Multifractal detrended cross-correlations and causations on financial markets 2014-2016

Marcin Wątorek

Complex Systems Theory Department
IFJ PAN

marcin.watorek@ifj.edu.pl

08.02.2017
Figure: Crude Light Oil 12.06.2014-05.02.2016 and its best fit (red) presented on Gródek 2016 conference.
Figure: Crude Light Oil 12.06.2014-03.02.2017.
Figure: Commodity currencies basket standardized (currbasket, blue), the inverse Crude Light Oil price standardized (CLinv, black) and the corresponding LPPL best fits: fit - currbasket (orange).
In my presentation I would like to answer on following question:

- What is the cross-correlation level between oil and other financial instruments?
- Are these cross-correlations multifractal?
- Are these relations stable in sub-periods?
- Are there any causal relationships between oil and rest financial instruments?
Figure: Standardized WTI Crude oil in USD per barrel, currencies expressed in US dollar, gold futures in USD, SP500 futures in period 01.2014-01.2017.
Figure: Log-log plot of the cumulative distributions of normalized absolute returns $|r(t)|$. 

Marcin Wątorek (IFJ PAN)  Gródek 2017  08.02.2017  7 / 27
Multifractal detrended cross-correlation analysis with sign preserving (MFCCA) [3] is a consistent generalization of the detrended cross-correlation approach (DCCA)[4]. A fundamental quantity for the detrended fluctuation analysis is the variance (covariance) $f_{ZZ}^2 (f_{XY}^2)$ of the detrended signals $X, Y$ ($Z$ stands for either $X$ or $Y$). Let us consider a pair of time series $x(i)_{i=1,...,T}$ and $y(i)_{i=1,...,T}$ divided into $2M_s$ separate boxes of length $s$ (i.e., $M_s$ boxes starting from the opposite ends). A detrending procedure consists of calculating in each box $\nu$ ($\nu = 0, ..., 2M_s - 1$) the residual signals $X, Y$ equal to the difference between the integrated signals and the $m$th-order polynomials $P^{(m)}$ fitted to these signals:

\[
X_{\nu}(s, i) = \sum_{j=1}^{i} x(\nu s + j) - P^{(m)}_{X,s,\nu}(j),
\]

\[
Y_{\nu}(s, i) = \sum_{j=1}^{i} y(\nu s + j) - P^{(m)}_{Y,s,\nu}(j).
\]
The covariance and the variances of $X$ and $Y$ in a box $\nu$ are defined as:

$$f_{XY}^2(s, \nu) = \frac{1}{s} \sum_{i=1}^{s} X_\nu(s, i) Y_\nu(s, i),$$ (3)

$$f_{ZZ}^2(s, \nu) = \frac{1}{s} \sum_{i=1}^{s} Z_\nu^2(s, i),$$ (4)

where $Z$ again means either $X$ or $Y$. These quantities can be used to define a family of the so-called fluctuation functions of order $q[1]$.

$$F^q_{XY}(s) = \frac{1}{2M_s} \sum_{\nu=0}^{2M_s-1} \text{sign} \left[ f_{XY}^2(s, \nu) \right] |f_{XY}^2(s, \nu)|^{q/2},$$ (5)

$$F^q_{ZZ}(s) = \frac{1}{2M_s} \sum_{\nu=0}^{2M_s-1} \left[ f_{ZZ}^2(s, \nu) \right]^{q/2}.$$ (6)
The signum function enters the equation for $F_{XY}^q(s)$ in order to preserve signs of the covariances $f_{XY}^2(s, \nu)$ that would otherwise be lost in moduli that are necessary to secure real values of $F_{XY}^q(s)$. The real parameter $q$ plays the role of a filter, by amplifying or suppressing the intra-box variances and covariances in such a way that for $q \gg 2$ only the boxes (of size $s$) with the highest fluctuations contribute substantially to the sums while for $q \ll 2$ only the boxes with the smallest fluctuations do that. The special case of $q = 2$ allows us to reduce the above formulas to a form:

$$F_{XY}^2(s) = \frac{1}{2M_s} \sum_{\nu=0}^{2M_s-1} f_{XY}^2(s, \nu),$$

(7)

$$F_{ZZ}^2(s) = \frac{1}{2M_s} \sum_{\nu=0}^{2M_s-1} f_{ZZ}^2(s, \nu),$$

(8)

