwartz (1997) models the dispersion term as a Brownian motion. This implies that the price fluctuations are normally distributed. We argue that this assumption is flawed when considering the natural gas market because of the relatively high frequency of sudden price spikes. We will therefore need a distribution accentuating this feature. The normal-inverse Gaussian distribution (NIG) is a continuous probability distribution with tails decreasing more slowly than the normal distribution. Consequently, by accounting for fat tails, the NIG distribution yields a superior fit for the noise increments compared to that of a Gaussian distribution. Therefore we analyse the empirical risk premia in the UK and US using a geometric model with NIG-distributed noise.

We calibrate the model using spot prices and conventional time series econometrics. Because the stochastic model is intended to replicate the actual evolution of prices, we attune the model on daily spot price data, rather than the previously used aggregated monthly prices. The model is based on the model presented by Benth et al. (2008, pp. 129:146).

6.1 Deriving the model

Let the dynamics of the spot price be described as

$$S_t = \Lambda_t e^{X_t} \qquad (19)$$

Taking the natural logarithm we write Eq. (19) as

$$X_t = ln\ (S_t) - ln\ (\Lambda_t)$$

where

$$ln\, \Lambda(t) = a_0 + a_1 t + a_2 \cos\left(\frac{2\pi(t-a_3)}{260}\right) \qquad (20)$$

and

$$dX_t = -\alpha X_t dt + dI_t \qquad (21)$$

Formally, Eq. (21) is an Ornstein-Uhlenbeck process with a long-run mean equal to zero and dispersion term modelled as a Lévy process[9]. Hence, expresses the speed of mean reversion. Eq. (20) represents the average level gas prices revert to, assuming 260 trading days per year. Hence, X_t is the remainder of the log spot price after deducting the mean, trend and seasonal effects. Really, X_t can be interpreted as the remaining “noise” in the log spot prices.

9 The definition of a Lévy process is provided in Appendix 3.

We now turn to the estimation procedure using daily logarithmic spot prices ranging from 1 January 1999 till 13 September 2010. Our sample contains 3047 price quotes. Fig. 6.1 shows the logarithm of the daily gas spot prices. In order to reduce the impact of different units of measurement, the NBP prices are on the left y-axis while HH prices are on the right y-axis.

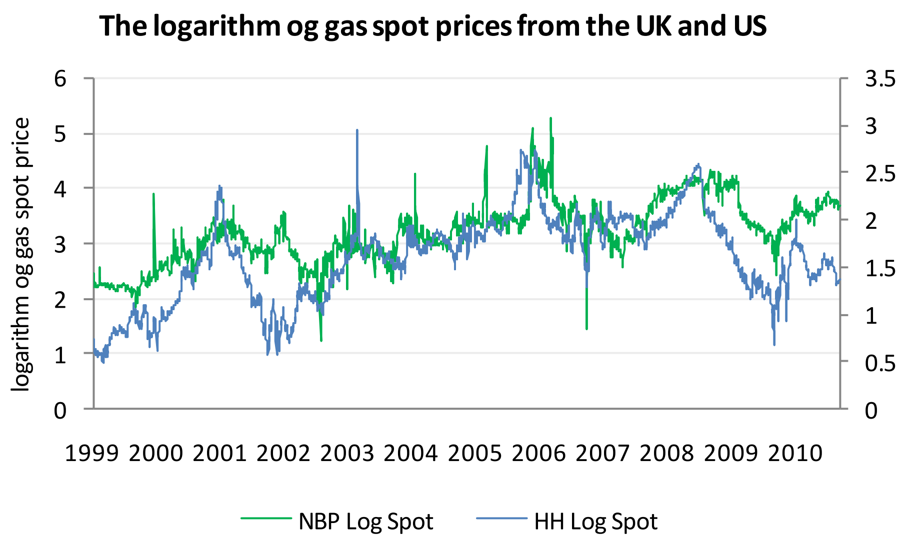

Fig. 6.1 Daily log spot prices from 1 January 1999 through 13 September 2010

Fig. 6.1 shows that the series contain large price spikes. Moreover, the price spikes are fairly symmetrical in both markets. The positive spikes amount to 52 percent and 55 percent of the total number of spikes for NBP and HH, respectively. As pointed out by Benth et al. (2008), the presence of outliers may strongly influence the estimation of the mean level function $\ln \ _t$. Therefore, we choose to remove these outliers before proceeding to the estimation of the parameters of $\ln \Lambda_t$. Using a Kolmogorov-Smirnov test on the daily changes in the log spot prices, we strongly reject the null hypothesis of normality. We use the following procedure to check for outliers when data are not normally distributed; Let Q1 and Q3 be the lower and upper quartiles, and IQR be the difference between Q3 and Q1. An outlier is defined as an observation smaller than Q1 – 3 x IQR, or larger than Q3 + 3 x IQR. Applying this filter to the series, we obtain 29 and 86 outliers for HH and NBP, respectively. We replace the outliers in the series by the average of the two closest observations. Table 6.1 shows the results from estimating the mean level function $\ln \Lambda_t$[10].

Coefficient	HH		NBP	
	Estimate	P-value	Estimate	P-value
a0	1.18	<0.01***	2.51	<0.01***
a1	2.71E-04	<0.01***	4.59E-04	<0.01***
a2	-0.04	<0.01***	-0.18	<0.01***
a3	143.00	<0.01***	133.20	<0.01***

Table 6.1 Parameters of the mean level function ($ln\ \Lambda_t$)

All the estimated parameters are significant at the 1 percent level. Therefore we conclude that there has been significant seasonality and a steady increase in spot prices throughout the period. Fig. 6.2 shows the original logarithmic spot prices series after subtracting the mean level function ln $_t$.

