EconomiX

http://economix.fr

Document de Travail

Working Paper

2012-28

Testing for crude oil markets globalization during extreme price movements

Bertrand Candelon
Marc Joëts
Sessi Tokpavi

UMR 7235

Université de Paris Ouest Nanterre La Défense
(bâtiment G)
200, Avenue de la République
92001 NANTERRE CEDEX

Tél et Fax : 33.(0)1.40.97.59.07
Email : nasam.zaroualete@u-paris10.fr

université Paris Ouest Nanterre La Défense

Testing for crude oil markets globalization during extreme price movements

Bertrand Candelon[*], Marc Joëts[†], Sessi Tokpavi[‡]

March, 2012

Abstract

This paper investigates the global crude oil market dependence during extreme price movements. To this aim we extend the univariate Granger causality test in extreme risk developed by Hong et al. (2009) in a multivariate context. Asymptotic as well as finite sample properties are delivered. Applying this test for 32 crude oil markets, it turns out that extreme price movements are governed by non-OPEC crude oil markets rather than OPEC ones. More precisely, WTI and Brent crude oils are price setters in both extreme downside and upside price movements. More surprisingly, Mediterranean Russian Urals and Europe Forcados (resp. Ecuador Oriente) rather than Dubai Fateh act as additional benchmarks in periods of extreme price falls (resp. rises). Moreover, the integration process between crude oil markets seems to decrease during extreme price movements making diversification strategies more feasible.

Keywords: Crude oil markets; Risk transmission; Globalization; Distribution tails; Granger-causality test.

[*]b.candelon@maastrichtuniversity.nl. Maastricht University, Department of Economics. The Netherlands.

[†]marc.joets@u-paris10.fr, EconomiX-CNRS, University of Paris Ouest, France.

[‡]Corresponding Author. sessi.tokpavi@u-paris10.fr, EconomiX-CNRS, University of Paris Ouest, France.

1 Introduction

The oil market constitutes without any doubt the most strategic row commodity market. Periods of extreme high energetic price (often labeled as oil shock) are usually associated with recession and/or inflationnary pressure (Sadorsky (1999), Hamilton (2003), and Kilian (2008), among others). Hence, understanding how oil price is fixed and evolved is a key issue for policy makers in order to implement adequate economic stabilization policies.[1]

Unfortunately this issue is not simple as the oil market is not homogenous and is composed by numerous local markets, sometimes organized into cartels (the most famous being the OPEC[2]) and trading different oil qualities (depending on the API[3] and the sulfur content). Different types of crude oils fetch distinct prices, and these prices are usually set as a discount or premium to a marker or reference crude oil according to their characteristics (Mabro (2005), Fattouh (2006, 2010, 2011), among others).[4] Many observers consider the world oil market as 'one great pool' (Adelman (1984)) in the sense that supply and demand shocks that affect prices in one region are transmitted into other regional markets. Several papers have therefore tested the integration hypothesis of the different crude oil markets (see inter alii Weiner (1991), Gülen (1997, 1999), Kleit (2001), Milonas and Henker (2001), Lanza et al. (2003), Hammoudeh et al. (2008), Fattouh (2010)) which assumes that same quality crude oil prices should be nearly identical or at least co-move in different regions implying that their price differentials would be more or less constant. This perspective has strong implications in terms of energy policy and market efficiency.

Nevertheless, as a consensus is not reached since marker crudes suffer from seri-

[1] Oil prices are determined by many factors such as supply versus demand, macroeconomic and financial shocks, etc.

[2] See Brémond et al. (2011).

[3] The American Petroleum Institute is a measure of how heavy or light a petroleum liquid is compared to water.

[4] The expansion of the crude oil market allowed the development of market-referencing pricing off spot crude oil markers such as WTI, Brent, and Dubai which are theoretically considered as benchmarks due to their ownership diversification properties (see, Horsnell and Mabro (1993)).

ous doubts about their ability to generate a marker price[5], Wlazlowski et al. (2011) (hereafter WHM) prefer to analyze global market dependencies, finding out if a particular crude oil market can be regarded as benchmark or follower. It is then possible to draw a distinction between price taker markets, which are affected by the variation on other local markets, and price setter markets, which give the pace for price changes. This distinction is therefore essential for policy makers, to evaluate for example the price consequences of an embargo on an oil producer. If this market is price setter (resp. taker) it should (resp. should not) impact the prices on the other local markets. Besides, WHM distinguish 4 qualities and 32 crude oil markets, concluding that widely used benchmarks such as WTI and Brent are indeed in fact global price setters joined by a third crude, the Mediterranean Russian Urals. The Asia Dubai Fateh and the Oman Blend finally act as benchmarks for their segment.

Cook (1998)[6] stresses that integration hypothesis is especially and almost uniquely important when crude oil price movements (upward or downward) are extreme indicating tension either on the demand or supply side. The general feeling is that price differentials would tend to widen across the markets during extreme upward movements, and decrease otherwise. Thus, the diversification strategy aiming at limiting the impact of an oil shock would be more efficient during extreme prices periods, whereas it would be more difficult and less beneficial in "regular" times. Indeed, the empirical justification of such a theory separating regular and extreme times would have strong policy implications.

This paper proposes to investigate this issue by analyzing the global market dependence during extreme crude oil price movements. To this aim we extend the univariate Granger causality approach in the tail distribution proposed by Hong et al. (2009) in order to investigate risk spillover between financial markets. This testing approach consists in checking whether a large downside (resp. upside) risk in

[5]See Fattouh (2006).

[6]Confirmed by several reports of the BMO Commodity Derivatives Group. In particular the one published in 2004 entitled "Managing Heavy Oil Price Risk" and available at corporate.bmo.com/cm/market/cdcom/images/Managing_Heavy_Oil_Price_Risk.pdf.

one market will Granger causes a large downside (resp. upside) risk in another market. Their approach considers risk transmission of two time series at a given quantile level α which is relatively restrictive because it doesn't check causality between distribution tails. Moreover, according to Engle and Manganelli (2004), dynamics of downside (upside) risk can vary considerably across risk levels. To overcome these constraints and improve the power properties of the Granger causality approach, we extent this setup by testing simultaneous Granger causality in downside (upside) risk for multiple risk levels across tail distributions. In other words, our testing procedure can be seen as an extension in multivariate context of the approach in Hong et al. (2009) and uses all tail distribution information.

Anticipating on our findings, we find that WTI and Europe Brent crude oil markets are dominant both in periods of extreme downside and upside price movements, with WTI being the leading benchmark in the first configuration. Asia Dubai Fateh, often used as third crude oil market benchmark appears to be price taker whatever the distribution side considered. Furthermore, contrary to WHM, we find that Oman Blend is not a price setter and that Mediterranean Russian Urals and Europe Forcados (resp. Ecuador Oriente) can be considered as benchmarks in extreme price falls (resp. rises). Moreover, outcomes support the view that integration between crude oil markets tends to decrease during the periods of extreme price movements.

The rest of the paper is organized as follows. Section 2 describes the concept of Granger causality in risk, presents our new test in multivariate context and analyzes its asymptotic properties. Section 3 presents our data and studies the international crude oil markets globalization. Section 4 concludes the paper.

2 Granger causality in distribution tails

In this section we develop a framework to test for Granger causality in distribution tails, that is, whether the occurrence of any tail event for a given time series can help predict the occurrence of any tail event for another time series. The section is divided into two parts. In the first part we describe the econometric environment, give an overview of our testing approach and present the test statistics, while in the second part we simulate its finite sample properties via Monte Carlo studies.

2.1 Econometric environment and testable hypotheses

We consider a stochastic process $X \equiv \left\{X_t : \Omega \to \mathbb{R}^2, t = 1, ..., T\right\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ where $\mathcal{F} \equiv \{\mathcal{F}_t, t = 1, ..., T\}$ and $\mathcal{F}_t$ is the σ-field $\mathcal{F}_t = \{X_s, s \leq t\}$. We partition the observed vector X_t as $X_t = (X_{1,t}, X_{2,t})$ where both $X_{1,t}$ and $X_{2,t}$ are continuous random variables of interest. The information set available at time t has the following structure $\mathcal{F}_t \equiv \{\mathcal{F}_{1,t}\} \cup \{\mathcal{F}_{2,t}\}$ with $\mathcal{F}_{1,t} = \{X_{1,s}, s \leq t\}$ and $\mathcal{F}_{2,t} = \{X_{2,s}, s \leq t\}$. Our test is related to the concept of Granger causality defined in terms of the entire conditional distribution (Granger (1980), Granger and Newbold (1986)). Using our notations, $X_{2,t}$ does not Granger-cause $X_{1,t}$ in distribution if and only if

$$\Pr\left[X_{1,t} < x \,|\mathcal{F}_{t-1}\right] = \Pr\left[X_{1,t} < x \,|\mathcal{F}_{1,t-1}\right] \text{ a.s. for all } x. \tag{1}$$

In this case, past values of $X_{2,t}$ in the information set $\mathcal{F}_{t-1}$ do not carry any useful information that helps predict the conditional distribution of $X_{1,t}$. This definition is rather broad since in many practical situations, a user with a specific objective may be concerned with whether causality occurs or not in particular regions of the distributions of both variables. For example, in the context of downside risk monitoring and diversification, risk managers are usually aware of whether a loss for a business line in their managed portfolio will exceed a fixed large value given that a large loss for another business line has occurred. In the international crude oil markets, prices have experienced strong fluctuations affecting the profile of risk. For investors with long (resp. short) positions in these energy assets, measuring the associated downside (resp. upside) risks and their spillover effect is primordial. From a macroprudential point of view, the recent episode of market turmoil gives many evidence that regulators should also take care about downside risk spillover between financial institutions.

