Mild explosivity in recent crude oil prices

Isabel Figuerola-Ferretti [a], J. Roderick McCrorie [b, *], and Ioannis Paraskevopoulos [c]

[a] *ICADE Business School, Universidad Pontificia Comillas, Calle Alberto Aguilera 23, 28015 Madrid, Spain & Research Centre for Energy Management, ESCP Europe Business School*

[b] *School of Economics and Finance, University of St Andrews, St Salvator's College, St Andrews, KY16 9AR, Scotland, UK*

[c] *Capital Markets, Bankia, Paseo de Castellana 189, 28046 Madrid, Spain*

Version: 31th October 2016

This paper uses the new, mildly explosive/multiple bubbles technology proposed by Phillips, Shi and Yu (2015, *International Economic Review* 56(4), 1043-1133) to assess whether crude oil prices exhibited departures from martingale trend behavior over the last decade and to explore whether any such departures can be explained by fundamentals or other proxy variables. The test dates two significant time periods in both Brent and WTI nominal and real front-month futures prices: a mildly explosive episode during the 2007-08 spike, prior to the peak of the Global Financial Crisis; and a negative such episode during the recent price decline, whose commencement is dated around a key OPEC meeting in November 2014. Evidence using other commodity prices points to explanatory factors beyond commodity markets. A demand-side fundamental is found to be decisive in explaining the episode in mid-2008, in a way that above-ground inventories and excess speculation are not. U.S. fracking, although a contributor to the post-June 2014 price decline, was not decisive in the rejection of random walk behavior in oil prices in late-2014. In spite of some recent work tying the CBOE Volatility Index (VIX) to oil futures prices, we find no evidence that the VIX affected price levels during the sample period. The results, shown to be robust to a changing dollar exchange rate and to non-stationary volatility in the shocks driving the oil prices, are compared and contrasted with those obtained by Baumeister and Kilian (2016, *Journal of the Association of Environmental and Resource Economists* 3, 131-158) using a forecasting approach based on a structural vector autoregressive model without financial variables.

JEL classification: Q41; G13; C22

Keywords: crude oil; oil prices; Global Financial Crisis; fundamentals; CBOE Volatility Index (VIX); mildly explosive process; generalized sup ADF test

---

* Corresponding author at: School of Economics and Finance, University of St Andrews, St Andrews, Scotland, UK

E-mail addresses: ifiguerola@der.upcomillas.es; mccrorie@st-andrews.ac.uk; iparaskevopoulos@bankia.com

1. Introduction

Crude oil prices over the last decade have exhibited behavior over two notable periods that suggest a standard model of martingale behavior over the whole sample may not appropriately represent their salient features. Towards the end of a sustained rise in crude oil prices from 2003, prices rose from around $50 per barrel at the beginning of 2007 to nearly $150 around the high point of the Global Financial Crisis (GFC) in mid-2008, before collapsing to around $40 towards the year end. Then, after rising in the aftermath of the GFC, with prices enjoying a period of three or so years of stability at just over $100, there was a recent, steep decline in prices of around 70% from June 2014 to January 2016. During the same period, commodity prices declined by nearly 50% from peaks obtained in the aftermath of the GFC.

While there is now a considerable amount of literature on the oil price spike during the 2007-08 period (e.g. Hamilton, 2009, and the discussion therein), there is still ongoing controversy as to whether excess speculation played a role in the run-up in prices towards their high point in mid-2008 (e.g. Kilian and Murphy, 2014; Singleton, 2014). In the same vein, there is as yet no consensus regarding the recent, steep post-June 2014 oil price decline and, in particular, whether there was a principal driver or multiple causes. For example, Azeki and Blanchard (2014) have conjectured the price decline resulted from consequences – actual and expected – of supply increases from the U.S. shale oil revolution; others such as Baffes et al. (2015) have pointed additionally to weakening global demand, a shift in the policy of the Organization of the Petroleum Exporting Countries (OPEC), and an appreciating U.S. dollar. Notable among observable oil supply shocks in the second half of 2014 was the shock to expectations in late-November 2014 when OPEC, rather than agreeing on production cuts, announced it would maintain its production ceiling in spite of an observable glut in supply. The purpose of this paper is to consider whether the 2007-08 spike and/or the post-2014 price decline represented statistically significant departures from random walk behavior and, if so, to assess the role of potential oil price drivers during those periods and at other times using a strategy based on formal statistical testing.

In a recent analysis of crude oil prices, the U.S. Energy Information Administration (EIA, 2016) discussed seven main, broadly-agreed drivers of oil prices in general. These involve the determinants of the four elements of supply and demand of crude oil by OPEC and non-OPEC countries; the relationship between spot and futures prices; market balance (which depends on inventories); and the nature of trading on financial markets. There was not an attempt, however, to assess how decisive each element had been in determining the recent oil price decline. In such an assessment using a forecasting approach based on a structural vector autoregressive (VAR) model, Baumeister and Kilian (2016) showed that more than half the price decline was in principle predictable in real time as of June 2014. Their findings suggest that part of the price decline can be attributed to adverse demand shocks reflecting a slowing global economy in the first half of 2014, and showed the remaining fall could be associated with positive oil supply shocks that occurred prior to July 2014. The strength of their approach, which embodies important methodological advances by Kilian (2009) and is backed by the structural model of Kilian and Murphy (2014), is in its ability to take account of oil price supply and demand shocks through the role it gives to expectations in the formation of oil prices. This means that, realistically, changes in expectations can affect oil prices even when fundamentals are not adjusting. Nevertheless, their approach also comes with some costs. Firstly, the lag structure is fixed *ex ante* as a structural VAR model of order 24, chosen simply to reflect two years' worth of monthly data, even although statistical inference in VAR approaches generally changes if a different lag structure is used.[1] Secondly, while there are attempts through the design of the structural model to ensure the VAR is stationary, this is not tested. Indeed, we show below that it is the very aspect of *non-stationarity* in crude oil prices that is the salient feature during two key time periods of interest. Thirdly, their structural model is based on strong informational assumptions in which agents directly observe structural shocks, to support an inventory-led strategy to detect speculative effects that is based on the classical theory of storage. Sockin and Xiong (2015) model informational frictions agents may face in the global oil market which, if true, challenge the classical theory of storage and would render the structural model misspecified. Fourthly, and not necessarily least, the structural model does not include some financial variables that potentially reflect changes in the oil market that occurred around and in the aftermath

[1] For this reason, lag structures in VAR approaches are often selected on the basis of the actual data using statistical procedures based on information criteria.

of the GFC. For example, Cheng, Kirilenko and Xiong (2015) empirically document recent "convective risk flows" between commodity futures prices and the positions of various trader groups to changes in the Chicago Board Options Exchange Volatility Index (CBOE VIX).

