

Křišťoufek, Ladislav

Working Paper

Are the crude oil markets really becoming more efficient over time? Some new evidence

IES Working Paper, No. 07/2018

Provided in Cooperation with:

Charles University, Institute of Economic Studies (IES)

Suggested Citation: Křišťoufek, Ladislav (2018) : Are the crude oil markets really becoming more efficient over time? Some new evidence, IES Working Paper, No. 07/2018, Charles University in Prague, Institute of Economic Studies (IES), Prague

This Version is available at:

<https://hdl.handle.net/10419/203186>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Are the Crude Oil Markets Really Becoming More Efficient over Time? Some New Evidence

Ladislav Kristoufek

IES Working Paper: 07/2018



Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague

[UK FSV – IES]

Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
<http://ies.fsv.cuni.cz>

Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze

Opletalova 26
110 00 Praha 1

E-mail : ies@fsv.cuni.cz
<http://ies.fsv.cuni.cz>

Disclaimer: The IES Working Papers is an online paper series for works by the faculty and students of the Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Czech Republic. The papers are peer reviewed. The views expressed in documents served by this site do not reflect the views of the IES or any other Charles University Department. They are the sole property of the respective authors. Additional info at: ies@fsv.cuni.cz

Copyright Notice: Although all documents published by the IES are provided without charge, they are licensed for personal, academic or educational use. All rights are reserved by the authors.

Citations: All references to documents served by this site must be appropriately cited.

Bibliographic information:

Kristoufek L. (2018): "Are the Crude Oil Markets Really Becoming More Efficient over Time? Some New Evidence" IES Working Papers 07/2018. IES FSV. Charles University.

This paper can be downloaded at: <http://ies.fsv.cuni.cz>

Are the Crude Oil Markets Really Becoming More Efficient over Time? Some New Evidence

Ladislav Kristoufek^a

^aInstitute of Economic Studies, Faculty of Social Sciences, Charles University
Opelatalova 21, 110 00, Prague, Czech Republic
Email (corresponding author): LK@fsv.cuni.cz

March 2018

Abstract:

We replicate the study of Tabak & Cajueiro (2007): „Are the crude oil markets becoming weakly efficient over time? A test for time-varying long-range dependence in prices and volatility“ published in *Energy Economics* 29, pp. 28-36. The results have been mostly confirmed. Specifically, we have confirmed that the crude oil markets efficiency could be rejected up till approximately 1994 and this holds for both studied crude oil commodities – Brent and WTI. Analyzing an extended dataset up till June 2017, we find that the markets remained efficient (at least with respect to long-range dependence) until the outbreak of the Global Financial Crisis in 2008. This is confirmed by all three applied methods – the rescaled range analysis used in the original study, and the detrended fluctuation analysis and the Geweke-Porter-Hudak estimator which were added for stronger validity of the results. The markets returned back to efficiency around 2012 and remained there until 2015 when the Hurst exponent started another rally and stayed rather high until the end of the examined sample. Comparing the two markets, the Brent crude oil shows stronger signs of inefficiency during the inefficient periods compared to the WTI crude oil. This is in hand with the results reported in the original study. Apart from rerunning the analysis on an extended dataset and using two additional methods, we also provide a firmer validity check using the moving-block bootstrap procedure, which outperforms the original shuffling procedure in the provided forecasting study.

Keywords: crude oil, efficient market hypothesis, forecasting, long-range dependence, replication

JEL: G14, Q40

1. Introduction

We replicate the results of Tabak and Cajueiro (2007) presented in “*Are the crude oil markets becoming weakly efficient over time? A test for time-varying long-range dependence in prices and volatility*”. Tabak and Cajueiro (2007) analyze the returns and volatility of the Brent and West Texas Intermediate (WTI) crude oil spot prices using the rescaled range analysis of Hurst (1951) with adjustments according to Lo (1991). The main result of the paper is that the crude oil markets become more efficient in time as the estimated Hurst exponent H converges towards the level of $H = 0.5$ suggesting no persistence (long-term memory), existence of which would suggest exploitable profitable trading strategies as argued already by Mandelbrot and van Ness (1968).

Apart from the actual results concerning the efficient market hypothesis (Fama, 1965a,b, 1970, 1991), an important contribution of the paper lies in the fact that it was one of the pioneering papers presenting results of interdisciplinary research in a mainstream economics journal – *Energy Economics* – and drawing attention of the economics and finance community. Until then, the interdisciplinary research in the energy economics and energy finance journals had been rather scarce. Panas and Ninni (2000) studied the petroleum markets with the use of correlation dimension, various entropies and Lyapunov exponent and found strong evidence of chaotic behavior. Adrangi et al. (2001) examined the crude oil, heating oil and gasoline prices to find strong non-linear dependencies which were, however, inconsistent with the chaos hypothesis. And Serletis and Andreadis (2004) uncovered multifractal structures and turbulent behavior in the WTI crude oil series.

Tabak and Cajueiro (2007) drew attention of many researchers and following their steps, various results with a broad range of topics have been published in economics and finance outlets. Only until 2010, several interesting papers building on the original paper of Tabak and Cajueiro (2007) were published. Alvarez-Ramirez et al. (2008) studied the Brent, Dubai and WTI crude oil markets using the detrended fluctuation analysis (DFA) and argued that the markets are inefficient in the short-term perspective but efficient in the long-term one. Kang et al. (2009) analyzed the same set of markets as the previous study and showed that the fractionally integrated generalized autoregressive heteroskedasticity model (FIGARCH) fits volatility better than the generally used GARCH model. This way, they showed that long-term memory (a specific case of fractional integration) plays an important role in the crude oil markets. Charles and Darné (2009) utilized the variance ratio test to argue for inefficiency of the Brent and WTI markets between 1994 and 2008. Aloui and Mabrouk (2010) added persistence, fat tails and asymmetry to their Value-at-Risk (VaR) and expected shortfall (ES) models to show that including these properties into the model improves the forecasting performance. Cunado et al. (2010) ran a battery of tests to show that many energy futures are persistent in volatility (but no persistence in returns was found). Alvarez-Ramirez et al. (2010) applied the lagged detrended fluctuation analysis to find deviations from market efficiency in crude oil via the multiscaling of Hurst exponent. And Wang and Liu (2010) found that crude oil markets are moving towards market efficiency in time but the trend is not stable. After 2010, there are many more examples of a successful utilization of methods from outside of economics and finance on

economic, financial, and specifically also energy problems (for the recent ones, see Kristoufek and Vosvrda (2014), Sensoy and Hacihasanoglu (2014), Kristoufek and Lunackova (2015), Lubnau and Todorova (2015), Gu and Zhang (2016) and Li et al. (2016)).

In our replication, we follow the steps of the original study and apply the rescaled range analysis on the series of the Brent and WTI crude oil returns and volatility (approximated by the absolute returns). Apart from extending the analyzed time series towards June 2017, we also apply two additional methods to estimate the Hurst exponent – the detrended fluctuation analysis and the Geweke-Porter-Hudak estimator. We also provide an extra discussion on statistical validity of the results and use an additional moving-block bootstrap procedure. As a further extension, we also provide a detailed forecasting study to see whether the detected inefficiencies can be exploited for successful forecasting models. Our results in general support the results of the original study in the sense that the crude oil returns had showed signs of persistence, which goes against the efficient market notion, but the persistence vanished towards the original study sample end (July 2004). However, the prolonged sample uncovers that the dynamics did not stabilize much after 2004, quite the contrary. Therefore, our study confirms the important message of the original one – statistical and dynamic properties of the crude oil prices change in time markedly and it is necessary to study local properties and their time dynamics rather than only global properties with limited information content.

The paper is organized as follows. Section 2 summarizes the utilized methodology, specifically the estimators of Hurst exponent – the rescaled range analysis applied in the original study and the detrended fluctuation analysis and the Geweke-Porter-Hudak estimator as two additional robustness check methods – and the approach towards statistical inference of the estimators. Section 3 presents the analyzed dataset and Section 4 shows the results of the original as well as the extended data sample. Section 5 presents the results of a forecasting study and Section 6 discusses the results and concludes.

2. Methodology

In its classical form, the efficient market hypothesis (EMH) assumes fair markets, which translates into inability of reaching long-term above-average risk-adjusted profits with respect to the given information set (Fama, 1970, 1991). Specifics of the information set yield three forms of EMH – weak (historical data), semi-strong (publicly available information), and strong (private information). Then if, for example, we are able to make the above-described profits using only the historical data, we say that the weak-form EMH is violated. The weak-form efficiency is studied most frequently as the historical data are usually easily available and the trading strategies are thus testable, which is not necessarily true for the other two forms.