Hong et al. (2009) introduced a formal statistical procedure to test for Granger causality in downside risk quantified by Value-at-Risk (VaR), the most popular metric of risk in the banking and financial industry. The VaR of a time series at the risk level $\alpha \in (0, 1)$ is defined as the α-quantile of the conditional distribution of the given time series. For the two time series we thus have

$$\Pr\left[X_{1,t} < Q_{1,t}\left(\theta_{1,\alpha}\right) | \mathcal{F}_{1,t-1}\right] = \alpha, \tag{2}$$

$$\Pr\left[X_{2,t} < Q_{2,t}\left(\theta_{2,\alpha}\right) | \mathcal{F}_{2,t-1}\right] = \alpha, \tag{3}$$

with $Q_{1,t}\left(\theta_{1,\alpha}\right)$ and $Q_{2,t}\left(\theta_{2,\alpha}\right)$ the anticipated VaR of $X_{1,t}$ and $X_{2,t}$ respectively at time $t-1$, $\theta_{1,\alpha}$ and $\theta_{2,\alpha}$ two finite-dimensional parameter from the specification of the dynamics of both variables. Consider the following two tail-events time series

$$Z_{1,t}\left(\theta_{1,\alpha}\right) = \begin{cases} 1 & \text{if} \quad X_{1,t} < Q_{1,t}\left(\theta_{1,\alpha}\right) \\ 0 & \text{else,} \end{cases} \tag{4}$$

$$Z_{2,t}\left(\theta_{2,\alpha}\right) = \begin{cases} 1 & \text{if} \quad X_{2,t} < Q_{2,t}\left(\theta_{2,\alpha}\right) \\ 0 & \text{else.} \end{cases} \tag{5}$$

In Hong et al. (2009), the time series $\{X_{2,t}\}$ does not Granger-cause the time series $\{X_{1,t}\}$ in downside risk at level α if the following hypothesis holds[7]

$$\mathbb{H}_0 : \mathbb{E}\left[Z_{1,t}\left(\theta_{1,\alpha}\right) | \mathcal{G}_{t-1}\right] = \mathbb{E}\left[Z_{1,t}\left(\theta_{1,\alpha}\right) | \mathcal{G}_{1t-1}\right], \tag{6}$$

where the two information sets are defined as

$$\mathcal{G}_t = \left\{\left(Z_{1,s}\left(\theta_{1,\alpha}\right), Z_{2,s}\left(\theta_{2,\alpha}\right)\right), s \leq t\right\}, \tag{7}$$

$$\mathcal{G}_{1t} = \left\{Z_{1,s}\left(\theta_{1,\alpha}\right), s \leq t\right\}. \tag{8}$$

[7] Note that both causality in downside and upside risk can be handled in the framework of Hong et al. (2009). In the former case the risk level or coverage rate is set to a small value (for e.g., $\alpha = 1\%$, 5% or 10%). In the latter case, a high value is retained (for e.g., 90%, 95% or 99%) and the tail-events time series are properly defined as follows

$$Z_{i,t}\left(\theta_{i,\alpha}\right) = \begin{cases} 1 & \text{if} \quad X_{i,t} > Q_{i,t}\left(\theta_{1,\alpha}\right), \;\; i = 1, 2, \\ 0 & \text{else.} \end{cases}$$

Hence, Granger causality in downside risk for the two-time series $\{X_{1t}\}$ and $\{X_{2t}\}$ is equivalent to Granger causality in mean for the two tail-events time series $\{Z_{1,t}(\theta_{1,\alpha})\}$ and $\{Z_{2,t}(\theta_{2,\alpha})\}$. It is worth noting that (6) is not a testable hypothesis since the two tail-events time series which depend on the unknown VaRs, $Q_{i,t}(\theta_{i,\alpha})$, $i = 1, 2$, are not observable. Hence a model is required for both series, to generate the in-sample VaRs and the corresponding tail-events time series. Hong et al. (2009) rely on the Conditional Autoregressive Value-at-Risk (CAViaR) model introduced by Engle and Manganelli (2004) in which the VaRs are estimated directly using an autoregressive specification for the quantiles rather than inverting a conditional distribution as usual in a purely parametric framework (for e.g., a GARCH model under a Student-t distribution). More precisely, the following specifications are retained to estimate the VaRs

$$Q_{i,t}(\theta_{i,\alpha}) = \theta_{i,\alpha}^{(0)} + \theta_{i,\alpha}^{(1)} Q_{i,t-1}(\theta_{i,\alpha}) + \theta_{i,\alpha}^{(2)} (X_{i,t-1})^{+} + \theta_{i,\alpha}^{(3)} (X_{i,t-1})^{-}, \tag{9}$$

where $(X_{i,t})^{+} = \max(X_{i,t}, 0)$, $(X_{i,t})^{-} = \min(X_{i,t}, 0)$, $\theta_{i,\alpha} = \left(\theta_{i,\alpha}^{(0)}, \theta_{i,\alpha}^{(1)}, \theta_{i,\alpha}^{(2)}, \theta_{i,\alpha}^{(3)}\right)$, $i = 1, 2$. Note that the autoregressive nature of the CAViaR model captures (directly) in the tails of the distributions some stylized facts in empirical finance with many compelling evidence, such as autocorrelation in daily returns arising from market microstructure biases and partial price adjustment (Boudoukh et al. (1994), Eom, Hahn and Joo (2004), Ahn et al. (2002)), volatility clustering (Engle (1982), Bollerslev (1986)), and time-varying skewness and kurtosis (Hansen (1994), Harvey and Siddique (1999, 2000), Jondeau and Rockinger (2003)). Moreover the asymmetric specification in (9) adresses the asymmetric response of volatility to news (Black (1976), Christie (1982)). The parameters of the CAViaR model are estimated by minimizing with respect to the unknown parameters the "check" loss function of Koenker and Bassett (1978), i.e.,

$$\widehat{\theta}_{i,\alpha} = \arg\min_{\theta_{i,\alpha}} \frac{1}{T} \sum_{t=2}^{T} \left[\alpha - \mathbb{I}\left(u_{i,t} < 0\right)\right] u_{i,t}, \quad i = 1, 2, \tag{10}$$

$$u_{i,t} = X_{i,t} - Q_{i,t}(\theta_{i,\alpha}), \tag{11}$$

with $\mathbb{I}(.)$ the usual indicator function and T the estimation sample length. The testable hypothesis of non-Granger causality in downside risk can thus be written as

$$\mathbb{H}_0 : \mathbb{E}\left[Z_{1,t}\left(\widehat{\theta}_{1,\alpha}\right)|\mathcal{H}_{t-1}\right] = \mathbb{E}\left[Z_{1,t}\left(\widehat{\theta}_{1,\alpha}\right)|\mathcal{H}_{1t-1}\right], \tag{12}$$

with the observable information sets

$$\mathcal{H}_{1t} = \left\{Z_{1,s}\left(\widehat{\theta}_{1,\alpha}\right), s \leq t\right\}, \tag{13}$$

$$\mathcal{H}_{t} = \left\{\left(Z_{1,s}\left(\widehat{\theta}_{1,\alpha}\right), Z_{2,s}\left(\widehat{\theta}_{2,\alpha}\right)\right), s \leq t\right\}. \tag{14}$$

Hong et al. (2009) adopted a kernel-based nonparametric test which checks for the nullity of the standardized sample cross covariances between the two processes $\left\{Z_{1,t}\left(\widehat{\theta}_{1,\alpha}\right)\right\}$ and $\left\{Z_{2,t}\left(\widehat{\theta}_{2,\alpha}\right)\right\}$, under the hypothesis of non-Granger causality in downside risk.[8] The test statistics has a standard asymptotic distribution under the null hypothesis which is not affected by parameter uncertainty in the estimated CAViaR models.

The test developed by Hong et al. (2009) is suitable to check for the existence of Granger causality in extreme movements of two time series, but at a given risk level α. Our objective in the sequel is to extend this setup, by testing simultaneously Granger causality in downside risk for multiple risk levels across the distribution tails. Two main reasons motivate our extension. First, it is apparent that when focusing on downside risk spillover effect, what really counts is to check whether causality exists between the left tail distributions of the two time series, and not between quantiles for a single risk level α. Second, estimation results of CAViaR models in Engle and Manganelli (2004) show that the process governing the dynamics of VaRs can vary remarquably across risk levels. Hence, application of the Hong et al. (2009) test can lead to contradictory results with respect to the risk levels, for example at 1%, 5% or 10%. In such a case, it is more suitable to make inference jointly for the three

[8] Note that Hong et al. (2009) also consider in their paper a regression-based approach to test for Granger causality in downside risk. As we will see in the sequel, our Granger causality test in distribution tails is a multivariate extension of the latter approach.

risk levels. From a statistical point of view, this strategy will improve the power properties of the Granger causality test as more information is exploited.