The above comments on the Baumeister and Kilian (2016) approach are not meant to take away from the essential aspect of their contribution – the provision they make for shocks – but they do motivate considering an alternative, reduced-form approach to reassess or to consolidate their findings. Reduced-form models are useful precisely when there are (actual or potential) questions over the specification of a structural model, or to provide empirical evidence against which its specification can be tested. Until recently, directly testing for potential trend departures in an explosive direction, of the sort that would be appropriate for the features in the crude oil data described above, would have represented a considerable challenge because conventional tests based on a model embodying martingale dynamics are very sensitive to explosive departures from the null. The mildly explosive/bubbles detection technology recently proposed by Phillips, Shi and Yu (2015a&b: PSYa&b) offers a natural way forward, given its efficacy in testing for departures from an unobservable stochastic trend such as random walk behavior, and departures from fundamental value, in the direction of positive or negative *mildly explosive process* alternatives (to be defined below). Because such processes dominate other statistical properties of a time series, such departures from trend represent salient features in the data to be explained. The PSY technology offers a statistically consistent basis upon which to date the origination and collapse of such episodes and provides a way of formally testing whether individual factors played a decisive role in any departure from an underlying stochastic trend. Here, we shall show that in both main benchmark crude oil price measures, the PSY procedure dates a positive mildly explosive episode in mid-2008 which associates with the run-up in oil prices towards their high point, during a period in which the GFC had already begun unfolding; and a negative such episode from late November/early December 2014 for two or so months, whose commencement therefore occurs around or just after the OPEC meeting in late November 2014. That such date-stamping results are found when applying the more stringent, wild bootstrap procedure of Harvey et al. (2016) to the benchmark nominal price series indicates the results are robust to non-stationary volatility in the shocks driving the oil prices. We also formally test whether a small number of key fundamental proxy variables and financial variables may have had a decisive role in explaining the detected departures from trend, and provide evidence that the drivers of the two episodes were different in nature.

The paper is organized as follows. Section 2 gives some background on the main crude oil price benchmarks and briefly presents their chronology over the last decade. Section 3 summarizes the PSY mildly explosive/bubbles testing methodology. In Section 4, the test is applied to the nominal front-month futures price series for the two main crude oil benchmark price series, and the same two series deflated by the U.S. Consumer Price Index (CPI) and by the currency value of the Special Drawing Rights (SDR) basket to control for aspects of the changing dollar numeraire.[2] In Section 5, we provide some interpretation of our results, focussing on the potential drivers of recent crude oil prices. Here, nominal prices are deflated by natural proxies for two traditional fundamental variables, global economic activity and above-ground inventory supply, and then by a financial variable, the CBOE VIX, as a means of assessing the extent of the wider role that oil futures now play in portfolio management and in calculations of systemic market risk. Section 6 concludes.

2. Crude oil prices over the last decade

There are two principal crude oil benchmarks based on quality and location – Brent Blend and West Texas Intermediate (WTI) – that are used by buyers of crude oil and speculators (who never actually take delivery of it). Around two-thirds of all crude contacts reference Brent Blend,[3] making it the most important marker price by volume. WTI refers to oil extracted from wells in the U.S. that is piped to Cushing, Oklahoma and is the main benchmark for oil consumed in the U.S. WTI crude is lighter and sweeter (has a lower sulphur content) than Brent Blend which, ceteris

[2] The currency value of the SDR is determined using a basket of major currencies (the U.S. dollar, the euro, Japanese yen and pound sterling) based on market exchange rates. Adjusting nominal oil price series by the U.S. CPI index and the SDR currency value confers some robustness on our approach but is not definitive because the oil price is a *world* price denominated in dollars: prices can rise in one currency and fall in another. Deflating by the U.S. CPI, we follow the convention that has been standard in the literature since the paper by Deaton and Laroque (1996).

[3] Brent Blend originally referred to oil extracted from the Brent oil field, but now refers to oil from four different fields: Brent, Forties, Oseberg and Ekofisk.

paribus, would make it the more expensive crude oil but supplies of the former are landlocked and are relatively expensive to ship to other parts of the World. Some important recent instances of spreads attributed to factors other than quality and location have been reported (e.g. Büyükşahin et al., 2013) and are briefly discussed below. We focus on front-month crude oil futures data, representing the settlement prices of the contracts that are closest to delivery with the highest open interest, and against which bid/offer prices and deals for physical cargos of crude oil are referenced against. The two contracts are traded on different markets: Brent futures are traded on the Intercontinental Exchange (ICE, into which the International Petroleum Exchange has been incorporated) and WTI contracts are sold principally on the New York Mercantile Exchange (NYMEX). Many oil futures traders follow strategies in which they never make or take delivery of physical oil: some unwind their positions and extract monetary gain or loss by selling the futures contract previously bought; others, in a strategy called rolling, repeatedly roll over their position by selling a current futures contract before expiry, and thereafter purchase the next month's contract. The futures market therefore supports both hedging, where net gains and losses are used to offset fluctuations in operating earnings; and speculation, where actual net gains are counted as profit or loss.

Hamilton (2008) and more recently Alquist et al. (2013) discuss whether it is preferable to employ nominal or real prices in an analysis of oil prices. Having sympathy with an argument in the former paper – that the CPI deflator can have as much influence on the ratio as the nominal price when it is influenced by effects seemingly unrelated to oil – we focus our results on nominal prices. Where relevant, however, we shall also report results for real prices. When considering proxy variables or indices, we assess their properties against the nominal price series and then either against the real series or by examining ratios of the nominal prices to each variable and index. Using ratios obviates the need to deflate variables in order to discuss real effects, although some care needs to be applied given the potential for the underlying statistical inference to be distorted. In reporting the above results, we are therefore able naturally to follow the pioneering strategy by Phillips, Wu and Yu (2011, PWY) and Phillips and Yu (2011, PY) in their treatment of fundamentals based on formal statistical testing, and at the same time we can make direct comparisons with the results obtained by Baumeister and Kilian (2016) which pertain to real oil spot prices.[4]

Our chosen sample begins in January 2003, immediately prior to the beginning of the sustained rise in oil prices during the last decade and, anticipating the analysis in Section 5, in line with the CBOE's revised, model-free measure of the VIX that has been calculated since then. The PSY test relies on a choice of an initial fraction of the data and the various test statistics and critical values are functions of that choice. Our chosen sample end-date is the end of April 2016 which, following a small recovery in oil prices in early 2016, removes potential end-of-sample dating issues and facilitates comparison of our results with those by Fantazzini (2016) who uses the same end-date. We use weekly data to facilitate testing for potentially relevant, shorter periods of mild explosivity than monthly data would permit, which will be appropriate for assessing aspects of the so-called financialization of oil, but the sampling frequency it is still low enough to permit a complementary analysis based on more traditional fundamentals. Our primary data source is the Bloomberg International database.