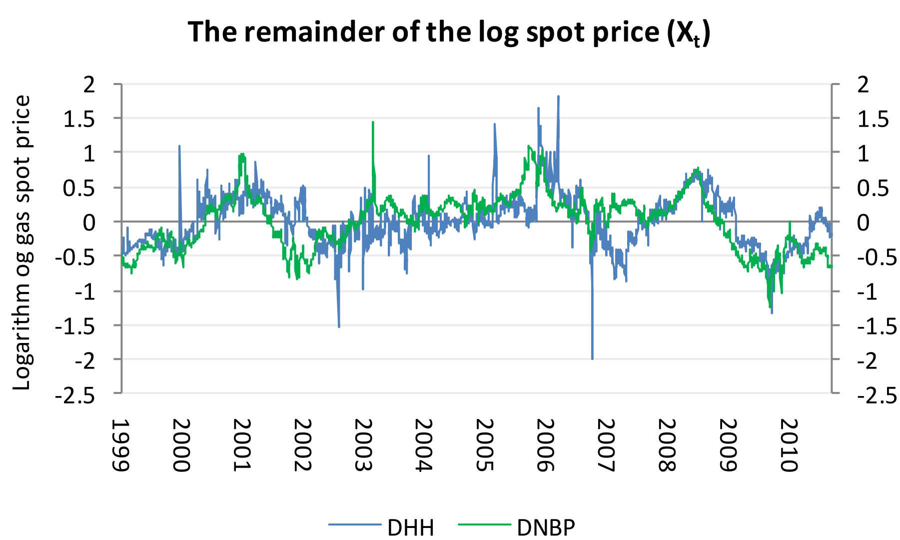

Fig. 6.2 The remainder of the log spot price (X_t)

10 We applied the "nls" function in R.

From Fig. 6.2, we observe that the remainder of the log spot price X_t is centred around zero. When prices deviate from zero, they revert back with the speed of mean reversion . However, the speed of mean reversion appears to be very low. In fact, prices can deviate for several years. This seems reasonable as gas supply is relatively price inelastic. Producers require long lead time to increase capacity, varying somewhere between six months and ten years (Natural Gas Supply Association, 2010). Also, several economic factors are preventing producers to ease production in spite of falling prices. As mentioned earlier, if production of natural gas from a well is halted, it may not be possible to restore the well's production due to reservoir and wellbore characteristics (Natural Gas Supply Association, 2010).

We now turn the procedure of fitting an appropriate time-series model to the detrended and deseasonalised logarithmic spot prices X_t. Box and Jenkins (1976) develop a methodology for estimating autoregressive integrated moving-average (ARIMA) models. In this procedure, the ACF and PACF serve as essential tools in indentifying and estimating the appropriate ARIMA model. Fig. 6.3 shows the ACF and PACF for the detrended and deseasonalised logarithmic spot prices, DNBP (UK) and DHH (US).

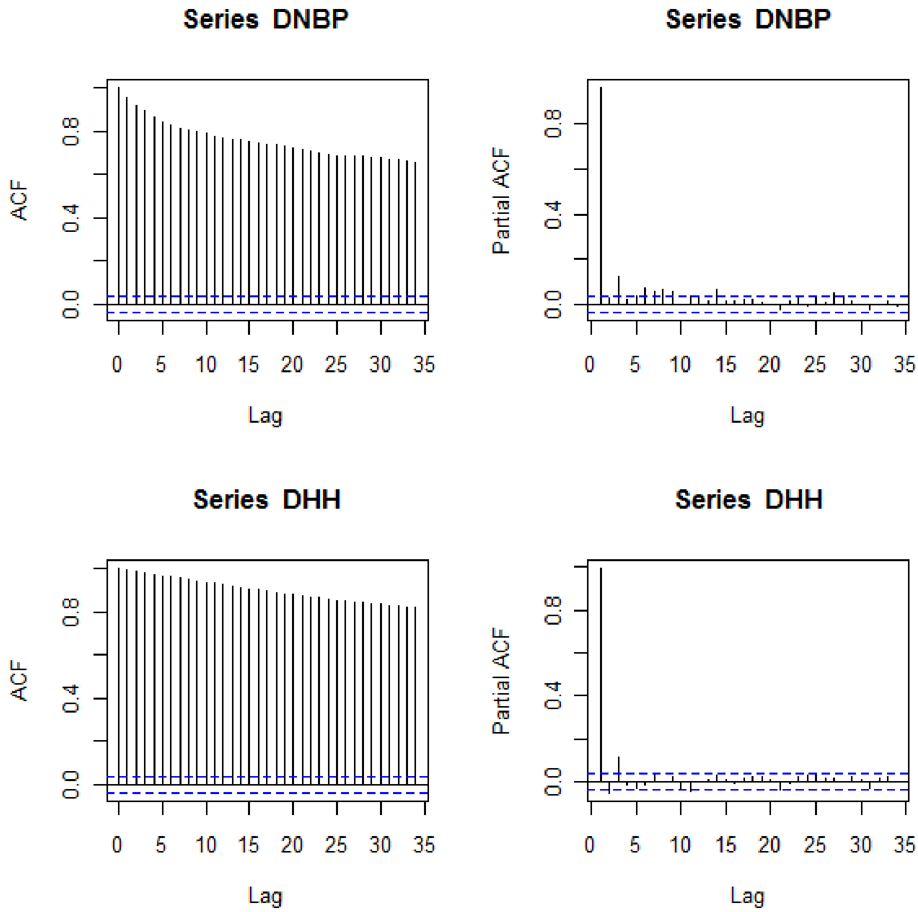

Fig. 6.3 ACF and PACF plots for the remainder of the log spot price (X_t)

The slowly decaying pattern of the ACF indicates an AR-coefficient close to unity for NBP. Combined with one large spike at lag 1 in the sample PACF, an AR(1) process seems to be a reasonable choice. However, several spikes in the PACF are in fact significant, also suggesting the presence of a moving-average component in the series. Indeed, for the AR(1) we fail to reject the null hypothesis of zero autocorrelation in the residuals, strengthening the latter argument. Albeit the ACF plot for HH is fairly similar to that of NBP, the PACF is somewhat different. The PACF has three significant lags suggesting an AR(3) process. However, the oscillatory decaying pattern of the PACF starting at lag 1 favours an ARMA process. We therefore choose to estimate an ARMA(1,2) process. Conveniently it turns out that an ARMA(1,2) model fits both the series, i.e. we fail to reject the null hypothesis of autocorrelated residuals.

We now need to link the ARMA (1,2) process to the presumed dynamics of the remainder of the log spot price X_t. We perform a first order Euler discretisation of Eq. (21) to obtain

$$X_{t+1} \approx (1-\alpha)X_t + \Delta I_t \tag{22}$$

$$= \rho_1 X_t + \epsilon_{t+1}$$

where $\rho_1 = (1-\alpha)$ and $\{\epsilon_t\}_{t=0}^{T}$ is a sequence of uncorrelated random variables. We recognise Eq. (22) as an AR(1) process where ρ_1 is equal to the first order autocorrelation coefficient. In order to obtain an expression for the speed of mean reversion we solve for the first order autocorrelation coefficient of the fitted ARMA(1,2) process. After some tedious algebra it is possible to show that

$$\rho_1 = {}^{\gamma_1}/_{\gamma_0} \tag{23}$$

$$\gamma_1 = a\gamma_0 + \sigma^2[\beta_1 + \beta_2(a + \beta_2)]$$

$$\gamma_0 = \sigma^2 \left[\frac{a\beta_1 + a\beta_2(a+\beta_1) + 1 + \beta_1(a+\beta_1) + \beta_2(a^2 + a\beta_1 + \beta_2)}{1-a^2}\right]$$

where σ^2 is the variance of the error term, *a* is the autoregressive coefficient and β_1 and β_2 are the moving-average coefficients from the estimated ARMA(1,2) process. After solving for ρ_1, we obtain the following estimates of the speed of mean reversion in Eq. (22)

$$\hat{\alpha}_{DNBP} = 0.04$$

$$\hat{\alpha}_{DHH} = 0.005$$

Next, we consider the residuals from the estimated ARMA(1,2) models. When performing a Kolmogorov-Smirnov test for normality, we strongly reject the null hypothesis of the residuals being normally distributed. Fig. 6.4 compares the density function of the estimated residuals to that of the normal distribution with the same mean and standard deviation.