Our testing procedure is based on the multivariate extension of the classical Granger causality test in mean, where the purpose is to make inference on interactions that take place among groups of variables (Gelper and Croux (2007), Barret et al. (2010)). To present the methodology, let $A = \{\alpha_1, ..., \alpha_m\}$ be a discrete set of m risk levels, strictly between 0 and 1 and considered as relevant for downside risk analysis. For $i = 1, 2$, let $W_{i,t}(\theta_{i,A}) = [Z_{i,t}(\theta_{i,\alpha_1}), ..., Z_{i,t}(\theta_{i,\alpha_m})]$ be the vector of dimension $(m, 1)$ collection of the tail-events variables $Z_{i,t}(\theta_{i,\alpha_k})$ associated to these m risk levels at time t, where $\theta_{i,A} = \left(\theta'_{i,\alpha_1}, ..., \theta'_{i,\alpha_m}\right)'$ is the vector of dimension $(4m, 1)$ with elements the parameters of the m CAViaR models, each at the risk level α_k, $k = 1, ..., m$. The null hypothesis of our non-Granger causality test in distribution tails can be stated as follows

$$\mathbb{H}_0 : \mathbb{E}\left[W_{1,t}(\theta_{1,A}) | \mathcal{I}_{t-1}\right] = \mathbb{E}\left[W_{1,t}(\theta_{1,A}) | \mathcal{I}_{1,t-1}\right], \tag{15}$$

where the sets $\mathcal{I}_{1,t}$ and $\mathcal{I}_t$ correspond respectively to

$$\mathcal{I}_{1,t} = \left\{W_{1,s}(\theta_{1,A}), s \leq t\right\}, \tag{16}$$

$$\mathcal{I}_t = \left\{\left(W'_{1,s}(\theta_{1,A}), W'_{2,s}(\theta_{2,A})\right)', s \leq t\right\}. \tag{17}$$

If the null hypothesis holds, this means that whatever the risk levels α_k, $k = 1, ..., m$, spillover of extreme downside movements (from X_{2t} to X_{1t}) does not exist. Hence in our setup, Granger causality in distribution tails is nothing but Granger causality in mean for the two multivariate processes $W_{i,t}(\theta_{i,A})$, $i = 1, 2$. Following Gelper and Croux (2007) and Barret et al. (2010), the test statistic is easily built by considering the following multivariate linear regression model[9]

$$W_{1,t}(\theta_{1,A}) = \psi_0 + \psi_1 W_{2,t-1}(\theta_{2,A}) + ... + \psi_p W_{2,t-p}(\theta_{2,A}) + \varepsilon_{1t}, \tag{18}$$

[9]It is worth noting that we do not include lagged values of $W_{1,t}(\theta_{1,A})$ in the regression equation (18), because under the null hypothesis, the m components of $W_{1,t}(\theta_{1,A})$ are independent, each following an i.i.d. Bernoulli distribution. This latter property is usually used to backtest Value-at-Risk models (see, Christoffersen (1998), Engle and Manganelli (2004), Berkowitz et al. (2011), Candelon et al. (2011), etc.).

where ψ_0 is a vector $(m, 1)$ of constants, ψ_s, $s = 1, ..., p$, are (m, m) matrices of parameters, and ε_{1t} the $(m, 1)$ residuals vector with covariance matrix Σ_1. The null hypothesis of non-Granger causality in distribution tails corresponds to

$$\mathbb{H}_0 : \psi_1 = \psi_2 = ... = \psi_p = 0. \tag{19}$$

When this null hypothesis holds, the multivariate regression in (18) reduces to

$$W_{1,t}\left(\theta_{1,A}\right) = \psi_0 + \varepsilon_{2t}, \tag{20}$$

with ε_{2t} the $(m, 1)$ residuals vector with covariance matrix Σ_2. As a consequence, the multivariate likelihood ratio test statistic defined as follows

$$\text{LR} = [\text{T} - (mp + 1)] \left[\log\left(\left|\varepsilon_2' \varepsilon_2\right|\right) - \log\left(\left|\varepsilon_1' \varepsilon_1\right|\right)\right], \tag{21}$$

can be used to test for the null hypothesis of non-Granger causality in distribution tails as stated in (15) or equivalently in (19). This test statistic follows under the null hypothesis a chi-squared distribution with degree of freedom equal to pm^2. Let us remark that the above testing approach is not computationally feasible, because the two multivariate processes $W_{1,t}\left(\theta_{1,A}\right)$ and $W_{2,t}\left(\theta_{2,A}\right)$ depend respectively on the unknown vector of the CaViaR models parameters $\theta_{1,A}$ and $\theta_{2,A}$. An operational test can be conducted by considering the following null hypothesis

$$\mathbb{H}_0 : \xi_1 = \xi_2 = ... = \xi_p = 0 \tag{22}$$

in the multivariate regression:

$$W_{1,t}\left(\widehat{\theta}_{1,A}\right) = \xi_0 + \xi_1 W_{2,t-1}\left(\widehat{\theta}_{2,A}\right) + ... + \xi_p W_{2,t-p}\left(\widehat{\theta}_{2,A}\right) + \varepsilon_{1t} \tag{23}$$

where the true vector of parameters $\theta_{i,A}$, $i = 1, 2$, are replaced by their respective consistent estimators $\widehat{\theta}_{i,A}$. However, uncertainty about the values of $\widehat{\theta}_{i,A}$, $i = 1, 2$, could affect the distribution of the test statistic. This problem is referred to as parameter uncertainty in the framework of hypothesis testing. With the problem of parameter uncertainty at hand, two different solutions can be adopted. First, one can proceed as if the problem of parameter uncertainty is not important and makes inference using the chi-squared asymptotic distribution. The relevance of this solution can be measured by the extent to which the parameter uncertainty problem is not of concern. Second, inference can be performed through robust methods such as Monte Carlo tests which are exact, in the sense that the actual probability of Type I error is equal to the nominal significance level of the test.

Formally, Monte Carlo tests are performed by generating M independent realizations of the test statistic - say S_i, $i = 1, ...M$ - under the null hypothesis. If we denote S_0 the value of the test statistic obtained for the original sample, as shown by Dufour (2006) in a general case, the Monte Carlo critical region is obtained as $\widehat{p}_M\left(S_0\right) \leq \eta$ with $1 - \eta$ the confidence level and $\widehat{p}_M\left(S_0\right)$ defined as

$$\widehat{p}_M\left(S_0\right) = \frac{M\widehat{G}_M\left(S_0\right) + 1}{M + 1}, \tag{24}$$

where

$$\widehat{G}_M\left(S_0\right) = \frac{1}{M}\sum_{i=1}^{M}\mathbb{I}\left(S_i \geq S_0\right), \tag{25}$$

when $\Pr\left(S_i = S_j\right) \neq 0$, and otherwise

$$\widehat{G}_M\left(S_0\right) = 1 - \frac{1}{M}\sum_{i=1}^{M}\mathbb{I}\left(S_i \leq S_0\right) + \frac{1}{M}\sum_{i=1}^{M}\mathbb{I}\left(S_i = S_0\right) \times \mathbb{I}\left(U_i \geq U_0\right). \tag{26}$$

Variables U_0 and U_1 are uniform draws from the interval $[0, 1]$. In our framework, application of the Monte Carlo test procedure of Dufour (2006) requires simulating the two multivariate processes $W_{i,t}\left(\theta_{i,A}\right)$ $i = 1, 2$, under the null hypothesis of non-Granger causality in distribution tails, in order to compute the M independent realizations of the test statistic LM_i, $i = 1, ..., M$, under $\mathbb{H}_0$. This task can be achieved very easily noting that for well-specified CAViaR models, each element of $W_{i,t}\left(\theta_{i,A}\right)$, $i = 1, 2$, i.e., the tail-event variable $Z_{i,t}\left(\theta_{i,\alpha_k}\right)$, $k = 1, ..., m$, follows an *i.i.d.* Bernoulli distribution with a success probability equal to α_k.

2.2 Analysis of finite sample properties

This section is devoted to Monte Carlo simulations studies with the objective of evaluating the small sample properties of our Granger causality test in distribution tails. We evaluate both inference with the asymptotic chi-squared critical region and the Monte Carlo critical region of Dufour (2006). This will allow us to quantify how much our testing procedure is affected by parameter uncertainty in the CAViaR models.

