Economics Bulletin

Volume 31, Issue 1

A wavelet analysis of oil price volatility dynamic

François Benhmad
LAMETA

Abstract

In this the paper we investigate the oil price volatility, by studying the causal relationships between different volatilities captured at different time scales. We first decompose the oil price volatility at various scales of resolution or frequency ranges by using wavelet analysis. We then explore the causalities between absolute returns of oil prices at different time scales. As traditional Granger causality test, designed to detect linear causality, is ineffective in uncovering certain nonlinear causal relationships, we use the nonlinear causality test introduced by Péguin-Feissolle and Teräsvirta (1999) and Péguin-Feissolle, Strikholm and Teräsvirta (2008). Our results confirm the fact that the vertical dependence is a strong stylised fact of oil returns volatility. But, the main finding consists on the presence of a feed- back effect from high frequency traders to low frequency traders. In contrast to Gençay et al. (2010), we prove that high frequency shocks could have an impact outside their boundaries and reach the long term traders.

Citation: François Benhmad, (2011) "A wavelet analysis of oil price volatility dynamic", *Economics Bulletin*, Vol. 31 no.1 pp. 792-806.
Submitted: Oct 18 2010. **Published:** March 14, 2011.

1. Introduction

The oil price shocks can be a major source of macroeconomic variability. Therefore, modelling and forecasting volatility of oil price is a high research topic. The fundamentals (supply and demand) can explain the dynamics of oil prices, but the increasing of speculative behavior and the oil market heterogeneity have made oil-market prices harder to predict (Bacon and Kojima (2008)). The goal of this paper is thus to investigate the oil price volatility, by studying the causal relationships between different volatilities captured at different time scales. To attain this goal, we first decompose the oil price volatility at various scales of resolution or frequency ranges by using wavelet analysis. We then explore the causalities between absolute returns of oil prices at different time scales. Some studies based on wavelet analysis explored linear causal relationships between economic and financial variables (Ramsey and Lampart (1998), Almasri and Shukur (2003), Kim and In (2003), Zhang and Farley (2004), Dalkır (2004), Mitra (2006), In and Kim (2006), and Cifter and Ozun (2007)). However, given the evidence on the nonlinear dynamics of economic and financial time series, some authors argue that the traditional Granger causality test, designed to detect linear causality, is ineffective in uncovering certain nonlinear causal relationships and recommend the use of nonlinear causality tests. (Baek and Brock (1992), Hiemstra and Jones (1994), Bell, Kay, and Malley (1996), Hiemstra and Kramer (1997), Skalin and Teräsvirta (1999), Chen, Rangarjan, Feng, and Ding (2004), Li (2006).We thus use the nonlinear causality test introduced by Péguin-Feissolle and Teräsvirta (1999) and Péguin-Feissolle, Strikholm and Teräsvirta (2008).

Our results confirm the fact that the vertical dependence is a strong stylised fact of oil returns volatility. But, the main finding consists on the presence of a feed- back effect from high frequency traders to low frequency traders. In contrast to Gençay et al. (2010), we prove that high frequency shocks could have an impact outside their boundaries and reach the long term traders.

The rest of the paper is organized as follows. The heterogeneous market hypothesis is presented in section 2. The section 3 introduces the Wavelet analysis and the nonlinear causality tests. The section 4 displays and comments the empirical study of the oil market volatility. The section 5 summarizes and concludes the paper.

2. Heterogeneous market hypothesis: literature review

According to heterogeneous market hypothesis, there is a presence of heterogeneity in the traders (Müller et al. (1997)). The market participants may differ in their beliefs, their expectations, risk profiles, informational sets.....etc. These differences translate to their sensitivity to different time horizons. Thus, the traders are categorized according to their characteristic time horizons or dealing frequencies. On the side of high frequency trading, we find intraday speculators and market makers. The central banks and institutional investors like pensions funds represent the low frequency trading.

De Long and al. (1990a) and (1990b) make a clear distinction between two categories of market participants. The first one is constituted by the fundamentalists who are well informed, more rational, risk-averse and base their trading rule on the fundamental value of asset prices determined by the dividend discount model. The second category is composed by the noise traders or chartists who are less informed, irrational, and less risk-averse; their trading rule is based on technical analysis and consists on extrapolating recent trend of asset prices. The relative portion of each trader type evolves during time and there exists a permanent shift between the two trading strategies according to their past relative performance. Price fluctuations are exacerbated by an interaction between a stabilizing force and a destabilizing one. The stabilizing force progressively pushes prices toward their fundamental values when the market is dominated by fundamentalists. When the noise traders are the dominant category, they constitute a destabilizing force which cause securities large deviation away from their fundamental value and lead to excessive volatility (Kyrtsou, Labys, Terraza, (2004)). This interaction can create complex price behaviour and a possible route to chaos (Hommes (2004)).