By way of a brief market chronology over our sample period, as shown in Figure 1, both Brent and WTI crude oil benchmark prices began a steady year-on-year rise from below $30 per barrel at the beginning of the sample, to over $50 in 2005, over $60 in 2005, and reached around $75 in mid-2006 before dropping back to $60 in early-2007. Prices were over $90 by October 2007, around $100 by late-November 2007 and rose sharply to a peak seen in the ICE Brent price of just under $147 in mid-July 2008.

[Figures 1 and 2 around here]

The beginning of the 2008 price collapse, which coincided with a low point in the U.S. dollar, can be tied to the announcement by U.S. President George W. Bush on July 14 that ended an executive ban on offshore drilling. The price collapsed below

[4] Conducting a forecasting evaluation exercise based on the Kilian and Murphy (2014) model, Baumeister and Kilian work with spot prices. The variability of the futures price about the spot price of oil renders the former inaccurate as a *predictor* of the latter; indeed Alquist and Kilian (2010) show the futures price can be less accurate as a mean-square predictor than is a no-change forecast. Here, we conduct statistical tests to evaluate past evidence, and use benchmark futures prices in financial markets to do so. The spot price of oil contains forward-looking elements based on expectations that cannot be captured by past data. It is natural that market participants trade on this information and, in the oil market, this is predominantly done through the purchase and sale of futures contracts rather than through the purchase, storage and sale of physical oil. Under perfect arbitrage, the two activities are equivalent, with spot and futures prices being jointly determined and responding to the same fundamentals. In the sequel, we compare our results with recent results by Fantazzini (2016) using spot prices, which allows for an assessment of how well arbitrage works in the oil market.

$100 in the late summer of 2008, below $70 in mid-October, below $60 in November and below $40 in December, during the time the GFC was unfolding. Two salient features are therefore the 2003-08 run-up in prices, during which there was a period of renewed growth in the Organization for Economic Co-operation and Development (OECD) countries and rapid industrial growth in Asia, and particularly in China; and the 2007-08 price spike during whose run-up phase prices rose significantly faster than before. A similar pattern was found in some other commodity prices in 2007-08. In 2009, prices began to recover, finally exceeding $100 in 2011 and then remaining around this level for two or more years. Figure 1 indicates a divergence between Brent and WTI nominal prices during this time, which can be attributed to U.S. supply bottlenecks at Cushing, Oklahoma (e.g. Büyükşahin et al., 2013; Robe and Wallen, 2016). The other salient feature in oil prices during the sample period is the post-June 2014 price decline, which occurred around a time that U.S. shale oil production had increased and China and Europe's demand for oil had decreased. A significant plunge in prices to a four-year low of just over $70 was seen in late-November 2014, around the time of an OPEC meeting which did not reinforce the cartel's longstanding policy of defending prices. Crude oil prices continued to fall in 2015, reaching a price of just over $40 in mid-August that had not been seen since February 2009. In December 2015, the price at below $34 hit an eleven-year low; it then dropped below $30 in January 2016 and reached a low of around $26 in early-February. Prices had recovered to around $45 by the end of the sample period.

3. The PSY mildly explosive/bubble-testing methodology

The PSY (2015a&b) strategy represents a breakthrough in the search for a statistically rigorous procedure to test for temporary regime shifts of exuberance and collapse that are embedded in a time series evolving as a stochastic trend. The test is based upon the notion of a *mildly explosive process* introduced by Phillips and Magdalinos (2007), which facilitates constructing appropriate distribution theory in autoregressive (AR) models whose AR parameter is locally above unity. The property of mild explosivity dominates other trend components in a way that makes such processes very different statistically from those exhibiting the random walk or martingale behavior commonly found in the empirical literature. It therefore represents a salient feature of the time series to be explained.

The PSY procedure is based on three steps:

- testing the null hypothesis that there are no mildly explosive periods in the sample against the alternative that there is at least one such period;
- if the test rejects, date-stamping the mildly explosive period(s) in the sample;
- assessing whether detected mildly explosive episodes may be due to the behavior of fundamentals or other variables.

Under the null hypothesis of the test, the time series of interest follows a unit root process with an asymptotically negligible drift:

$$x_t = dT^{-\eta} + x_{t-1} + \varepsilon_t, \quad \eta > \tfrac{1}{2}, \tag{1}$$

where $x_0 = O_p(1)$, d is a parameter and η is a localizing coefficient that controls the magnitude of the drift as $T \to \infty$. Starting from a fraction r_1 and ending at a fraction r_2 of the total sample, with window size $r_w = r_2 - r_1$, we fit the regression model

$$\Delta x_t = \alpha_{r_1,r_2} + \beta_{r_1,r_2} x_{t-1} + \sum_{i=1}^{k} \psi_{r_1,r_2}^{i} \Delta x_{t-i} + \varepsilon_t, \tag{2}$$

where k is the lag order chosen on sub-samples using the Schwarz Bayesian Information Criterion (BIC), and $\varepsilon_t \sim$ i.i.d. $(0, \sigma_{r_1,r_2}^2)$. The number of observations in the regression is $T_w = \lfloor T r_w \rfloor$, where $\lfloor . \rfloor$ is the floor function, and we denote the ADF-statistic (*t*-ratio) of the coefficient of x_{t-1} based on this regression by $ADF_{r_1}^{r_2}$.

PSY (2015a&b) introduce two statistics, the backward sup ADF (BSADF) statistic and the generalized sup ADF (GSADF) test. They are defined as:

$$BSADF_{r_2}(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \{ADF_{r_1}^{r_2}\}. \tag{3}$$

$$GSADF(r_0) = \sup_{r_2 \in [r_0, 1]} \{BSADF_{r_2}(r_0)\}, \tag{4}$$

where the endpoint of the sample is fixed at r_2 and the window size is allowed to expand from an initial fraction r_0 of the total sample to r_2. PSY (2015a) propose r_0 be chosen according to the rule $r_0 = 0.01 + 1.8/\sqrt{T}$, where T is the sample size. This procedure defines a particular BSADF statistic. The GSADF statistic is then constructed through repeated implementation of the BSADF procedure for each

$r_2 \in [r_0, 1]$. Critical values are obtained by simulation. PSY (2015a&b) provide limiting distribution theory and small sample simulation evidence.

The null test of no mildly explosive periods is based on the GSADF statistic. Date-stamping mildly explosive periods is achieved through the BSADF statistic: the origination and termination points of a first mildly explosive period, $r_{1,e}$ and $r_{1,f}$, are estimated, subject to a minimum duration condition, by

$$\hat{r}_{1,e} = \inf_{r_2 \in [r_0,1]} \left\{ r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\}, \tag{4}$$

$$\hat{r}_{1,f} = \inf_{r_2 \in [\hat{r}_e + \delta \log(T)/T,\, 1]} \left\{ r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\}, \tag{5}$$

where $scv_{r_2}^{\beta_T}$ is the $100(1 - \beta_T)\%$ right-sided critical value of the BSADF statistic based on $\lfloor T r_2 \rfloor$ observations and δ is a tuning parameter that can be chosen on basis of sampling frequency. A tuning parameter of unity, reflecting a standard application of the test, implies a minimum duration condition of log(T) observations. A mildly explosive period is declared if and when the BSADF statistic has been above its critical value for at least $\lfloor T \hat{r}_{1,e} \rfloor + \lfloor \log(T) \rfloor$ observations. Conditional on a first mildly explosive period having been found and estimated to have terminated at $\hat{r}_{1,f}$, the procedure is then repeated in search of a second and possibly subsequent periods. PSY (2015b) show that, subject to rate conditions, the sequential procedure provides consistent estimates of the origination and termination dates of one, two and three (and, in principle, more) bubbles.