2.2.1 Finite sample size analysis

To illustrate the size performance of our test, we follow Hong et al. (2009) simulating the two time series $X_{1,t}$ and $X_{2,t}$ using the following data generating process (DGP):

$$\begin{cases} X_{i,t} = 0.5X_{i,t-1} + u_{i,t}, \;\; i = 1, 2, \\ u_{i,t} = \sigma_{i,t} v_{i,t}, \\ \sigma_{i,t}^2 = 0.1 + 0.6\sigma_{i,t-1}^2 + 0.2u_{i,t-1}^2, \\ v_{i,t} \sim m.d.s.\ (0, 1) \, . \end{cases}$$

Hence, each time series $X_{i,t}$, $i = 1, 2$, follows an AR(1)-GARCH model. The two processes are independent and there is no Granger causality in distribution tails between them. We simulate the size of the test considering three different sample sizes (T = 500, 1000, 1500), which correspond roughly to two, four and six years of daily data. For a given value of T, and for each simulation, CAViaR models are estimated to compute the two multivariate tail-events variables $W_{i,t}\left(\widehat{\theta}_{i,A}\right)$, $i = 1, 2$, with A the discrete set of the m VaRs risk levels, $A = \{\alpha_1, ..., \alpha_m\}$. With the two multivariate processes $W_{i,t}\left(\widehat{\theta}_{i,A}\right)$, $i = 1, 2$, we test the null hypothesis of non-Granger causality in tail distributions checking via the LM statistic the restriction (22) in the multivariate regression (23).

Table 1 in Appendix reports the empirical sizes of our multivariate LM test statistic (over 500 simulations) for different values of $p \in \{5, 10, 15\}$ the lag order in the regression equation (23). The set A of the m VaRs risk levels is set to $A = \{1\%, 5\%, 10\%\}$. These values correspond to the usual risk levels considered when focusing on downside risk analysis.[10] For each simulation, the null hypothesis of non-Granger causality in distribution tails is rejected relying on the asymptotic chi-squared critical region, with two different nominal risk levels $\eta = 5\%$, 10%. Results in Table 1 indicate that our Granger causality test in distribution tails is oversized whatever the sample size T and the value of the lag-order parameter p. For example, with a nominal risk level $\eta = 5\%$ and two years of daily data ($T = 500$), the rejection frequency of the null hypothesis is around 15% when $p = 5$ and 17% with $p = 10$. These results show that our regression testing procedure is affected by parameter uncertainty. The problem seems to be more prominent in small samples where the estimated parameters in the CAViaR models fail to converge to the correct model parameters because of data scarcity. The failure of convergence should be more acute at the 1% VaRs risk level compared to the other two VaRs risk levels $(5\%, 10\%)$. Therefore, the parameter uncertainty problem which affects the empirical sizes of our test should come mainly from the estimation errors of the CAViaR models at the 1% VaRs risk level. To confirm this analysis, we report in Table 2 (see Appendix) the empirical sizes of the LM test statistic with $A = \{5\%, 10\%\}$. The presentation is similar to Table 1. We observe that the reported rejection frequencies of the null hypothesis are much closer to the nominal risk levels $\eta = 5\%$, 10%.

The above results suggest that for our testing procedure, inference using the asymptotic chi-squared distribution should be conducted only for moderate VaRs risk levels in the left-tail distribution. More precisely, one should not include the 1% VaRs risk level in the set A. Nevertheless, in the analysis of spillover effect in downside movements, considering the extreme case of 1% risk level is crucial, because in financial markets, market prices movements are more stronger at this risk

[10] Of course, one can extend the set A by considering more risk levels, for example $A = \{1\%, 2.5\%, 5\%, 7.5\%, 10\%\}$. The advantage of this extension is to consider more information in the inferential procedure. However, when the size of the set A increases, the considered risk levels are more closer, and there is a non zero probability to face a problem of multicollinearity in the multivariate regression (23). This reason also motivates our choice of the set A.

level, with major consequences on the values of assets and the solvability of assets owners. Hence, we propose to make inference with the three VaRs risk levels, i.e., $A = \{1\%, 5\%, 10\%\}$, simulating the critical region through the Monte Carlo approach of Dufour (2006). As already stressed, this testing procedure helps to alleviate the problem of parameter uncertainty by simulating via Monte Carlo experiments, the exact distribution of the test statistic under the null hypothesis. Table 3 in Appendix displays the empirical sizes with the Monte Carlo critical region, where the parameter M (see equations 24-26) is set to $9,999$. The overall picture from Table 3 is that our LM test statistic, used in conjonction with the Monte Carlo procedure of Dufour (2006), is correctly sized. For each sample size T, the choice of the lag-order parameter p has little impact on the size of the test.

2.2.2 Finite sample power analysis

We now investigate the power of the test. Since causality in distribution tails or in extreme movements is mainly due to causality in mean, variance or higher order moments such as kurtosis and skewness, we assume the following DGPs for the two time series $X_{1,t}$ and $X_{2,t}$, in order to generate data under the alternative hypothesis:

$$\begin{cases} X_{2,t} = 0.5X_{2,t-1} + u_{2,t}, \\ u_{2,t} = \sigma_{2,t}v_{2,t}, \\ \sigma^2_{2,t} = 0.1 + 0.6\sigma^2_{2,t-1} + 0.2u^2_{2,t-1}, \end{cases} \tag{27}$$

$$\begin{cases} X_{1,t} = 0.5X_{1,t-1} + 0.2X_{2,t-1} + u_{1,t}, \\ u_{1,t} = \sigma_{1,t}v_{1,t}, \\ \sigma^2_{1,t} = 0.1 + 0.6\sigma^2_{1,t-1} + 0.2u^2_{1,t-1} + 0.7u^2_{2,t-1}, \end{cases} \tag{28}$$

where both $v_{1,t}$ and $v_{2,t}$ are martingale difference sequences with mean 0 and variance 1. Under this setting, the time series $X_{2,t}$ Granger causes the time series $X_{1,t}$ in distribution tails via causality in both mean and variance. The empirical powers of our multivariate LM test statistic are computed over 500 simulations, for different values of p the lag-order parameter, and for two nominal risk levels $\eta = 5\%$, 10%.

The results are reported in Tables 4 and 5 for $A = \{1\%, 5\%, 10\%\}$. To stress the relevance of our multivariate approach, we also display in these tables the power of the univariate testing approach of Hong et al. (2009), where A is reduced to the sets $\{1\%\}$, $\{5\%\}$ and $\{10\%\}$ respectively. The rejection frequencies are computed using the Monte Carlo critical region of Dufour (2006), with the parameter M set to $9,999$. Our multivariate test displays fairly good power properties. For example, with $T = 1000$ and $p = 5$, the test rejects the null of non-Granger causality in distribution tails 82% (resp. 89%) of time when $\eta = 5\%$ (resp. $\eta = 10\%$). As expected, the power increases as the sample size T increases. As usual in the setting of parametric Granger causality test, increasing the lag-order parameter p lowers the power of the test. Finally and importantly, the advantage of the multivariate approach over the univariate testing procedure of Hong et al. (2009) is clear-cut. Indeed, for a given value of the sample T and the lag-order parameter p, the multivariate test rejects more strongly the null hypothesis of non-Granger causality in distribution tails. For instance, with $\eta = 5\%$, and $(T, p) = (500, 10)$, the rejection frequency is equal to 51% for $A = \{1\%, 5\%, 10\%\}$, whereas it is only equal to 15%, 41%, and 41% for A equal to $\{1\%\}$, $\{5\%\}$, and $\{10\%\}$ respectively.

3 Empirical application

3.1 Data

Following WHM we consider the weekly prices of 32 crude oils extracted from the Energy Information Agency for the period April 21, 2000 to October 20, 2011. They are reported in Table 6. Out of the total of 32 crude oils in our sample, 15 fall into the jurisdiction of the OPEC bloc, whereas 17 are not part of it (non-OPEC countries). Each crude oil is also characterized by its quality defined both by its density and its sulphur content. The density of any crude oil is measured in degrees API. The higher the API degree, the lighter (and the better) the crude. Crudes with API higher than 35° are considered light, API between 26° and 35° are medium, whereas all API smaller than 26° are considered heavy. Sweet and Sour refer to the sulphur content of the crude. Crudes with high content in sulphur are said to be sour and are generally avoided, as they produce more pollution. Crudes are considered to be sweet when the sulphur content does not exceed 0.5% and sour when they do.

3.2 Results

We consider weekly returns and implement our Granger causality test in distribution tails. Our goal is to statistically identify which crude oil market (Granger) causes in distribution tails the others. More precisely, for each couple of crude oils among the 32 crude oils considered, with time series returns $x_t^{(i)}$ and $x_t^{(j)}$, $i \neq j$, we apply our Granger causality test in distribution tails to check (a) whether $x_t^{(i)}$ Granger-causes $x_t^{(j)}$, and (b) whether $x_t^{(j)}$ Granger-causes $x_t^{(i)}$. Results are displayed in Table 7 for both left and right tails, corresponding respectively to extreme downside and upside movements. For each side (downside or upside), the first column presents the proportion of time a market Granger-causes other markets.[11] Symmetrically, the second column presents the proportion of time a market is Granger-caused by other ones. The last column displayed the difference between the values reported in the first and the second column. The results can be analyzed as follows: a crude oil is identified as benchmark or exhibits price setter characteristics in extreme movements, if it causes other crude oils without being caused (or weakly caused) reciprocally. These crude oils are thus highly sensitive to oil market shocks, i.e., they respond to oil market news. On the contrary, crude oils with high price taker and low price setter characteristics follow the trend of the global market, and are less sensitive to oil market shocks. Lastly, crude oils with both high setter and taker dynamics are intermediate between leaders and followers, and can be considered as perfectly integrated. It is worth noting that these three categories can be easily identified focusing on the dif-

[11] Inferences are conducted at the 5% nominal risk level. Following WHM, we set the value of the lag-order parameter p to 16 (4 months). Results available from the authors upon request show that our foundings are robust with respect to p.

ference between the proportions displayed in the third column: markets with a large positive (resp. negative) difference in proportions are benchmarks (resp. followers), whereas small absolute values indicate that markets are well integrated in the general market.

