A dependence model for pairs of commodity forward curves with an application to the US natural gas and oil markets

Steve OHANA Birkbeck, University of London*

October 2006

Abstract

The goal of this paper is to present and calibrate a model for the joint evolution of correlated commodity forward curves. The main originality of the model is that it *captures both the local and global dependence structures* of two forward curves, through an error-correcting term in the risk-premia of the forward price returns. The model is applied here to the US oil and gas forward markets, which have strong economic relations, from the demand and supply sides.

JEL classification: G12, G13

Keywords: commodity forward curves, crude oil, natural gas, cointegration

1 Introduction

The modeling of the co-movements of commodity forward curves has so far received very little attention in the financial literature. Yet, this is a subject of considerable importance for the pricing, risk management,

*Comments should be addressed to Steve OHANA; Address: Birkbeck, University of London, Malet Street Bloomsbury, London WC1E 7HX, Tel +4 4 (0)20 7631 0813, Fax +44 (0)20 7631 6416, Email: steve.ohana@gmail.com

I gratefully acknowledge Helyette Geman for her careful reading of the manuscript and her advice. I thank Gaz de France for providing me with the data, Céline Jerusalem for helping me to treat them, Gregory Benmenzer, Olivier Bardou, and Jean-Jacques Ohana for stimulating discussions

and optimization of portfolios composed of multi-commodity assets such as gas-fired power plants, oil-indexed natural gas contracts, or oil refineries. Indeed, the financial value of a multi-commodity asset is a function of the entire forward curves and the hedging strategies for multi-commodity portfolios are based on futures contracts rather than spot transactions. As a consequence, a model describing the evolution of commodity spot prices only, provides a partial view of the risks/value entailed in such portfolios and of the possible actions of the portfolio manager. A model describing the joint evolution of two commodity forward curves should capture at the same time their global and local dependence structures. The *local dependence structure* describes the volatilities, the marginal densities and the correlations of the daily forward curve moves. A framework of analysis for this type of dependence was described in Clewlow and Strickland (2000), who propose to extend the classical PCA on one commodity forward curve to a PCA on the returns of two commodity forward curves, thus obtaining several types of co-movements of the two forward curves. By contrast with the local dependence structure, the *global dependence structure* describes the long-term relations existing between commodity prices[1]. Much attention has been devoted to the study of cointegration between series of different spot/futures commodity prices[2], with a view to describing *the interaction between several particular points* in the same forward curve or in different forward curves (for example the relations between the front-month prices of a pair of commodities or the relations between the spot and front-month prices of the same commodity). There is extensive work also on the evolution of a single interest rate or commodity forward curve, either for forecasting (see Diebold and Li (2003)) or VaR calculation (see e.g. Brooks (2001)). But no work, to our knowledge, has ever proposed a framework to simulate the evolution of *two entire commodity forward curves*, describing the way the two curves "revert to each other". The retained approach for this problem follows Pilipovic (1997), Manoliu and Tompaidis (2002), Schwartz and Smith (2000), and Geman and N'Guyen (2005), who decompose the daily deformations of a forward curve into a short-term shock, affecting only the first maturities,

[1]two frequent examples of long-term interactions between commodity markets are the possibility to use a given commodity to product another one (natural gas to produce power, crude oil to produce heating oil...) or to use a given commodity as a substitute to another one (e.g. heating oil instead of natural gas for heating, coal instead of natural gas to produce power)

[2]see e.g. Alexander (1999) for a study of the cointegration between gas/oil spot and futures prices on the NYMEX, Ates and Wang (2005) for an analysis of the relations between spot and first-near by natural gas prices in the US, Siliverstovs et al. (2005) for an analysis of cointegration between Japanese, European, and North American gas prices, Nguyen (2002) for the analysis of the cointegration between the futures prices of metals on the London Metal Exchange, Pekka and Antti (2005) for the study of cointegration between spot and futures electricity prices on the NordPool

and a long-term shock, consisting of an overall translation of the forward curve. Regarding the local dependence structure, the model captures, on the one hand, the causal relations between the daily short-term and long-term shocks of the two commodities, and on the other hand, the time-dependent volatilities of the four comovements (see e.g. Geman and Nguyen (2005), Richter and Sorensen (2000), and Duffie (2002), for evidence of stochasticity of the volatility of commodity prices, and Blix (2003) for evidence of seasonality of natural gas implicit volatility), and their possibly non Gaussian dependence structure. The approach to capture the long-term relations between two forward curves can be viewed as an *extension of the concept of cointegration to forward curves.* The decomposition of the forward curve daily moves translates into a decomposition of the *shape* of the forward curve into a seasonal term, slope[3] and level. The long-term relationships between the two commodity forward curve slopes and levels are looked for and the deviations to these equilibriums become predictive variables for the future relative evolution of the two curves. The model is applied here to the US natural gas and crude oil markets during the period January 1999-July 2005. These two markets, in spite of their differences, are intertwined by economic relations, from the consumption side and the production side. Regarding the local dependence structure, we find evidence of causal relations between natural gas and crude oil shocks, stochastic volatility for the different shocks, seasonal volatility for natural gas short-term shocks only, and positive correlations between the co-movements of oil and gas forward curves. Regarding the global dependence structure, our analysis highlights the existence of a strong long-term relationship between the levels of natural gas and oil (with two break points occurring in the beginning of year 2000 and in the middle of year 2003), and of a weaker long-term relationship between gas and oil slopes. The analysis of the temporal stability of the model parameters reveals that the correlations between the daily co-movements of oil and gas forward curves have increased significantly throughout the period 1999-2005.

I view the contribution of this paper as twofold: from an economic standpoint, the presented forward curve model sheds light on the relations between the natural gas and oil markets in the US, in particular the lead and lag properties between the two energies; from a statistical standpoint, the model proposed here opens a new avenue for the modeling of the joint evolution of several correlated forward curves, giving a simple way to capture in a single arbitrage-free model the long-term relations between the shapes of different forward curves and the local statistical relations between their daily co-movements.

The rest of this paper is organized as follows. In section 2, we describe the economic relations between oil

[3]depending on the sign of the slope, the curve will be said to be in contango or in backwardation

and natural gas markets in the US, from the demand side and the offer side. In section 3, we present the two-factor model and describe the global and local dependence structures between oil and gas forward prices in the US. In section 4, the model is precisely calibrated and the temporal stability of the model parameters is studied. Section 5 contains concluding comments.

2 The economic relations between oil and natural gas in the US

Even though the natural gas market is a competitive and local market whereas the crude oil market is an oligopolistic and global market, the natural gas and oil prices are intertwined by strong economic relations, emanating from both the demand and supply sides. Industry represents approximately 30% and power generation 20% of the global US gas consumption[4]. On the whole, the global available switching potential represents around 5% of the natural gas consumption in the US, 30% coming from industry, the rest from power generation. Around 4.3% of the natural gas consumption of the industrial sector is switchable: these customers are equipped with dual-fuel capacity (essentially boilers and process heaters), allowing them to switch from gas to oil (generally distillate or residual fuel oil) depending on the market prices of the two energies[5]. As regards power generation, the fuel-switching potential represents 20% of the gas consumption, but is expected to decline due to the progressive replacement of dual-fuel steam boilers by gas-fired combined cycles facilities. Fuel-switching implies a dependence between oil and gas prices which is both in the very short term (due to existing switching capacities), and in the medium-long-term (due to technology changes following a sustained period of abnormally high natural gas or oil prices).
Because industrials or electricity producers often lock in their margins using the forward markets, we expect positive correlations not only between oil and gas spot prices but between oil and gas forward prices as well. This convergence between gas and oil forward markets is reinforced by the current behavior of hedge funds and financial investors, who tend more and more to consider the different commodity markets as a unified asset class (see Geman (2005)).