The data generation process (DGP) under the alternative hypothesis exhibits K mildly explosive episodes in the sample period, represented in terms of sample fraction intervals $B_i = [\tau_{i,e}, \tau_{i,f}]$ (i = 1, 2 , . . . , K), within periods of prevailing martingale-type behavior in the intervals $N_0 = [1, \tau_{1,e}]$, $N_j = [\tau_{j-1,f}, \tau_{j,e}]$ (j = 1, 2 , . . . , $K - 1$) and $N_K = [\tau_{K,f}, T]$, as follows:

$$\begin{aligned} x_t &= (x_{t-1} + \varepsilon_t)\, 1(t \in N_0) + (\delta_T\, x_{t-1} + \varepsilon_t)\, 1(t \in B_i) \\ &\quad + \sum_{i=1}^{K} \left(\sum_{l=\tau+1}^{t} \varepsilon_l + x_{\tau_{i,f}}^{*} \right) 1(t \in N_i) \qquad (t = 1, \ldots, T) \end{aligned} \tag{6}$$

$$\delta_T = 1 + c/T^{\alpha}, \quad c > 0, \quad \alpha \in (0, 1). \tag{7}$$

Under the given conditions on c and α, the autoregressive parameter δ_T is greater than unity and is what Phillips and Magdalinos (2007) called a mildly integrated root on the explosive side of unity, or simply a *mildly explosive process*.[5] Because the statistics remain the same if the time series is reversed, such processes can be rising or falling, and for our purpose can therefore be applied to test for a departure in trend in both the 2007-08 oil price spike and the price decline beginning in mid-2014. The PSY test has sometimes been presented as a test for bubbles, although here we have avoided this nomenclature given its many different connotations in the Finance literature. Also, we refer to testing the recent price decline for departure in trend as being in the (negative) mildly explosive direction rather than being in a mildly implosive direction because "implosion" has typically been used in the literature to refer to a collapse period immediately following a bubble (e.g. Phillips and Shi, 2014). Implosion in the PSY test is modelled, for each i, by $x^*_{\tau_{i,f}} = x_{\tau_{i,e}} + x^*_i$, where $x^*_i = O_p(1)$. For a positive mildly explosive episode, this therefore entails a collapse to the value of the last pre-bubble observation plus an $O_p(1)$ perturbation, from which the process resumes its trend. There is ongoing work, including by Phillips and Shi (2014), which seeks to model bubble implosion in a richer way than in PSY (2015a).

An important aim of this paper is to investigate whether any detected mildly explosive episodes in oil prices can be linked to fundamentals or financial variables. Following the approach by PWY (2011), we apply the PSY test to assess whether the mildly explosive periods in the nominal price coincide with such periods in the fundamental or financial variable and, using ratios of the same, whether the degree of non-stationarity in the nominal price is greater than the degree of non-stationarity in the particular fundamental or financial variable.

The importance of the PSY (2015a) strategy is exemplified by the number and diversity of applications that have already appeared, in spite of its recent inception. For example, it has been applied to the S&P500 index (PSY, 2015a); residential property markets (e.g. Greenaway-McGrevy and Phillips, 2016); food commodity markets (e.g. Etienne, Irwin and Garcia, 2015); and metals markets (e.g. Figuerola-Ferretti et al. 2015). Indeed it has already been applied to crude oil markets by Caspi

[5] Under this construct, they were able to establish a basis for statistical inference on the explosive side of unity.

et al. (2016) who look at historical periods of oil price explosivity since 1876; Tsvetanov et al. (2016) who study the shape of the oil futures forward curve; and by Fantazzini (2016) who predominantly used the results of the second-stage of the test applied to Brent and WTI spot prices as one of two disparate approaches to consider whether the recent oil price decline was a negative bubble.

Our work is distinguished from the approaches above through its unified treatment of both the 2007-08 oil price shock and the post-2014 price decline; in its use of all aspects of the PWY/PSY testing strategy to support an evidence-based approach that reaches beyond simply dating structural breaks to assess the drivers of recent oil prices; in its robustness to a potentially changing dollar numeraire and to possible non-stationary volatility in the shocks driving the oil prices; in its collation and interpretation of empirical evidence as a way of generating stylized facts for structural models to explain; and, not least, in the clarity it gives to the assessment of evidence regarding the salient features of oil prices around and in the aftermath of the GFC.

4. Test results

In this section, we conduct the first two stages of the PSY test using both Brent and WTI front-month futures benchmark series, documenting any detected departures from normal martingale trend behavior in the direction of mildly explosive alternatives. We shall also examine the robustness of the results to a changing dollar numeraire and to possible non-stationary volatility in the shocks driving the oil prices. In Section 5 below, we provide an interpretation of the salient features we find in the data using the third stage of the PSY procedure along with specifically chosen proxy variables for global economic activity, above-ground inventories and market risk.

Formally, the GSADF statistic tests in the direction of at least one episode of mildly explosive behavior in the sample. Tables 1A and 1B give the GSADF statistics corresponding to the Brent and WTI nominal front-month futures prices observed on a weekly basis, and the same series deflated by the SDR and an interpolated U.S. CPI series, given that the raw CPI series is available only at a monthly frequency. Critical values are generated for the given sample size and the rule-based value of r_0.

[Tables 1A and 1B around here]

In every case, the GSADF statistic rejects the null hypothesis of there being no mildly explosive periods during 2003-2011 at the 1% significance level. With such corroboratory evidence for mild explosivity in the sample, we move to the second, date-stamping stage of the PSY procedure, which uses the BSADF statistic, noting that because of the imposition of a minimum duration condition, it is possible that no mildly explosive period is actually dated here. Figures 2A and 2B show the respective BSADF sequences at the 5% significance level.

[Figures 2A and 2B around here]

We date mildly explosive episodes by the periods during which the BSADF sequence is above the 5% critical value line such that the minimum duration condition is satisfied. These are reported in Tables 2A and 2B, where we find essentially the *same two* mildly explosive episodes dated for *both* benchmark Brent and WTI nominal *and* real (U.S.-CPI-deflated) prices. These are the central results of this paper: a positive mildly explosive episode is dated between mid-May and mid-July 2008 (when nominal prices are rising) and a negative such episode is dated between late November/early December 2014 and mid-February 2015 (when nominal prices are falling). Given the weekly sampling frequency, we report dates identified up to the Friday of a given week. In all cases, the reported end-date of the first episode is the Friday immediately following President Bush's announcement on 14 July 2008 on offshore drilling, reported in Section 2, which marked the high point in oil prices.