Apart from the profit-oriented specification of EMH, the hypothesis has also statistical implications towards the dynamic properties of the price and thus also returns series of assets. Specifically, the (logarithmic) prices of an efficient asset should follow a random walk (Fama, 1965b) or a martingale (Samuelson, 1965). Both specifications have a common implication for the connected returns – there should be no serial correlation

(auto-correlation). Violation of the “no serial correlation” assumption is usually considered to be a proof of the EMH rejection. From the statistical perspective, there is a fundamental separation between short-term dependence and long-term dependence. The short-term dependence is characterized by non-zero but rapidly (usually exponentially) vanishing auto-correlation structure. Such dependence is standardly modeled using the autoregressive moving average models (ARMA) and its various forms. The long-range dependent (long-term memory, persistent) processes have a slowly (usually hyperbolically) decaying auto-correlation function. These processes are standardly modeled via the fractionally integrated ARMA models (FARIMA/ARFIMA) and fractional Gaussian noise (or fractional Brownian motion for integrated processes). Both types of dependence cause troubles towards the formal EMH. Yet, as noted by Timmermann and Granger (2004), short-term dependencies change rapidly over time and investors need to adjust their trading strategies over time as well. There is thus a big difference between finding historical short-term dependence and actually being able to build a trading strategy and utilize it profitably. However, long-term memory in returns is inconsistent with continuous time stochastic processes used in option and futures pricing (Lo, 1991) and more importantly, its existence suggests extensive opportunities for long-lasting profitable trading strategies (Mandelbrot and van Ness, 1968).

Formally, the long-range dependent processes can be defined in both time and frequency domain. In the time domain, persistence is characterized through a power-law decaying auto-correlation function. Asymptotically, the auto-correlation function $\rho(k)$, with time lag k , of a persistent process decays as $\rho(k) \propto k^{2H-2}$. Hurst exponent H is a crucial parameter of long-range dependent processes. For $H > 0.5$, the process is persistent and it is reminiscent of a locally trending process. For $H < 0.5$, we have an anti-persistent process with an excessive (compared to a serially uncorrelated process) switching. And for $H = 0.5$, there is no long-term memory. In the frequency domain, persistent processes have a divergent at origin spectrum so that for spectrum $f(\lambda)$ with frequency λ , we have $f(\lambda) \propto \lambda^{1-2H}$ for $\lambda \rightarrow 0+$. These properties have further implications such as non-summable auto-correlations and diverging covariance of partial sums which are used as basis for estimators of Hurst exponent H (Beran, 1994; Taqqu et al., 1995; Taqqu and Teverovsky, 1996; Robinson, 1995; Geweke and Porter-Hudak, 1983; Barunik and Kristoufek, 2010; Teverovsky et al., 1999).

Tabak and Cajueiro (2007) build their analysis on the above-mentioned connection between persistence and violation of EMH. Using the rescaled range analysis, they show that the crude oil returns changed their dynamic properties between 1983 and 2004. In our study, we extend the portfolio of estimators to three, utilizing the detrended fluctuation analysis and the Geweke-Porter-Hudak estimator in addition to the rescaled range analysis. We also prolong the analyzed dataset to June 2017. In the following parts, we introduce the utilized estimators, discuss how to approach the results on statistical basis and provide details on the moving window estimation procedure.

2.1. Hurst exponent estimators

Rescaled range analysis

The rescaled range analysis procedure (RS) (Hurst, 1951; Mandelbrot and Wallis, 1968) is based on the asymptotic power-law scaling behavior of ranges of persistent integrated series. In the procedure, a time series x_t of length T is split into N adjacent sub-periods of length v , i.e. $Nv = T$. In each sub-period I_n ($n = 1, \dots, N$), a rescaled range of a profile $X_{t,I_n} = \sum_{i \in I_n} (x_i - \bar{x}_{I_n})$ is calculated as R_{I_n}/S_{I_n} , where $R_{I_n} = \max_{I_n}(X_{t,I_n}) - \min_{I_n}(X_{t,I_n})$ is a range of the corresponding profile and S_{I_n} is a standard deviation of the corresponding series x_t . This rescaled range is calculated for each sub-period of length v and then averaged over all such sub-periods to get $(R/S)_v$ (Peters, 1994; Taqqu et al., 1995; Weron, 2002). The Hurst exponent is obtained via the least squares regression on the log-log transformation of the scaling law

$$(R/S)_v \propto v^H. \quad (1)$$

The scaling law is estimated over a range of sub-period lengths. In our application¹, we set $v_{min} = 10$ and $v_{max} = T/5$ with a step of 10.

Detrended fluctuation analysis

The detrended fluctuation analysis (DFA) (Peng et al., 1994; Kantelhardt et al., 2002) builds on the asymptotic power-law scaling of variance of persistent integrated series. Similarly to the rescaled range analysis, a time series x_t of length T is divided into sub-periods of length v and a profile is constructed in each such sub-period. A polynomial (usually, and also in our case, linear) fit $X_v(t)$ of the profile is estimated in each sub-period and a detrended signal $Y_v(t)$ is constructed as $Y_v(t) = X(t) - X_v(t)$. The fluctuation $F_{DFA}^2(v)$, calculated as an average mean squared error from the polynomial trend over all sub-periods of the same length, follows

$$F_{DFA}^2(v) \propto v^{2H}. \quad (2)$$

In the same logic as for RS, minimum and maximum sub-period lengths need to be set. Here, we keep the same ones as for RS so that $v_{min} = 10$ and $v_{max} = T/5$ with a step of 10 as proposed by Einstein et al. (2002), Alvarez-Ramirez et al. (2005) and Matos et al. (2008).

GPH estimator

The Geweke-Porter-Hudak (GPH) estimator (Geweke and Porter-Hudak, 1983) is based on a full functional specification of the underlying process as the fractional Gaussian noise. Such specification implies the following form of the spectrum:

$$\log f(\lambda) \propto -(H - 0.5) \log(4 \sin^2(\lambda/2)) \quad (3)$$

¹Selection of the scaling range is not discussed in the original paper of Tabak and Cajueiro (2007). However, we follow the standard procedure of selecting rather low v_{min} , and v_{max} such that the average rescaled range is based on at least five values for high v .

As an estimator of the spectrum of series x_t , we use the periodogram defined as $I(\lambda_j) = \frac{1}{T} \sum_{t=1}^T \exp(-2\pi i t \lambda_j) x_t$ with $j = 1, 2, \dots, m$, where $m \leq T/2$ and $\lambda_j = 2\pi j/T$, and Hurst exponent is estimated using the ordinary least squares on

$$\log I(\lambda_j) \propto -(H - 0.5) \log(4 \sin^2(\lambda_j/2)). \quad (4)$$

The GPH estimator is consistent and asymptotically normal (Beran, 1994) so that

$$\sqrt{T}(\hat{H} - H^0) \rightarrow_d N(0, \pi^2/6). \quad (5)$$

The GPH estimator performs very well asymptotically, assuming that the underlying process is in fact the fractional Gaussian noise (fGn). However, this is not precisely the case in financial and economic time series which are usually a combination of short-term memory series (such as autoregressive moving average – ARMA – processes of various specifications) and long-term memory series (such as the fractional Gaussian noise of fractionally integrated ARMA, i.e. ARFIMA). The GPH estimator becomes biased for these processes. Such bias can be avoided or at least mitigated by basing the estimation only on a part of the spectrum (periodogram) close to the origin, i.e. at low frequencies which reflect potential long-term memory. In our case, we follow Phillips and Shimotsu (2004) and use $m = T^{0.6}$.

2.2. Statistical inference

The rescaled range analysis utilized in the original study of Tabak and Cajueiro (2007) has several weaknesses. One of the most important ones is its sensitivity to short-range dependence. Specifically, the method can confuse strong short-term memory with long-term memory (Lo, 1991; Teverovsky et al., 1999; Kristoufek, 2012). Tabak and Cajueiro (2007) consider this issue and apply a specific shuffling procedure – they split the time series into blocks of 20 observations and in each block, they shuffle the data. The idea behind is to destroy the short-range serial correlations while preserving the long-range ones. This procedure is repeated several times (the number is not disclosed in the original study) and confidence intervals of the estimated Hurst exponent are reported as quantiles of estimates based on this shuffling procedure. The biggest issue of this approach is that it has only weak theoretical basis. Even though it is based on a firm evidence of Lo (1991), who shows that a short-term correlated series can be identified as a long-range correlated, the procedure does not ensure that the standard errors or confidence intervals based on it are the correct ones. Therefore, apart from reporting the results based on this “box shuffling” procedure (in our case with 100 repetitions), we also provide results based on the more standard moving-block bootstrap procedure (Lahiri, 1993; Davison and Hinkley, 1997; Srinivas and Srinivasan, 2000).