in which all the boxes contribute to the sums with the same weights.
Multifractal cross-correlation is expected to manifest itself in the power-law dependence of $F^q_{XY}(s)$ and the following relation is fulfilled:

$$F^q_{XY}(s)^{1/q} = F_{XY}(q, s) \sim s^{\lambda(q)},$$

where $\lambda(q)$ is an exponent that quantitatively characterizes fractal properties of the cross-covariance. For the monofractal cross-correlation, the exponents $\lambda(q)$ is constant. In the case of multifractal cross-correlation, $\lambda(q)$ varies with $q$. 
The $q$-dependent detrended cross-correlation coefficient $\rho_q(s)$ is defined by means of the $q$th order fluctuation functions [2]:

$$\rho_q(s) = \frac{F^q_{XY}(s)}{\sqrt{F^q_{XX}(s)F^q_{YY}(s)}}.$$  \hspace{1cm} (10)

For $q = 2$ Eq. (10) reduces to the definition of $\rho_{\text{DCCA}}$ [6]. The filtering ability of $\rho_q(s)$ manifests itself in such a way that the more deviated from the value $q = 2$ the exponent $q$ is, the more extreme fluctuations in the corresponding boxes contribute to the coefficient $\rho_q(s)$. For $q \geq 0$, values of $\rho_q$ fit within the range

$$-1 \leq \rho_q \leq 1.$$ \hspace{1cm} (11)

As in the case of the Pearson and the $q_{\text{DCCA}}$ coefficient, $\rho_q = 1$ indicates a perfect correlation, $\rho_q = 0$ indicates independent signals, and $\rho_q = -1$ indicates a perfect anti-correlation.
Data set consist 13 financial instruments: WTI Crude Oil futures (CL), SP500 futures (SP500), Gold futures (GOLD) and 10 currencies expressed in US dollar: Australian dollar (AUD/USD), Canadian dollar (CAD/USD), euro (EUR/USD), Pound sterling (GBP/USD), Japanese yen (JPY/USD), Mexican peso (MXN/USD), Norwegian krone (NOK/USD), Polish zloty (PLN/USD), Russian ruble (RUB/USD), South African rand (ZAR/USD). Calculations were conducted on 5 min intraday data from Swiss forex bank Dukascopy in 02.01.2014-30.12.2016. For each considered financial instrument, the corresponding time series represents the logarithmic price increments (returns) \( r(t) = \log(p(t+1)) - \log(p(t)) \). Weekend gap and rolling day gap were removed, after these operations we got approximately N=150000 observations for each time series. Calculations were performed for \( q \in (-0.5, 4) \) (for the negative qs, \( F^q_{xy}(s) \) fluctuates around zero), minimum scale s=10 (50min), max scale s=N/40 (13 trading days), number of scales=40.
Figure: Family of the \( q \)-th-order fluctuation (cross-covariance) functions \( F_{xy}(q,s) \) for different values of \( q \in \langle 0.5, 4 \rangle \) (the topmost one represents \( q = 4 \)) calculated for CL vs AUD/USD, CAD/USD, RUB/USD, SP500, MXN/USD, NOK/USD, GOLD, ZAR/USD, EUR/USD, USD/JPY, PLN/USD and GBP/USD.
Multifractal cross-correlation scaling exponents $\lambda(q)$ and the average of the generalized Hurst exponents $h_{xy}$

$$F_{XY}^q(s)^{1/q} = F_{XY}(q, s) \sim s^{\lambda(q)},$$
$\lambda(q)$ is an exponent that quantitatively characterizes fractal properties of the cross-covariance.

$$h_{xy}(q) = \left( h_x(q) + h_y(q) \right) / 2$$ [1], where $h_x(q)$ and $h_y(q)$ refer to fractal properties of individual time series. Relation between $\lambda(q)$ and $h_{xy}(q)$ depends on temporal organization of the signals as determined by their generalized Hurst exponents. The smaller the difference between $\lambda(q)$ and $h_{xy}(q)$, the more the two series are similar to each other.