[Tables 2A and 2B around here]

Though other applications of the PSY test to various crude oil series have a different focus from ours it is, nonetheless, important we compare our date-stamping results with such work where there is potential overlap, to check for robustness. The issue of robust date-stamping is important: for example, our reduced-form results suggest there was a statistically significant departure from random walk behavior in late November/early December 2014 around or just after the meeting when OPEC producers maintained their production ceiling when faced with observed supply increases elsewhere. Arzeki and Blanchard (2014) and Baffes et al. (2015) cite this meeting as a source of a supply shock that was a key component of the post-2014 price decline. This conclusion was disputed by Baumeister and Kilian (2016) who, on the basis of their structural approach, inferred that there was a negative demand shock around this time, which they attributed to an unexpected weakening of the global economy. We shall first establish the robustness of date-stamped episodes above and then, in Section 5 below, address the issue of their drivers.

Caspi et al. (2016) report a longer mildly explosive episode in real prices, from October 2007 to August 2008, right at the end of their study of historical oil prices using *monthly data*. This was only the thirteenth such episode they detected in the oil price since 1876, which points to the significance of the departure in trend, but it should be noted (in a way not to detract from their contribution to the historical evidence base) that their results were obtained using an earlier version of the PSY test,[6] and that with their sampling frequency, the dating algorithm is not as sharp as it could be with higher frequency data. We note, for example, that real oil prices had begun their steep decline before the end of their detected period. Using WTI weekly data and date-stamping to the nearest month, Tsvetanov et al. (2016) detect a mildly explosive period in the front-month contract (*ibid*., Figure 5, Contract 1) between April and July 2008, beginning a week or two earlier than ours.[7] Using *daily data*, Fantazzini (2016) reports a (negative) mildly explosive period between early October 2014 and late February 2015 in Brent nominal and real *spot* prices, and between early December 2014 and late March 2015 in WTI nominal and real spot prices. In robustness checks using weekly data with a longer sample, however, an earlier episode from mid-May 2008 to mid-July 2008 is reported, synchronous across Brent and WTI real and nominal spot prices (*ibid*., Table 2, p. 392). The later episode is dated in each case from October 2014 to April 2015, with *exact* synchronicity reported between Brent real and nominal prices, and between WTI real and nominal prices.[8] Corroboratory evidence exists outside the PSY framework for the first mildly explosive or similar type of episode during 2008. For example, Shi and Arora (2012) applied a regime-switching model with the oil price decomposed into deterministic and bubble components, and found the estimated probability of being in a bubble surviving regime rose just prior to a spike in the estimated probability of being a bubble collapsing regime in late-2008.

---

[6] Their application is based on Phillips et al. (2013) which uses a different value for the initial sample fraction, r_o, than the rule-based value in the final version (PSY, 2015a) which was calibrated by sample size.

[7] Their graphical method of date-stamping, however, obscures the role of the minimum duration condition and so, even although our data are the same, we prefer to take our own result at face value.

[8] This corrects typographical errors in the third and eight lines of the last column of Table 2.

Three conclusions emerge from the above. First, that both spot prices and the benchmark front-month futures prices have seen two significant departures from martingale trend behavior in crude oil prices over the last decade. Second, that the episode detected by Fantazzini in spot prices in 2008 is *exactly synchronous* with the episode we detected using front-month futures prices, whether using Brent or WTI, or real or nominal prices. And third, that the trend departure in 2014 may have begun earlier in spot prices than in front-month futures prices.

We now provide a further robustness check of our results using SDR-adjusted prices which control for a changing dollar exchange rate. The GSADF test statistic again rejects the null of no mildly explosive episode in the sample for both benchmark prices at the 1% significance level. The second-stage application of the test indicates that for Brent prices at the 5% significance level, there is a mildly explosive episode detected from the beginning of December synchronous with those detected using nominal and real prices. An episode that is similarly synchronous with nominal and real prices is detected in mid-2008, but only at the 10% significance level. The results for SDR-adjusted WTI prices are similar: here both episodes are detected at the 5% significance level, although the episode in 2008 is slightly longer, beginning in late-April rather than mid-May. This result is not especially significant given that dating identification is to the nearest week. All in all, the synchroneity of the date-stamping using SDR-adjusted prices on one hand, and nominal and real prices on the other, indicates that the depreciating dollar in mid-2008 or the appreciating dollar during 2014 were not decisive in the movements in oil prices around then that led to the rejection of random walk null behavior. Evidence from the BSADF sequence points to the actual appreciation in the dollar numeraire of just under 10% after June 2014 as having been mildly contributing to the post-June 2014 oil price decline, corroborating the conclusion reached by Baumeister and Kilian (2016) on the basis of their structural approach, rather than its having been a strongly contributing factor, as envisioned by Baffes et al. (2015).

Simulation evidence presented by Harvey et al. (2016), applied to the PWY (2011) prototype of the PSY test, suggests the efficacy of the test may be affected by size distortion arising from changes in the unconditional variance of the shocks across the different regimes. This can be informally assessed by an examination of a plot of the first differenced series.

[Figures 3A and 3B around here]

Figures 3A and 3B show that both the Brent and WTI benchmark nominal series do seem to exhibit some increase in volatility during and in the immediate aftermath of the price spike in 2008, namely around the price rise during our first detected mildly explosive period and in the price collapse in its immediate aftermath. Indeed, this increase is observed generally, but less markedly, during the whole aftermath of the GFC, and may point to a weak but nevertheless non-negligible permanent volatility shift in oil prices in the aftermath of the 2007-08 financial crisis. Though weaker than is suggested by Figure 2 in Figuerola-Ferretti and McCrorie (2016) for the precious metals gold, silver and platinum over the same period, we feel the evidence is strong enough to warrant conducting a formal statistical test. Accordingly, following Harvey et al. (2016), we apply the wild bootstrap resampling scheme to the first differenced data, which controls for the (potential) problem by replicating the pattern of volatility in the resampled data that is present in the original innovations.[9] This makes for an unambiguously more stringent test than is provided by a standard application of the PSY test.