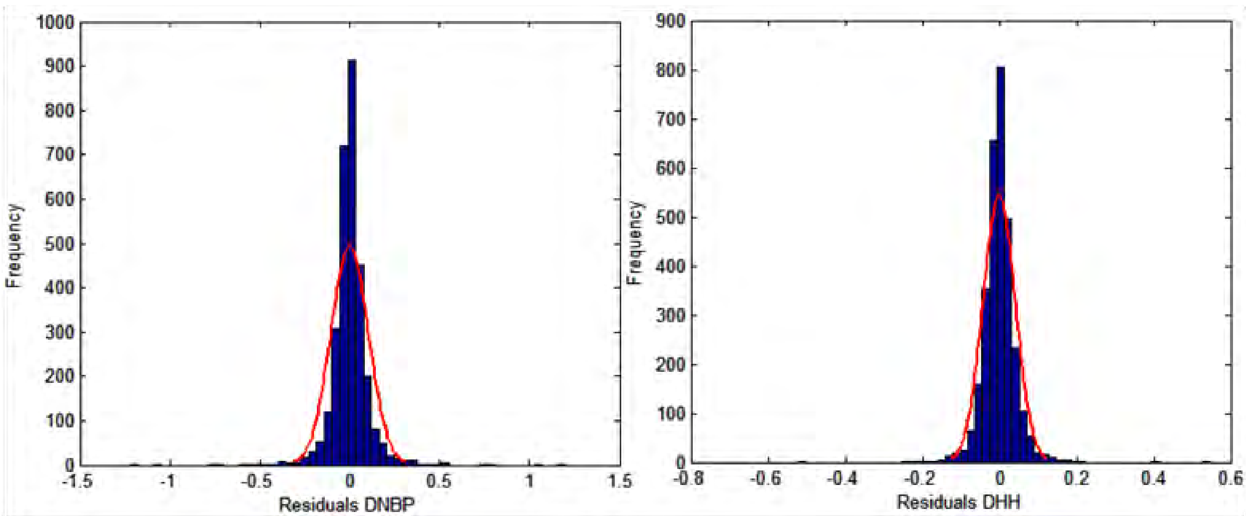

Fig. 6.4 Density function of the estimated residuals vs the normal distribution

Fig. 6.4 confirms that the residuals are far from normal. In particular, they exhibit significant leptokurtic behavior. Following Benth and Saltyte-Benth (2004) we therefore proceed with the Lévy process I_t where I_1 is NIG distributed, i.e. $I_1 \sim NIG(\zeta, \beta, \delta, \mu)$. Here, ζ, , and denotes tail-heavyness, skewness, scale and location, respectively. We fit the NIG distribution to the residuals by using maximum likelihood estimation (MLE)[11]. Hence, we take the leptokurtic behavior of the residuals into account. Table 6.2 reports the ML estimates of the parameters.

Parameter	HH	NBP
ζ	19	4.91
β	0.6604	0.1972
δ	0.03078	0.04856
μ	-0.00109	-0.00191

Table 6.2 NIG parameters

11 We applied the nigFit function in R.

Finally, we derive the theoretical forward price. In order to do this, we must first solve the stochastic differential equation given in Eq. (21). First, we define

$$Y_t = X_t e^{\alpha t} \tag{24}$$

Using Eq. (24) and (21), by Itô's lemma we have that

$$dY_t = \{\alpha Y_t + e^{\alpha t}(-\alpha X_t)\}dt + e^{\alpha t} dI_t = e^{\alpha t} dI_t \tag{25}$$

Y_t is now given by

$$Y_t = X_0 + \int_0^t e^{\alpha u} dI_u \tag{26}$$

Thus, we obtain the solution to Eq. (21) as

$$X_t = X_0 e^{-\alpha t} + \int_0^t e^{-\alpha(t-u)} dI_u \tag{27}$$

Assuming that a buy-and-hold strategy is possible and the absence of arbitrage, we must have that

$$0 = e^{-r(T-t)} E_t[\ {}_tF_T - S_T] \tag{28}$$

Using the conditional expectation operator $\mathrm{E_t}$ and substituting Eq. (19) into (28) we get the following expression for the forward price ${}_t\mathrm{F_T}$

$${}_tF_T = \Lambda(\mathrm{T})\mathrm{E_t}(\mathrm{e}^{\mathrm{X_T}}) \tag{29}$$

At time T, Eq. (27) becomes

$$X_T = X_t e^{-\alpha(T-t)} + \int_t^T e^{-\alpha(T-u)} dI_u \tag{30}$$

Now, substituting Eq. (19) and (30) in (29) yields

$$ {}_tF_T = \Lambda(\mathrm{T})\mathrm{E_t}(\mathrm{e}^{\ln(\frac{S_t}{\Lambda(t)})e^{-\alpha(T-t)}+\int_t^T e^{-\alpha(T-u)}dI_u}) \quad (31) $$

$$ = \Lambda(\mathrm{T})\left(\frac{\mathrm{S_t}}{\Lambda(\mathrm{t})}\right)^{\mathrm{e}^{-\ (\mathrm{T-t})}} \mathrm{E_t}(\mathrm{e}^{\int_t^T e^{-\alpha(T-u)}dI_u}) $$

By calculating the conditional expectation using Baye's formula, it is possible to show that Eq. (31) can be written as

$$ {}_tF_T = \Lambda(\mathrm{T})\left(\frac{\mathrm{S_t}}{\Lambda(\mathrm{t})}\right)^{\mathrm{e}^{-\ (\mathrm{T-t})}} \Theta(\mathrm{t},\mathrm{T};\theta) \quad (32) $$

where

$$ \ln\Theta(\mathrm{t},\mathrm{T};\theta) = \int_{\mathrm{t}}^{\mathrm{T}} \psi\left(-\mathrm{ie}^{-\alpha(\mathrm{T-u})}\right)\mathrm{du} \quad (33) $$