Results in Table 7 indicate that both WTI and Europe Brent behave as benchmarks (they cause other crude oils and are weakly affected reciprocally), and it turns out that WTI is the leader in downside movements. Furthermore, unlike WHM's analysis, we find that Mediterranean Russian Urals is a benchmark only in extreme downside movements. Europe Forcados also appears as benchmark in downside movements. Ecuador Oriente, and to a lesser extent, Mexico Isthmus and Mexico Maya, turn out to be benchmarks in extreme upside movements. Colombia Cano Limon, Malaysia Tapis, Saudi Arabia Saudi Light, Saudi Arabia Arab Medium, and Ecuador Oriente are followers in periods of large price decrease, whereas Mediterranean Russian Urals and Kuwait Blend show the same characteristic in periods of large price increase. Besides, and importantly, Asia Dubai Fateh and Oman Blend which are considered in practice as benchmarks for their segment, do not exhibit price setter characteristics in the universe of the 32 crude oils considered.

Is the crude oil markets less or more integrated in periods of extreme movements? This question adresses the issue of diversification in international crude oil markets, and is important for policy makers and portfolio managers. In the literature, the intuition is that crude oil markets are less integrated in extreme situations. The rationale of this claim (see, Cook (1998), BMO (2004), Bacon and Tordo (2005)) is that if the demands for all petroleum products increased proportionately, and all product prices and the general crude price also increased proportionately, then crudes with the largest proportion of high value products would increase in price relative to crudes with a lower proportion of high value products, with the result that price differential would tend to widen across the crude oil markets. The symmetric reasoning holds in the case of general fall in demands and prices. To confirm this analysis, we report in Table 8 the same statistics as in Table 7, with the difference that we consider Granger causality test in mean rather than in distribution tails. In both tables, we measure the level of markets integration by the mean of the absolute value of the differences between the two percentages (setter and taker). The lower the value of this statistic, the more integrated are crude oil markets. Our results confirm the intuition that crude oil markets are less integrated in extreme situations. Hence, the possibility of diversification turns out to be enhanced during the periods of extreme movements in crude oil prices.

To go deeper beyond these results, we implement our analysis conditional to the quality segment of crude oils. Following WHM, we consider three quality segments: light & sweet, medium & sweet, and medium & sour. From table 6, it is easy to see that the light & sweet group has 9 crude oils, the medium & sour group contains 13 crude oils, while the medium & sweet group has 6 crude oils. Further potential groups, in particular those involving sour crudes, were discarded given limitations of the sample size. Results are displayed in Tables 9, 10 and 11. Regarding to the light density and sweet crude oils (Table 9), WTI and Europe Brent are leaders in extreme price falls, WTI being the dominant crude oil, while in extreme price rises, only WTI behaves as benchmark. All other crude oils in this quality segment can be considered as followers (low setter and high taker proportions) or integrated to different extent (high setter and taker proportions), in periods of extreme downside movements. Malaysia Tapis appears clearly as a follower in extreme upside movements. Concerning the medium density and sweet crude oils (Table 10), Europe Forcados (resp. Colombia Cano Limon) can be considered as benchmark in extreme downside (resp. upside) movements. Finally, for the medium density and sour crude oils (Table 11), two markets, that is, Mediterranean Russian Urals and Mediterranean Seri K Iran Light appear to be preponderant in extreme downside movements, whereas Ecuador Oriente and Mexico Isthmus are the leaders in extreme upside movements. These findings are different from those reported by Montepeque (2005) and WHM about the lack of leading benchmarks and a high degree of integration in this group.

Our analysis goes on, distinguishing OPEC from non-OPEC members. Table 12 gathers the outcomes of the Granger-causality test and indicates a very high degree of integration in periods of positive extreme movements, with a lack of leading benchmark in the group of OPEC countries. For negative extreme movements, Europe Forcados, Algeria Saharan Blend, and Europe Libyan Es Sider are the dominant markets, whereas Asia Dubai Fateh and Saudi Arabia Arab Medium are clearly followers. When considering the non-OPEC members, WTI and Europe Brent are dominant in extreme downside movements, with WTI being the leader. Both markets dominate other crude oils equally in extreme upside movements. These distinct results depending on the side of the distribution (upside or downside), would be the consequence of fundamental and speculative specificities of each market. Indeed, unlike WTI crude oil, which prices are largely reflected by market fundamentals, Europe Brent oil market is relatively opaque (Miller et al. (2010)), with inherent lack of transparency and illiquidity in price determination processes. Consequently, the market could become unhinged from physical factors by action of market participants. Moreover, since several years, Brent and more generally North Sea crude oils have known a sharp decline in production, and more of the supply is now mainly absorbed locally in Europe. Therefore, Brent has become disconnected from US and Asian markets (Miller et al. (2010)). In this context, two types of extreme risk could exist in international oil market depending on downside and upside circumstances: "speculative risk" and "fundamental risk". On one hand, in periods of price decreases, fundamental mechanisms would dominate speculative ones. The fundamental mechanisms would be based on the international oil demand from North American and Asian emerging countries on NYMEX rather than IPE markets leading to the dominance of WTI crude oil. On the other hand, in periods of price increases, fundamental and speculative mechanisms would operate equally in oil markets, where financial investors without any physical interests could influence benchmarks through speculative purpose. This makes both crude oil markets to be dominant to the same extent. Note that in the group of non-OPEC countries, Mediterranean Russian Urals (resp. Ecuador Oriente) appears as a third benchmark in extreme price falls (resp. rises). These results are consistent with those reported in Table 7 where all the 32 crude oils are considered.

4 Conclusion

This paper proposes to investigate oil markets dependence during extreme price movements. To this aim we extend the univariate Granger causality approach in downside risk developed by Hong et al. (2009) in multivariate context. Asymptotic as well as finite sample properties are proposed. The new causality test is then applied to investigate the hypothesis of crude oil market globalization in periods of extreme price movements.

Several interesting results can be drawn: extreme crude oil prices are governed by non-OPEC markets rather than OPEC ones. More precisely, WTI and Brent crude oils are price setters both in downside and upside price movements, due to the fundamental and speculative components of each market. Surprisingly, Mediterranean Russian Urals and Europe Forcados (resp. Ecuador Oriente) also act as benchmarks in periods of extreme downside (resp. upside) price movements. Asia Dubai Fateh and Oman Blend, the acclaimed crude oil benchmarks act as followers rather than leaders. Besides, we observe that the integration level between crude oil markets tends to decrease during extreme periods.

These results highlight the leading role played by the US and UK markets in the determination of crude oil prices. Understanding and forecasting crude oil price evolutions in periods of extreme price occurences would require a precise analysis of these two high quality markets. Nevertheless, attention should be payed to additional leading markets which have lower quality: Mediterranean Russian Urals, Europe Forcados, and Ecuador Oriente. This paper also paves the way to important advices for energy policy as it indicates that diversification strategies are the more relevant in periods of sharp variations in crude oil prices.

Appendix: Tables

Table 1: Empirical sizes with the asymptotic critical region: $A = \{1\%, 5\%, 10\%\}$

	$p = 5$	$p = 10$	$p = 15$
	Nominal risk level $\eta = 5\%$		
$T = 500$	0.1503	0.1703	0.1523
$T = 1000$	0.1122	0.1222	0.1022
$T = 1500$	0.1142	0.1002	0.1122
	Nominal risk level $\eta = 10\%$		
$T = 500$	0.2144	0.2044	0.2064
$T = 1000$	0.1523	0.1743	0.1583
$T = 1500$	0.1743	0.1663	0.1764

Notes: The table displays the empirical rejection frequencies of the multivariate LM statistic under the null of non-Granger causality in distribution tails. The statistics are reported for different sample sizes, values of the lag order p in the multivariate regression (23), and nominal risk level η. The rejection frequencies are computed using the chi-squared asymptotic distribution.

Table 2: Empirical sizes with the asymptotic critical region: $A = \{5\%, 10\%\}$

	$p = 5$	$p = 10$	$p = 15$
	Nominal risk level $\eta = 5\%$		
$T = 500$	0.0802	0.0701	0.0782
$T = 1000$	0.0782	0.0661	0.0621
$T = 1500$	0.0441	0.0621	0.0701
	Nominal risk level $\eta = 10\%$		
$T = 500$	0.1343	0.1303	0.1343
$T = 1000$	0.1222	0.1162	0.1222
$T = 1500$	0.0922	0.1162	0.1263

Notes: The table displays the empirical rejection frequencies of the multivariate LM statistic under the null of non-Granger causality in distribution tails. The statistics are reported for different sample sizes, values of the lag order p in the multivariate regression (23), and nominal risk level η. The rejection frequencies are computed using the chi-squared asymptotic distribution.