As a consequence, this evidence of market heterogeneity leads to a presence of different dealing frequencies, and thus different reactions to the same news in the same market. Each market component has its own reaction time to information, related to its time horizon and characteristic dealing frequency (Dacorogna et al. (2001)). Thus, the volatility process has scaling behaviour. We can distinguish the low frequency volatility (coarse) which captures the perceptions and actions of long term horizon traders, and a high frequency volatility (fine) which captures the expectations and decisions of short term traders (Gençay, Gradojevic, Selcuk and Whitcher (2010)). To further examine the volatility multi-frequency structure and identify the relative presence of market components, Müller et al. (1997) introduce a heterogeneous GARCH model (HARCH) which differs from all other ARCH-type processes in the unique property of considering the volatilities of returns over different time horizons. We assume that the time –frequency analysis of wavelet transform is a pertinent statistical tool for modelling financial markets heterogeneity and price dynamics induced with influence from different types of investors characterized by different time horizons.

Some authors (Müller et al. (1997) and Dacorogna et al. (2001)) show that the asymmetry comes from the fact that coarse volatility predicts fine volatility better than the other way around (see also Zumbach (2007) and Borland et al. (2008)). The explanation is that, when the coarse volatility increases or decreases, the short-term traders modify their trading activity and thus change the level of the fine volatility; besides, the level of fine volatility does not influence the long-term traders (Müller et al. (1997)). The presence of this information flow from large to short time scales motivates a cascade model of volatility (Zumbach and Lynch (2001)). Arneodo et al. (1998) show that the nature of correlations that are implied by this cascade across scales has a profound implication on the market risk. Gençay, Gradojevic, Selçuk and Whitcher (2010) show that in heterogeneous markets, a low-frequency shock to the system penetrates through all layers to the short-term traders, while high frequency shocks appear to be short lived and may have no impact outside their boundaries.

3. Wavelets analysis

The wavelet analysis was introduced to overcome the Fourier transform limitations. Indeed, Fourier series requires that the time series under study must be periodic. In addition, it assumes that frequencies do not evolve in time. The inadequacy of this stationary assumption in dealing with economic and time series stems from the fact that theses time series are subject to structural breaks, regime switching, GARCH effects,outliers. Although the short-time Fourier transform and Gabor transform tried to deal with the stationary assumption by using a single fixed window, they have the disadvantage of capturing more and more cycles within the analysis window as frequency increases. The innovation of wavelet transform is that its window is adjusted automatically to the high or low frequency as it uses short window for high frequency and long window at low frequency by employing time compression or dilatation rather than a variation of frequency in the modulated signal. This is achieved by dividing the time axis into a sequence of successively smaller segments (Percival and Walden (2000)). The discrete wavelet transform (DWT) transforms a time series by dividing it into segments of the time domain called ''scales '' or frequency ''bands'' (Priestley (1996)). The scales; from the shortest to the largest; represent progressively high and low frequency fluctuations.

There are two types of wavelets; father wavelets ϕ and mother wavelet ψ. The father wavelet

$$\int \phi(t)dt = 1,\ \int \psi(t)dt = 0 \tag{1}$$

The father wavelets represent the smooth or low frequency parts of a signal, and the mother wavelets capture the details or high-frequency components. Thus, father wavelets and mother wavelets capture respectively the signal trend components and all deviations from this trend. A lot of wavelets families have been introduced. The most used empirically are orthogonal wavelets such as the Haar, Daublets, Symmlets and coiflets (Daubechies (1992).

Wavelets consist on a two-scale dilatation equation. The dilatation equation of father wavelet $\phi(x)$ can be expressed as follows:

$$\phi(x) = \sqrt{2}\sum_k l_k \phi(2x-k) \tag{2}$$

The mother wavelet $\psi(x)$ can be derived from the father wavelet by the following formula:

$$\psi(x) = \sqrt{2}\sum_k h_k \phi(2x-k). \tag{3}$$

The coefficients l_k and h_k are called respectively the low-pass and high-pass filter coefficients. They can be expressed as:

$$l_k = \frac{1}{\sqrt{2}}\int \phi(t)\phi(2t-k)dt \tag{4}$$

$$h_k = \frac{1}{\sqrt{2}}\int \psi(t)\phi(2t-k)dt. \tag{5}$$

Thus, a wavelet representation of a signal or a function $f(t)$ in $L^2(\mathbf{R})$ consists on a sequence of projections onto father and mother wavelets through scaling (stretching and compressing) and translation.