Proving Eq. (33) is beyond the scope of this paper. Interested readers can turn to Benth and Saltyte-Benth (2004). In Eq. (33)   is the cumulant function and $\left(-\mathrm{ie}^{-\ (\mathrm{T-t})}\right)$ becomes the logarithm of the moment generating function. Finally, by the NIG Lévy process we have that

$$ \left(-\mathrm{ie}^{-\ (\mathrm{T-t})}\right) = \ \mathrm{e}^{-\ (\mathrm{T-t})} + \ \left[\left(\sqrt{\zeta^2 - \ ^2}\right) - \left(\sqrt{\zeta^2 - (\mathrm{e}^{-\ (\mathrm{T-t})} + \ )^2}\right)\right] \quad (34) $$

where   is the implied market price of risk to be estimated. We solve for the theoretical forward price given by Eq. (32) by integrating Eq. (33) numerically[12].

12 Because T appears both inside the integral and in the upper limit we had to make a change in the variable to solve Eq. (40). We define a new variable x = T-u so that dx = -du. Hence we have that $\int_{\mathrm{t}}^{\mathrm{T}} \ \left(-\mathrm{ie}^{-\ (\mathrm{T-u})}\right)\mathrm{du} = -\int_{\mathrm{T-t}}^{0} \ \left(-\mathrm{ie}^{-\ (\mathrm{x})}\right)\mathrm{dx}$.

6.2 Results

We now wish to compare the theoretical forward prices to the actual forward prices observed in the market. To facilitate comparison with actual prices, we use spot prices aggregated to monthly resolution in Eq. (32). Moreover, in order to compute the theoretical forward price, we need to decide on a value for the market price of risk . Fig. 5.5 compares the theoretical 1-month ahead forward prices to the actual observed prices by choosing equal to zero. Hence, a finding that the deviations between the two series are large will imply that forward prices are not unbiased predictors of future spot prices.

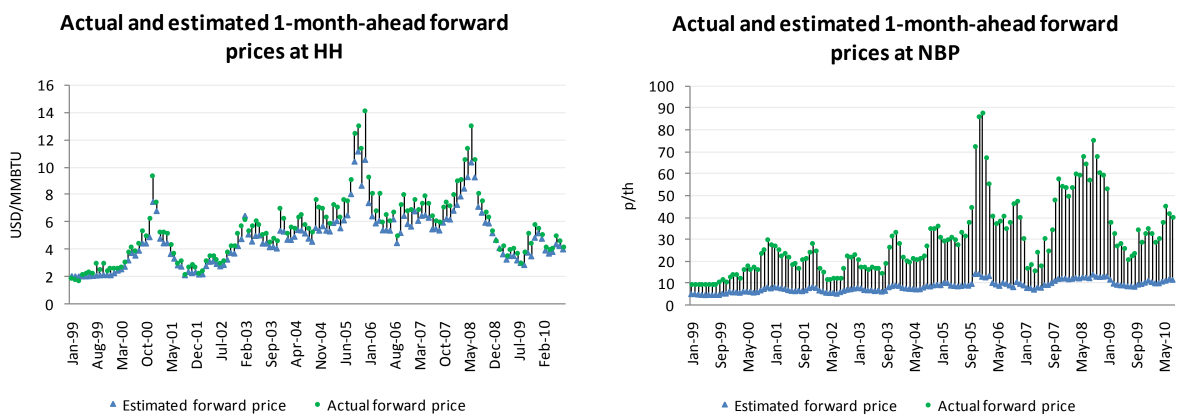

Fig. 6.5 Comparison of theoretical prices with a market price of risk equal to zero and actual observed prices

Our estimated forward prices seem to trace the observed prices rather well in the US. Thus, the risk premium seems to be very small, which can explain why we failed to reject our initial test of unbiasedness for HH in section 5.2. Nevertheless, we see that the theoretical prices are consistently below the actual observed prices, indicating the existence of a risk premium. Clearly, the differences between actual and estimated forward prices are too large to be explained by any reasonable risk premium for NBP. Most likely, our simple one-factor model is unable to capture the complex relationship between spot and forward prices. Without putting too much emphasis on the results for NBP, we still find that the observed forward prices are larger than the estimated risk-neutral prices throughout the whole period. Because a long investor pays more than what the risk-neutral price dictates, the short position is rewarded the premium. Hence, these findings are consistent with the notion of contango found in most of the tests in section 5.

Healthy scepticism is warranted when comparing the deviations in the US to those in the UK. Firstly, we have just argued that our model seems to perform poorly on UK data. Secondly, the model involves several estimation procedures, all subject to potential estimation errors. Although the deviations in the two markets are not directly comparable, we still find it noteworthy that the spread is systematically larger in the UK. This and previous findings indicate that the risk premium could be greater in the former market. We know that UK inventories are more seasonal than their US counterparts. This is also reflected by the probability of sudden price spikes being higher in the UK than in the US as measured by the number of outliers. Moreover, from our investigation of the theory of storage we know that inventories are highly significant in explaining the contemporaneous basis. Because of its significant effect on spot prices we argue that the variation in storage levels is likely to be an influential factor in determining risk premia as well. When inventory levels are high, the spot price tends to be low, and the inventory will act to even out variations in the spot price. On the other hand, when inventories are low, they obviously cannot act as a buffer against price variations, causing price volatility to increase, thereby driving up a potential risk premium. When commodity forwards are used to insure against price volatility, the cost of insurance should be increasing in the amount of risk (volatility). The lower inventory levels are, the higher the probability of sudden price spikes, and thus higher downside risk for the speculator providing insurance. Consequently, the speculator should on average require a higher risk premium on contracts that expire when the inventory is expected to be low.

We now wish to examine whether the size of the market price of risk is directly related to the variation in storage. To investigate this we calculate the implied values of by requiring that the theoretical risk-neutral forward price, given by Eq. (32), must equal the actual forward price for a given point in time. We solve the equation using Newton's search algorithm for zero points with Matlab's *fsolve* function. For simplicity we restrict the analysis to only two months, namely October 2009 and March 2010. These months are chosen to reflect the difference in the market price of risk due to seasonality in inventory levels. In particular, we analyse whether contracts maturing in the winter when inventories are low are associated with a higher market price of risk than contacts with delivery in the summer months. The results are shown in Table 6.3.