Table 3: Empirical sizes with the Monte Carlo critical region: $A = \{1\%, 5\%, 10\%\}$

	$p = 5$	$p = 10$	$p = 15$
	Nominal risk level $\eta = 5\%$		
$T = 500$	0.0441	0.0401	0.0501
$T = 1000$	0.0461	0.0501	0.0641
$T = 1500$	0.0721	0.0661	0.0721
	Nominal risk level $\eta = 10\%$		
$T = 500$	0.1022	0.1102	0.1182
$T = 1000$	0.1082	0.1242	0.1222
$T = 1500$	0.1242	0.0962	0.1303

Notes: The table displays the empirical rejection frequencies of the multivariate LM statistic under the null of non-Granger causality in distribution tails. The statistics are reported for different sample sizes, values of the lag order p in the multivariate regression (23), and nominal risk level η. The rejection frequencies are computed relying on the Monte Carlo critical region.

Table 4: Empirical powers with the Monte Carlo critical region: nominal risk level = 5%

	$p = 5$	$p = 10$	$p = 15$
		$A = \{1\%, 5\%, 10\%\}$	
$T = 500$	0.5331	0.5150	0.5030
$T = 1000$	0.8216	0.7355	0.7094
$T = 1500$	0.9319	0.9238	0.8617
		$A = \{1\%\}$	
$T = 500$	0.2164	0.1543	0.1182
$T = 1000$	0.3166	0.2946	0.2565
$T = 1500$	0.4790	0.5190	0.4569
		$A = \{5\%\}$	
$T = 500$	0.4910	0.4148	0.3607
$T = 1000$	0.6774	0.6814	0.6232
$T = 1500$	0.8737	0.8297	0.7996
		$A = \{10\%\}$	
$T = 500$	0.5230	0.4128	0.3928
$T = 1000$	0.7255	0.7054	0.6112
$T = 1500$	0.8978	0.8657	0.8176

Notes: The first panel of the Table displays the empirical rejection frequencies of the multivaraite LM statistic under the alternative of Granger causality in distribution tails. The statistics are reported for different sample sizes, values of the lag order p in the multivariate regression (23), and nominal risk level η. For comparison, the following panels present the same statistics for the univariate test of Hong et al. (2009). The rejection frequencies are computed relying on the Monte Carlo critical region.

Table 5: Empirical powers with the Monte Carlo critical region: nominal risk level = 10%

	$p = 5$	$p = 10$	$p = 15$
	$A = \{1\%, 5\%, 10\%\}$		
$T = 500$	0.6874	0.6754	0.6072
$T = 1000$	0.8938	0.8637	0.8196
$T = 1500$	0.9739	0.9519	0.9259
	$A = \{1\%\}$		
$T = 500$	0.3868	0.2525	0.2585
$T = 1000$	0.5311	0.4890	0.3868
$T = 1500$	0.6032	0.6052	0.5792
	$A = \{5\%\}$		
$T = 500$	0.6393	0.5772	0.4890
$T = 1000$	0.7896	0.7756	0.7315
$T = 1500$	0.9158	0.9038	0.8517
	$A = \{10\%\}$		
$T = 500$	0.6553	0.5731	0.5210
$T = 1000$	0.8056	0.8036	0.7255
$T = 1500$	0.9399	0.9178	0.8818

Notes: The first panel of the Table displays the empirical rejection frequencies of the multivariate LM statistic under the alternative of Granger causality in distribution tails. The statistics are reported for different sample sizes, values of the lag order p in the multivariate regression (23), and nominal risk level η. For comparison, the following panels present the same statistics for the univariate test of Hong et al. (2009). The rejection frequencies are computed relying on the Monte Carlo critical region.

Table 6: Details of Crudes analyzed

Crude	API	Sulphur (%)
Non-OPEC		
WTI Cushing	40°-light	0.2-sweet
Europe Brent	38°-light	0.4-sweet
Europe Norwegian Ekofisk	43°-light	0.1-sweet
Canadian Par	40°-light	0.3-sweet
Canada Lloyd Blend	22°-heavy	3.1-sour
Mexico Isthmus	35°-medium	1.5-sour
Mexico Maya	22°-heavy	3.3-sour
Colombia Cano Limon	30°-medium	0.5-sweet
Ecuator Oriente	29°-medium	1.0-sour
Angola Cabinda	32°-medium	0.2-sweet
Cameroon Kole	35°-medium	0.3-sweet
Egypt Suez Blend	32°-medium	1.5-sour
Oman Blend	34°-medium	0.8-sour
Australia Gippsland	45°-light	0.1-sweet
Malaysia Tapis	44°-light	0.1-sweet
Mediterranean Russian Urals	32°-medium	1.3-sour
China Daqing	33°-medium	0.1-sweet
OPEC		
Saudi Arabia Saudi Light	34°-medium	1.7-sour
Saudi Arabia Arab Medium	31°-medium	2.3-sour
Saudi Arabia Saudi Heavy	28°-medium	2.8-sour
Asia Murban	40°-light	0.8-sour
Asia Dubai Fateh	32°-medium	1.9-sour
Qatar Dukhan	40°-light	1.2-sour
Mediterranean Seri Kerir Iran Light	34°-medium	1.4-sour
Mediterranean Seri Kerir Iran Heavy	31°-medium	1.6-sour
Kuwait Blend	31°-medium	2.5-sour
Algeria Saharan Blend	44°-light	0.1-sweet
Europe Nigerian Bonny Light	37°-light	0.1-sweet
Europe Forcados	30°-medium	0.3-sweet
Europe Libyan Es Sider	37°-light	0.4-sweet
Indonesia Minas	34°-medium	0.1-sweet
Venezuela Tia Juana	31°-medium	1.1-sour

Table 7: Results of Granger causality test in distribution tails

Crude	Left Tail			Right Tail		
	Setter (1)	Taker (2)	(1)-(2)	Setter (1)	Taker (2)	(1)-(2)
WTI	0.9677	0.0000	0.9677	0.9677	0.3226	0.6452
Europe Brent	0.9677	0.0968	0.8710	0.9677	0.3548	0.6129
Europe Norwegian Ekofisk	0.8387	0.7097	0.1290	0.0968	0.1290	−0.0323
Canadian Par	0.6452	0.8710	−0.2258	0.1935	0.0645	0.1290
Canada Lloyd Blend	0.0323	0.1613	−0.1290	0.0323	0.0645	−0.0323
Mexico Isthmus	0.5484	0.8710	−0.3226	0.5484	0.1613	0.3871
Mexico Maya	0.5806	0.8387	−0.2581	0.4839	0.1290	0.3548
Colombia Cano Limon	0.2581	0.9355	−0.6774	0.3871	0.1290	0.2581
Ecuador Oriente	0.5484	0.9355	−0.3871	0.7419	0.0645	0.6774
Angola Cabinda	0.8065	0.6774	0.1290	0.1613	0.0968	0.0645
Cameroon Kole	0.9355	0.7419	0.1935	0.0323	0.1935	−0.1613
Egypt Suez Blend	0.8710	0.9032	−0.0323	0.0968	0.0968	0.0000
Oman Blend	0.9032	0.8065	0.0968	0.0645	0.3871	−0.3226
Australia Gippsland	0.7097	0.6452	0.0645	0.1935	0.2903	−0.0968
Malaysia Tapis	0.4516	0.9677	−0.5161	0.5484	0.8065	−0.2581
Mediter. Russian Urals	0.9355	0.4516	0.4839	0.2581	0.6774	−0.4194
China Daqing	0.7742	0.9032	−0.1290	0.0323	0.3226	−0.2903
Saudi Arabia Saudi Light	0.3548	0.7742	−0.4194	0.0968	0.1290	−0.0323
Saudi Arabia Arab Medium	0.4194	0.8065	−0.3871	0.0968	0.2258	−0.1290
Saudi Arabia Saudi Heavy	0.6452	0.6774	−0.0323	0.0968	0.1290	−0.0323
Asia Murban	0.6774	0.8065	−0.1290	0.0968	0.2903	−0.1935
Asia Dubai Fateh	0.6129	0.9355	−0.3226	0.0645	0.3548	−0.2903
Qatar Dukhan	0.8065	0.7419	0.0645	0.1290	0.4516	−0.3226
Mediter. Seri K Iran Light	0.6129	0.4194	0.1935	0.0968	0.0968	0.0000
Mediter. Seri K Iran Heavy	0.4194	0.6129	−0.1935	0.1935	0.2903	−0.0968
Kuwait Blend	0.8387	0.8710	−0.0323	0.1290	0.5806	−0.4516
Algeria Saharan Blend	0.8065	0.5806	0.2258	0.0968	0.1613	−0.0645
Europe Nigerian Bonny Light	0.8387	0.7742	0.0645	0.1613	0.0968	0.0645
Europe Forcados	0.8387	0.3548	0.4839	0.1290	0.0968	0.0323
Europe Libyan Es Sider	0.9032	0.6452	0.2581	0.1290	0.0645	0.0645
Indonesia Minas	0.8065	0.5806	0.2258	0.0968	0.2903	−0.1935
Venezuela Tia Juana	0.6452	0.9032	−0.2581	0.3226	0.1935	0.1290
Mean absolute value			0.2782			0.2137

Note: For each crude oil, the table displays the proportion of time the granger-causality test in distribution tails rejects the null of no causality for the system of pair markets. Nominal size is set to 5 percent.