To establish level of similarity between WIT Crude oil (CL) and other time series we calculate difference $\lambda(q) - h_{xy}(q)$ for each pair CL vs x. $\lambda(q)$ varies with $q$ for all cases, which is the sign of multifractal cross-correlations.
Figure: Differences between multifractal cross-correlation scaling exponents $\lambda(q)$ and the average generalized Hurst exponents $h_{xy}(q)$ estimated for $1 \leq q \leq 4$. 
Figure: The $q$-dependent detrended cross-correlation coefficient $\rho_q$ between CL and rest instruments as a function of the temporal scale $s$ for $q = -1, q = 1, q = 2$ and $q = 3$. For $q = -1$, $\rho_q$ is fluctuating around zero, which is the effect that for the negative $q$, $F_{xy}^q(s)$ fluctuates around zero.
Figure: The $q$-dependent detrended cross-correlation coefficient $\rho_q$ between SP500 vs rest instruments as a function of the temporal scale $s$ for $q = 2$. Euro, yen and gold inverted (USD/EUR, USD/JPY, USD/gold to get positive cross-correlations).
Figure: The $q$-dependent detrended cross-correlation coefficient $\rho_q$ between gold vs rest instruments as a function of the temporal scale $s$ for $q = 2$. 
**Table:** The $q$-dependent detrended cross-correlation coefficient between CL and rest instruments $\bar{\rho}_q$ averaged on $q \in \langle 1, 4 \rangle$ - $q$ and averaged on $s \in \langle 1, 34 \rangle$ in sub-periods.

<table>
<thead>
<tr>
<th></th>
<th>14_1</th>
<th>14_2</th>
<th>15_1</th>
<th>15_2</th>
<th>16_1</th>
<th>16_2</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUD</td>
<td>0.18</td>
<td>0.39</td>
<td>0.53</td>
<td>0.60</td>
<td>0.65</td>
<td>0.53</td>
</tr>
<tr>
<td>CAD</td>
<td>0.15</td>
<td>0.50</td>
<td>0.65</td>
<td>0.75</td>
<td>0.74</td>
<td>0.69</td>
</tr>
<tr>
<td>EUR</td>
<td>0.10</td>
<td>0.31</td>
<td>0.47</td>
<td>-0.28</td>
<td>0.18</td>
<td>0.31</td>
</tr>
<tr>
<td>GBP</td>
<td>0.15</td>
<td>0.35</td>
<td>0.48</td>
<td>0.38</td>
<td>0.48</td>
<td>0.39</td>
</tr>
<tr>
<td>JPY</td>
<td>-0.28</td>
<td>-0.34</td>
<td>0.33</td>
<td>-0.46</td>
<td>-0.48</td>
<td>-0.27</td>
</tr>
<tr>
<td>MXN</td>
<td>0.22</td>
<td>0.48</td>
<td>0.54</td>
<td>0.65</td>
<td>0.71</td>
<td>0.55</td>
</tr>
<tr>
<td>NOK</td>
<td>0.20</td>
<td>0.48</td>
<td>0.53</td>
<td>0.53</td>
<td>0.60</td>
<td>0.59</td>
</tr>
<tr>
<td>PLN</td>
<td>-0.12</td>
<td>0.34</td>
<td>0.47</td>
<td>0.18</td>
<td>0.39</td>
<td>0.38</td>
</tr>
<tr>
<td>RUB</td>
<td>0.22</td>
<td>0.45</td>
<td>0.58</td>
<td>0.82</td>
<td>0.82</td>
<td>0.71</td>
</tr>
<tr>
<td>ZAR</td>
<td>0.23</td>
<td>0.38</td>
<td>0.48</td>
<td>0.55</td>
<td>0.60</td>
<td>0.43</td>
</tr>
<tr>
<td>SP500</td>
<td>0.37</td>
<td>0.50</td>
<td>0.47</td>
<td>0.63</td>
<td>0.69</td>
<td>0.51</td>
</tr>
<tr>
<td>GOLD</td>
<td>0.21</td>
<td>0.35</td>
<td>0.50</td>
<td>0.25</td>
<td>-0.25</td>
<td>0.26</td>
</tr>
</tbody>
</table>
Multifractal causation analysis

Multifractal causation analysis (MCA) algorithm description: We have two cross-correlated time series $s_1, s_2$

1. Move $s_1$ by $\tau$ in time backward, synchronize $s_2$ with $s_1$. We get $s_1\tau b = s_1(1 + \tau : end)$, $s_2\tau b = s_2(1 : end - \tau)$. Series $s_2$ leads now.