The dependence between oil and natural gas prices in the US is also originating in the supply side. In this case, two effects play in opposite directions. On the one hand, as natural gas is a co-product of oil, a rise in crude oil prices provokes an increase of the supply of crude oil, which in turn leads to an increased

[4] All data are found in American Gas Foundation (2003)

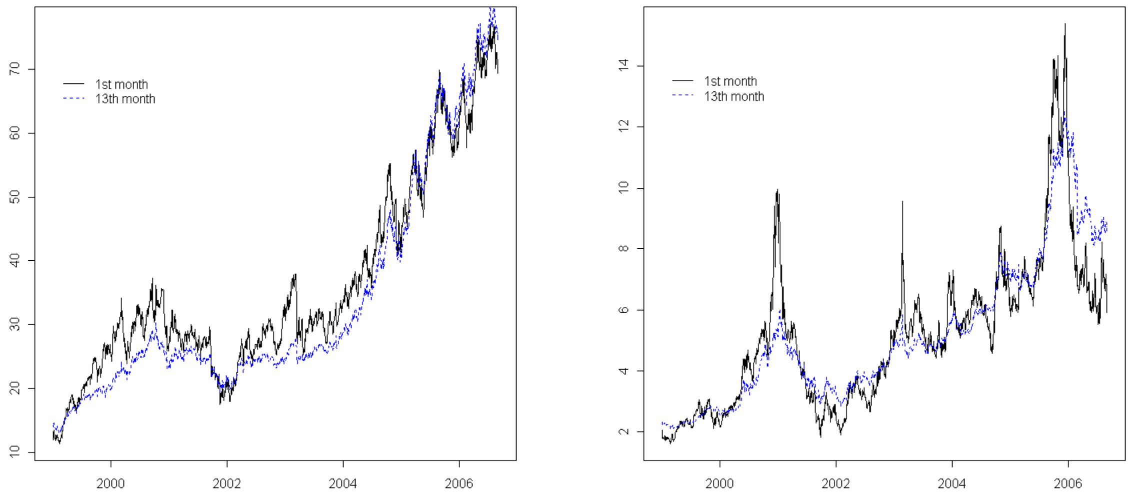
[5] Note that the recent environmental regulations, imposing air pollutant emission constraints or costs to industrials, tend to prevent them from using distillate fuel or coal as substitutes to natural gas

production of natural gas, thus putting a downward pressure on natural gas prices; on the other hand, as the Gulf of Mexico concentrates major gas and oil fields, gas processing plants, and oil refineries, the supply of natural gas and oil distillate products in the US are both strongly dependent on the frequent natural events striking this region: for example, when Katrina, Rita, and Wilma made landfall, they affected at the same time the production of natural gas, Crude Oil, and refined products in the US, thus causing a simultaneous rise in prices of the two energies.

3 Empirical observation of the dependence between oil and gas forward curves in the US

3.1 Data description

The data used here are the NYMEX daily futures prices of natural gas and crude oil from January 1999 to the end of August 2006. For the three energies, the prices are the 1st month, 2nd month,...,15th month futures prices. Concerning natural gas, the price is based on delivery at the Henry Hub in Louisiana. The futures prices are expressed in dollars per Million British Thermal Units (MMBtu). For crude oil, the NYMEX futures contracts's delivery point is Cushing, Oklahoma, and the prices are expressed in dollars per barrel.



(a) crude oil futures prices in dollars/Barrel (b) natural gas futures prices in dollars/MMBtu

Figure 1: Price trajectories from January 1999 to August 2006

Figure (1) represents the trajectories of 1st month and 13th month futures prices for the two energies:

- the trends of natural gas and oil long term prices display a parallel direction
- even though the 1st month natural gas futures price exhibits much larger moves than oil within the period, oil and gas approximately share the same backwardation and contango periods[6][7]
- the period 1999-2004 can be separated into several subperiods:
  - from January 1999 to end of 2001, the two trajectories display a "bump": they first follow an upward trend until the end of 2000, and then a decay until the end of 2001
  - in the years 2002-2003, gas prices start rising while oil prices remain stable
  - from the beginning of 2004 to now on, the two energies display a very clear surge
  - during the period August 2005[8]-February 2006, the natural gas short-term and long-term prices displayed a pronounced spike, which was observed also but to a much lesser extent on the crude oil market

[6]Remark that the effect of the seasonality in the gas forward curve is filtered out here since the delivery periods of the two observed contracts are distant from one year

[7]There are a few notable exceptions to this rule such as the summer 2004, when the gas forward curve was in contango and the crude oil curve was backwardated

[8]time of Hurricane Katrina's landfall in the Gulf of Mexico

From now on, we will restrict our analysis of the dependencies between oil and gas to the period beginning in January 4, 1999 and ending in July 29, 2005, based on the observation that Katrina's landfall provoked a temporary disconnection in the normal long-term relations between oil and gas markets.

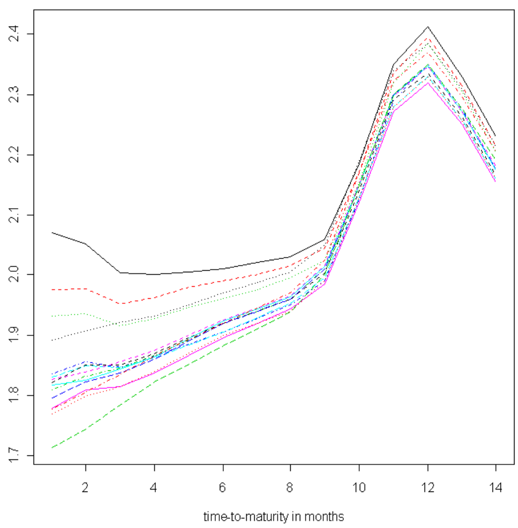
3.2 Decomposition of daily forward curve moves into short term and long-term shocks

3.2.1 Justification and interpretation of the decomposition

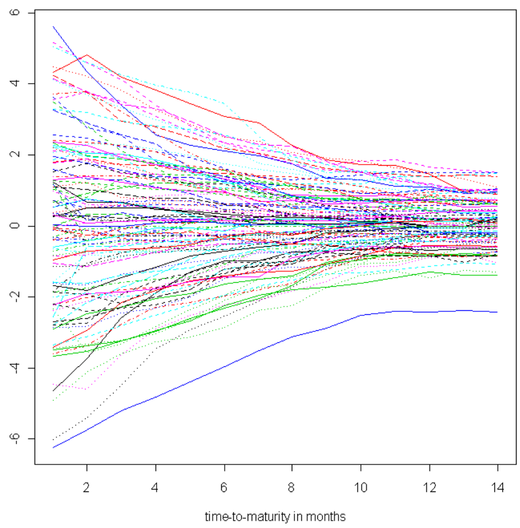
In Figure (2), it appears that forward curve moves decompose into a long-term shock, which provokes a global upward or downward translation of the forward curve, and a short term shock, which only impacts the short term futures prices, with an amplitude that decays with time-to-maturity. In economic terms, the interpretation of the decomposition is the following:

- the short term shock refers to events that are expected to affect the market for a limited period of time (temperature change, transitory supply shortage or transportation congestion...)[9]
- the long-term shock relates to events or news that potentially impact the long-term energy price (news about the likelihood of a war or political instability in an oil producing country, disclosure of lower than expected reserves...)

[9] One could wonder why events of weekly time scale such as a temperature drop or a bottleneck in the transportation system should affect the prices of the contracts delivering in the following months; this link between spot and forward markets is explained by the storability of the three considered energies. Indeed, tensions in the day-ahead market prompt utilities and distribution companies to pump on their reserves in order to take advantage of high spot prices or be able to deliver their firm clients; this in turn creates a situation of scarcity in the medium term, which, as explained by the theory of storage, has a direct impact on the slope of the monthly forward curve



(a) natural gas futures prices (in \$/MMBtu) as a function of time to maturity (in months) from January, 4th to January 19th, 1999



(b) Natural gas futures prices returns as a function of time to maturity (in months) from January 5th to May 27th, 1999

Figure 2: Decomposition of returns into a short and long term shocks

3.2.2 Mathematical formulation of the decomposition

We denote $F^e(t,T)$ the futures price at time t of a futures contract with settlement date T written on the energy product e. We assume the following arbitrage-free daily evolution model for the forward curve of energy e:

$$\begin{aligned} \frac{\Delta F^e(t,T)}{F^e(t,T)} &= e^{-k_e(T-t)}\Delta X^e_t + \Delta Y^e_t \qquad (1)\\ \Delta X^e_t &= \alpha^{e,X}_t + \sigma^{e,X}_t \eta^{e,X}_t \\ \Delta Y^e_t &= \alpha^{e,Y}_t + \sigma^{e,Y}_t \eta^{e,Y}_t \end{aligned}$$

where:

- $(\alpha^{e,X}_t)$ and $(\alpha^{e,Y}_t)$ are $(\mathcal{F}_t)$-adapted processes representing the drifts
- $(\sigma^{e,X}_t)$ and $(\sigma^{e,Y}_t)$ are $(\mathcal{F}_t)$-adapted processes representing the volatilities
- $(\eta^{e,X}_t)$ and $(\eta^{e,Y}_t)$ are correlated processes formed of i.i.d variables
- $\frac{1}{k_e}$ represents the characteristic time of the short term shock

3.2.3 Calculation of the short term and long-term shocks

Assuming that the short term shock does not affect the 14th month return[10], the short term and long-term shocks can be readily derived from the observed short term and long-term returns of energy e:

$$\begin{aligned} \Delta Y_t^e &= \left(\frac{\Delta F^e(t,T_{14})}{F^e(t,T_{14})}\right)_{obs} \\ \Delta X_t^e &= e^{k_e(T_1-t)}\left(\frac{\Delta F^e(t,T_1)}{F^e(t,T_1)} - \frac{\Delta F^e(t,T_{14})}{F^e(t,T_{14})}\right)_{obs} \end{aligned} \tag{2}$$

where the variable T_i denotes the last trading day of the i-th month futures contract observed at date t.