The projections give the wavelet coefficients $s_{J,k}$, $d_{J,k}$,….., $\mathrm{d}_{1,k}$:

$$s_{J,k} \approx \int \phi_{J,k}(t) f(t)dt \tag{6}$$

$$d_{j,k} \approx \int \psi_{j,k}(t) f(t)dt, \text{ for j=1,2,……. } J. \tag{7}$$

The coefficients $s_{J,k}$ (smooth) represent the smooth behaviour of the signal at the coarse scale 2^J (trend). The coefficients $d_{j,k}$ (details) coefficients represent deviations from the trend; $d_{J,k}$, $d_{J-1,k}$,...., $d_{1,k}$ capture the deviations from the coarsest to finest scale and $d_{J-1,k}$,...., $d_{1,k}$

The wavelet representation can be expressed as follows:

$$f(t) = \sum_k s_{J,k}\phi_{J,k}(t) + \sum_k d_{J,k}\psi_{J,k}(t) + \sum_k d_{J-1,k}\psi_{J-1,k}(t) + ... + \sum_k d_{1,k}\psi_{1,k}(t) \tag{8}$$

where J is the number of multiresolution levels, and k ranges from 1 to the number of coefficients in each level.

Assuming that:

$$S_J(t) = \sum_k s_{J,k}\phi_{J,k}(t) \tag{9}$$

and

$$D_j(t) = \sum_k d_{j,k}\psi_{J,k}(t) \text{ for j=1, 2,…, } J \tag{10}$$

where S $_J$ (t) refers to the decomposed time series using scaling function at scale J and $D_j(t)$ refers to the decomposed time series using wavelet function at scales j up to scale J , the equation (8) can be expresses as :

$$f(t) = S_J(t) + D_J(t) + D_{J-1}(t) + ... + D_1(t). \tag{11}$$

As each term in (11) represents an orthogonal component of the signal $f(t)$ at different resolutions (scales or frequency ranges). Thus, equation(11) is called a multiresolution analysis (Mallat (1989)).

4. Empirical evidence:

4.1. Data description:

The data set consists on daily data of the WTI oil prices ranging from September 8, 1992 to December 31, 2008. The returns of oil prices in a continuous compound basis are calculated as $r_t = \log\left(\frac{P_t}{P_{t-1}}\right)$ where P_t and P_{t-1} are respectively the prices for day t and t-1.

The descriptive statistics for return series are summarized in Table I.

Table 1. Descriptive statistics	
	Oil returns
Number of obs.	4096
Mean	7,60E-05
Standard.dev	0,010549
Skewness	-0,225042
Kurtosis	8,032999
Jarque-Bera	4357,74
Jarque-Bera(p-value)	0.0000000

We take the oil price absolute returns as a proxy of the volatility. Figures 1, 2, and 3 present respectively the plot of oil prices, oil returns, and oil absolute returns(see appendix 2).

4.2. Oil price absolute returns wavelet decomposition:

In order to perform a wavelet decomposition of oil price absolute returns in a set of six orthogonal components D1,D2, . . . , D6, that stand for different dealing frequencies in the oil market, we choose the Symmlet basis LA(8). This wavelet is orthogonal, near symmetric and have a compact support and good smoothness properties. Figure 4 presents the wavelet decomposition plot of oil price absolute returns (see appendix 2). Figure 5 and table II show the time scale interpretation of wavelet multiresolution analysis; each time scale corresponds to a specific dealing frequency of a category of traders at the oil market.

Figure 5. Dealing frequencies according to wavelet decomposition

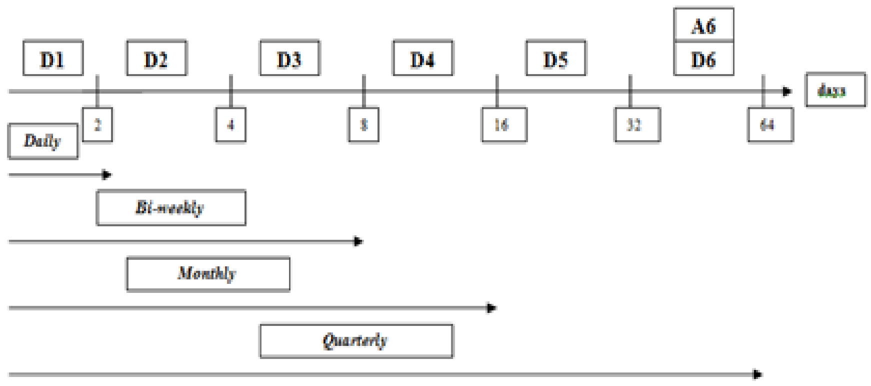

Table II. Frequency interpretation of MRA scales

Scale	trading frequency
D1	1-2 days
D2	2-4 days
D3	4-8 days
D4	8-16 days
D5	16-32 days
D6	32-64 days