Contract date, t	Delivery period, T	Implied $\tilde{\theta}$		Contract date, t	Delivery period, T	Implied $\tilde{\theta}$	
		Henry Hub	NBP			Henry Hub	NBP
Mar-10	Apr-10	1.23	2.15	Oct-09	Nov-09	3.46	2.39
	May-10	1.49	1.98		Dec-09	3.11	2.17
	Jun-10	1.56	1.81		Jan-10	2.77	1.99
	Jul-10	1.59	1.66		Feb-10	2.52	1.81
	Aug-10	1.58	1.54		Mar-10	2.32	1.66
	Sep-10	1.55	1.42		Apr-10	2.19	1.55

Table 6.3 Implied market price of risk

We observe that all the implied s have the same sign. Because these are the market prices of risk necessary to equate actual and theoretical risk-neutral forward prices, we are again able to confirm the notion of contango for these particular months. Interestingly, we find that contracts maturing when inventories are expected to be low yield a relatively higher risk premium, reflected in the higher market price of risk, supporting the above reasoning. Note that the market price of risk does not have a meaningful unit-interpretation, but still provides essential information about the risk premium. Firstly, a positive market price of risk implies that forward contracts trade at a premium compared to the expected future spot price (contango). Secondly, everything else equal, a higher market price of risk dictates a higher forward risk premium.

However, for all but the far left series, the market price of risk is decaying with time to maturity. Intuitively, we would expect the opposite because a speculator requires a higher compensation the longer he is exposed to risk. For contracts with delivery in the summer months we believe this contradiction stems from inventory levels being much more important in explaining risk premia than the time to maturity. Still, we would expect to find a market price of risk increasing with time to maturity for contracts with delivery in the winter when inventories are relatively lower. One possible explanation to this is that although the NIG distribution fits the residuals in the spot price dynamics quite well, it might not be able to appropriately model the jump risk. Obviously, the risk of large jumps increases with the contract length. The jump frequency measured by the amount of outliers is approximately 3 times larger at NBP than at HH. Consequently, the estimated market prices of risk for the UK are less reliable. Furthermore, when choosing only two observations from the sample, the risk of picking abnormal market situations must be taken into account.

In this section, disregarding the possibility of model misspecification, we have again found evidence of contango in both markets. Moreover, by choosing two arbitrary dates from our data sample, we find a strong link between the market price of risk and the expected level of inventory at delivery. In fact, we find that inventories at a contract's maturity seem to be a more important determinant of the risk premium than the contractual length.

7 Inventory levels and risk premia

We conclude this paper by further examining the relationship between inventories and the risk premium. Ekern (2009) defines the risk premium as the difference between the expected payoff and the certainty equivalent. In other words, the risk premium is the payoff which makes a person indifferent between accepting an uncertain payoff and a guaranteed amount. To earn a risk premium when markets are in contango, an investor without interest in the physical commodity can short a forward contract at time t – k with delivery at time t. At time t – 1, he locks in his payoff by entering a long forward position with the same delivery date t. This strategy yields a payoff of ${}_{t-k}F_t - {}_{t-1}F_t$ at time t for a month $k \in [2,6]$. The payoff at time t has an interesting interpretation. In fact, it is the maximum value a risk adverse market participant is willing to pay to eliminate his price risk, i.e. the k -1 risk premium. The payoff can therefore be considered the compensation, paid by a hedger, to a speculator for providing him with insurance. The price of insurance is the exact value making the hedger indifferent between a certain and an uncertain outcome. In other words, it is the definition of the certainty equivalent.

We now proceed by comparing the risk premium from the aforementioned trading strategy for lag k = 6 and the inventory level. As previously posited, we expect to find a negative relationship between the risk premium on forward contracts maturing when storage is scarce and the expected level of inventory at delivery. Obviously, as maturity approaches, more information will be available to the investor. Hence, we choose to apply the longest forward contracts in our sample because preliminary studies show that the relationship between the risk premium and inventories is less clear cut for shorter contracts.

The risk premium is calculated as the annualised 5-month return earned at delivery by a speculator following the above trading strategy, i.e. $({}_{t-6}F_t - {}_{t-1}F_t) / {}_{t-1}F_t$. The results are shown in Table 7.1 below.

	Arithmetic Mean p.a.		Median p.a.		Standard Deviation p.a.	
Delivery month	NBP	HH	NBP	HH	NBP	HH
January	46%	22%	10%	1%	68%	79%
February	71%	45%	19%	27%	64%	70%
March	130%	38%	46%	8%	83%	64%
April	67%	13%	36%	-8%	143%	48%
May	55%	-3%	8%	-30%	75%	53%
June	17%	1%	-15%	-35%	85%	59%
July	12%	-27%	-11%	-42%	56%	70%
August	-4%	-11%	-19%	-24%	50%	71%
September	11%	33%	3%	19%	55%	49%
October	16%	66%	9%	46%	39%	89%
November	36%	35%	24%	21%	47%	62%
December	49%	53%	10%	56%	49%	67%

The sample period extends from January 2000 through December 2009.

Table 7.1 Annualised mean and median 5-month risk premia

Clearly, the mean return is heavily skewed due to the presence of large outliers. Hence, we choose to proceed with the median returns rather than the mean returns. In order to facilitate comparison with the inventory level at delivery we compute a normalised measure of the inventory (henceforth knowns as NDI). This is because we want to disregard any growth in the storage capacity. The NDI is computed by dividing the detrended inventory level by the average detrended inventory level. The trend was removed by running a linear regression of the inventory level on time. Fig. 7.1 plots both the annualised median 5-month risk premium and the estimated NDIs at delivery for NBP. The former is given by the left y-axis, whereas the latter is found on the right y-axis.