3.2.4 Estimation of the short term characteristic time for the three energies

To estimate k_e for the two energies, we minimize the root mean squared errors (RMSE) i.e., the root of the mean squared differences between the observed returns and the model implied returns

$$\left(\frac{\Delta F^e(t,T_i)}{F^e(t,T_i)}\right)_{model} = \left(\frac{\Delta F^e(t,T_{14})}{F^e(t,T_{14})}\right)_{obs} + e^{-k_e(T_i-T_1)}\left(\frac{\Delta F^e(t,T_1)}{F^e(t,T_1)} - \frac{\Delta F^e(t,T_{14})}{F^e(t,T_{14})}\right)_{obs} \tag{3}$$

Therefore, we solve, for each energy e, the following programme:

$$\underset{k_e}{Min} RMSE = \sqrt{\frac{1}{N\times 14}\sum_{t=1}^{N}\sum_{i=1}^{14}\left[\left(\frac{\Delta F^e(t,T_i)}{F^e(t,T_i)}\right)_{obs} - \left(\frac{\Delta F^e(t,T_i)}{F^e(t,T_i)}\right)_{model}\right]^2} \tag{4}$$

where N is the number of observations.

	crude oil	natural gas
k_e	2.60	3.30
$\frac{1}{k_e}$ (in months)	4.61	3.64
$\frac{(RMSE)^2}{(RMSR)^2}$ (in %)	1.41	5.56

Table 1: k_e and unexplained variances for the two energies; $RMSE$ (resp. $RMSR$) stands for the root mean squared errors (resp. returns)

Table (1) reports the short term characteristic times of the three energies. As a first observation, these characteristic times are compatible with the assumption $3 \times \frac{1}{k_e} < 14$ months, which helped us calculate the short term and long-term shocks. In addition, we observe that the short term characteristic time of natural gas is significantly smaller than the one of crude oil. The economic interpretation is that the short

[10] This is equivalent to the assumption $3 \times \frac{1}{k_e} < 14$ months

term shocks in the natural gas *local* market are linked to very short-lived events (e.g., sudden drop of temperature in the US, bottleneck in the gas transportation system etc...) whereas the short term shocks in the *global* crude oil market correspond to events with a longer time scale (e.g. disclosure of a lower than expected world inventory).

In addition, Table (1) reveals that the performance of the model in explaining the variance of the observed returns is significantly lower for natural gas (with an unexplained variance of 5.56%) than for oil (with an unexplained variance of 1.41%). A plausible explanation is that the relative importance of "twist" moves (which are not accounted for in the two factor model) in the global forward curve volatility is more pronounced for natural gas than for oil. This is confirmed by a Principal Component Analysis on the 14 series of forward curve returns, whose results are displayed in table (2):

	crude oil	natural gas
1st factor	96.64%	93.21%
2nd factor	2.91%	4.69%
3rd factor	0.33%	1.18%

Table 2: Proportion of overall variance explained by the 1st (translation), 2nd (rotation), and 3rd factors (twists) for the two energies

3.3 Slope and level: two state variables for the shape of the forward curve

The evolution model (1) implies a *forward curve shape model*. Indeed, if we neglect the second-order terms:

$$\Delta lnF^e(t,T) \approx \frac{\Delta F^e(t,T)}{F^e(t,T)} = e^{-k_e(T-t)}\Delta X_t^e + \Delta Y_t^e$$

and we obtain the following expression for the shape of the forward curve at date t:

$$lnF^e(t,T) \quad = \quad lnF^e(0,T) + \sum_{s=0}^{t} e^{-k_e(T-s)}\Delta X_s^e + \sum_{s=0}^{t} \Delta Y_s^e \tag{5}$$

Let us assume that the shape of the initial forward curve is of the type:

$$lnF^e(0,T) = Q(T) + e^{-k_e T}\bar{X}_0^e + Y_0^e \tag{6}$$

where T takes integer values representing months and Q is a function of period one year and zero mean. Then, equation (5) leads to:

$$lnF^e(t,T) = Q(T) + e^{-k_e(T-t)}\bar{X}_t^e + Y_t^e \tag{7}$$

with:

$$\bar{X}_t^e = \bar{X}_0 e^{-k_e t} + \sum_{s=1}^{t} e^{-k_e(t-s)}\Delta X_s^e \tag{8}$$

$$Y_t^e = Y_0 + \sum_{s=1}^{t} \Delta Y_s^e \tag{9}$$

Equation (7) shows that, under model (1), the shape of the forward curve at any date t is the superposition of a seasonal function $Q(T)$, a slope $\bar{X}_t$, and a level Y_t. The slope and level can be derived from the daily shocks $(\Delta X_t^e, \Delta Y_t^e)$ via (8)-(9). The slope follows a mean-reverting process driven by the short term shocks and the level a random walk driven by the long-term shocks:

$$\bar{X}_t^e = \bar{X}_{t-\Delta t}^e e^{-k_e \Delta t} + \Delta X_t^e \tag{10}$$

$$Y_t^e = Y_{t-\Delta t}^e + \Delta Y_t^e \tag{11}$$

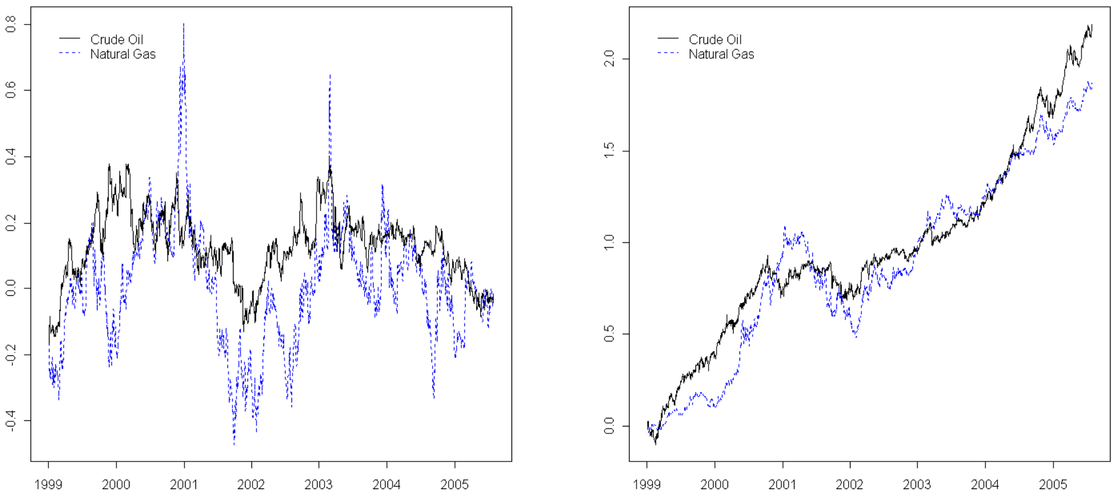
The forward curve model (7) has very classical economic interpretations: the seasonality of the forward curve is explained by a structural imbalance between winter and summer consumptions and by the small number of market participants having access to storage reservoirs; the level is related to the long-term price of the commodity and the slope to the benefit (classically referred to as the "convenience yield") of holding the physical commodity vs holding a contract for future delivery. Fama and French (1988) in particular use the slope of the forward curve as a proxy for the inventory level.

The level at any date is computed using formula (11), the level at the first date being initialized at 0[11]. The slope is estimated by "inversion" of formula (7) using the observed 1st month and 13th month log futures prices at date t:

$$\bar{X}_t^e = e^{k_e(T_1-t)} ln(F_e(t,T_1)/F_e(t,T_{13}))$$

From now on, we will refer to the relations between slopes and levels as the global dependence structure and to the correlation and causal relations between daily shocks as the local dependence structure.

[11] the level reflects the gains of an investor holding a constant sum of \$1 in the 15-th month futures contract and rolling over his position at each last trading day



(a) Slopes (b) Levels

Figure 3: Slopes and levels of the two energies

3.4 Analysis of global dependence structure

3.4.1 Stationarity properties of the slopes and levels

Table (3) reports the results of the Phillips-Perron unit root tests on the slopes and levels of the two energies. Not surprisingly, the hypothesis of a unit root can be rejected for the slopes but not for the levels.

	Dickey-Fuller	Lag Parameter	p-value
Crude oil slope			
PP	-3.833	8	0.017
ADF	-3.552	11	0.037
Natural gas slope			
PP	-3.713	8	0.023
ADF	-3.319	11	0.0674
Crude oil level			
PP	-0.704	8	0.970
ADF	-0.57	11	0.979
Natural gas level			
PP	-1.835	8	0.648
ADF	-1.674	11	0.717

Table 3: Phillips-Perron (PP) and Augmented-Dickey-Fuller (ADF) unit-root tests on the slopes and levels of the two energies; the test-statistics, truncation lag parameters, and p-values of the tests are reported

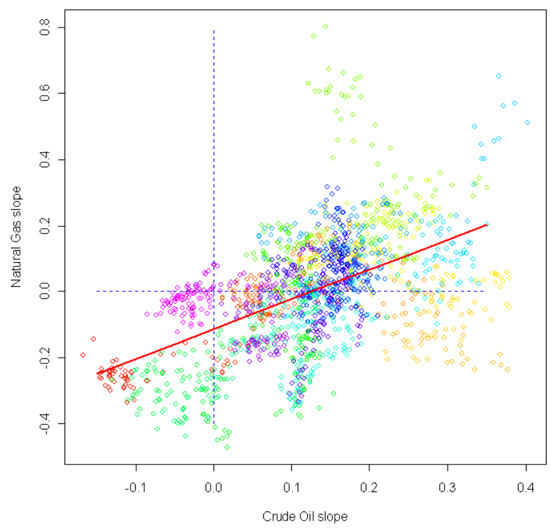
3.4.2 Long-term relation between forward curve slopes and levels

Figure (4(a)) displays the relation between natural gas and crude oil slopes. We see that, when the natural gas forward curve is in backwardation (positive slope), the oil forward curve is also in backwardation[12]. However, a backwardated oil curve does not necessarily imply a backwardated natural gas forward curve. In particular, year 2002 experienced a backwardated oil curve and a natural gas forward curve in contango. The results of the linear regression of natural gas slope on crude oil slope are reported on table (4). Note that the regression coefficients is very close to 1, the regression R^2 being around 30%.