4.3 Granger causality tests:

We investigate the causal relationships between different oil prices volatilities captured at different frequency bands by using the nonlinear causality test introduced by Péguin-Feissolle and Teräsvirta (1999) and Péguin-Feissolle, Strikholm and Teräsvirta (2008).
The standard linear Granger causality test works best when the true causal relationship is linear, but loses a lot of power when this is no longer the case. To overcome this drawback, Péguin-Feissolle and Teräsvirta (1999) and Péguin-Feissolle, Strikholm and Teräsvirta (2008) introduced non causality tests built on a general non linear framework. Two of these tests (general and additive) are based on a Taylor expansion of the nonlinear model around a given point in a sample space. Another test is based on artificial networks, and puts more

restrictions on the functional form of a potential causal relationship between two variables than the others. The causal relationship can be represented by regression functions as TAR, ESTAR, LSTAR, SETAR, general nonlinear models,…..etc.

None of the tests dominates the other. Their behavior depends on the non linear functional form of the relationship between two variables. Thus, one test (general, semi-additive, neural network based) can strongly dominates the others according to a specific causality type but yields poor results when there is a change in the causality functional form**.** We must use the test which seems to capture most of the relationship as the functional form of the causal relationship strongly affects the outcome of the tests. Thus, applying the general test when the relationship is semi-additive may result in a substantial loss of power compared to the power of the additive test.Therefore, using linear, general, additive, or neural tests is very useful at approximating all potential causal relationships that may exist between two variables.

The results of the causality tests between oil price absolute returns decomposed into six frequency bands are reported in Table III (see appendix 1)

It is worth noting that, in many cases, nonlinear causality tests give different results than the standard Granger non causality test; for instance for D6, the null hypotheses of no causality from D2 to D1, D3, D6, are accepted by the linear causality test and rejected by the nonlinear tests; this conclusion is the same when we consider the causality from D4 to D3, D5. Therefore, the nonlinear causality tests, i.e. the two tests based on a Taylor series approximation as well as the test based on the artificial neural network, may detect causality that would be ignored by the linear Granger test.

Our main finding can be resumed as follows:

- The null hypothesis of no causality from all the frequency bands to D1 are not accepted by at least one causality test. Thus, there is a causal relationship from all the frequency bands towards D1. On the opposite, there is no causal relationship from D1 to D4, D5, D6 and D7. Therefore, considering the frequency band D1, we may assert that all frequency bands (low, middle, and high) linearly or nonlinearly cause the highest frequency band D1 which corresponds to the trading behavior of intraday-traders or noise traders.

- There are strong bidirectional causal relationships between the first three highest frequency bands, i.e. D1, D2 and D3. They correspond to an investment horizon less than 10 days; the time horizon imposed by regulation authorities to financial institutions in order to compute Value at Risk. Thus, on can consider theses frequency ranges as representing an homogeneous category of traders corresponding to high frequency component of oil prices volatility.

- When we crowd out the frequency band D1 corresponding to noise traders whom investment horizon is less than 2 days, there is a bidirectional causal relationship between all frequency bands which reveals a strong feed-back effect in the oil volatility process. Thus, every frequency bands which corresponds to a specific class of traders in the oil market is able to impact the others frequency ranges, i.e the other categories of traders.

5. Conclusion

We investigate the causal relationships between different oil prices volatilities captured at different time scales by using wavelet analysis and the nonlinear causality test introduced by Péguin-Feissolle and Teräsvirta (1999) and Péguin-Feissolle, Strikholm and Teräsvirta (2008).

Our results confirm the fact that the vertical dependence is a strong stylised fact of oil returns volatility. However, our main finding consists on the presence of a feed- back effect from high frequency traders to low frequency traders. In contrast to Gençay, Gradojevic, Selçuk and Whitcher (2010), we prove that high frequency shocks could have an impact outside their boundaries and reach the long term traders. In contrast to (Müller et al. (1997)), the level of fine volatility may have a strong influence on the long-term traders. The motivation of a cascade model of volatility (Zumbach and Lynch (2001)) comes under question as there is presence of information flow not only from large to short time scales but also the reverse hypothesis has been proved.

References

Almasri, A., Shukur, G. (2003) “An Illustration of the Causality Relationship Between Government Spending and Revenue Using Wavelets Analysis on Finnish Data”, *Journal of Applied Statistics*, 30, 5, 571-584.

Arneodo, A., Bacry, E., Manneville, S., Muzy, J. F. (1998) “Analysis of Random Cascades Using Space-Scale Correlation Functions**”,** *Physical Review Letters*, 80, 4, 708-711.