Table 8: Results of Granger causality test in mean

Crude	Setter (1)	Taker (2)	(1)-(2)
WTI	1.0000	0.8065	0.1935
Europe Brent	1.0000	0.0323	0.9677
Europe Norwegian Ekofisk	0.8065	0.7742	0.0323
Canadian Par	0.9355	0.9677	−0.0323
Canada Lloyd Blend	0.8387	0.6129	0.2258
Mexico Isthmus	0.9032	0.9355	−0.0323
Mexico Maya	0.9355	0.9677	−0.0323
Colombia Cano Limon	0.9032	0.9355	−0.0323
Ecuador Oriente	0.8387	0.8710	−0.0323
Angola Cabinda	0.7742	0.8065	−0.0323
Cameroon Kole	0.7097	0.6774	0.0323
Egypt Suez Blend	0.7097	0.7742	−0.0645
Oman Blend	0.8387	0.9355	−0.0968
Australia Gippsland	0.9355	0.9677	−0.0323
Malaysia Tapis	0.4839	0.9677	−0.4839
Mediter. Russian Urals	0.9677	0.9677	0.0000
China Daqing	0.9355	0.9677	−0.0323
Saudi Arabia Saudi Light	0.9355	0.9677	−0.0323
Saudi Arabia Arab Medium	0.9355	0.9355	0.0000
Saudi Arabia Saudi Heavy	0.9355	0.8710	0.0645
Asia Murban	0.8710	0.8387	0.0323
Asia Dubai Fateh	0.8387	0.8387	0.0000
Qatar Dukhan	0.8387	0.9032	−0.0645
Mediter. Seri K Iran Light	0.9355	0.8710	0.0645
Mediter. Seri K Iran Heavy	0.9355	0.8065	0.1290
Kuwait Blend	0.8065	0.8387	−0.0323
Algeria Saharan Blend	0.7097	0.8065	−0.0968
Europe Nigerian Bonny Light	0.8065	0.9032	−0.0968
Europe Forcados	0.9032	0.8710	0.0323
Europe Libyan Es Sider	0.7097	0.8710	−0.1613
Indonesia Minas	0.9355	0.9677	−0.0323
Venezuela Tia Juana	0.6452	1.0000	−0.3548
Mean absolute value			0.1109

Note: For each crude oil, the table displays the proportion of time the granger-causality test in mean rejects the null of no causality for the system of pair markets. Nominal size is set to 5 percent.

Table 9: Results of Granger causality test in distribution tails: Light density and sweet

Crude	Left Tail			Right Tail		
	Setter (1)	Taker (2)	(1)-(2)	Setter (1)	Taker (2)	(1)-(2)
WTI	0.8750	0.0000	0.8750	0.8750	0.1250	0.7500
Europe Brent	0.8750	0.1250	0.7500	0.8750	0.5000	0.3750
Europe Norwegian Ekofisk	0.7500	0.8750	−0.1250	0.1250	0.3750	−0.2500
Canadian Par	0.7500	1.0000	−0.2500	0.2500	0.2500	0.0000
Australia Gippsland	0.7500	0.8750	−0.1250	0.2500	0.2500	0.0000
Malaysia Tapis	0.7500	1.0000	−0.2500	0.2500	0.8750	−0.6250
Algeria Saharan Blend	0.6250	1.0000	−0.3750	0.2500	0.3750	−0.1250
Europe Nigerian Bonny Light	0.7500	1.0000	−0.2500	0.1250	0.3750	−0.2500
Europe Libyan Es Sider	0.7500	1.0000	−0.2500	0.3750	0.2500	0.1250
Mean absolute value			0.3611			0.2778

Note: For each crude oil, the table displays the proportion of time the granger-causality test in distribution tails rejects the null of no causality for the system of pair markets. Nominal size is set to 5 percent.

Table 10: Results of Granger causality test in distribution tails: Medium density and sweet

Crude	Left Tail			Right Tail		
	Setter (1)	Taker (2)	(1)-(2)	Setter (1)	Taker (2)	(1)-(2)
Colombia Cano Limon	0.2000	1.0000	−0.8000	0.4000	0.0000	0.4000
Angola Cabinda	0.8000	0.8000	0.0000	0.2000	0.0000	0.2000
Cameroon Kole	1.0000	1.0000	0.0000	0.0000	0.0000	0.0000
China Daqing	0.8000	0.8000	0.0000	0.0000	0.4000	−0.4000
Europe Forcados	1.0000	0.4000	0.6000	0.0000	0.0000	0.0000
Indonesia Minas	1.0000	0.8000	0.2000	0.0000	0.2000	−0.2000
Mean absolute value			0.2667			0.2000

Note: For each crude oil, the table displays the proportion of time the granger-causality test in distribution tails rejects the null of no causality for the system of pair markets. Nominal size is set to 5 percent.

Table 11: Results of Granger causality test in distribution tails: Medium density and sour

Crude	Left Tail			Right Tail		
	Setter (1)	Taker (2)	(1)-(2)	Setter (1)	Taker (2)	(1)-(2)
Mexico Isthmus	0.8333	0.9167	−0.0833	0.5833	0.0833	0.5000
Ecuador Oriente	0.5000	1.0000	−0.5000	0.8333	0.0000	0.8333
Egypt Suez Blend	1.0000	1.0000	0.0000	0.1667	0.0000	0.1667
Oman Blend	1.0000	1.0000	0.0000	0.0000	0.4167	−0.4167
Mediter. Russian Urals	1.0000	0.2500	0.7500	0.2500	0.6667	−0.4167
Saudi Arabia Saudi Light	0.5000	0.7500	−0.2500	0.1667	0.0833	0.0833
Saudi Arabia Arab Medium	0.5000	0.6667	−0.1667	0.1667	0.2500	−0.0833
Saudi Arabia Saudi Heavy	0.8333	0.5833	0.2500	0.0833	0.0833	0.0000
Asia Dubai Fateh	0.5000	1.0000	−0.5000	0.0000	0.4167	−0.4167
Mediter. Seri K Iran Light	0.8333	0.3333	0.5000	0.1667	0.0000	0.1667
Mediter. Seri K Iran Heavy	0.5833	0.4167	0.1667	0.3333	0.2500	0.0833
Kuwait Blend	0.8333	0.9167	−0.0833	0.0833	0.7500	−0.6667
Venezuela Tia Juana	0.7500	0.8333	−0.0833	0.3333	0.1667	0.1667
Mean absolute value			0.2564			0.3077

Note: For each crude oil, the table displays the proportion of time the granger-causality test in distribution tails rejects the null of no causality for the system of pair markets. Nominal size is set to 5 percent.

Table 12: Results of Granger causality test in distribution tails: OPEC

Crude	Left Tail			Right Tail		
	Setter (1)	Taker (2)	(1)-(2)	Setter (1)	Taker (2)	(1)-(2)
Saudi Arabia Saudi Light	0.2857	0.8571	−0.5714	0.0714	0.0000	0.0714
Saudi Arabia Arab Medium	0.2857	0.7857	−0.5000	0.0714	0.1429	−0.0714
Saudi Arabia Saudi Heavy	0.5714	0.5000	0.0714	0.0714	0.0000	0.0714
Asia Murban	0.7143	0.8571	−0.1429	0.0000	0.0714	−0.0714
Asia Dubai Fateh	0.5000	1.0000	−0.5000	0.0000	0.1429	−0.1429
Qatar Dukhan	0.8571	0.7857	0.0714	0.0714	0.2857	−0.2143
Mediterranean Seri K Iran Light	0.5714	0.2857	0.2857	0.0714	0.0000	0.0714
Mediterranean Seri K Iran Heavy	0.4286	0.5714	−0.1429	0.2143	0.0000	0.2143
Kuwait Blend	0.9286	0.9286	0.0000	0.0714	0.6429	−0.5714
Algeria Saharan Blend	0.9286	0.5000	0.4286	0.0714	0.0000	0.0714
Europe Nigerian Bonny Light	0.8571	0.7143	0.1429	0.2143	0.0714	0.1429
Europe Forcados	0.8571	0.2857	0.5714	0.0714	0.0000	0.0714
Europe Libyan Es Sider	1.0000	0.5714	0.4286	0.0714	0.0000	0.0714
Indonesia Minas	0.7143	0.5000	0.2143	0.0000	0.0714	−0.0714
Venezuela Tia Juana	0.5714	0.9286	−0.3571	0.3571	0.0000	0.3571
Mean absolute value			0.2952			0.1524

Notes: For each crude oil, the table displays the proportion of time the granger-causality test in distribution tails rejects the null of no causality for the system of pair markets. Nominal size is set to 5 percent.