[Figures 4A and 4B around here]

Figures 4A and 4B show WTI front-month nominal and real prices and their respective BSADF sequences overlaid with a number of wild-bootstrap-adjusted critical value lines computed using 5,000 replications.[10] In both cases, there are spikes in the BSADF sequences that either touch or go above the 5% critical value line for a period, around the time of the first episode in mid-2008 detected using standard PSY critical values. In both cases, similar spikes go above the 1% critical value line from early December 2014, around the time of the second episode detected previously. As reported in Table 2C, at the 5% significance level with imposed minimum duration

[9] Etienne et al. (2015) and Figuerola-Ferretti and McCrorie (2016) also employ the wild bootstrap in the PSY setting and, although Harvey et al. (2016) point to reasons why the approach should naturally extend to this case, its validity has yet to be formally justified. The wild bootstrap controls for a wide variety of permanent volatility changes, whether single or multiple, abrupt or smooth. A step-by-step algorithm for computing wild bootstrapped critical values in the current context is provided, for example, by Figuerola-Ferretti and McCrorie (2016, p. 724).

[10] This is the same number of replications that PSY (2015a) used in the generation of critical values in the standard case and ensures comparability with their approach.

condition, there are mildly explosive periods dated only for the second episode, that were shorter than were detected using standard PSY critical values.

[Table 2C around here]

Evidence using a 10% significance level reveals slightly longer episodes beginning in December 2014 that, crucially, are closer to the late-November 2014 OPEC meeting, and now a six-week episode in WTI nominal prices is 2008, dated between mid-June and again the Friday in mid-July immediately after President Bush's announcement on 14 July that marked their high point. In the latter episode, the minimum duration condition is just satisfied. In contrast, it fails for WTI real prices and so, even although the BSADF sequence touches the 5% critical value line and is above the 10% line for a few weeks, indicating that the test decision is marginal, no mildly explosive episode in mid-2008 is actually dated.

To summarize all of the evidence presented in this section, there is compelling evidence on the basis of a standard application of the PSY test that points to there having been two significant departures from random walk behavior over the last decade in both the nominal and real prices of both crude oil benchmarks, the first occurring between the second and third quarters of 2008 and the second beginning in late-November/early December 2014 and lasting for two or so months. These results are robust to a potentially changing dollar exchange rate. There is convincing evidence on the basis of applying the wild bootstrap to the generation of critical values that the second episode in both nominal and real prices is robust to non-stationary volatility in the shocks driving the oil prices, and can be persuasively tied to the period around or in the immediate aftermath of the late-November OPEC meeting, given the tolerance level in date-stamping such episodes using weekly data. There is significant evidence that the first episode in nominal prices is also robust, being date-stamped in the final stages of the run-up in prices during the 2007-08 spike. The evidence for such an episode in real prices, however, is notably weaker. Taking account of the aspect that applying the wild bootstrap imposes a stringency beyond the standard application of the PSY test, other than for the first episode for WTI real prices, we cannot reasonably attribute the detected mildly explosive episodes to non-stationary volatility in the innovations driving the oil prices, and proceed on the basis of the episodes being salient oil price features to be explained.

5. Fundamentals and the role of financial variables

We now apply the third stage of the PSY strategy, which uses the test to assess whether the detected mildly explosive periods can be related to fundamentals or other proxy variables. In so doing, we can formally assess various conjectures and results about the drivers of oil prices that have appeared recently. Even today, economic theory is surprising silent on general principles of choosing or identifying the fundamentals for a given asset or commodity and this means we are forced to follow the convention of choosing fundamental and other proxy variables on an ad hoc basis. Given the reduced-form nature of our approach, we can only provide corroborative evidence: we cannot assert causation between two variables because there is always the possibility that a third variable is acting on both the oil price and the proxy variable. It is also possible that the same evidence can be consistent with two or more competing hypotheses. We can, however, use results from the PSY test to seek evidence *against* causation given that the mild explosivity property dominates other features in the series, and in this way we can build up an evidence base that allows the implications of potentially different structural model specifications to be assessed.

In this section, we firstly view oil through a pure commodity lens and ask whether recent oil price behavior was mirrored in other commodities or commodity indices. Finding the evidence points to factors specific to oil, we then consider whether the detected mildly explosive episodes pertaining to the 2007-08 oil price spike and the post-2014 price collapse were consistent with having been driven by oil fundamentals. We use two traditional fundamentals: global economic activity (on the demand side), which we proxy by the Baltic Dry Index; and inventories (on the supply side) which, following Phillips and Yu (2011) and others, can be proxied by U.S. crude oil stocks, which conveniently are made available by the EIA as end-of-week data. We then assess recent conjectures regarding the financialization of oil, where large numbers of financial actors, specifically investment banks, hedge funds and index investors, have become involved in oil and commodity futures markets. Financialization has undeniably changed the composition of oil markets but whether it has been the *major driver* of oil prices is much more controversial.[11] We consider evidence using the

[11] See Fattouh et al. (2013) and Cheng and Xiong (2014) for surveys that apply generally to commodity markets.

PSY strategy by taking two perspectives: first, to consider whether there was any significant, wider role that might have been played by oil prices as an element in financial portfolios or in calculations of systemic market risk, particularly around and in the aftermath of the GFC; and second, to consider whether there is evidence of there having been excess speculative activity in oil. The former is assessed using the equity VIX and the latter using Working's (1960) *T*-index computed using U.S. Commodity Futures Trading Commission (CFTC) position data.

5.1. Oil and other commodities

In the pioneering paper, PWY (2011) used the NASDAQ stock price index and a natural fundamental – dividend yield – to test whether there was a non-fundamentals-based mildly explosive departure from trend in NASDAQ stock prices. Finding evidence of a period of mildly explosive behavior in the former but not the latter constituted evidence that the NASDAQ stock prices were in this period not fundamentals-driven. While convenience yield plays the same role for commodity prices as dividend yield plays for stock prices, it is not suitable when using the PSY test for assessing whether the commodity prices such as oil prices are fundamentals-driven: for perfectly natural reasons, the convenience yield can change in a way that makes it uninformative.[12] We shall instead consider whether other commodities exhibited their own statistically significant departures from a random walk behavior that were synchronous with our detected periods in oil, formally testing the same for copper and zinc – which Azeki and Blanchard (2014) cited as exemplars – and the Bloomberg Commodity Index (BCI), the preferred index used in the EIA (2016) in its summary of oil price drivers. This represents a sharper approach than simply making an informal appeal to a seemingly apparent observation of a coincident boom and bust in certain commodity markets in 2007-08.

Figuerola-Ferretti et al. (2015) apply the PSY test to weekly data on six London Metals Exchange (LME) three-month futures prices, reflecting the most liquid futures contract traded.

[Table 1C around here]

[12] For example, Morana (2013) documents a switch in the oil market in 2005 from backwardation to contango. See Figuerola-Ferretti and McCrorie (2016) for further discussion.

Table 1C reports results of the GSADF statistic for copper and zinc where the parameters (the sample start and end dates, and the initial sample fraction r_0) are chosen to be exactly as above, and so include the recent period of commodity price falls, and show the null of no mildly explosive period is rejected at the 1% significance level. The BSADF sequence dates mildly explosive episodes in copper between March and June 2006 and in zinc between December 2015 and February 2016, and March and June 2006, at the 5% level but *not* during the two detected episodes for oil.[13] These results echo those for all six non-ferrous metals in Table IV in Figuerola-Ferretti et al. (2015) obtained on the basis of a different sample.