The moving-block bootstrap procedure for the Hurst exponent estimators is based on splitting the time series into boxes of the same length (we keep the same length, i.e. 20 observations, as for the box shuffling of the original study) and shuffling these boxes (but not the observations inside the boxes). This way, the new series keeps the short-term correlation structure (within the blocks) and distributional properties of the original series

but the long-range dependence is removed (due to shuffling). The Hurst exponent is thus estimated on a (possibly) short-range correlated series with a specific distribution. After repeating the procedure several times (here again 100 times), we obtain the distribution of the Hurst exponent estimator under the null hypothesis of no long-range dependence (but possibly under short-range dependence and given the distributional properties). If the estimated Hurst exponent for the original series falls outside of the critical levels based on the bootstrapping procedure, we have evidence of long-range dependence (after controlling for possible short-range dependence and distributional biases). This gives us two ways how to comment on the results on the statistical basis.

2.3. Moving window estimation

Time variation of persistence and thus potential trends in crude oil market efficiency are essential contributions of the original study of Tabak and Cajueiro (2007). To provide the dynamics of the Hurst exponents, the estimates are obtained from a time window of limited length that moves in time. This moving window or rolling window procedure is now a standard way of analyzing changes in dynamical properties of series in time. In the original study, the authors fix the time series length to 1004 observations, estimate the Hurst exponent using the rescaled range analysis with the box shuffling procedure and then slide the window by a fixed number of observations. However, the specific step length is not provided in the original study and based on a visual inspection of the resulting charts, it seems higher than one day. In our analysis, we fix the time series length to 1,000 observations (so that it is divisible by the box length of the box shuffling and bootstrap procedures, i.e. 20 observations). For this sample, we estimate the Hurst exponents based on all three methods and using both shuffling procedures. The sample window is then moved by 5 observations and repeated until the end of the sample is reached.

3. Data

As in the original study, we analyze the daily series of the Brent and WTI crude oil spot prices. We study the publicly available data provided by the U.S. Energy Information Administration (EIA)². The WTI prices are available between 2 January 1986 and 30 June 2017 (7945 observations) and the Brent prices between 20 May 1987 and 30 June 2017 (7645 observations). Compared to the original dataset, the series start approximately 3 years later (the original dataset is not publicly available) but are prolonged by 13 years. Evolution of prices is illustrated in Fig. 1 where we can see the split between the original series and the extended one. Interestingly, the examined period of the original study ends just at the beginning of a sharp price increases culminating in 2008 when the crude oil prices broke the \$140 barrier (compared to prices around \$30 between years 2000 and 2004). The price dynamics is rather interesting in the extended sample which promises interesting results from the efficient market hypothesis perspective.

²The spot prices are available at https://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm.

4. Results

We estimate the time-varying Hurst exponent on a moving window of 1,000 days with a step of 5 days using three estimators – the rescaled range analysis utilized in the original study of Tabak and Cajueiro (2007), and the detrended fluctuation analysis and the Geweke-Porter-Hudak estimator. The long-term memory parameter is estimated both for the returns series and the volatility series, which is approximated by the absolute returns as in the original study. Statistical validity of the results is provided by two shuffling procedures described in the previous sections.

Figs. 2 and 4 show the results for returns of the Brent and WTI crude oil, respectively. For the Brent returns (Fig. 2), Hurst exponent is estimated firmly above the level of 0.5 for the RS³ and DFA methods at the beginning of the analyzed period (until approximately 1994) and then slowly converges towards 0.5 until the end of the original study time frame. This interpretation is validated by the bootstrapped confidence intervals as well. Implications based on the GPH estimator are not as straightforward as the estimator apparently has a much higher variance than the other two.

In the extended period, i.e. after 2004, the series remain quite stable around 0.5 but increase markedly after 2008. This is true for all three estimators. The box shuffling confidence intervals suggest that the Brent crude oil was not efficient between 2008 and 2012. Such claim comes mostly from the DFA and GPH estimates, the estimates of RS stay close to the EMH null hypothesis. This is again supported by the bootstrapped confidence intervals, and now even for the GPH estimator. Between 2012 and 2014, the Brent market seems efficient but from 2015 onwards, the Hurst exponent increases markedly again. Market efficiency is then rejected by RS and DFA, while the GPH evidence is not convincing mainly due to the high estimator variance. Overall, the Brent crude oil market has evidently undergone rather wild dynamics after the analyzed period of the original study and it has not stayed efficient as one might have guessed based on the data until 2004.

The results are rather similar for the WTI returns (Fig. 4). As in the original study, the market is very close to efficiency at the beginning of the analyzed period. Based on RS, the market is on the edge of inefficiency until approximately 1994. This is confirmed by the bootstrapped confidence intervals both for RS and DFA. Estimated Hurst exponents are high for GPH as well but not significantly different from 0.5 due to high estimator variance again. In the following periods and during the extended period after 2004 as well, the dynamics of the WTI market is quantitatively similar to the one of the Brent market but the evidence is statistically weaker. Even though there are periods of higher Hurst exponents between 2008 and 2012, and then from 2015 onwards, the deviations are on the

³Quite an important omission of the original study is the fact that the expected value of the RS-based Hurst exponent for uncorrelated series is not 0.5 but it depends on the time series length as the values of rescaled ranges depend on the period lengths, i.e. there is a finite sample effect/bias (Peters, 1994; Couillard and Davison, 2005). For the time series of 1,000 observations, the expected value of the Hurst exponent is approximately 0.55 instead of 0.5.

edge of statistical significance.

The original study also reports the time evolution of the Hurst exponent of the volatility series, which are represented by the absolute returns. These are reported only as kind of a byproduct of the analysis as potential persistence of volatility has no direct effect on the market efficiency. In the same vein, we report these in Figs. 3 and 5 for Brent and WTI, respectively. We observe strong temporal variation for both series. For the Brent crude oil, the estimated Hurst exponent starts at around 0.85 and quite stably decreases to 0.5 in around 2006 to bounce back starting from 2009. This result is quite stable across the three methods. However, at least based on the bootstrapped confidence interval of RS and DFA, a big part of deviation of H from 0.5 might be due to strong short-range dependence and distribution of the volatility series. The volatility persistence is more variable in time for WTI, even though the qualitative results of Brent hold here as well. It again needs to be noted that the identified persistence of the volatility series does not go against the efficient market hypothesis (Lo and MacKinlay, 1988).

5. Implications for forecasting

To further illustrate importance and utility of the results of the time-varying Hurst exponent, we now focus on their implications for possible trading strategies and how their performance depends on Hurst exponent itself and detected inefficiency. As the volatility persistence is of only little importance for returns forecasting, we focus only on the results for the returns series as described in the previous section and illustrated in Figs. 2 and 4. We utilize three models to assess usefulness of analyzing the Hurst exponent evolution over time:

- autoregressive process (AR) with a maximum number of 5 lags (the optimal number of lags is based on the Akaike Information Criterion)
- autoregressive moving average process (ARMA) with a maximum number of 5 lags for both parts of the process, i.e. AR and MA (the optimal number of lags is based on the Akaike Information Criterion)
- autoregressive fractionally integrated moving average process (ARFIMA) with a maximum number of 5 lags for both parts of the process, i.e. AR and MA (the optimal number of lags is based on the Akaike Information Criterion)

We keep the same setting as for the estimation of Hurst exponent as presented in the previous sections. This means that on the moving window of the size of 1,000 observations, the three models are estimated and a one-step ahead forecast is calculated. Such forecasts are compared with the true realizations of the returns series and an absolute error is calculated. This is repeated for both Brent and WTI returns series with a step of 5 days.

From the perspective of usefulness of the Hurst exponent and efficiency discussion for successful forecasting, we are interested in possible dependence of model performance on Hurst exponent and/or (in)efficiency arguments based on Hurst exponent and its deviation from expected values.