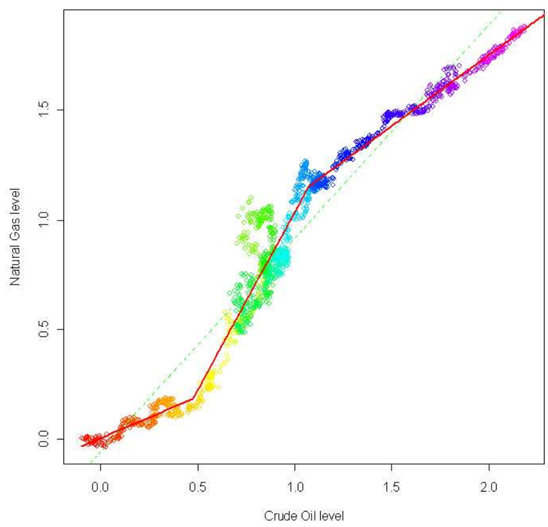
[12]note however that there are outliers in the linear relation: for instance, during the winters 2000-2001 and 2002-2003, the natural gas slope was very high while the oil slope was mildly positive

	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
b_{crude}	-0.113	0.00594	-19.12	$< 2.10^{-16}$	***
a_{crude}	0.906	0.0365	24.80	$< 2.10^{-16}$	***
					$R^2 = 27.25\%$

Table 4: Linear regression of natural gas slope on crude oil slope: a denotes the regression coefficient and b the intercept; the estimated coefficients, standard deviations, t-statistics, and two-sides p-values are reported; *** indicates significance at the 0.1% level, ** at the 1% level, * at the 5% level, and · at the 10% level



(a) natural gas slope in terms of crude oil slope



(b) natural gas level in terms of crude oil level

Figure 4: Left: natural gas slope in terms of crude oil slope; the linear fit is displayed in dotted lines; Right: natural gas level in terms of crude oil level; the linear fit is displayed in dotted lines and the best three-lines fit is displayed in bold lines; in both graphs, the sequence of colors red, yellow, green, blue, purple, red marks the passage of time from January 1999 to July 2005

Figure (4(b)) displays the relation between natural gas and crude oil levels: a piecewise-linear relation appears, with two break points occurring at the beginning of year 2000, where gas long-term futures prices surge[13], and in the middle of year 2003[14], where oil long-term futures prices start rising sharply.

Tables (5) and (6) report the results of the classical linear regression and the piecewise-linear regression of gas level on crude oil level. First, we see that the R^2 is much higher than for the regression on the slopes: the long-term equilibrium between the levels is much stronger than the long-term relation between the slopes. Second, the piecewise-linear regression coefficients are significant, which confirms the validity of the piecewise linear model, and of the same negative sign, causing the gas long-term price to be less sensitive to the variations of oil long-term price above the up-threshold $\bar{Y}$ and below the down-threshold $\underline{\mathrm{Y}}$. Lastly, Table (7) shows that the unit-root hypothesis H_0 can be rejected by the Phillips-Perron test but not by the Augmented-Dickey-Fuller test for the residuals of the piecewise linear relation between gas and oil levels; in addition, H_0 cannot be rejected by either test for the residuals of the linear relation between gas and oil levels. As a conclusion, only the piecewise linear relation allows one to obtain the desired stationary residuals.

	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
b_{crude}	-0.0548	0.00714	-7.672	2.89^{-14}	***
a_{crude}	0.972	0.00654	148.568	$< 2.10^{-16}$	***
					$R^2 = 93.09\%$

Table 5: Linear regression of natural gas level on crude oil level: a denotes the regression coefficient and b the intercept; estimated coefficients, standard deviations, t-statistics, and two-sided p-values are reported; *** indicates significance at the 0.1% level, ** at the 1% level, * at the 5% level, and · at the 10% level

[13]Several events triggered this run-up: the oil price rise, setting a higher backstop price for natural gas, the lack of drilling activity in the previous years due to low gas prices, the hot weather in the Southwest and reduced hydroelectric generation, and lastly the resumed growth of gas consumption in the industrial sector

[14]During this period, corresponding to the invasion of Iraq, the oil market spare capacity declined due to the loss of production capacity in Iraq and Venezuela and to the growing international demand

	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
b_{crude}	-0.586	0.0139	-42.28	$< 2.10^{-16}$	***
a_{crude}	1.627	0.0160	101.72	$< 2.10^{-16}$	***
a^{-}_{crude}	-0.982	0.0222	-44.29	$< 2.10^{-16}$	***
a^{+}_{crude}	-1.250	0.0387	-32.32	$< 2.10^{-16}$	***
		$\underline{Y}_{crude} = 0.474$	$\bar{Y}_{crude} = 1.0668$	$R^2 = 96.86\%$	

Table 6: Piecewise-linear regression of natural gas level on crude oil level; the regression variables are Y_e, $(Y_e - \underline{Y}_e)^- = Min(0; Y_e - \underline{Y}_e)$, and $(Y_e - \bar{Y}_e)^+ = Max(0; Y_e - \bar{Y}_e)$, with e=crude oil; a denotes the different regression coefficients and b the intercepts; the thresholds $\underline{Y}_e$ and $\bar{Y}_e$ are determined by the minimization over the couples $(\underline{Y}_e, \bar{Y}_e)$ of the sum of squared residuals of the regression of Y_{gas} on the variables Y_e, $(Y_e - \underline{Y}_e)^-$, and $(Y_e - \bar{Y}_e)^+$; the estimated coefficients, standard deviations, t-statistics, and two-sided p-values are reported; *** indicates significance at the 0.1% level, ** at the 1% level, * at the 5% level, and $\cdot$ at the 10% level

Piecewise linear relation			
	Dickey-Fuller	Lag Parameter	p-value
PP	-3.831	8	0.0174
ADF	-2.902	11	0.197
Linear relation			
	Dickey-Fuller	Lag Parameter	p-value
PP	-2.031	8	0.565
ADF	-1.586	11	0.753

Table 7: Phillips-Perron and Augmented-Dickey-Fuller unit root tests on the residuals of the piecewise linear (up) and linear (down) relations between gas and oil levels; the test-statistics, truncation lag parameters, and p-values of the test are reported

4 A new dependence model for pairs of commodity forward curves

4.1 Formulation of the model

We want to introduce an error-correction mechanism on the levels and on the slopes between the energies e and e'. Therefore, we postulate that the drifts are the sums of a constant part, a term expressing dependence on past returns (with a maximal lag of one day[15], and an error-correction term:

$$\begin{pmatrix} \Delta X_t^e \\ \Delta X_t^{e'} \\ \Delta Y_t^e \\ \Delta Y_t^{e'} \end{pmatrix} = \begin{pmatrix} \mu_{X,e} \\ \mu_{X,e'} \\ \mu_{Y,e} \\ \mu_{Y,e'} \end{pmatrix} + \Gamma \begin{pmatrix} \Delta X_{t-1}^e \\ \Delta X_{t-1}^{e'} \\ \Delta Y_{t-1}^e \\ \Delta Y_{t-1}^{e'} \end{pmatrix} + \Pi \begin{pmatrix} \bar{X}_{t-l_X}^e \\ \bar{X}_{t-l_X}^{e'} \\ R_{t-l_Y}^Y \end{pmatrix} + \begin{pmatrix} \epsilon_t^{X,e} \\ \epsilon_t^{X,e'} \\ \epsilon_t^{Y,e} \\ \epsilon_t^{Y,e'} \end{pmatrix} \tag{12}$$

$$R_t^Y = Y_t^e - f_Y^{e,e'}(Y_t^{e'})$$

In the model (12):