Comparison of our oil results with those obtained by Figuerola-Ferretti and McCrorie (2016) in their application of the PSY test to precious metals are less clear. Episodes of mild explosivity in gold and platinum front-month futures prices were identified in 2008 in the midst of the GFC, although in the two or three months *before* the detected episode in oil.[14] Etienne et al. (2015), using the PSY test with daily data, detected some mildly explosive episodes in grain futures prices in 2007 and 2008 but each lasted only a few days in length, and although there were two such episodes in soybeans in June 2008, there is nothing substantive in their results to tie them to oil prices. In the literature, only the sequential model proposed by Caballero et al. (2008) comes close to explaining why a mildly explosive episode in gold prices might have occurred before a similar episode in oil prices. In this model, episodes of bubble creation and collapse migrate across markets, from the U.S. subprime market, through certain commodities markets including precious metals and oil, to the U.S. bond market. We note en passant that it was the high point of oil prices, not of gold prices, that coincided with the low point of the U.S. dollar in July 2008.

We also applied the PSY test to the BCI, an index measure of a basket of commodities where the combined weighting of WTI and Brent prices is 15%.

[13] Only weekly data has been used in this paper. The imposition of the minimum duration condition, as shown in expression (5), means there is the potential for a temporary blip in prices to split the detection of one longer mildly explosive episode into two shorter episodes. It could well be that one, longer mildly explosive episode would have been reported in zinc from December 2005 to June 2006 had lower frequency data been used.

[14] Further assessment reported in Figuerola-Ferretti and McCrorie (2016) indicated the platinum price was supported by an electricity strike in early 2008 that affected mines in South Africa (which at that time were producing 85% of world production) but using gold lease rates – the interest that can be earned by lending the physical metal at various maturities – as a fundamental proxy variable, there was some evidence to support a short departure in the gold price from its fundamental value.

Interestingly, the GSADF statistic reported in Table 1C now only rejects the null of there being no mildly explosive episode at the 10% significance level. As reported in Table 3, the second-stage of the test using the BSADF statistic and minimum duration condition again dates the very same two mildly explosive episodes as before, but using a 10% critical value line. This higher significance level can be explained by combining the fact that the BCI dilutes the role of Brent and WTI prices with our lack of strong-enough evidence to reject a random walk null in the other commodities around the same time as in oil. We cannot therefore provide decisive evidence in support of the negative demand shock in December 2014 inferred by Baumeister and Kilian (2016), although we note that correlations between falling commodity prices at the time offer weaker corroboratory evidence. Our overall conclusion strongly supports the conclusion of Azeki and Blanchard (2014) and Baumeister and Kilian (2016) that factors beyond the recent behavior of commodity prices underlie recent oil prices, notably in the run-up to the high point in prices during the GFC, and during the recent oil price decline.

The above conclusion leads us now to consider a wider group of variables, including fundamental proxy variables on the demand and supply side.

5.2. Global economic activity

The traditional proxy variable for global economic activity, the OECD measure of global industrial production, was discontinued in 2012 and so, rather than constructing later data points by aggregating country-specific data and then creating a weekly series by interpolation, we follow the general standpoint of Killian (2009, p. 1005) who constructed a bespoke index for the analysis of oil prices based on dry cargo single ocean freight rates. This type of measure anticipates and captures demand shifts for industrial commodities in global markets. Here, we use the Baltic Dry Index (BDI) to proxy global economic activity, a similar such index that is available on the Bloomberg International Database.[15] The GSADF statistics of both the BDI reported in Table 1E and the ratio of WTI nominal prices to the BDI reported in Table 1F both reject the null of there being no mildly explosive episode at the 1% significance level.

[15] The BDI is an economic indicator issued daily by the London-based Baltic Exchange, and provides an assessment of the price of moving raw materials by sea. It measures the demand for shipping capacity versus the supply of dry bulk carriers and indirectly measures the global supply and demand for the commodities shipped aboard dry bulk carriers.

[Tables 1E and 1F around here]

The second-stage of the test reported in Table 4 dates two mildly explosive periods in the BTI index itself, between mid-April until the end of May 2007 and end-August to end-December 2007,[16] while the test based on the WTI/BDI ratio as reported in Table 5 dates a two-month episode in 2005, and a six-week-or-so episode during October and November 2008 when the BDI is falling.

[Tables 4 and 5 around here]

Evidence obtained on the basis of the PSY test to support global economic activity as having been a key driver of the post-June 2014 oil price decline is therefore non-existent; however, the extent to which it might explain the behavior of oil prices during the 2007-08 price spike depends on the extent to which the BDI can be viewed as a leading indicator. The supply of cargo ships is generally tight and inelastic and so marginal increases in the demand for shipping capacity can quickly push up the BDI. Because dry bulk primarily consists of materials that function as raw material inputs to the production of intermediate or final goods, the index is indeed an indicator of future growth and production. The lag time involved is plausibly of a length similar to the time difference between the episodes detected in the BDI and in nominal oil prices, providing support for global economic activity – a traditional oil fundamental – as having been an important driver in the run-up in oil prices during the 2007-08 oil price spike. Our results therefore provide support for the explanations of the run-up in oil prices in 2007-08 advanced by Hamilton (2009), Kilian (2009) and Kilian and Hicks (2013) as being due to a surge in demand in the world economy.[17] The later, negative mildly explosive episode seen in the ratio of WTI nominal prices to the BDI (when nominal prices are falling) indicates that in the immediate aftermath of the high point of the GFC in mid-September, when the global financial services firm Lehman Brothers Holdings Inc. filed for bankruptcy, WTI oil prices fell faster than can be explained by global economic activity alone. Finally, and not insignificantly, the BSADF sequence of the BDI level series, as shown in Figure 6,

[16] For the same reason as in Footnote 13, this evidence may be consistent with one, longer mildly explosive episode being detected from mid-April to mid-Dececmber 2007 on the basis of lower frequency data.

[17] Indeed, on the basis of a forecasting approach, Kilian and Hicks (2013) showed that the run-up in real oil prices over the period from mid-2003 to mid-2008 was attributable to repeated demand shocks. Only during the 2007-08 period, however, was the run-up significant enough for the null of random walk behavior to be rejected.

indicates that in 2014 the BDI was falling, consistent with the evidence provided by Bauermeister and Kilian (2016) that lower global demand played *some* role in the post-June 2014 price decline.

[Figure 6 around here]

The BSADF sequence, being always below the 5% critical value line around this time, indicates the BDI was not, however, anticipatory of any later shift to a mildly explosive regime (in contrast to 2007-08), leading to the conclusion that demand factors were not decisive in the post-June 2014 oil price decline. This offers prima facie evidence that the drivers of the two detected mildly explosive episodes over the last decade were different in character.