We first focus on the dependence of the model forecasting performance with respect to the estimated Hurst exponent. Note that the analysis encompasses both the original shuffling procedure and the moving-block bootstrapped approach proposed in the current replication. Fig. 6 shows the dependence between the average Hurst exponent and the mean absolute forecasting error (MAFE/MAE) implied by the three used models (AR, ARMA and ARFIMA). For the shuffling procedure, the average Hurst exponent is the average estimate based on the three used methods (RS, DFA and GPH). Note that these estimates are directly based on the shuffling procedure as proposed in the original paper. For the bootstrap methods, the average Hurst exponent is simply an average of the original three methods (the bootstrapping is used for the efficiency testing and not for the Hurst exponent estimation). Inspecting Fig. 6, we observe that the results are more dependent on the analyzed series than the method. For Brent, we see a nice pattern of increasing mean absolute error between the Hurst exponent levels of 0.44 and 0.5. Note that $H = 0.5$ is the level of a serially uncorrelated process that should be hard to forecast, i.e. the mean absolute error should be high. Above $H = 0.5$, the processes should be easier to forecast so that the mean absolute error should decrease. We can see that MAE is lower for $H > 0.5$ compared to $H = 0.5$ but there is no apparent (decreasing) trend. The results are very similar for all three models. For WTI, the results are weaker. It seems that the critical level for serially uncorrelated process is at $H = 0.45$, which might be induced by distributional properties of the series. For the bootstrapping procedure, we can see that MAE is lower both below and above $H = 0.45$ with the exception of extremely high Hurst exponents around 0.7. For the shuffling procedure, the results are even more volatile.

At the first sight, such results might seem confusing and unsupportive of the original claims that Hurst exponent connects to possible inefficiency and predictability of the series. However, coming back to the results presented in Figs. 2-5, we can see that not only the estimates but also the confidence intervals based on either method vary considerably. At some point, $H = 0.5$ seems like a good description of serially uncorrelated process, but for some other points, it goes as low as $H = 0.45$ or even less (this is mainly visible for the WTI returns). This possibly explains the big difference between Brent and WTI. And as for the surprising results for the high values of Hurst exponent, there are two likely reasons. First, Hurst exponent varies considerably in time and this is true also for the shuffled and bootstrapped values. In many cases, $H = 0.65$ is found for serially uncorrelated (efficient) process⁴. And second, the Hurst exponent estimates based on the GPH method are very volatile. For the shuffling procedure, we find levels of the efficient process to possibly be as low as $H = 0.15$ but also as high as $H = 0.85$. For the bootstrapping procedure, the estimates of the efficient market null hypothesis generally vary between $H = 0.2$ and $H = 0.75$. This suggests that looking at the Hurst exponent by itself is not informative enough, which is in fact represented by poor outcomes of the relationship between Hurst exponent and MAE, and we should rather inspect the periods when Hurst exponent uncovers inefficiency in more detail.

⁴This is the case for RS and DFA. For GPH, the extreme values are even higher.

As we have three methods of the Hurst exponent estimation, three forecasting models and two ways to test whether the analyzed time series is efficient, there are many results. Fig. 7 summarizes the mean absolute (forecasting) errors with respect to these options. We compare five scenarios with respect to market efficiency implied by the Hurst exponent levels (either based on the confidence intervals implied by the shuffling procedure of the original paper or the critical levels implied by the bootstrapping procedure) – no inefficiency detected, inefficiency detected by one method (either RS or DFA or GPH), inefficiency detected by two methods, inefficiency detected by all three methods, and inefficiency detected by RS and DFA. The first four scenarios are quite straightforward as we would suspect that when more methods suggest inefficiency, the forecasting models will perform better. The last scenario – focusing on the situations when RS and DFA detect inefficiency – is included due to quite problematic features of the remaining methods (GPH) as described in the previous paragraph. There is a real risk that GPH could spoil the results.

The results differ markedly for the two analyzed series, even though the main qualitative outcomes are the same. We also observe discrepancies between the shuffling and bootstrapping procedures. For the Brent series, the pattern is very similar for both methods – the best forecasting performances of all three models (AR, ARMA, and ARFIMA) are found for the cases when the RS and DFA methods suggest inefficiency and the worst performance is observed for the case when all three methods find inefficiency. The GPH method is thus evidently not appropriate here. An important finding arises when comparing the performance of methods using the shuffling and bootstrapping approach towards detecting inefficiency. For the shuffling procedure, the difference between the case when the RS and DFA methods find inefficiency and the case when no inefficiency is found is very small. But for the bootstrapping procedure, the difference is clearly visible and MAE decreases by more than 30% between these two cases. Note that the forecasting performance of the models (AR, ARMA, and ARFIMA) does not differ much even though ARFIMA performs the best in the best case scenarios. For the WTI series, the performances are much closer which is quite in hand with the results presented in Fig. 6 but the pattern is again quite clear. The performance is very different for the shuffling procedure and the bootstrapping procedure. The former one finds no advantage in identifying inefficiency whereas the latter does. Even though the superiority of the RS and DFA detection is not so dominant as in the case of Brent, such scenario wins again. ARFIMA gives the best forecasting performance of the three models here as well.

Shortly summarizing the main implications of this forecasting performance analysis, there are several strong results. First, Hurst exponent needs to be used in combination with confidence intervals to identify inefficient periods where model-based forecasting pays off. Simply using the Hurst exponent variation in time is not enough. Second, the bootstrapping approach towards identifying inefficient periods performs much better than the shuffling approach of the original paper, which is in hand with our earlier suspicions about the shuffling procedure. Third, the GPH estimator is so volatile that it is not very useful as an inefficiency detection tool. And fourth, the ARFIMA model performs the best in the best case scenarios which suggests that the long-term memory in fact plays its role here.

6. Discussion and conclusions

We have replicated the study of Tabak and Cajueiro (2007) and their results have been mostly confirmed. Specifically, we have confirmed that the crude oil markets efficiency could be rejected up till approximately 1994 and this holds for both studied markets – Brent and WTI. After that, the markets remained efficient (at least with respect to long-range dependence) until the outbreak of the Global Financial Crisis in 2008. This is confirmed by all three applied methods. The markets returned back to efficiency around 2012 and remained there until 2015 when the Hurst exponent started another rally and stayed rather high until the end of the examined sample, i.e. June 2017. Comparing the two markets, the Brent crude oil shows stronger signs of inefficiency during the inefficient periods compared to the WTI crude oil. This is in hand with the results reported in the original study. As an additional extension, we have provided a detailed forecasting study which confirmed that the detected inefficiencies can be exploited for successful forecasting models. Supremacy of the bootstrapping procedure over the shuffling one has been clearly demonstrated.

Even though the results of the original study are mostly confirmed, the statistical validity is not as strong as claimed in the original study. This is mainly due to three reasons.

First, the original study estimates the Hurst exponent using the rescaled range analysis and bases the implications about efficiency on comparison of the estimated exponents with a breaking value of $H = 0.5$. This value signalizes no long-range dependence in the series asymptotically. However, the rescaled range analysis has been repeatedly shown to be biased for finite samples both analytically (Peters, 1994; Couillard and Davison, 2005) and numerically (Taqqu et al., 1995; Teverovsky et al., 1999; Weron, 2002; Barunik and Kristoufek, 2010; Kristoufek, 2012). For the specific case of the original study, i.e. estimates being made based on approximately 1,000 observations, the expected value of the Hurst exponent for a serially uncorrelated process is 0.55 rather than 0.5. This means that the found persistence is in fact weaker and the proof against EMH is in turn weaker as well. Nevertheless, there are still rather long periods of time when the estimated Hurst exponent is well above the correct theoretical value and EMH could be challenged.

Second, even though the study claims that the long-range dependence is based on a novel approach of the within-block shuffling, this procedure has only weak theoretical validity and background. No reference to a theoretical paper supporting the method and describing its theoretical properties is given so that even though the method has its intuitive appeal and logic, it remains a heuristic rather than a proper statistical procedure. To validate the results, we add the moving-block bootstrapping procedure that is well established (Lahiri, 1993; Davison and Hinkley, 1997; Srinivas and Srinivasan, 2000) for the long-range dependence framework. Qualitatively, the original claims are supported, even though the evidence is weaker as the estimated Hurst exponents barely pass outside of the bootstrapped critical values even for the periods of apparent inefficiency. The division between these two approaches is clearly visible in the provided forecasting study where the bootstrapping procedure clearly outperforms the shuffling procedure of the original study.

And third, we report that most of the results are confirmed qualitatively. This is due to the fact that some important parameters of the utilized procedure are not disclosed in the original study. Namely, these are the minimum and maximum sub-period lengths v in the rescaled range procedure, an unexplained indivisibility of the moving window length (1004) by the shuffled block size (20), a number of repetitions for the shuffling procedure, and the step size of the moving windows. In our replication and extension, we disclose all necessary information about the testing procedures for possible future replications.

Overall, the results up to 2004 (the final year of the original study) have been confirmed but the crude oil markets have not stayed efficient the whole time afterwards. On the contrary, two rather long periods of inefficiency have been identified. Our results highlight lively dynamics of the crude oil markets and its interconnectedness with the economic reality of the last decade, i.e. the Global Financial Crisis and the evolution of the world economy afterwards.

Acknowledgements

The research leading to these results was supported by the European Union's Horizon 2020 Research and Innovation Staff Exchange programme under the Marie Skłodowska-Curie grant agreement No 681228. The author further acknowledges financial support from the Czech Science Foundation (grant number 16-00027S).