Table 13: Results of Granger causality test in distribution tails: Non-OPEC

Crude	Left Tail			Right Tail		
	Setter (1)	Taker (2)	(1)-(2)	Setter (1)	Taker (2)	(1)-(2)
WTI	0.9375	0.0000	0.9375	0.9375	0.4375	0.5000
Europe Brent	0.9375	0.1250	0.8125	0.9375	0.4375	0.5000
Europe Norwegian Ekofisk	0.8125	0.7500	0.0625	0.1250	0.1875	−0.0625
Canadian Par	0.6250	0.8750	−0.2500	0.1875	0.1250	0.0625
Canada Lloyd Blend	0.0000	0.2500	−0.2500	0.0625	0.1250	−0.0625
Mexico Isthmus	0.5625	0.8125	−0.2500	0.5625	0.3125	0.2500
Mexico Maya	0.5625	0.8750	−0.3125	0.3750	0.2500	0.1250
Colombia Cano Limon	0.3125	0.8750	−0.5625	0.3750	0.2500	0.1250
Ecuador Oriente	0.6250	0.8750	−0.2500	0.7500	0.1250	0.6250
Angola Cabinda	0.7500	0.6875	0.0625	0.2500	0.1875	0.0625
Cameroon Kole	0.8750	0.7500	0.1250	0.0625	0.3750	−0.3125
Egypt Suez Blend	0.8125	0.8125	0.0000	0.1250	0.1875	−0.0625
Oman Blend	0.8125	0.6875	0.1250	0.1250	0.5000	−0.3750
Australia Gippsland	0.6875	0.6875	0.0000	0.1875	0.3750	−0.1875
Malaysia Tapis	0.3750	0.9375	−0.5625	0.5000	0.6875	−0.1875
Mediterranean Russian Urals	0.8750	0.5000	0.3750	0.1875	0.6875	−0.5000
China Daqing	0.7500	0.8125	−0.0625	0.0625	0.5625	−0.5000
Mean absolute value			0.2941			0.2647

Notes: For each crude oil, the table displays the proportion of time the granger-causality test in distribution tails rejects the null of no causality for the system of pair markets. Nominal size is set to 5 percent.

References

ADELMAN, M. A. (1984), "International oil agreements", *The Energy Journal,* 5, 1-9.

AHN, D., BOUDOUKH, J., RICHARDSON, M., WHITELAW R.F. (2002), "Partial Adjustment or Stale Prices? Implications from Stock Index and Futures Return Autocorrelations", *Review of Financial Studies,* 15, 655-689.

BACON, R. AND TORDO, S. (2005), "Crude Oil Price Differentials and Differences in Oil Qualities: A Statistical Analysis", ESMAP Technical Paper, 081, 1-31.

BARRETT, A. B, BARNETT, L. AND SETH, A. K (2010), "Multivariate Granger Causality and Generalized Variance", *Physical Review E*, 81:041907, 2-14.

BERKOWITZ, J. , CHRISTOFFERSEN, P. F. AND PELLETIER, D. (2011), "Evaluating Value-at-Risk Models with Desk-level Data", *Management Science*, 57, 12, 2213-2227.

BLACK, F. (1976), "The Pricing of Commodity Contracts", *Journal of Financial Economics*, 3, 167-179.

BMO COMMODITY DERIVATIVES GROUP (2004), "Managing Heavy Oil Price Risk", corporate.bmo.com/cm/market/cdcom/images/Managing_Heavy_Oil_Price_Risk.pdf.

BOLLERSLEV, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics,* 31, 3, 307–327.

BOUDOUKH, J., RICHARDSON, M. P. AND WHITELAW, R. F. (1994), "A Tale of Three Schools: Insights on Autocorrelations of Short-Horizon Returns", Review of Financial Studies, 7, 3, 539-573.

BREMOND, V., HACHE, E. AND MIGNON, V. (2011), "Does OPEC Still Exist as a Cartel? An Empirical Investigation", forthcoming in *Energy Economics*.

CANDELON B., COLLETAZ, G., HURLIN C. AND TOKPAVI S. (2011), "Backtesting Value-at-Risk: A GMM Duration-Based Test", *Journal of Financial Econometrics*, 9, 2, 314-343.

CHRISTIE, A. (1982), "The Stochastic Behavior of Common Stock Variances - Value, Leverage, and Interest Rate Effects", *Journal of Financial Economics*, 10, 407–432.

CHRISTOFFERSEN, P. F. (1998), "Evaluating Interval Forecasts", *International Economic Review,* 39, 841-862.

COOK J. (1998), “California Crude Oil Prices”, Mimeo www.eia.doe.gov/pub/

DUFOUR, J.-M. (2006), "Monte Carlo Tests with Nuisance Parameters: a General Approach to Finite Sample Inference and Nonstandard Asymptotics", *Journal of Econometrics*, 27, 2, 443-477.

ENGLE, R. F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation", *Econometrica* 50, 987-1008.

Engle, R. F. and Manganelli S. (2004), "CAViaR: Conditional autoregressive Value at Risk by regression quantile", *Journal of Business and Economic Statistics*, 22, 367-381.

Eom, K. S., Hahn, S. B. and Joo S. (2004), "Partial Price Adjustment and Autocorrelation in Foreign Exchange Markets", Preprint, University of California at Berkeley.

Fattouh, B. (2006), "Middle East crude pricing and the Oman crude oil futures contract: a critical assessment", *Oxford Energy Comment*, August, 1-7.

Fattouh, B. (2010), "The dynamics of crude oil price differentials", *Energy Economics*, 32, 334-342.

Fattouh, B. (2011), "An Anatomy of the crude oil pricing system", *The Oxford Institute for Energy Studies*, WPM 40.

Gelper, S. and Croux, C. (2007), "Multivariate out-of-sample tests for Granger causality", *Computational Statistics & Data Analysis*, 51, 7, 3319-3329.

Granger, C .W. J. (1980), "Testing for causality: A personal view", *Journal of Economic Dynamics and Control*, 2, 329 352.

Granger, C. W. J. and Newbold, P. (1986), Forecasting Economic Time Series Academic Press (2nd edition).

Gülen, S. G. (1997), "Regionalization in the World Crude Oil Market", *The Energy Journal,* 18, 109-126.

Gülen, S. G. (1999), "Regionalization in the World Crude Oil Market: Further Results", *The Energy Journal*, 20, 125-139.

Hamilton, J. D. (2003), "What Is an Oil Shock?", *Journal of Econometrics*, 113, 363-398.

Hammoudeh, S., Thompsion, M., and Ewing, B., (2008), "Threshold cointegration analysis of crude oil benchmarks", *The Energy Journal*, 29, 79-95.

Hansen, B. E. (1994), "Autoregressive Conditional Density Estimation", *International Economic Review*, 35, 3, 705-30.

Harvey, C. R. and A. Siddique A. (1999), "Autoregressive Conditional Skewness", *Journal of Financial and Quantitative Analysis,* 34, 4, 465-488.

Harvey, C. R. and Siddique A. (2000), "Conditional Skewness in Asset Pricing Tests", *Journal of Finance,* 55, 1263-1295.

Hong, Y., Liu, Y., Wang, S. (2009), "Granger Causality in Risk and Detection of Extreme Risk Spillover between Financial Markets", *Journal of Econometrics*, 150, 2, 271–287.

Horsnell, P and Mabro, R. (1993), Oil Markets and Prices: the Brent Market and the formation of world oil prices, Oxford: Oxford University Press.

Jondeau, E. and Rockinger, M. (2003), "Testing for Differences in the Tails of Stock-Market Returns", *Journal of Empirical Finance*, 10, 5, 559-581.

Kilian, L. (2008), "A Comparison of the Effects of Exogenous Oil Supply Shocks on Output and Inflation in the G7 Countries", *Journal of the European Economic Association*, 6, 1, 78-121.

Kleit, A. N. (2001), "Are Regional Oil Markets Growing Closer Together? An Arbitrage Cost Approach", *The Energy Journal*, 22, 2, 1-15.

Koenker, R. and Bassett, G. (1978), "Regression Quantiles", *Econometrica*, 46, 1, 33-50.

Lanza, A., Manera M. and Giovannini M. (2003), "Oil and Product Price Dynamics in International Petroleum Markets", Working Papers 2003.81, Fondazione Eni Enrico Mattei.

Mabro, R. (2005), "The international oil price regime: origins, rationale and assessment", *The Journal of Energy Literature*, 10, 1, 3-20.

Miller, K. D., Chevalier, M. T., and Leavens, j. (2010), "The Role of WTI as a crude oil Benchmark", *Purvin & Gertz Inc.*

Milonas, N. and Henker, T. (2001), "Price spread and convenience yield behaviour in the International oil market", *Applied Financial Economics*, 11, 23-36.

Montepeque, J. (2005), "Sour crude pricing: a pressing global issue, *Middle East Economic Survey*, 48, 1-42.

Sadorsky, P. (1999), "Oil Price Shocks and Stock Market Activity", *Energy Economics,* 2, 449-469.

Weiner, R. J. (1991), "Is the World Oil Market 'One Great Pool': Comment ", *The Energy Journal*, 12, 3, 95-107.

Wlazlowski, S., Hagstromer, B. and Giulietti M. (2011), "Causality in Crude Oil Prices", *Applied Economics*, 43, 24, 3337-3347.