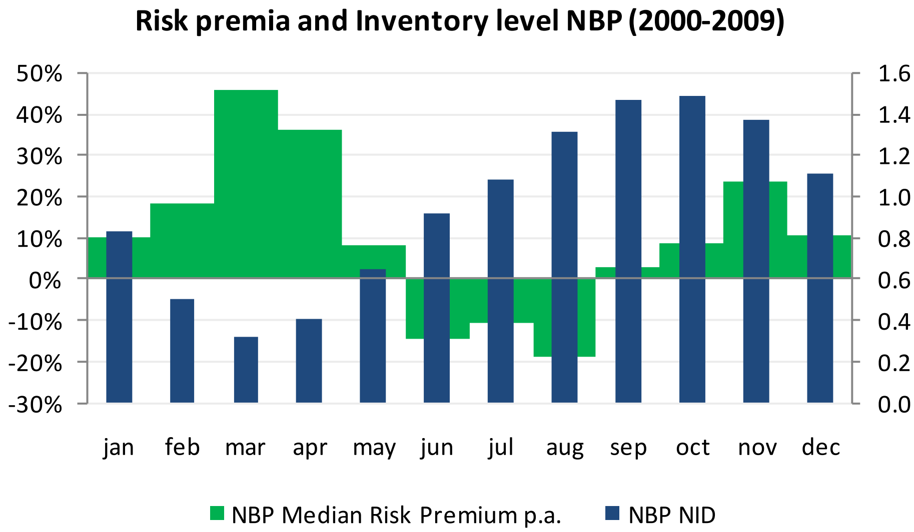

Fig. 7.1 UK 5-month median risk premium and normalised inventory levels

From Fig 7.1 we find that contracts maturing in the winter, when inventories are generally low, contain a relatively high risk premium. Moreover, the prices of forward contracts with delivery in the summer when inventories are full contain lower risk premia. Hence, we are able to confirm that the UK risk premium is indeed negatively correlated with the level of inventory at delivery. Again, we argue that this pattern is more pronounced for longer-term forward contracts because of greater uncertainty with respect to future inventory levels. Moreover, it is interesting to observe that the risk premium seems to change sign for some months. This indicates a complex structure of the risk premium and that it is highly variable. Fig. 7.2 shows the same trading strategy applied to the US market.

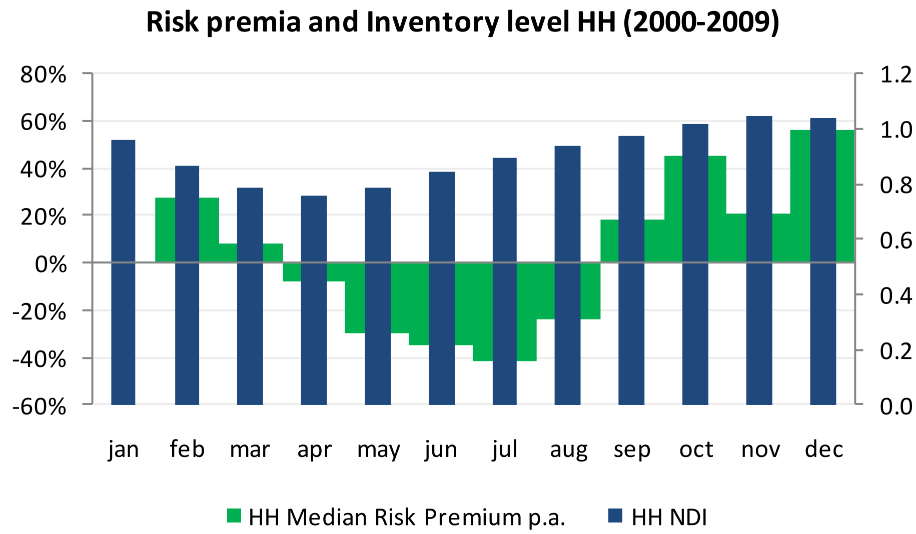

Fig. 7.2 US 5-month median risk premium and normalised inventory levels

From Fig. 7.2 we immediately find that the proposed relationship between the risk premium and inventories at delivery does not seem to hold in the US. In fact, it is hard to observe any intuitive pattern between the two at all. However, we know that the NDIs have been approximately 4 times as volatile in the UK as in the US during the last decade. Higher variability in inventories reflects both a higher price and delivery risk. Thus, the risk inherent in US forward contracts with respect to the inventory is lower in the US than in the UK. In particular, it seems that the level of inventory at delivery is a less important factor in explaining the US forward risk premium.

An interesting feature in Fig. 7.2 is that the risk premium becomes highly negative from April till August. One possible explanation could be related to the US market structure. US inventories reach their lowest level in April which marks the start of a period of net gas inflow. Since the US supply side is highly fragmented, one could argue that producers might be willing to sell forward contracts maturing in the summer below the expected future spot price to ensure delivery in a period of relatively low demand.

8 Concluding remarks

8.1 Reliability of results

Throughout this paper we have made several assumptions regarding our models, some of which may be flawed. A small paragraph discussing the most essential assumptions is therefore warranted.

It is well-known that economies experience sporadic unexpected changes. Some of these have significant effects on the state of the economic system. Such changes not only lead to difficulties in economic forecasting, but also in the formulation of economic models. One recent example is the financial crisis. Another is the emergence of shale gas. Both events could have induced structural breaks implying that earlier data is no longer relevant to the current market situation. As such, the reliability of our results may be lowered trough our use of data both prior to and after these events.

The stochastic model in section 6 relies on the assumption that the mean reversion coefficient is constant over a period of 11 years. This is clearly erroneous. Moreover, we would expect the mean reversion coefficient to be different during periods of frequent price spikes and periods of relative tranquillity. The latter could be circumvented by estimating a multifactor model which allows for various speeds of mean reversion using the Kalman filter.

Finally, this study applies a large body of statistical tests and software. It is possible that calculation mistakes are present. Furthermore, we cannot rule out that mistakes were made when writing scripts used in Matlab and R.

8.2 Conclusion

Our findings suggest that there has been a positive trend in natural gas prices, probably a result of the period from 2003 to 2007 being one of high economic growth. Moreover, price volatility appears to be time-varying. In particular, the emergence of unconventional gas has served to increase volatility in recent years, possibly imposing a structural break in both markets. We also observe a pronounced seasonality in the prices of natural gas. Not surprisingly, this feature is more evident in the UK, where natural gas, to a larger extent, is used for heating. Our unit-root tests show that all the series are covariance-stationary. In particular, we find that the UK series are trend-stationary, i.e. exhibiting reversion to an upward trend. This finding is consistent with the widely-held view that commodity prices should be mean-reverting, the "mean" being the real marginal cost of production.

We find that the basis has been positive on average in the period analysed, reflecting a positive cost-of-carry. However, in both markets a negative basis occurs rather frequently, indicating a relatively high convenience yield. Moreover, the average basis is higher in the UK than in the US, possibly a result of a larger concentration of suppliers keeping forward prices "artificially" high.

We only find partial support for the theory of storage, where the cost of capital is not particularly relevant in predicting the basis. On the contrary, the level of storage is very important. Indeed, we find that a concave relationship between the basis and the level of inventory seems appropriate, implying that the basis is increasing at a decreasing rate as inventories are being filled.

We argue that the UK basis seems to be larger than any reasonable cost of carry would predict. We believe that such a disparity would be eliminated in an efficient market. This could indicate that only a few agents are in fact able to take advantage of the arbitrage opportunities available to them. We therefore proceed by testing whether the markets are efficient in the sense that forward prices are unbiased predictors of future spot prices, i.e. the unbiasedness hypothesis. Our findings suggest that UK forward prices are biased. In particular, the UK market structure, with a pronounced hedging-demand for forward contracts and supply-side market power, dictates that forward prices are consistently above expected spot prices (contango). Moreover, the bias appears to be time-varying.