References

- Adrangi, B., C. Arjun, K. Dhanda, and K. Raffiee (2001). Chaos in oil prices? evidence from future markets. *Energy Economics* 23(4), 405–425.
- Aloui, C. and S. Mabrouk (2010). Value-at-risk estimations of energy commodities via long-memory, asymmetry and fat-tailed garch models. *Energy Policy* 38(5), 2326–2339.
- Alvarez-Ramirez, J., J. Alvarez, and E. Rodriguez (2008). Short-term predictability of crude oil markets: A detrended fluctuation analysis approach. *Energy Economics* 30(5), 2645–2656.
- Alvarez-Ramirez, J., J. Alvarez, and R. Solis (2010). Crude oil market efficiency and modeling: Insights from the multiscaling autocorrelation pattern. *Energy Economics* 32(5), 993–1000.
- Alvarez-Ramirez, J., E. Rodriguez, and J. Echeverria (2005). Detrending fluctuation analysis based on moving average filtering. *Physica A* 354, 199–219.
- Barunik, J. and L. Kristoufek (2010). On hurst exponent estimation under heavy-tailed distributions. *Physica A* 389(18), 3844–3855.
- Beran, J. (1994). *Statistics for Long-Memory Processes*, Volume 61 of *Monographs on Statistics and Applied Probability*. New York: Chapman and Hall.
- Charles, A. and O. Darné (2009). The efficiency of the crude oil markets: Evidence from variance ratio tests. *Energy Policy* 37(11), 4267–4272.
- Couillard, M. and M. Davison (2005). A comment on measuring the hurst exponent of financial time series. *Physica A* 348(15 March), 404–418.
- Cunado, J., L. Gil-Alana, and F. Perez de Gracia (2010). Persistence in some energy futures markets. *Journal of Futures Markets* 30(5), 490–507.
- Davison, A. and D. Hinkley (1997). *Bootstrap Methods and Their Application*. Cambridge University Press.
- Einstein, A., H.-S. Wu, and J. Gil (2002). Detrended fluctuation analysis of chromatin texture for diagnosis in breast cytology. *Fractals* 10(1)(4), 19–25.
- Fama, E. (1965a). The behavior of stock market prices. *Journal of Business* 38(1), 34–105.
- Fama, E. (1965b). Random walks in stock market prices. *Financial Analysts Journal* 21(5), 55–59.
- Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25(2), 383–417.

- Fama, E. (1991). Efficient capital markets: Ii. *Journal of Finance* 46(5), 1575–1617.
- Geweke, J. and S. Porter-Hudak (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4(4), 221–238.
- Gu, R. and B. Zhang (2016). Is efficiency of crude oil market affected by multifractality? evidence from the WTI crude oil market. *Energy Economics* 53(January), 151–158.
- Hurst, H. (1951). Long term storage capacity of reservoirs. *Transactions of the American Society of Engineers* 116, 770–799.
- Kang, S., S.-M. Kang, and S.-M. Yoon (2009). Forecasting volatility of crude oil markets. *Energy Economics* 31(1), 119–125.
- Kantelhardt, J., S. Zschiegner, E. Koscielny-Bunde, A. Bunde, S. Havlin, and E. Stanley (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A* 316(1-4), 87–114.
- Kristoufek, L. (2012). How are rescaled range analyses affected by different memory and distributional properties? A Monte Carlo study. *Physica A* 391(17), 4252–4260.
- Kristoufek, L. and P. Lunackova (2015). Rockets and feathers meet Joseph: Reinvestigating the oilgasoline asymmetry on the international markets. *Energy Economics* 49(May), 1–8.
- Kristoufek, L. and M. Vosvrda (2014). Commodity futures and market efficiency. *Energy Economics* 42(March), 50–57.
- Lahiri, S. (1993). On the moving block bootstrap under long range dependence. *Statistics & Probability Letters* 18(5), 405–413.
- Li, D., Y. Nishimura, and M. Men (2016). Why the long-term auto-correlation has not been eliminated by arbitragers: Evidences from NYMEX. *Energy Economics* 59(September), 167–178.
- Lo, A. (1991). Long-term memory in stock market prices. *Econometrica* 59(5), 1279–1314.
- Lo, A. and C. MacKinlay (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies* 1, 41–66.
- Lubnau, T. and N. Todorova (2015). Trading on mean-reversion in energy futures markets. *Energy Economics* 51(September), 312–319.
- Mandelbrot, B. and J. van Ness (1968). Fractional brownian motions, fractional noises and applications. *SIAM Review* 10(4), 422–437.
- Mandelbrot, B. and J. Wallis (1968). Joah, joseph and operational hydrology. *Water Resources Research* 4(5), 909–918.

- Matos, J., S. Gama, H. Ruskin, A. Sharkasi, and M. Crane (2008). Time and scale hurst exponent analysis for financial markets. *Physica A* 387(15), 3910–3915.
- Panas, E. and V. Ninni (2000). Are oil markets chaotic? a non-linear dynamic analysis. *Energy Economics* 22(5), 549–568.
- Peng, C., S. Buldyrev, S. Havlin, M. Simons, H. Stanley, and A. Goldberger (1994). Mosaic organization of dna nucleotides. *Physical Review E* 49(2), 1685–1689.
- Peters, E. (1994). *Fractal Market Analysis – Applying Chaos Theory to Investment and Analysis*. New York: John Wiley & Sons, Inc.
- Phillips, P. C. B. and K. Shimotsu (2004). Local whittle estimation in nonstationary and unit root cases. *The Annals of Statistics* 32(2), 659–692.
- Robinson, P. (1995). Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics* 23(5), 1630–1661.
- Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review* 6, 41–49.
- Sensoy, A. and E. Hacıhasanoglu (2014). Time-varying long range dependence in energy futures markets. *Energy Economics* 46(November), 318–327.
- Serletis, A. and I. Andreadis (2004). Random fractal structures in North American energy markets. *Energy Economics* 26(3), 389–399.
- Srinivas, V. and K. Srinivasan (2000). Post-blackening for modeling dependent annual streamflows. *Journal of Hydrology* 230(1-2), 86–126.
- Tabak, B. and D. Cajueiro (2007). Are the crude oil markets becoming weakly efficient over time? a test for time-varying long-range dependence in prices and volatility. *Energy Economics* 29(1), 28–36.
- Taqqu, M., W. Teverosky, and W. Willinger (1995). Estimators for long-range dependence: an empirical study. *Fractals* 3(4), 785–798.
- Taqqu, M. and V. Teverovsky (1996). On Estimating the Intensity of Long-Range Dependence in Finite and Infinite Variance Time Series. In *A Practical Guide To Heavy Tails: Statistical Techniques and Applications*.
- Teverovsky, V., M. Taqqu, and W. Willinger (1999). A critical look at lo’s modified r/s statistic. *Journal of Statistical Planning and Inference* 80(1-2), 211–227.
- Timmermann, A. and C. Granger (2004). Efficient market hypothesis and forecasting. *International Journal of Forecasting* 20(1), 15–27.

- Wang, Y. and L. Liu (2010). Is WTI crude oil market becoming weakly efficient over time?. new evidence from multiscale analysis based on detrended fluctuation analysis. *Energy Economics* 32(5), 987–992.
- Weron, R. (2002). Estimating long-range dependence: finite sample properties and confidence intervals. *Physica A* 312(1-2), 285–299.

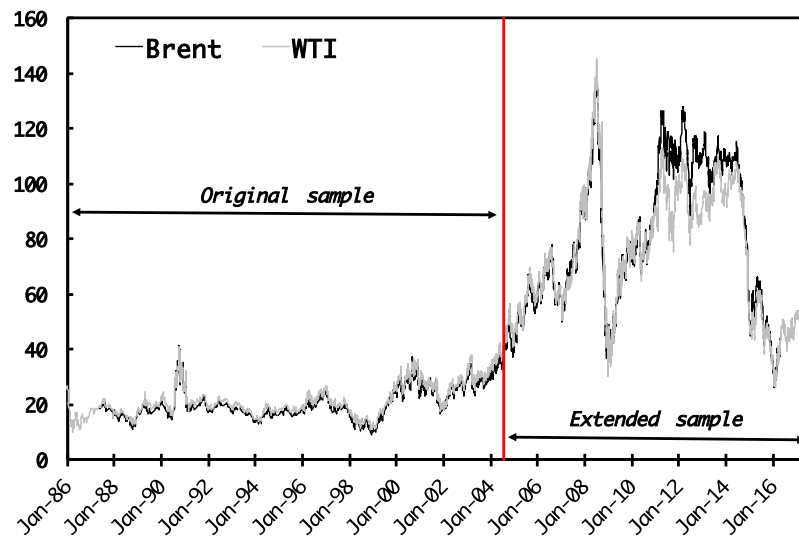


Figure 1: **Crude oil prices.** Prices of the Brent (black curve) and WTI (gray curve) crude oil in the U.S. dollars are shown for the original and extended sample. The samples are split by the red line.