- e stands for natural gas and e' stands for crude oil
- $\mu = (\mu_{X,e}, \mu_{X,e'}, \mu_{Y,e}, \mu_{Y,e'})$ is the 1×4 vector composed of the constant part of the drifts
- Γ is a 4×4 matrix expressing dependence on past returns
- $\bar{X}_t^e$ and Y_t^e denote the slope and level of the forward curve of the energy e
- $x \to f_Y^{e,e'}(x)$ is the relation between the levels of energy e and e' (in the case of gas and oil, f_Y is piecewise linear function)
- (R_t^Y) is the process composed of the deviations to the long-term relation between the levels
- Π is a 4×3 matrix expressing sensitivity to the slopes and deviations to the long-term relation between the levels
- l_X (resp. l_Y) refers to the lags between the observed slopes (resp. level deviations) and the corrections mechanisms
- the processes $(\epsilon_t^{X,e} = \sigma_t^{X,e}\eta_t^{X,e}, \epsilon_t^{X,e'} = \sigma_t^{X,e'}\eta_t^{X,e'}, \epsilon_t^{Y,e} = \sigma_t^{Y,e}\eta_t^{Y,e}, \epsilon_t^{Y,e'} = \sigma_t^{Y,e'}\eta_t^{Y,e'})$ follow *independent GARCH processes*; we include a seasonal component in the GARCH process followed by natural short-term shocks

[15]The inspection of the cross-correlation functions between the different shocks reveals that the shocks are not dependent on past shocks over a lag of one day

- the residual shocks $(\eta_t^{X,e}, \eta_t^{X,e'}, \eta_t^{Y,e}, \eta_t^{Y,e'})$ are assumed to be i.i.d

We use the 4×1 vector process $\Delta Z_t = (\Delta X_t^e, \Delta X_t^{e\prime}, \Delta Y_t^e, \Delta Y_t^{e\prime})'$ and the 3×1 *state vector process* $\xi_t = \left(\bar{X}^e_{t-l_X}, \bar{X}^{e'}_{t-l_X}, R^Y_{t-l_Y}\right)'$. A few comments are required here. First, keeping in mind equations (10) and (11), assuming linear function f_Y, and making abstraction of the dependence between $(\bar{X}_t^e, \bar{X}_t^{e'})$ and $(Y_t^e, Y_t^{e'})$ induced by the terms $\Gamma \Delta Z_{t-1}$ and $\Pi \xi_t$, the model (12) implies a vector autoregressive model (VAR) for the slopes and a vector error-correction model (VECM) for the levels, which makes sense from an economic standpoint. Second, we believe that the model (12) is sufficiently general to account for the evolution of any pair of related commodity forward curves, with an appropriate long-term relation $f_Y^{e,e'}$. However, as we choose to model the processes $(\sigma_t^{X,e}\eta_t^{X,e}, \sigma_t^{X,e'}\eta_t^{X,e'}, \sigma_t^{Y,e}\eta_t^{Y,e}, \sigma_t^{Y,e'}\eta_t^{Y,e'})$ as independent GARCH processes, we exclude from our scope the relations slope/volatility (which are studied by Ates and Wang (2005) in the US gas market) and the effect of *volatility transmission* between the two commodity prices, an effect which was highlighted before in the literature in the case of gas and oil markets (see Pindyck (2004) and Ewing et al. (2003)). Lastly, we assume a constant dependence structure between the residuals $(\eta_t^{X,e}, \eta_t^{X,e'}, \eta_t^{Y,e}, \eta_t^{Y,e'})$, thus neglecting the possible correlation clustering (see Eydeland and Wolyniec (2003)) and the potential relations between correlation and volatilities (see e.g. Goorbergh et al. (2005)).

4.2 Calibration of the model

To calibrate the model, we proceed in three steps: first, we find the lags l_X and l_Y and we estimate α and Γ by a linear regression of ΔZ_t on ΔZ_{t-1} and ξ_t; second, we apply independent GARCH models to the residuals of this linear regression; third, we study the dependence structure between the standardized residuals of the independent GARCH models. This decomposed procedure, which is also adopted by Ng and Pirrong (1994) and Ates and Wang (2005), was motivated by the high number of parameters to be estimated.

4.2.1 Estimation of l_X, l_Y, Γ and Π

Figure (5) represent the cross-correlations between the gas shocks and the state variables ξ_t. We therefore choose $l_X = 4$ and $l_Y = 6$. Tables (8)-(9) report the results of the four linear regressions. Regarding the cross-energy dependence on past shocks, we find that the causality generally runs from oil to natural gas and is negative[16]. Regarding the inter-temporal dependence on past shocks, the causality runs both ways between the short-term and the long-term, is positive (resp. negative) in the direction long-term $\hookrightarrow$ short-term for oil (resp. gas) and negative in the direction short-term $\hookrightarrow$ long-term for both energies[17]. Regarding the auto-correlation of shocks, we find that in the short-term, oil and gas markets tend to amplify the previous move, whereas in the long-term, they are more likely to correct it.

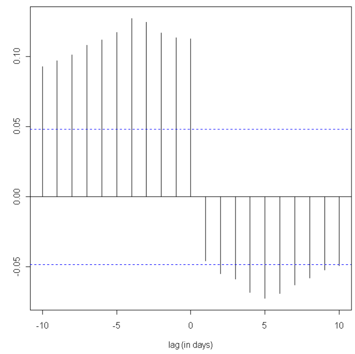
The most important results concern the reaction of the shocks to the state variables (ξ_t). As regards the impact of the slopes, we find that crude oil short-term shocks tend to correct the spread between gas and oil slopes and that gas long term shocks react positively to a positive spread between gas and oil slopes, thus having an amplification effect on the spread between the two curves. Concerning the impact of the levels, we find that the gas short-term shocks react negatively to an overvalued natural gas long term price[18], whereas the gas long-term shocks correct the deviations to the long-term equilibrium on the levels; note that the crude oil long-term shocks are not sensitive to the different state variables (ξ_t). We conclude that gas (resp. crude oil) plays the leading role in the slope (resp. long term price) discovery.

The low value of the R^2 in the four regressions shows that the forecasting power of the model is however relatively low as concerns daily shocks.

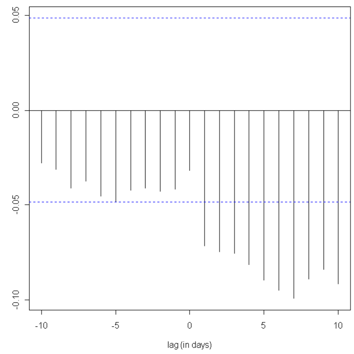
[16]Remember the previous remark on the negative causality originating from the supply side made in section 2

[17]This *substitution effect* between short and long term shocks has a stabilizing impact on the futures prices

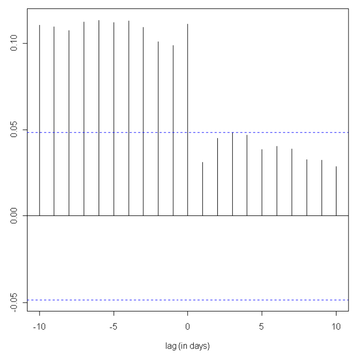
[18]therefore there is a feedback effect balancing the previous amplification effect



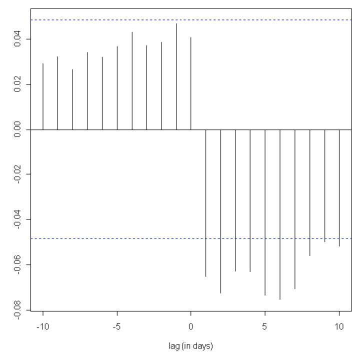
(a) gas slope/gas short-term shocks



(b) residuals of the long-term relation on the levels/gas short-term shocks



(c) gas slope/gas long-term shocks



(d) residuals of the long-term relation in the levels/gas long-term shocks

Figure 5: Cross correlation functions between the gas shocks and the state variables ξ_t with lags of 1 to 10 days; the cross correlation functions with lag i (resp. $-i$) represent the correlation between the state variables at time t and the shocks at time $t+i$ (resp. $t-i$)

	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
$\mu_{X,gas}$	-0.000250	0.000737	-0.339	0.735	
$\Gamma_{1,1}$	0.106	0.0284	3.728	0.000200	***
$\Gamma_{1,2}$	-0.208	0.0515	-4.044	0.0000549	***
$\Gamma_{1,3}$	-0.198	0.0565	-3.495	0.000486	***
$\Pi_{1,3}$	-0.0304	0.00803	-3.786	0.000159	***
					$R^2 = 2.82\%$

Wald test for gas

DF	F	$Pr(> F)$
3	0.969	0.407

	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
$\mu_{X,crude}$	0.00167	0.000623	2.668	0.00771	*
$\Gamma_{2,1}$	0.0237	0.0139	1.700	0.0893	.
$\Gamma_{2,2}$	0.104	0.0267	3.891	0.000104	***
$\Gamma_{2,3}$	-0.0824	0.0280	-2.943	0.00330	**
$\Gamma_{2,4}$	0.143	0.0275	5.195	$2.31.10^{-7}$	***
$\Pi_{2,1}$	0.00431	0.00233	1.847	0.0649	.
$\Pi_{2,2}$	-0.0114	0.00406	-2.814	0.00494	**
					$R^2 = 4.34\%$