The US market is more balanced in terms of hedging pressure. The supply-side is more competitive and large purchasers are subject to strict regulatory restrictions. Consequently, rejecting the unbiasedness hypothesis is somewhat harder in the US. Still, we are able to infer the notion of contango. This is surprising, considering that most previous studies find the market to be backwardated. We argue that the emergence of index speculators has increased the demand for forward contracts in the recent decade and possibly altered the direction of US risk premia. Moreover, the increased importance of shale gas could have changed the historical US spot-forward relationship.

Finally, we discover a strong link between the market price of risk and the expected level of storage at delivery in the UK. In fact, we find that the level of inventory at a contract's maturity appears to be a more important determinant of the UK risk premium than the contractual length. In the US however, it seems that the expected storage level at delivery is a less important factor in explaining the risk premium, probably due to US inventories being less seasonal.

9 References

Augustine, C.W., Broxson, B. & Peterson, S. (2006) *Understanding Natural Gas Markets*. Available from: < http://www.api.org/aboutoilgas/upload/UNDERSTANDING_NATURAL_GAS_MARKETS.pdf> *(Accessed* 2 November *2010).*

Benth, F.E., Benth, J. Š. & Koekebakker, S. (2008) *Stochastic Modelling of Electricity and related Markets.* London: World Scientific Publishing Co. Pte. Ltd.

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

Borenstein, S., Busse, M. & Kellogg, R. (2007) *Principal-agent Incentives, Excess Caution, And Market Inefficiency: Evidence from Utility Regulation.* Working Paper 13679, National Bureau of Economic Research.

Box, G.E.P. & Jenkins, G.M. (1970) *Time Series Analysis: Forecasting and Control.* San Francisco: Holden Day.

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

Brown, S.P.A. & Yücel, M.K. (2007) *What Drives Natural Gas Prices.* Working Papers 0703, Federal Reserve Bank of Dallas

Cartea, A. & Williams, T. (2007) *UK Gas Markets: The Market Price of Risk and Applications to Multiple Interruptible Supply Contracts*. Birkbeck Working Papers in Economics & Finance, pp. 1-29.

Chincarini, L.B. (2007) *The Amaranth Debacle: A Failure of Risk Measures or a Failure of Risk Management?* The Journal of Alternative Investments, Winter 2007, pp. 91-104.

Cleveland, R.B., Cleveland, W.S., McRae, J.E. & Terpenning I. (1990) *STL: A Seasonal-trend Decomposition Procedure based on Loess*. Journal of Official Statistics, 6(1), pp. 3-73.

Cohen, D. (2009) *A Shale Gas Boom?* Available from: < http://www.energybulletin.net/node/49342 > (Accessed 14 November 2010).

Considine, T. J. & Larson, D.F. (2001) *Uncertainty and the Convenience Yield in Crude Oil Price Backwardations*. Energy Economics, 23(5), pp. 533– 548.

Department of Energy and Climate Change (2010) *Energy Consumption in the UK: Domestic data tables, 2010 update.* Available from: < http://www.decc.gov.uk/publications/basket.aspx?filetype=4&filepath=Statistics%2fpublications%2fecuk%2f269-ecuk-domestic-2010.xls&minwidth=true> (Accessed 15 November 2010).

Economist Intelligence Unit (2010) *Commodities: Hards: Natural Gas.* Available from: < http://gfs.eiu.com/Article.aspx?articleType=cfh&articleId=1347136519 > (Accessed 17 November 2010).

Ekern, S. (2010) *ECO421 Lecture 2: Expected Utility and Risk Aversion.* NHH (26 August 2009)

Enders, W. (2009) *Applied Econometric Time Series.* 3rd ed. New Jersey: Wiley Publishing.

Fama, E.F. & French, K.R. (1987) *Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage*. The Journal of Business, 60(1), pp. 55– 73.

Gjølberg, O. & Brattested, T.L. (2010) *The Forecasting Performance of the Nord Pool Futures Price, 1995-2008.* Preliminary version presented at FIBE, 6 January 2010.

Gjølberg, O. & Johnsen, T. (2004) *The Pricing of Electricity Futures: Empirical Tests for Market Efficiency at Nord Pool.* Discussion Paper IOR, UMB.

Gorton, G. & Rouwenhorst K.G. (2004) *Facts and Fantasies about Commodity Futures*. Yale ICF Working Paper, 4(20).

Haff, I.H., Lindqvist, O & Løland, A. (2008) *Risk Premium in the UK Natural Gas Forward Market*. Energy Economics, 30(5), pp. 2420-2440.

Hull, J.C. (2008) *Options, Futures, and other Derivatives*. 7th ed. New Jersey: Pearson Prentice Hall.

ICIS Heren (2007) *What do falling spot prices mean for storage players?* Available from: < http://www.icis.com/heren/articles/2007/01/31/9293590/what-do-falling-spot-prices-mean-for-storage-players-.html > (Accessed 16 November 2010).

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

Keynes, J.M. (1930) *A Treatise on Money. Volume 2: The Applied Theory of Money.* London: Palgrave MacMillan.

Milonas, N.T. & Henker, T. (2001) *Price Spread and Convenience Yield Behaviour in the International Oil Market*. Applied Financial Economics, 11(1), pp. 23–36.

Modjtahedi, B. & Movassagh, N. (2005) *Natural-Gas Futures: Bias, Predictive Performance and the Theory of Storage*. Energy Economics, 27(4), pp. 617-637.

Movassagh, N. & Modjtahedi, B. (2005) *Bias and Backwardation in the Natural-gas Futures Prices*. The Journal of Futures Markets, 25(3), pp. 281– 308.

The National Petroleum Council (2007) *Energy Markets Grow Up Part 2: Who Trades Energy Now and How Much Does It Matter?* Reference Report #32. Available from: < http://www.npc.org/Study_Topic_Papers/32-MacroMarketsGrowUp2.pdf > (Accessed 20 October 2010).

Natural Gas Supply Association (2010) *Industry and Market Structure.* Available from: < http://www.naturalgas.org/business/industry.asp > (Accessed 15 September 2010).

Pindyck, R.S. (2001) *Volatility and Commodity Price Dynamics*. A more recent version appeared in (2004) The Journal of Futures Markets, 24(11), pp. 1029–1047.

Pirog, R. (2006) *CRS Report for Congress: Natural Gas Markets in 2006.* Available from: < http://www.fas.org/sgp/crs/misc/RL33714.pdf > (Accessed 17 November 2010).