Bacon, R., Kojima, M. (2008) “Coping with Oil Price Volatility”, Energy Sector Management Assistance Program, Special Report 005/08, The World Bank, Washington.

Baek, E., Brock, W. (1992) "A general test for nonlinear Granger causality: bivariate model", Working Paper, Iowa State University and University of Wisconsin Madison.

Bell, D., Kay, J., Malley, J. (1996) “A non-parametric approach to non-linear causality testing”, *Economics Letters*, 51, 7–18.

Bénassy-Quéré, A., Mignon, V., Penot, A. (2007) “China and the relationship between the Oil price and the Dollar”, *Energy Policy*, 35,5795-5805.

Chen, Y., Rangarjan, G., Feng, J., Ding, M. (2004) “Analyzing multiple nonlinear time series with extended Granger causality”, *Physics Letters A*, 324, 26–35.

Cifter, A., Ozun, A. (2007) “Multi-scale Causality between Energy Consumption and GNP in Emerging Markets: Evidence from Turkey”, MPRA Paper No. 2483.

Couderc, V., Mignon, V., Penot, A. (2008) “Oil price and the Dollar”, *Energy Studies Review,* 18, 2.

Dacorogna, M., Gençay, R., Müller, U., Pictet, O., Olsen, R. (2001) *An Introduction to High-Frequency Finance*, Academic Press: San Diego.

Dalkir, M. (2004) “A new approach to causality in the frequency domain”, *Economics Bulletin*, 3, 44, 1-14.

Gençay, R., Gradojevic, N., Selçuk, F., Whitcher, B. (2010) Asymmetry of information flow between volatilities across time scales, *Quantitative Finance*, forthcoming.

Ghysels, E., Santa-Clara, P., Valkanov, R. (2006) “Predicting volatility: how to get the most out of returns data sampled at different frequencies », *Journal of Econometrics*, 131, 59–95.

Golub, S.S., (1983) “Oil price and exchange rates”, *The Economic Journal*, 93, 371, 576-593.

Hiemstra, C., Jones, J. (1994) "Testing for linear and nonlinear Granger Causality in the Stock Price-Volume Relation", *Journal of Finance,* 5, 1639-1664.

Hiemstra, C., Kramer, C. (1997) " Nonlinearity and Endogeneity in Macro-Asset Pricing", *Studies in Nonlinear Dynamics and Econometrics*, 2, 3, 61–76.

In, F., Kim, S. (2006) “The Hedge Ratio and the Empirical Relationship Between the Stock and Futures Markets: A New Approach Using Wavelets”, *The Journal of Business*, 79, 799-820.

Kim, S., In, H. F. (2003) “The relationship between financial variables and real economic activity: evidence from spectral and wavelet analyses”, *Studies in Nonlinear Dynamics and Econometrics*, 7, 4, Article 4.

Krugman, P. (1983a) “Oil and the dollar”, In: Bhandari, J.S., Putman, B.H. (Eds), *Economic Interdependence and Flexible Exchange rates,* MIT Press, Cambridge, MA.

Krugman, P. (1983b) “Oil shocks and exchange rate dynamics”, In: Frenkel, J.A. (Ed.), *Exchange Rates and International Macroeconomics,* University of Chicago Press, Chicago.

Li, J. (2006) “Testing Granger Causality in the presence of threshold effects”, *International Journal of Forecasting*, 22, 771–780.

McGuirk, A.K. (1983) “Oil price changes and real exchange rate movements among industrial countries”, *IMF Staff Papers*, 30, 843-883.

Mitra, S. (2006) “A wavelet filtering based analysis of macroeconomic indicators: the Indian evidence “, *Applied Mathematics and Computation,* 175 1055–1079.

Péguin-Feissolle, A., Strikholm, B., Teräsvirta, T. (2008) “Testing the Granger non causality hypothesis in stationary non linear models of unknown functional form”, CREATES Research Paper 2008-19, School of Economics and Management, University of Aarhus.

Péguin-Feissolle, A., Teräsvirta, T. (1999) “A General Framework for Testing the Granger non-causality Hypothesis”, Stockholm School of Economics Working Paper Series in Economics and Finance No: 343.

Ramsey, J.B., Lampart, C. (1998) “Decomposition of Economic Relationships by Timescale Using Wavelets”, *Macroeconomic Dynamics*, 2, 1, 49–71.

Rogoff, K. (1991) “Oil productivity, government spending and the real yen-dollar exchange rate”, Federal Reserve Bank of San Francisco Working Papers 91-06.

Skalin, J., Teräsvirta, T. (1999) “Another look at Swedish business cycles, 1861-1988”, *Journal of Applied Econometrics*, 14, 359–378.