Wald test for crude

DF	F	$Pr(> F)$
1	2.12	0.146

Table 8: Linear regression of the natural gas (resp. crude oil) short-term shocks on $(\Delta Z_{t-1})_{1,2,3}$ and $(\xi_t)_3$ (resp. $(\Delta Z_{t-1})_{1,2,3,4}$ and $(\xi_t)_{1,2}$); the estimated coefficients, standard deviations, t-statistics, and two-sided p-values are reported; *** indicates significance at the 0.1% level, ** at the 1% level, * at the 5% level, and $\cdot$ at the 10% level; the Wald test for gas (resp. crude) tests the null hypothesis that $\Gamma_{1,4} = \Pi_{1,1} = \Pi_{1,2} = 0$ (resp. $\Pi_{2,3} = 0$)

	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
$\mu_{Y,gas}$	0.00229	0.000679	3.372	0.000765	***
$\Gamma_{3,1}$	−0.0500	0.0127	−3.939	0.0000855	***
$\Gamma_{3,2}$	−0.0525	0.0259	−2.028	0.0427	*
$\Pi_{3,1}$	0.0116	0.00286	4.060	0.0000513	***
$\Pi_{3,2}$	−0.00892	0.00450	−1.983	0.0475	*
$\Pi_{3,3}$	−0.0243	0.00480	−5.063	$4.59.10^{-7}$	***
					$R^2 = 2.72\%$

Wald test for gas

DF	F	$Pr(>F)$
2	1.083	0.339

	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
$\mu_{Y,crude}$	0.00150	0.000360	4.180	0.0000307	***
$\Gamma_{4,2}$	−0.105	0.0265	−3.977	0.0000729	***
$\Gamma_{4,4}$	−0.0907	0.0266	−3.411	0.000664	***
					$R^2 = 2.58\%$

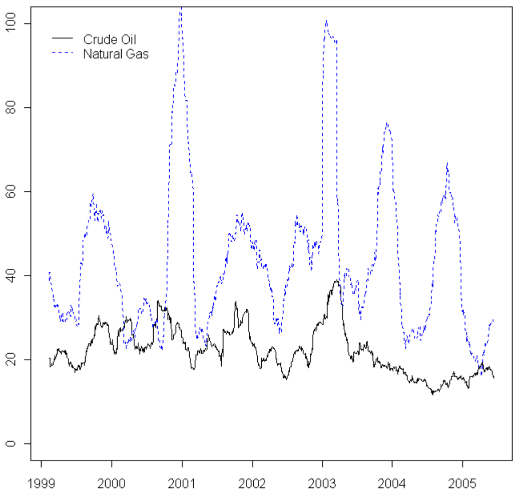
Wald test for crude

DF	F	$Pr(>F)$
5	0.148	0.981

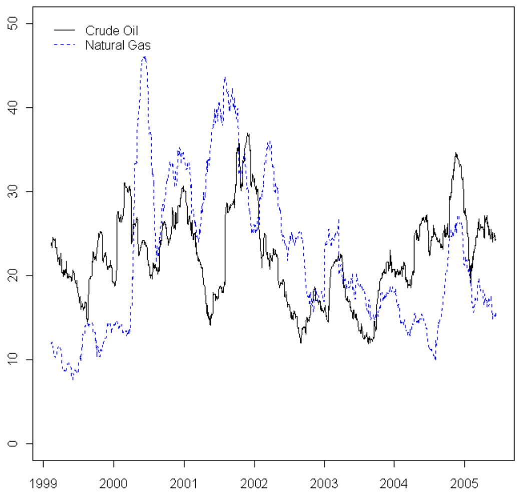
Table 9: Linear regression of the natural gas (resp. crude) long-term shocks on $(\Delta Z_{t-1})_{1,2}$ and (ξ_t) (resp. $(\Delta Z_{t-1})_{2,4}$); the estimated coefficients, standard deviations, t-statistics, and two-sided p-values are reported; *** indicates significance at the 0.1% level, ** at the 1% level, * at the 5% level, and · at the 10% level; the Wald test for gas (resp. crude) tests the null hypothesis that $\Gamma_{3,3} = \Gamma_{3,4} = 0$ (resp. $\Gamma_{4,1} = \Gamma_{4,3} = \Pi_{4,1} = \Pi_{4,2} = \Pi_{4,3} = 0$)

4.2.2 GARCH models for the volatilities

For the two energies, the volatilities of the short-term and long-term shocks are estimated by the standard deviation of the shocks within a 50-days sliding window. The obtained trajectories are displayed on figure (6): all shocks exhibit volatility clusters, jumps, and the natural gas short-term volatility follow a seasonal pattern, with high values in the winter (60 % in normal winters) and lower values in the summer (20 % on average). The phenomenon of stochastic volatility, observed in most commodity markets, is linked to the temporal variations of some key indicators of the supply flexibility, such as the deviation to "normal" storage level, and the proportion of spare production/refining capacity. Note also that the short-term volatility peaks correspond to periods of high positive forward curve slopes, an observation which is consistent with the theory of storage (Kaldor (1939)), and which was also made by Ates and Wang (2005) in the US gas market. The seasonal pattern of natural gas short-term volatility can be explained by the fact that the demand is more sensitive to the temperature during the heating season and that the demand and production shocks have more impact on the prices during the winter, when storage is part of the supply curve and the market is tight, than during the summer, when storage is part of the demand curve and the market is loose. The seasonal behavior of gas implicit volatilities was already observed by Blix (2003) in the US gas market.



(a) short-term volatilities



(b) long-term volatilities

Figure 6: Annualized short-term and long-term volatilities (in %) of the two energies estimated with a 50-days sliding window

In this section, we model the volatility processes of the residuals of the four previous regressions (i.e. the processes (ϵ_t) in model (12)). Box-Pierce tests on the processes (ϵ_t^2) show the heteroscedasticity of the different residuals:

data	X^2	df	p-value
$(\epsilon_t^{X,gas})^2$	19.424	1	$1.047.10^{-5}$
$(\epsilon_t^{Y,gas})^2$	17.276	1	$3.233.10^{-5}$
$(\epsilon_t^{X,crude})^2$	9.619	1	0.00193
$(\epsilon_t^{Y,crude})^2$	26.693	1	$2.385.10^{-7}$

Table 10: Box-Pierce tests on the processes $(\epsilon_t^{X,gas})^2$,$(\epsilon_t^{Y,gas})^2$,$(\epsilon_t^{X,crude})^2$, $(\epsilon_t^{Y,crude})^2$, $(\epsilon_t^{X,heat})^2$, and $(\epsilon_t^{Y,heat})^2$; the test-statistics, degrees of freedom of the approximate chi-square distribution of the test statistics, and p-values of the tests are reported

Moreover, figure (6) exhibits a significant seasonal component in the natural gas short-term volatility. The following seasonal GARCH model, proposed by Diebold (2003) for the modeling of temperature series, accounts for this phenomenon:

$$\begin{aligned}\epsilon_t &= \sigma_t\eta_t \\ \sigma_{t+1}^2 &= a_1\epsilon_t^2 + b_1\sigma_t^2 + (a_0 + Acos(2\pi t/252) + Bsin(2\pi t/252)) \\ (\eta_t) &\quad i.i.d\end{aligned} \tag{13}$$

Note that the volatility of volatility is itself seasonal since the volatility shocks $\epsilon_t^2 = \sigma_t^2\eta_t^2$ have different average winter and summer values. This characteristic is compatible with the observation of gas short-term volatility, which mostly cluster during the winters (see figure (6)). This model was calibrated by Quasi-Maximum Likelihood on natural gas short-term residuals $(\epsilon_t^{X,gas})$. The log-likelihood of the model in the case of Gaussian residuals (η_t) is:

$$LL = -\sum_{t=1}^{N}\left(\frac{\epsilon_t^2}{2\sigma_t^2} + log(\sigma_t)\right) - \frac{N}{2}log(2\pi)$$

The calibration of model (13) (available on request) reveals that the coefficient A is not significant. The estimation of the model imposing $A = 0$ is provided in Table (11). Table (12) reports the results of the

	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
a_1	0.155	0.0108	14.377	$< 2.10^{-16}$	***
b_1	0.752	0.0367	20.528	$< 2.10^{-16}$	***
a_0	0.0000884	0.0000228	3.883	0.000103	***
B	−0.0000491	0.0000152	−3.234	0.00122	**

Table 11: Quasi-Maximum-Likelihood estimation of a seasonal GARCH model on $(\epsilon_t^{X,gas})$; the estimated coefficients, standard deviations, t-statistics, and two-sided p-values are reported

Jarque-Bera (resp. Box-Pierce) tests on the residuals (resp. squared residuals) of the seasonal GARCH model. The Jarque-Bera tests allow us to reject the hypothesis of Gaussian residuals. By contrast, we cannot reject the hypothesis of independence for the squared residuals, which is an indication of the validity of the model. Figure (7) plots the trajectories of $(\epsilon_t^{X,gas})^2$ together with the variance $(\sigma_t)^2$ predicted by the seasonal GARCH models and the long-term seasonal variance functions $a_0 + Bsin(2\pi t/252)$.