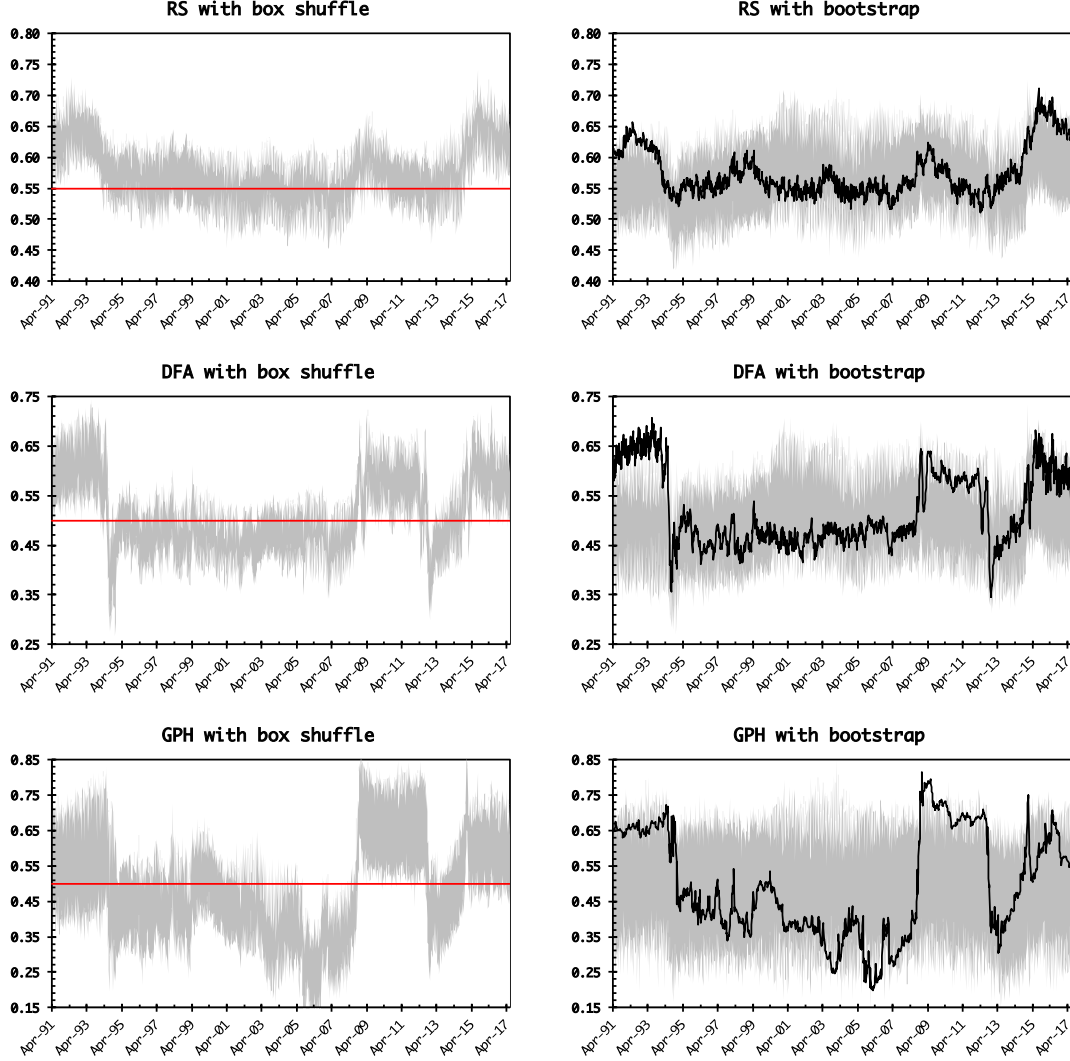


Figure 2: **Hurst exponent estimates for Brent returns.** Hurst exponents are estimated (y -axes) using the rescaled range analysis (first row), detrended fluctuation analysis (second row), and Geweke-Porter-Hudak estimator (third row). Both the original box shuffling procedure (left column) and the moving-block bootstrap procedure (right column) are shown. For each, the gray areas represent the 95% confidence intervals. For the box shuffling procedure, the confidence intervals represent the uncertainty of the estimates. For the moving-block bootstrap, the confidence intervals represent the Hurst exponent estimator distribution under the null hypothesis of no long-range dependence (controlling for the effects of short-range dependence and distributional properties of the analyzed series) and the black curve is the estimated Hurst exponent. For both methods, the box size is set to 20. The moving window size is set to 1,000 days with the step of 5 days. Dates on the x -axes represent the end date of a given window of 1,000 observations.

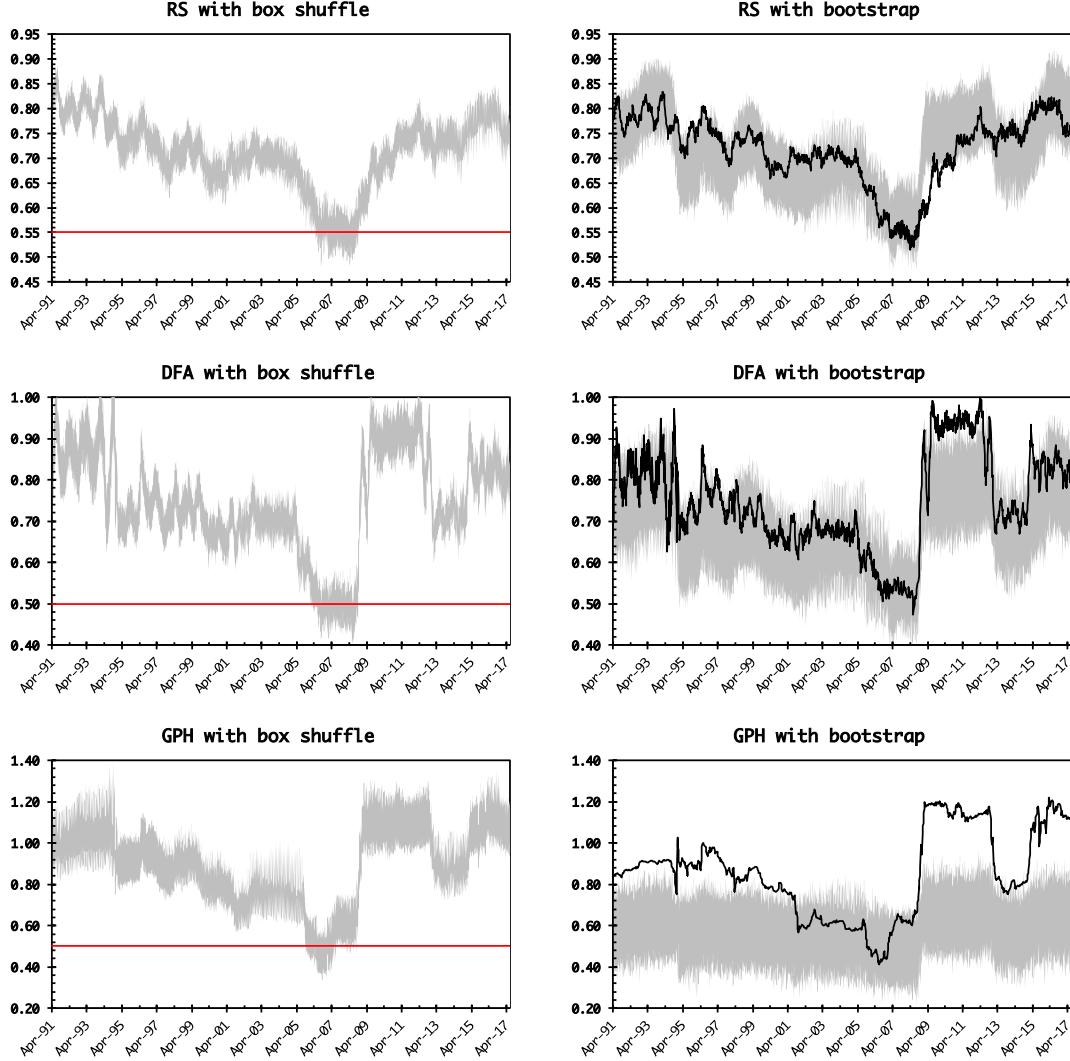


Figure 3: **Hurst exponent estimates for Brent absolute returns.** Hurst exponents are estimated (y -axes) using the rescaled range analysis (first row), detrended fluctuation analysis (second row), and Geweke-Porter-Hudak estimator (third row). Both the original box shuffling procedure (left column) and the moving-block bootstrap procedure (right column) are shown. For each, the gray areas represent the 95% confidence intervals. For the box shuffling procedure, the confidence intervals represent the uncertainty of the estimates. For the moving-block bootstrap, the confidence intervals represent the Hurst exponent estimator distribution under the null hypothesis of no long-range dependence (controlling for the effects of short-range dependence and distributional properties of the analyzed series) and the black curve is the estimated Hurst exponent. For both methods, the box size is set to 20. The moving window size is set to 1,000 days with the step of 5 days. Dates on the x -axes represent the end date of a given window of 1,000 observations.