Zhang, C., Farley, A. (2004) “A Multiscaling Test of Causality Effects Among International Stock Markets”, *Neural, Parallel and Scientific Computations*, 12, 1, 91-112.

Zumbach, G. (2007) “Time reversal invariance in finance”, Working paper, Available online at: SSRN: http://ssrn.com/abstract=1004992.

Zumbach, G., Lynch, P. (2001) “Heterogeneous volatility cascade in financial markets”, *Physica A*, 298, 521–529.

Appendix 1: Nonlinear causality testing

1. Noncausality testing based on a Taylor series approximation

The test is based on a Taylor expansion of the nonlinear function:

$$y_t = f^*(y_{t-1},\ldots,y_{t-q},x_{t-1},\ldots,x_{t-n},\theta^*)+\varepsilon_t \qquad (1)$$

where θ^* is a parameter vector and $\varepsilon_t \sim nid(0,\sigma^2)$; the sequences $\{x_t\}$ and $\{y_t\}$ are weakly stationary and ergodic. The functional form of f^* is unknown but we assume that is adequately represents the causal relationship between x_t and y_t. Moreover, we assume that f^* has a convergent Taylor expansion at any arbitrary point of the sample space for every $\theta^* \in \Theta$ (the parameter space). In order to apply (1) for testing noncausality hypothesis, it is stated that x_t does not cause y_t if the past values of x_t does not contain any information about y_t that is already contained in the past values of y_t itself. More specifically, under the noncausality hypothesis, we have:

$$y_t = f\left(y_{t-1},\ldots,y_{t-q},\theta\right)+\varepsilon_t. \qquad (2)$$

To test (2) against (1), following Péguin-Feissolle and Teräsvirta (1999), we linearize f^* in (1) by expanding the function into a *k*th-order Taylor series around an arbitrary fixed point in the sample space. After approximating f^*, merging terms and reparametrizing, we obtain:

$$\begin{aligned} y_t = {} & \beta_0 + \sum_{j=1}^{q}\beta_j y_{t-j} + \sum_{j=1}^{n}\gamma_j x_{t-j} + \sum_{j_1=1}^{q}\sum_{j_2=j_1}^{q}\beta_{j_1 j_2} y_{t-j_1} y_{t-j_2} + \sum_{j_1=1}^{q}\sum_{j_2=1}^{n}\delta_{j_1 j_2} y_{t-j_1} x_{t-j_2} \\ & + \sum_{j_1=1}^{n}\sum_{j_2=j_1}^{n}\gamma_{j_1 j_2} x_{t-j_1} x_{t-j_2} + \ldots + \sum_{j_1=1}^{q}\sum_{j_2=j_1}^{q}\cdots\sum_{j_k=j_{k}-1}^{q}\beta_{j_1\ldots j_k} y_{t-j_1}\ldots y_{t-j_k} \\ & + \ldots + \sum_{j_1=1}^{n}\sum_{j_2=j_1}^{n}\cdots\sum_{j_k=j_{k}-1}^{n}\gamma_{j_1\ldots j_k} x_{t-j_1}\ldots x_{t-j_k} + \varepsilon_t^* \end{aligned} \qquad (3)$$

where $\varepsilon_t^* = \varepsilon_t + R_t^{(k)}(x,y)$, $R_t^{(k)}$ being the remainder, and $n \le k$ and $q \le k$ for notational convenience. Expansion (3) contains all possible combinations of lagged values of y_t and x_t up to order k. The assumption that x_t does not cause y_t means that all terms involving functions of elements of lagged values of x_t in (3) must have zero coefficients. According to Péguin-Feissolle and Teräsvirta (1999), there are two practical difficulties related to equation (3). One is numerical and the other one has to do with the amount of information. The numerical problem arises because the regressors in (3) tend to be highly collinear if both k, q and n are large. The other problem is that the number of regressors increases rapidly with k, so that the number of degrees of freedom may become rather small. A practical solution to both problems consists in replacing some observation matrices by their largest principal components. First divide the regressors in (3) into two groups: those being the function of lags of y_t only and the rest. Replace the regressors in (3) by the first p^* principal components of each matrix of observations. The null hypothesis is that the principal components of the latter group have zero coefficients. This yields the test statistic:

$$General = \frac{(SSR_0 - SSR_1)/p^*}{SSR_1/(T-1-2p^*)} \quad (4)$$

where SSR_0 and SSR_1 are obtained as follows. Regress y_t on 1 and the first p^* principal components of the matrix of lags of y_t only, form the residuals $\hat{\varepsilon}_t$, t=1,...,T, and the corresponding sum of squared residuals SSR_0. Then regress $\hat{\varepsilon}_t$ on 1 and all the terms of the two principal components matrices, form the residuals and the corresponding sum of squared residuals SSR_1. The test statistic has approximately an *F*-distribution with p^* and $T-1-2p^*$ degrees of freedom.