Jarque-Bera			
data	X^2	df	p-value
$\eta_t^{X,gas}$	2354.378	2	$2.2.10^{-16}$
Box-Pierce			
data	X^2	df	p-value
$(\eta_t^{X,gas})^2$	0.0069	1	0.934

Table 12: Jarque-Bera and Box-Pierce tests on the residuals of the seasonal GARCH model for natural gas; the test statistics, degrees of freedom, and p-values of the tests are reported

We model the other series by classical GARCH models, whose implementation is not reported here.

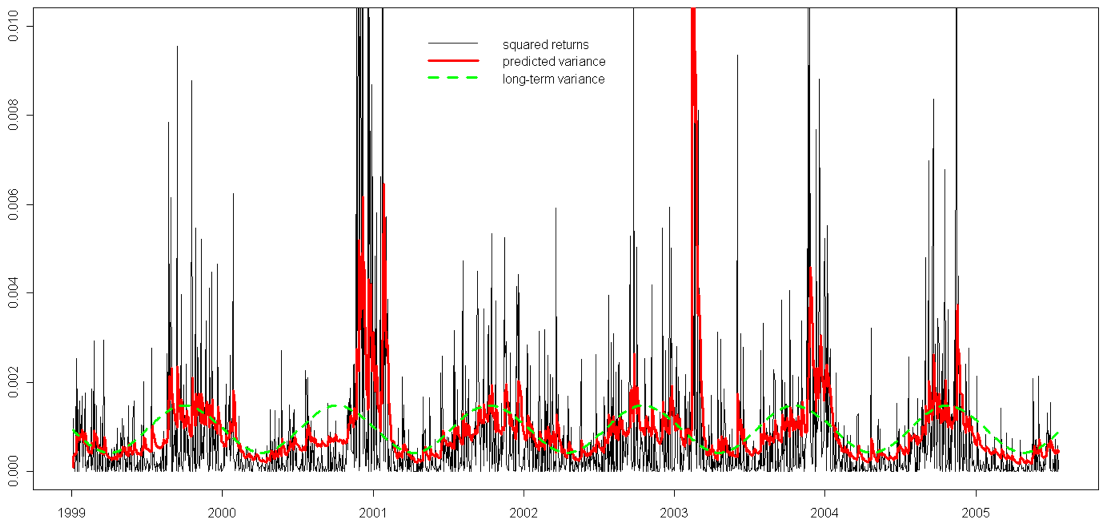


Figure 7: Trajectories of $(\epsilon_t^{X,gas})^2$ (black), variance σ_t^2 predicted by a seasonal GARCH model (red), and long-term variance $\frac{1}{1-a_1-b_1}(a_0 + Bsin(2\pi t/252))$ (green)

4.2.3 Dependence structure of the standardized co-movements

We model the dependence structure of the residuals (η) using the copula representation:

$$\mathbb{P}(\eta_t^{X,e} \leq z_1, \eta_t^{X,e'} \leq z_2, \eta_t^{Y,e} \leq z_3, \eta_t^{Y,e'} \leq z_4) = C(F^{X,e}(z_1), F^{X,e'}(z_2), F^{Y,e}(z_3), F^{Y,e'}(z_4)) \tag{14}$$

where the copula function C is defined in $[0;1]^4$ with values in $[0;1]$, and $(F^{X,e}, F^{X,e'}, F^{Y,e}, F^{Y,e'})$ denote the marginal distributions of the residuals $(\eta^{X,e}, \eta^{X,e'}, \eta^{Y,e}, \eta^{Y,e'})$. We will use here the Gaussian copula defined by:

$$C(u_1, u_2, u_3, u_4) = \Phi_\rho^4(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \Phi^{-1}(u_3), \Phi^{-1}(u_4)) \tag{15}$$

where Φ_ρ^4 is a 4-variate normal distribution of correlation matrix ρ, and Φ^{-1} is the inverse of the univariate standard normal distribution. As explained in Joe and Xu (1996), the calibration of the model (14) is done in two steps:

- we first fit the marginal densities $(f^{X,e}, f^{X,e'}, f^{Y,e}, f^{Y,e'})$ of the different shocks using the Skewed Generalized Error Distribution (SGED); the obtained parameters and goodness-of-fit results, which are not reported here, indicate the relevance of this representation

- once the marginal distributions are determined, the correlation matrix ρ is estimated by computing the empirical Kendall's τ[19] for each pair, which is linked to the correlation matrix of the Gaussian copula through the relation $\tau = \frac{2}{\pi}\, arcsin(\rho)$ (see Lindskog et al. (2003))

The estimates of matrix correlations ρ for the pair gas/crude oil are reported on tables (13). We note that all correlations are significantly positive.

	short-term gas	short-term crude	long-term gas	long-term crude
short-term gas	1	0.282	0.623	0.294
short-term crude	0.282	1	0.300	0.470
long-term gas	0.623	0.300	1	0.380
long-term crude	0.294	0.470	0.380	1

Table 13: Estimation of the correlation matrix of the Gaussian copula

4.3 Evolution of the correction mechanisms

Table (14) reports the separate estimations of matrix Π on the periods January 1999-December 2001, January 2002-December 2003 and January 2004-July 2005, and compares them with the global estimator on the whole period January 1999-July 2005: we observe first that the signs of the coefficients $\Pi_{1,3}$ and $\Pi_{3,3}$, expressing the correction of the deviations to the long-term relation between gas and oil levels, have been stable throughout the period under study; conversely, the pairs $(\Pi_{2,1}, \Pi_{2,2})$ and $(\Pi_{3,1}, \Pi_{3,2})$, expressing the sensitivity to the gas and oil slopes have had a different behavior on the most recent period, the former switching signs, and the latter becoming non-significant.

[19] the empirical Kendall's τ expressing dependance between two samples (X_t) and (Y_t) is computed by $\tau(X,Y) = \frac{1}{C_n^2}\sum_{1\leq i_1 \leq i_2 \leq N} sign(X_{i_1} - X_{i_2})(Y_{i_1} - X_{i_2})$

	Global	Jan 1999-Dec 2001	Jan 2002-Dec 2003	Jan 2004-Jul 2005
$\Pi_{1,3}$	-0.0304 ***	-0.0264 **	-0.0351 0.103	-0.0835 ·
$\Pi_{2,1}$	0.00431 ·	0.00657 *	0.0103 ·	-0.00921 ·
$\Pi_{2,2}$	-0.0112 **	-0.0167 **	-0.0319 **	0.00648 0.325
$\Pi_{3,1}$	0.0116 ***	0.0150 ***	0.0127 ·	0.00113 0.871
$\Pi_{3,2}$	-0.00892 *	-0.00948 0.157	-0.0203 ·	0.00436 0.571
$\Pi_{3,3}$	-0.0243 ***	-0.0241 ***	-0.0303 *	-0.0611 *

Table 14: Estimation of the error-correction parameters on the three periods; *** indicates significance at the 0.1% level, ** at the 1% level, * at the 5% level, and · at the 10% level; p-values above 10% are reported below the estimated parameters

4.4 Evolution of the correlations

The objective here is to study the stability of the dependence structure which was found between the forward curves co-movements. Figure (8) represents the temporal evolution of the correlation between gas and oil short-term (resp. long-term) shocks derived from Kendall's τ[20] (with a one-year sliding window). Both correlations display an upward trend on the period. Possible explanations for this observation are the correlation induced by the growing investment of hedge funds in the commodity asset class and the fact that, in commodity markets, in contrast to equity markets, correlation is generally bigger when prices are rising. To account for this trend in the dependence structure, we have estimated (following Rockinger and Jondeau (2001)) a bi-variate normal copula model with a quadratic trend in the correlation coefficient[21] for the two pairs gas short-term shocks/oil short-term shocks and gas long-term shocks/oil long-term shocks:

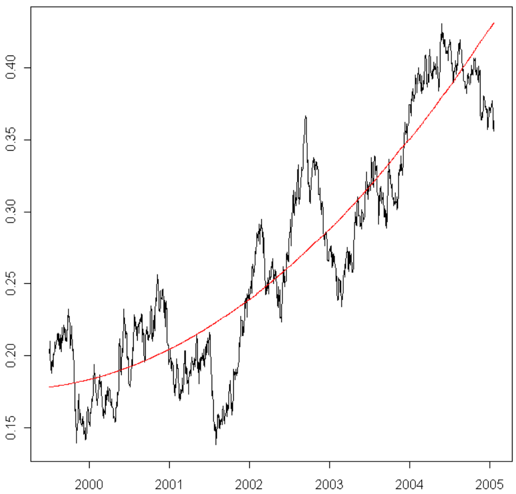
$$C_t(u_1, u_2) = \Phi^2_{\rho_t}(\Phi^{-1}(u_1), \Phi^{-1}(u_2)) \quad (16)$$

$$\rho_t = a + bt^2 \quad (17)$$

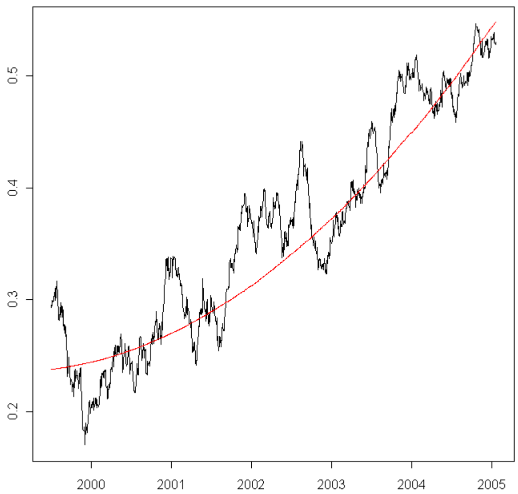
where Φ^2_ρ is the bivariate normal distribution with correlation coefficient ρ, and a, b are unknown parameters describing the temporal evolution of the correlation coefficient. The log-likelihood of the copula