2. Move $s_1$ by $\tau$ in time forward, synchronize $s_2$ with $s_1$. We get $s_1\tau f = s_1(1 : end - \tau)$, $s_2\tau f = s_2(1 + \tau : end)$. Series $s_1$ leads now.

3. Calculate $\rho_q(s)b$ between $s_1\tau b$ and $s_2\tau b$.

4. Calculate $\rho_q(s)f$ between $s_1\tau f$ and $s_2\tau f$.

5. Calculate the difference between $\rho_q(s)b$ and $\rho_q(s)f$ then average it on all $s$.

6. If the difference is significantly higher than 0 then detrended cross-correlations between $s_1\tau b$ and $s_2\tau b$ are higher than between $s_1\tau f$ and $s_2\tau f$ which implies that $s_2$ is a driver and $s_1$ is a target. If $\rho_q b - \rho_q f < 0$ then $s_1$ is a driver and $s_2$ is a target.
Figure: Differences between the $q$-dependent detrended cross-correlation coefficient $\rho_q \bar{\rho}_q b - \bar{\rho}_q f$ averaged on $q \in \langle 1, 4 \rangle$ and averaged on $s \in \langle 1, 39 \rangle$, time shifts $\tau=5m, 10m, 15m, 30m, 1h$. 
Figure: Differences between the $q$-dependent detrended cross-correlation coefficient $\bar{\rho}_q b - \bar{\rho}_q f$ averaged on $s \in \langle 1, 39 \rangle$ for $q = 1, 2, 3, 4$, time shifts $\tau = 5\text{m}, 10\text{m}, 15\text{m}, 30\text{m}, 1\text{h}$. 

Marcin Wątorek (IFJ PAN)
Table: Granger causality CL vs x, H0: x does not Granger cause y, red color H0 rejected, green H0 accepted, significance level=0.05 (F-crit=3.84), max lag=12 (60min).

<table>
<thead>
<tr>
<th>x</th>
<th>F-stat CL -&gt;x</th>
<th>H0</th>
<th>F-stat CL &lt;- x</th>
<th>H0</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUD</td>
<td>33,71</td>
<td>R</td>
<td>31,46</td>
<td>R</td>
</tr>
<tr>
<td>CAD</td>
<td>72,98</td>
<td>R</td>
<td>18,35</td>
<td>R</td>
</tr>
<tr>
<td>EUR</td>
<td>0,58</td>
<td>A</td>
<td>20,4</td>
<td>R</td>
</tr>
<tr>
<td>GBP</td>
<td>31,83</td>
<td>R</td>
<td>27,49</td>
<td>R</td>
</tr>
<tr>
<td>iJPY</td>
<td>0,11</td>
<td>A</td>
<td>1,99</td>
<td>A</td>
</tr>
<tr>
<td>MXN</td>
<td>70,94</td>
<td>R</td>
<td>40,23</td>
<td>R</td>
</tr>
<tr>
<td>NOK</td>
<td>125,17</td>
<td>R</td>
<td>21,01</td>
<td>R</td>
</tr>
<tr>
<td>RUB</td>
<td>173,37</td>
<td>R</td>
<td>22,04</td>
<td>R</td>
</tr>
<tr>
<td>PLN</td>
<td>46,45</td>
<td>R</td>
<td>16,22</td>
<td>R</td>
</tr>
<tr>
<td>ZAR</td>
<td>176,76</td>
<td>R</td>
<td>47,78</td>
<td>R</td>
</tr>
<tr>
<td>SP500</td>
<td>7,54</td>
<td>R</td>
<td>10,44</td>
<td>R</td>
</tr>
<tr>
<td>Gold</td>
<td>4,86</td>
<td>A</td>
<td>27,63</td>
<td>R</td>
</tr>
</tbody>
</table>
the highest similarity level and the highest cross-correlations were between oil vs SP500 futures and currencies of the countries which are oil producers

- negative cross-correlations between JPY/USD vs CL and SP500
- negative cross-correlations between SP500 vs JPY/USD, EUR/USD and gold, they could be treated as defensive instruments
- cross-correlations between oil and risky instruments (AUD/USD, CAD/USD, MXN/USD, NOK/USD, RUB/USD, ZAR/USD, SP500) increases with time, they rise from the 2nd half of 2014, after breaking 60$ per barrel until 2nd half of 2016
- driver/target relation between oil and Russian ruble, oil prices themselves are driven by euro and gold behaviour


Thank for your attention