The problem of degrees of freedom is less acute if we can assume that the general model is "semi-additive":

$$y_t = g(y_{t-1},\ldots,y_{t-q},\theta_g) + f(x_{t-1},\ldots,x_{t-n},\theta_f) + \varepsilon_t \quad (5)$$

where $\theta' = (\theta_g', \theta_f')'$ is the parameter vector; in this case, x_t does not cause y_t if $f(x_{t-1},\ldots,x_{t-n},\theta_f)$*=constant*. We linearize both functions into a kth-order Taylor series as before and we obtain the statistic called *Additive*.

2. Noncausality test based on artificial neural networks

The ANN-based noncausality tests is characterized by a single hidden layer network with a logistic neural function and related to model (5), that is, semi-additivity of the functional form has to be assumed before applying the test; $f(x_{t-1},\ldots,x_{t-n},\theta_f)$ in (5) can be approximated by

$$\theta_0 + \tilde{w}_t'\alpha + \sum_{j=1}^{p} \beta_j \frac{1}{1+e^{-\gamma_j' w_t}} \quad (6)$$

where $\theta_0 \in \Re$, $w_t = (1, \tilde{w}_t')'$ is a $(n+1)\times 1$ vector, $\tilde{\omega}_t = (x_{t-1},\ldots,x_{t-n})'$, $\alpha = (\alpha_1,\ldots,\alpha_n)'$ are $n\times 1$ vectors, and the $\gamma_j = (\gamma_{j0},\ldots,\gamma_{jn})'$, for *j=1,...,p*, are $(n+1)\times 1$ vectors. The sequences $\{x_t\}$ and $\{y_t\}$ are weakly stationary and ergodic. The null hypothesis of Granger noncausality, i.e. that x_t does not cause y_t, can be formulated as $H_{02}: \alpha = 0 \ and \ \beta = 0$ where $\beta = (\beta_1,\ldots,\beta_p)'$ is a *p×1* vector. The identification problem the γ_j under the null hypothesis is solved by generating γ_j, *j=1, ..., p*, randomly from a uniform distribution, following Lee, White and Granger (1993). Implementing a Lagrange multiplier type version of the test requires the computation of the *T×(n+p+m)* matrix R=[Z F] where Z is a *T×m* matrix containing all variables due to the k-th order Taylor expansion of g, and the t-th row of F has the form

$$[F]_t = \left(\tilde{w}_t', \frac{1}{1+e^{-\gamma_1^{*\prime}\omega_t}}, \ldots, \frac{1}{1+e^{-\gamma_p^{*\prime}\omega_t}} \right) \quad (7)$$

where γ_j^*, j=1,..,p, contain the randomly drawn values of the corresponding unidentified parameter vectors. As Lee, White and Granger (1993) pointed out, the elements of the second submatrix of *F* tend to be collinear with themselves and with the first part of *F*. Thus we conduct the test using the first principal components of the second submatrix of *F*. This leads to the test statistic called *Neural* where we generate the hidden unit weights, i.e. the different elements of the vectors γ_j, for j=1,...,p, randomly from the uniform distribution.

Table III. Results of linear and nonlinear causality tests (p-values) (non causality)

		D1	D2	D3	D4	D5	D6
	(days)	*(1-2)*	*(2-4)*	*(4-8)*	*(8-16)*	*(16-32)*	*(32-64)*
D1→Dj							
	Linear	x	0.0742	0.0812	0.5268	0.9201	0.7039
	General	x	0.0000	0.0000	0.4673	0.6806	0.2546
	Additive	x	0.2311	0.4971	0.7299	0.6930	0.2604
	Neural	x	0.0041	0.9858	0.9990	0.9965	0.9882
D2→Dj							
	Linear	0.2408	x	0.5443	0.0406	0.0264	0.7457
	General	0.0000	x	0.0000	0.0039	0.8999	0.0000
	Additive	0.0000	x	0.0078	0.0162	0.9168	0.0000
	Neural	0.3548	x	0.0016	0.8858	0.9967	0.0000
D3→Dj							
	Linear	0.4913	0.4926	x	0.0000	0.0007	0.9572
	General	0.0145	0.0000	x	0.0981	0.0000	0.0053
	Additive	0.0332	0.2924	x	0.0591	0.0000	0.0070
	Neural	0.6510	0.1239	x	0.1192	0.0025	0.2869
D4→Dj							
	Linear	0.9878	0.0905	0.1542	x	0.0661	0.0000
	General	0.0000	0.0100	0.0001	x	0.0000	0.0000
	Additive	0.9292	0.0700	0.0195	x	0.0000	0.0000
	Neural	0.9583	0.0698	0.0000	x	0.0000	0.0000
D5→Dj							
	Linear	0.9894	0.0814	0.0181	0.0000	x	0.0006
	General	0.0000	0.0000	0.0000	0.0014	x	0.0000
	Additive	0.2873	0.3036	0.0000	0.0070	x	0.0000
	Neural	0.5123	0.2547	0.0000	0.0062	x	0.0000
D6→Dj							
	Linear	0.9610	0.8785	0.9991	0.0678	0.0000	x
	General	0.0000	0.0008	0.0850	0.1428	0.0000	x
	Additive	0.3389	0.5505	0.3489	0.1983	0.0000	x
	Neural	0.7763	0.6740	0.3521	0.0061	0.0000	x