[20] $\tau = \frac{2}{\pi}\, arcsin(\rho)$

[21] The term expressing linear dependence with respect to time is not included here in the model as it was not found significant



(a) short-term gas/short-term crude



(b) long-term gas/long-term crude

Figure 8: $sin(\frac{\pi}{2}\tau)$ as a function of time (with a one-year sliding window) and quadratic trend $a + bt^2$ estimated by maximum-likelihood

model (16)-(17) is:

$$LL(\rho, F^e, F^{e'}) = \sum_{t=1}^{N} ln\left[c_t(F^e(\eta_t^e), F^{e'}(\eta_t^{e'}))\right] + \sum_{t=1}^{N}\left(ln\left[f^e(\eta_t^e)\right] + ln\left[f^{e'}(\eta_t^{e'})\right]\right) \tag{18}$$

In (18), $(f^e, f^{e'})$ are the univariate densities, $(F^e, F^{e'})$ are the univariate cumulative distributions, $(\eta_t^e)_{1\leq t\leq N}$ are the observations for energy e, and c_t is the density of the bi-variate normal copula of correlation coefficient $\rho_t = a + bt^2$:

$$\begin{aligned} c_t(u_1, u_2) &= \frac{\partial^2 C_t}{\partial u_1 \partial u_2} \\ &= \frac{\phi_{\rho_t}(\Phi^{-1}(u_1), \Phi^{-1}(u_2))}{\phi(\Phi^{-1}(u_1))\phi(\Phi^{-1}(u_2))} \end{aligned}$$

where $\phi_\rho(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2\sqrt{1-\rho^2}}\left[x^2 + y^2 - 2\rho xy\right]\right)$ is the bi-variate normal density for correlation coefficient ρ, and ϕ is the density of the $N(0, 1)$ distribution. The parameters a and b maximizing the term $\sum_{t=1}^{N} ln\left[c_t(F^e(\eta_t^e), F^{e'}(\eta_t^{e'}))\right]$ in (18) are reported on Table (15) and the obtained quadratic trend plotted in Figure (8).

gas/crude short-term shocks					
	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
a	0.176	0.0337	5.29	$1.25.10^{-7}$	***
b	$1.12.10^{-7}$	$2.64.10^{-8}$	4.25	$2.16.10^{-5}$	***
gas/crude long-term shocks					
	Estimate	Std. Error	t value	$Pr(>\|t\|)$	
a	0.235	0.0321	7.33	$2.34.10^{-13}$	***
b	$1.379.10^{-7}$	$2.19.10^{-8}$	6.30	$3.06.10^{-10}$	***

Table 15: Maximum-likelihood estimation of the quadratic trend in the correlation between oil and gas

5 Conclusion

This paper has presented a new dependence model for commodity forward curves. Like popular models on single commodity forward curves, it decomposes the forward curve moves into a short-term and a long-term shocks, with stochastic and possibly seasonal volatilities. The correlation between the shocks of the two curves is captured through a non-Gaussian dependence structure. The originality of the model is that, in addition to this local dependence structure, it accounts for the long-term relations between the commodity forward prices through an error-correction term in the risk-premia of the forward price returns. The long-term relations are based on the state variables describing the shape of a forward curve under the two-factor model, namely the slope and level. Our current research concerns the modeling of stochastic dependence structure, and the implications of the model for multi-commodity asset pricing, risk measurement, and portfolio optimization. As far as asset pricing is concerned, Duan and Pliska (2004) have shown that the combination of cointegration and stochastic volatility has an impact on asset prices: thus the model would lead to different pricing results than standard local dependence models without risk-premia. Regarding risk management, the model, because it captures the long-term relations between two curves, allows one to realistically simulate portfolios' Earning-at-Risk on a long-term perspective. With respect to portfolio optimization, our error-correction model allows the portfolio manager to forecast the relative evolutions of the two considered forward curves given their initial slopes and levels, a property which has numerous implications; hedge funds will be provided with directional strategies based on long/short positions on the two curves while physical portfolio managers will have a way to choose the best moments to lock in the margin of their assets with futures contracts.

6 References

Alexander C. (1999), *Correlation and Cointegration in Energy Markets* in *Managing Energy Price Risk* (2. nd. Edition) V. Kaminsky (ed.).

Ates A., Wang G. (2005), *Storage, Weather and Dynamics of Natural Gas Prices in Futures and Spot Markets*, Working Paper

American Gas Foundation (2003), *Natural Gas and Energy Price Volatility*, Energy and Environmental Analysis, Inc.

Blix M. (2003), *The volatility Structure of Gas Futures Contracts*, Working Paper, Stockholm School of Economics

Brooks R. (2001), *Value at Risk Applied to Natural Gas Forward Contracts*, WP01-08-01, Departmentof Economics, Finance and Legal Studies, University of Alabama

Campbell D., Diebold F. (2003), *Weather Forecasting for Weather Derivatives*, NBER Working Papers 10141, National Bureau of Economic Research, Inc.PIER Working Paper No. 01-031.

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

Diebold F., Li C. (2003) *Forecasting the Term Structure of Government Bond Yields*, NBER Working Papers 10048, National Bureau of Economic Research, Inc

Duan J., Pliska S. (2004), *Option Valuation with Co-integrated Asset Prices*, Journal of Economic Dynamics and Control, 28(4), 727-54

Duffie D. (1999), *Volatility in Energy Prices* in *Managing Energy Price Risk* (2. nd. Edition) V. Kaminsky (ed.).

Energy Information Administration (2001), *Impact of Interruptible Natural Gas Service on Northeast Heating Oil Demand*

Ewing B., Malik F., Ozfidan O. (2002), *Volatility Transmission in the Oil and Natural Gas Markets*, Energy Economics 24 (2002) 525-538

Eydeland A., Wolyniec K. (2003), *Energy and Power Risk Management: New Developments in Modeling,* *Pricing, and Hedging*, WileyEurope

Fama E., French K. (1988), *Business Cycles and the Behavior of Metals Prices*, Journal of Finance, Vol. 43, No. 5, pp. 1075-1093

Geman H. (2005), *Commodities and Commodity Derivatives : Modelling and Pricing for Agriculturals, Metals and Energy*, Wiley Finance

Geman H., Nguyen V. (2005), *Soybean Inventory and Forward Curves Dynamics*, Management Science, Fall 2005

Goorbergh R. van den, Genest C., Werker B. (2005), *Bivariate option pricing using dynamic copula models*, Insurance: Mathematics and Economics, 37, 101-114.

Joe H., Xu J. (1996), *The estimation method of inference functions for margins for multivariate models*, Technical Report no. 166., Vancouver: University of British Columbia, Department of Statistics.

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

Lindskog F., McNeil A., Schmock U. (2003), *Kendall's Tau for Elliptical Distributions*, In Credit Risk measurement, evaluation and management, Bol, Nakhaeizade et al., eds. PhysicaVerlag, Heidelberg, pp. 149156.

Manoliu M., Tompaidis S. (2002), *Energy Futures Prices: Term Structure Models with Kalman Filter Estimation* Applied Mathematical Finance, vol. 9, issue 1, pages 21-43

Ng V., Pirrong C. (1994), *Fundamentals and Volatility: Storage, Spreads and Dynamics of Metal Prices*, Journal of Buisness, 67 (2) 203-230

N'guyen V. (2004), *Etude des marchés de matières premières : Valorisation et couverture d'instruments dérivés*, PhD Thesis, University Paris Dauphine

Pekka M., Antti K. (2005), *Evaluating Multivariate GARCH models in the Electricity Markets*, Helsinki School of Economics, Working Paper, W-382

Pilipovic D. (1997), *Energy Risk: Valuing and Managing Energy Derivatives*, Mc Graw Hill

Pindyck R. (2004), *Volatility in Natural Gas and Oil Markets*, Journal of Energy and Development

Richter C., Sorensen C. (2000), *Stochastic Volatility and Seasonality in Commodity Futures and Options: The Case of Soybeans*, Department of Finance, Copenhagen Business School, Working Paper

Rockinger M., Jondeau E. (2006), *The Copula-GARCH Model of Conditional Dependencies: An International Stock-Market Application*, Journal of International Money and Finance

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

Siliverstovs B., Neumann A., LHégaret G., Hirschhausen C. (2005), *International Market Integration for Natural Gas? A Cointegration Analysis of Gas Prices in Europe, North America and Japan*, Energy Economics, Vol. 27, No. 4, 603-615.