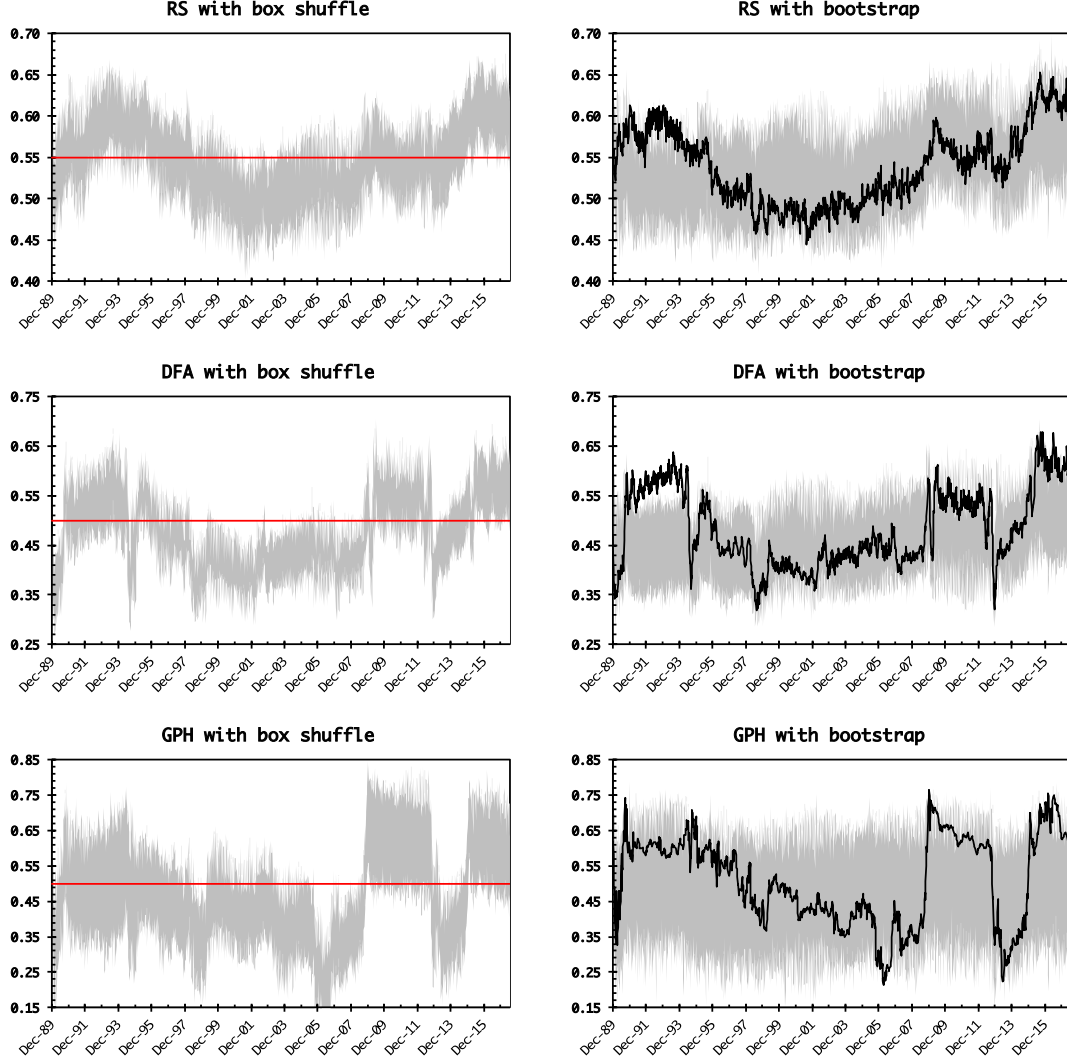


Figure 4: **Hurst exponent estimates for WTI returns.** Hurst exponents are estimated (y -axes) using the rescaled range analysis (first row), detrended fluctuation analysis (second row), and Geweke-Porter-Hudak estimator (third row). Both the original box shuffling procedure (left column) and the moving-block bootstrap procedure (right column) are shown. For each, the gray areas represent the 95% confidence intervals. For the box shuffling procedure, the confidence intervals represent the uncertainty of the estimates. For the moving-block bootstrap, the confidence intervals represent the Hurst exponent estimator distribution under the null hypothesis of no long-range dependence (controlling for the effects of short-range dependence and distributional properties of the analyzed series) and the black curve is the estimated Hurst exponent. For both methods, the box size is set to 20. The moving window size is set to 1,000 days with the step of 5 days. Dates on the x -axes represent the end date of a given window of 1,000 observations.

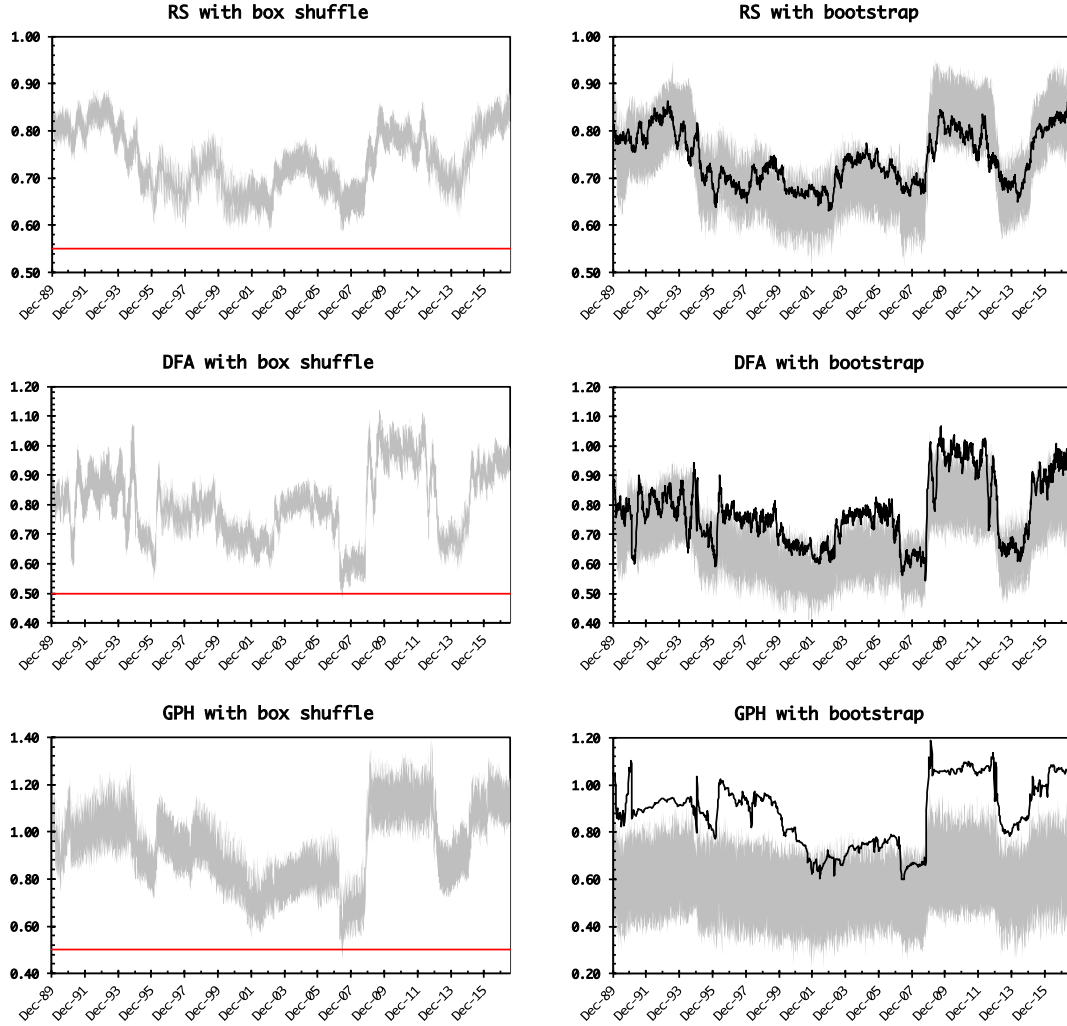


Figure 5: **Hurst exponent estimates for WTI absolute returns.** Hurst exponents are estimated (y -axes) using the rescaled range analysis (first row), detrended fluctuation analysis (second row), and Geweke-Porter-Hudak estimator (third row). Both the original box shuffling procedure (left column) and the moving-block bootstrap procedure (right column) are shown. For each, the gray areas represent the 95% confidence intervals. For the box shuffling procedure, the confidence intervals represent the uncertainty of the estimates. For the moving-block bootstrap, the confidence intervals represent the Hurst exponent estimator distribution under the null hypothesis of no long-range dependence (controlling for the effects of short-range dependence and distributional properties of the analyzed series) and the black curve is the estimated Hurst exponent. For both methods, the box size is set to 20. The moving window size is set to 1,000 days with the step of 5 days. Dates on the x -axes represent the end date of a given window of 1,000 observations.