Samuelson, P. (1965) *Proof that Properly Anticipated Prices Fluctuate Randomly*. Industrial Management Review, 6 (41).

Schwartz, E.S. (1997) *The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging*. The Journal of Finance, 52(3), pp. 923– 973.

Stock, J.H. & Watson, M.W. (1993) *A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems*. Econometrica, 61(4), pp. 783– 820.

Telser, L.G. (1958) *Futures Trading and the Storage of Cotton and Wheat. Journal of Political Economy*, 66, pp. 223-255.

UK Parliament (2008) *The Wholesale Gas Market.* Available from: < http://www.parliament.the-stationery-office.co.uk/pa/cm200708/cmselect/cmberr/293/29305.htm#note24 > (Accessed 17 November 2010).

U.S. Energy Information Administration (2010a) *U.S. Underground Natural Gas Storage by Storage Type.* Available from: < http://tonto.eia.doe.gov/dnav/ng/ng_stor_type_s1_m.htm > (Accessed 20 September 2010).

U.S. Energy Information Administration (2010b) *U.S. Natural Gas Consumption by End Use.* Available from: < http://www.eia.gov/dnav/ng/ng_cons_sum_dcu_nus_m.htm > (Accessed 15 November 2010).

U.S. Energy Information Administration (2010c) *Natural Gas Year – In – Review 2009.* Available from: <http://www.eia.gov/pub/oil_gas/natural_gas/feature_articles/2010/ngyir2009/ngyir2009.html> (Accessed 2 November 2010).

U.S. Energy Information Administration (2007) *An Analysis of Price Volatility in Natural Gas Markets.* Available from: <http://www.eia.doe.gov/pub/oil_gas/natural_gas/feature_articles/2007/ngprivolatility/ngprivolatility.pdf > (Accessed 5 October 2010).

Vasicek, O. (1977) *An Equilibrium Characterization of the Term Structure*. Journal of Financial Economics, 5(2), pp. 177-188.

Volmer, T. (2009) *A Robust Model of the Convenience Yield in the UK Natural Gas Market.* Available from: < *http://www.fmpm.ch/docs/13th/papers/E3c.pdf* > *(Accessed 20 October 2010).*

Wei, S.Z.C. & Zhu, Z. (2006) *Commodity Convenience Yield and Risk Premium Determination: The Case of the U.S. Natural Gas Market*. Energy Economics, 28(4), pp. 523–534.

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

Wooldridge, J.M. (2008) *Introductory Econometric: A Modern Approach*. 4th ed. United States: South-Western College Pub.

Working, H. (1949) *The Theory of the Price of Storage*. The American Economic Review, pp. 1254–1262.

Wray, R. L. (2008) *The Commodities Market Bubble: Money Manager Capitalism and the Financialization of Commodities.* Public Policy Brief -The Levy Economics Institute of Bard College , pp. 1-57.

10 Appendices

Appendix 1

Enders (2009) defines a stochastic process with a finite mean and variance as *covariance-stationary* if for all t and t – s

$$E(y_t) = E(y_{t-s}) = \mu \tag{1}$$

$$E[(y_t - \mu)(y_{t-s} - \mu)] = E\left[(y_{t-j} - \mu)(y_{t-j-s} - \mu)\right] = \gamma_s \tag{2}$$

where μ, σ_y^2 and γ_s are all constants. From Eq. (2) we also have that, for s = 0, the variance must be constant over time.

Appendix 2

It is possible to derive forecasts of any ARMA(p,q) model by means of an iterative technique. In order to keep the algebra simple, consider the ARMA(2,1) model along the lines of Enders (2009, p.83):

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \varepsilon_t + \beta_1 \varepsilon_{t-1} \tag{1}$$

Assuming that all coefficients are known and that $E_t \varepsilon_{t+j} = 0$ for all $j > 0$, the conditional expectation of Y_{t+1} is:

$$E_t Y_{t+1} = a_0 + a_1 Y_t + a_2 Y_{t-1} + \beta_1 \varepsilon_t$$

The one-step-ahead forecast error is:

$$e_t(1) = Y_{t+1} - E_t Y_{t+1} = \varepsilon_{t+1}$$

By forward iteration we can in a similar manner find the two-step-ahead forecast error:

$$Y_{t+2} = a_0 + a_1 Y_{t+1} + a_2 Y_t + \varepsilon_{t+2} + \beta_1 \varepsilon_{t+1}$$

$$E_t Y_{t+2} = a_0 + a_1 E_t Y_{t+1} + a_2 Y_t$$

$$e_t(2) = Y_{t+2} - E_t Y_{t+2} = a_1(Y_{t+1} - E_t Y_{t+1}) + \varepsilon_{t+2} + \beta_1 \varepsilon_{t+1} = (a_1 + \beta_1)\varepsilon_{t+1} + \varepsilon_{t+2} \quad (2)$$

When using multi-step-ahead forecasts, the forecasts may exhibit serial correlation. Updating Eq. (2) by one period yields the two-step-ahead forecast error for Y_{t+3}:

$$e_{t+1}(2) = (a_1 + \beta_1)\varepsilon_{t+2} + \varepsilon_{t+3} \quad (3)$$

It is now obvious that the two forecast errors in Eq. (2) and (3) are correlated. In particular we have that:

$$Cov[e_t(2), e_{t+1}(2)] = Cov[(a_1 + \beta_1)\varepsilon_{t+1} + \varepsilon_{t+2}, (a_1 + \beta_1)\varepsilon_{t+2} + \varepsilon_{t+3}]$$

$$= (a_1 + \beta_1)Var(\varepsilon_{t+2}) = (a_1 + \beta_1)\sigma^2$$

since $\{\varepsilon_t\}$ is a white noise sequence, i.e. having a mean value of zero, a constant variance and is uncorrelated with all other realizations. The point is that due to the presence of ε_{t+2} in both the prediction of Y_{t+2} and Y_{t+3} from the perspective of period t and t+1 respectively, both contain an error. Note however that for $k > 1$, $Cov[e_t(2), e_{t+k}(2)] = 0$ since there are no overlapping forecasts. Hence the autocorrelations of the 2-step-ahead forecast errors is reduced to zero after lag 1. It is the general version of this result that causes the j-step-ahead forecast errors to act as an MA(j-1) process.

Appendix 3

A Levy process $\{\widetilde{I_t}\}_{t\epsilon[0,T]}$, is a stochastic process on a probability space (Ω, Ƒ, P) with the following properties

I. $I_0 = 0$

II. Independent increments

III. Stationary increments