Note: Di →Dj is for the null hypothesis of no causality from Di to Dj, for I , j=1, …, 6.. *Linear* is the linear Granger causality test, *General* and *Additive* are the nonlinear causality tests based on a Taylor series approximation (*Additive* is the test statistic based on the "semi-additive" model), and *Neural* is the nonlinear causality test based on artificial neural networks.
We Assume that we have two weakly stationary and ergodic time series x_t and y_t. In order to compute each test statistic, the number of lagged values of y_t is $q=2$, the number of lagged values of x_t is $n=3$ and the order of Taylor expansion is $k=3$. In the neural network test, following Lee, White and Granger (1993), the number of hidden units is $p=20$ and we generate the different elements of the vectors γ_j , for $j=1,...,p$, randomly from the uniform $[-\mu,\mu]$ distribution with $\mu=2$. Moreover, the number of principal components is determined automatically in each case; this is done by including the largest principal components that together explain at least *80%* of the variation in the corresponding matrix.

Appendix 2. Figures 1, 2, 3 and 4

Figure 1 . The time series plot of WTI oil prices

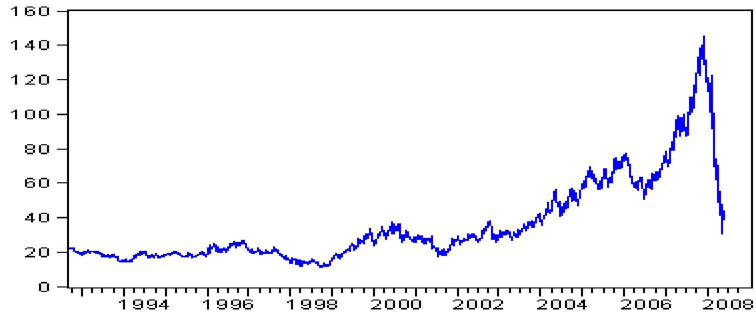

Figure 2. WTI oil prices returns

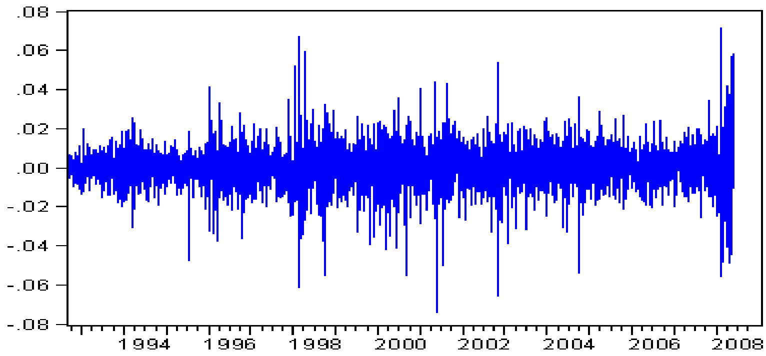

Figure 3. WTI oil absolute returns

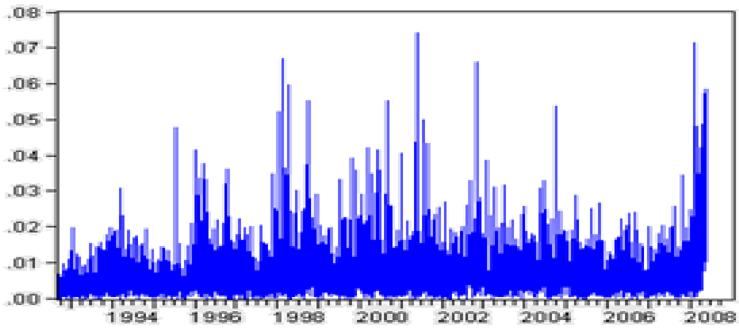

Figure 4. Oil absolute returns wavelet decomposition