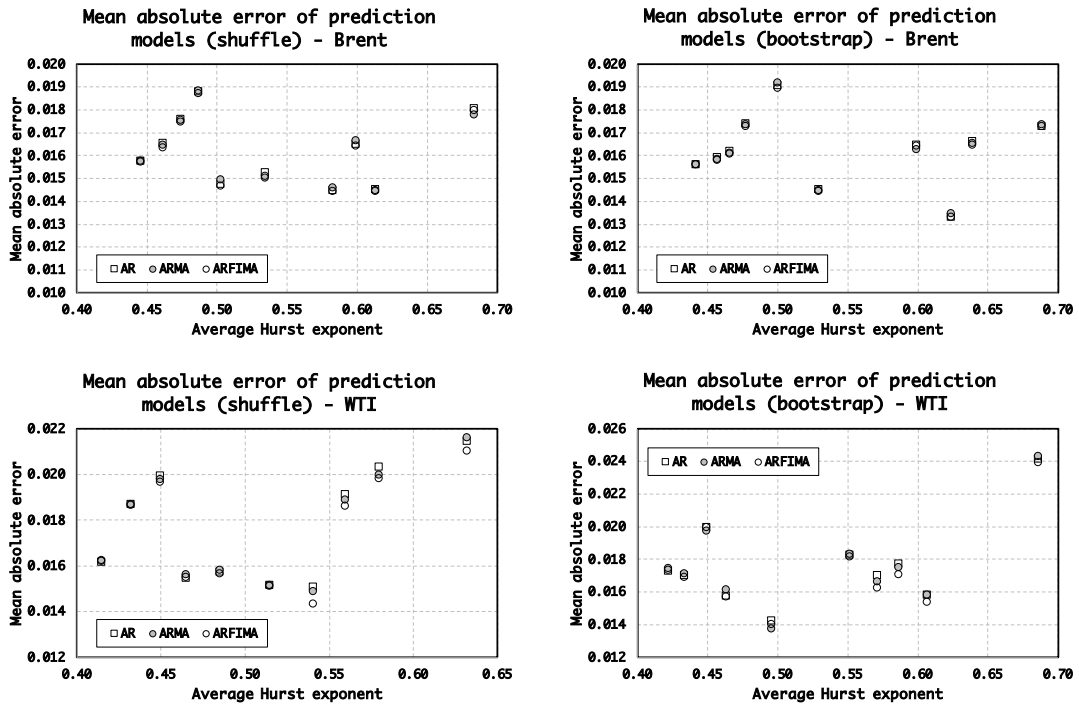


Figure 6: **Dependence of mean absolute (forecasting) error on levels of Hurst exponent.** Forecasting performance is shown for Brent (top) and WTI (bottom) returns. The models (AR, ARMA and ARFIMA) are estimated on the moving window of 1,000 observations with a step of 5 days and we are interested in a 1-day ahead forecasts. For each window of 1,000 observations, the three models are estimated, the Hurst exponents are estimated based on the shuffling (left) and bootstrapping (right) procedures.

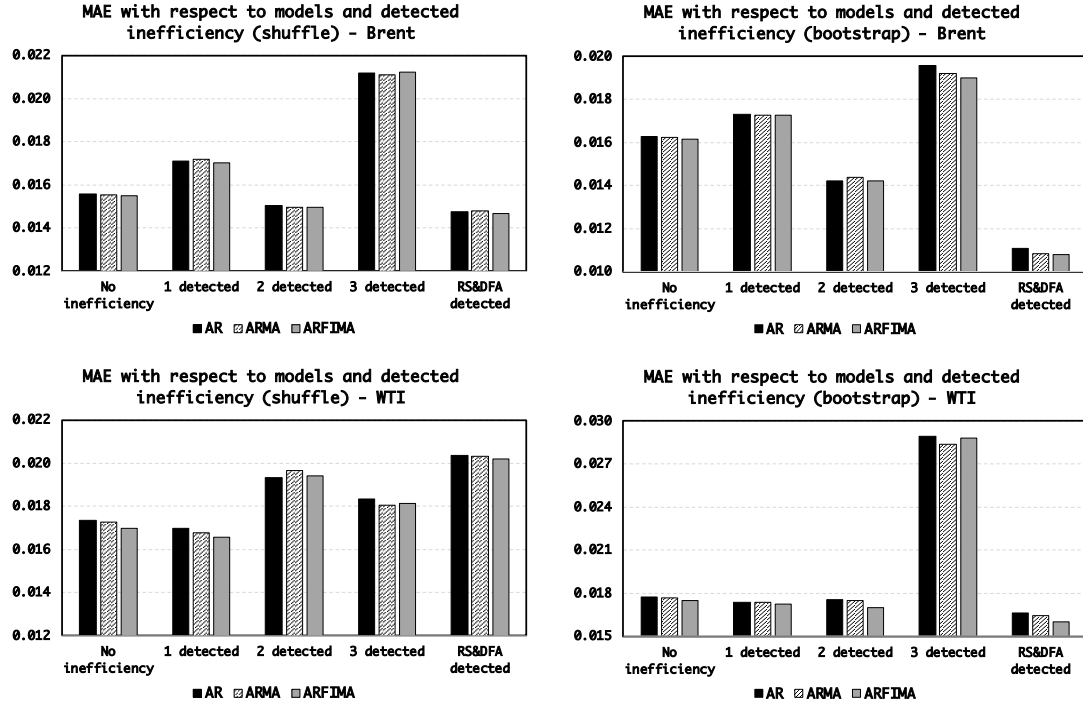


Figure 7: **Dependence of mean absolute (forecasting) error on detected inefficiency implied by Hurst exponent.** Forecasting performance is shown for Brent (top) and WTI (bottom) returns. The models (AR, ARMA and ARFIMA) are estimated on the moving window of 1,000 observations with a step of 5 days and we are interested in a 1-day ahead forecasts. For each window of 1,000 observations, the three models are estimated, the Hurst exponent procedures are performed (estimates, confidence intervals and critical levels) based on the shuffling (left) and bootstrapping (right) procedures. The decision whether the series is detected as inefficient is based on the 95% confidence level. The results are shown for no efficiency detected (meaning that none of RS, DFA and GPH detected inefficiency), 1-3 methods detected inefficiency and the situation when inefficiency is detected by both RS and DFA.

IES Working Paper Series

2018

1. Karel Janda, Martin Strobl: *Smoking Czechs: Modeling Tobacco Consumption and Taxation*
2. Karel Janda, Michaela Koscova: *Photovoltaics and the Slovak Electricity Market*
3. Simona Malovana, Dominika Kolcunova, Vaclav Broz: *Does Monetary Policy Influence Banks' Perception of Risks?*
4. Karolina Vozkova: *Why Did EU Banks Change Their Business Models in Last Years and What Was the Impact of Net Fee and Commission Income on Their Performance?*
5. Jan Malek, Lukas Recka, Karel Janda: *Impact of German Energiewende on Transmission Lines in the Central European Region*
6. David Svacina: *Devaluation with Exchange rate Floor in a Small Open Economy*
7. Ladislav Kristoufek: *Are the Crude Oil Markets Really Becoming More Efficient over Time? Some New Evidence*

All papers can be downloaded at: <http://ies.fsv.cuni.cz>



Univerzita Karlova v Praze, Fakulta sociálních věd

Institut ekonomických studií [UK FSV – IES] Praha 1, Opletalova 26

E-mail : ies@fsv.cuni.cz

<http://ies.fsv.cuni.cz>