5.3. Inventories

Deaton and Laroque (1992) argue that the classical fundamental in a commodity price setting is the stock or inventory variable. Caspi et al. (2016) and Fantazzini (2016) follow PY (2011) in applying the PWY/PSY strategy to crude oil nominal spot prices deflated by the traditional above-ground inventory supply proxy of U.S. WTI stocks. Using the prototype test by PWY (2011), PY (2011) detected a mildly explosive period in the ratio between March and July 2008 using monthly data. Here, on the basis of weekly data, we apply the PSY test – which PSY (2015b) shows supersedes the PWY (2011) test – to the U.S. stocks themselves and then to the ratio of nominal WTI front-month futures prices to the same. The GSADF statistic for stocks reported in Table 1E shows the null hypothesis of no mildly explosive period is not rejected at the 10% significance level. We do not therefore proceed to the second, date-stamping stage of the PSY test. Consistent with this evidence, the GSADF statistic of the ratio of WTI nominal prices to U.S. stocks as reported in Table 1F is significant at the 1% level and, as can be seen in Figure 7 and is reported in Table 5, a mildly explosive period is dated between mid-May and mid-June 2008 and between mid-November 2014 and mid-March 2015.

[Figure 7 around here]

These are coincident with the same such periods in WTI nominal prices alone, as reported earlier, but begins slightly later than in the detected period by PY (2011) based on monthly spot price data.

For the ratio of monthly data on WTI nominal spot prices to U.S. stocks, Caspi et al. (2016, Table 2) report a longer mildly explosive period at the end of their sample, between July 2007 and September 2008, than PY (2001) on the basis of the PWY (2011) test.[18] For the same ratio using weekly spot price data, Fantazzini (2016, Table 3) dates a positive episode between October 2007 and September 2008, and a negative episode between mid-December 2014 and mid-March 2015, the first encompassing the first episode and the second coincident with the second episode detected using WTI nominal prices alone.

Taking the evidence altogether, and unlike on the demand side, our fundamental proxy does *not* offer prima facie evidence that above-ground inventories drove oil prices either during the run-up in prices in 2008 or the post-2014 price decline. Our results therefore corroborate the conclusion of Kilian and Lee (2014) that above-ground inventories were not decisive in explaining the 2007-08 price spike. Their latter conclusion was reached using two inventory measures and, on the basis of the underlying Kilian and Murphy (2014) structural model they used, which embodies the theory of storage, the effect of speculation was quantified as raising the real oil price by between \$5 and \$14 between March and July 2008. Taken at face value, their estimates provide a counterfactual that is insufficient to reverse the conclusions of the PSY test.

The traditional argument that U.S. stocks proxy above-ground inventory supply is based on their being inventories in the country in the world economy that has the largest consumption of oil. It neglects other important elements of world supply, notably the role of OPEC producers. Kilian (2016) examines the consequences of the recent emergence of U.S. hydraulic fracturing (or "fracking") on Arab crude oil producers and estimates that the price was around \$10 lower in mid-2014, falling to about \$5 lower in mid-2015. Taken at face value, the estimates provide a counterfactual under which its marginal impact can be assessed using the PSY test by adjusting prices upwards, but is not enough to prevent the declaration of a mildly explosive episode at the 10% significance level. This evidence suggests that while U.S. fracking played a role in the falling oil price from June 2014, it was not decisive in contributing to the rejection of random walk null behavior in oil prices in late-2014.

[18] The samples of PY (2011) and Caspi et al. (2016) do not encompass the post-June 2014 price decline.

Where our work can offers some insight and complementarity to the forecasting-based approaches of Baumeister and Kilian (2016) and Kilian (2016) is in the *date-stamping* of short, mildly explosive regime shifts. The results in this paper for nominal and real futures prices and similar results reported by Fantazzini (2016) for spot prices point strongly and robustly to a move towards (negative) mild explosivity in late-2014, at the point or in the immediate aftermath of the OPEC meeting on 27 November when OPEC agreed to maintain their production ceiling. On the basis of an observed negative forecast error in their proxy for global real activity, Baumeister and Kilian (2016) instead inferred that there was a *negative demand shock* in December 2014, which they attributed to an unexpected weakening of the global economy. As discussed above, we were unable to demonstrate similarly significant coincident price falls in other commodities around this time.[19] Baumeister and Kilian's analysis suggests there were two major errors in forecasting the price of oil in the second-half of 2014: one in July that accounted for a decline by \$9 per barrel; and the other in December that accounted for a decline by \$13. Azeki and Blanchard (2014) and Baffes et al. (2015) have argued strongly that the OPEC meeting in November 2014 represented a watershed moment: with OPEC agreeing to maintain their production ceiling at 30 million barrels per day, it signalled a change in their policy from targeting an oil price band to maintaining market share. This would represent a *supply shock* where, rather than agreeing on production cuts to maintain prices, OPEC signalled a fundamental change in expectations about future global oil supply and prices (with clear implications for the profitability of emergent and continuing U.S. fracking activity). Our results for both real and nominal, Brent and WTI benchmark crude oil prices, based on a formal statistical test, date a short (negative) mildly explosive episode as beginning around the same time of the OPEC meeting, or in its immediate aftermath. While our reduced-form approach can never offer prima facie evidence *to prove* such a linkage, it can produce significant empirical evidence to be explained. The supporting evidence presented on the behavior of other commodities

[19] The discrepancy found by Baumeister and Kilian between their prediction at June 2014 of a \$27 decline in oil prices and their attribution of only around \$10 to a slowdown in economic activity is consistent, as the authors note, with the disparity later seen in the larger decline in crude oil prices from around then compared with other commodities. The issue here, however, is that if there had been a negative demand shock in December 2014 that was decisive in moving oil prices away from random walk behavior, its impact should have been observed across other commodities, or otherwise it would have to be derived from at least one other source.

around the same time, and on the relatively small contribution of U.S. fracking to oil prices, consolidates a supply-side explanation. And from today's standpoint, a negative forecast error in global economic activity around December 2014 would not automatically be seen as evidence of a significant, negative demand shock given the weak transmission mechanism from falling and low oil prices to economic activity that has recently been seen.

Comparison of our results with Fantazzini's suggest the possibility of there having been a discrepancy between spot and futures prices during the 2007-08 spike, at a point when the GFC had begun its course. An interesting element of the Kilian and Murphy (2014) structural model is that it does not contain a futures price as a variable. As discussed above, this does not rule out speculation: it is defined as an inventory demand shock in the spot market that follows changes in expected fundamentals. Perfect arbitrage would then imply full pass-through from speculation in the futures market to the spot market via changing inventories.[20] There are plausible reasons recently propounded in the literature, however, to suggest that perfect arbitrage might not hold in oil or other commodities markets, with the consequence that inventory data may not be sufficient to capture speculative activity within spot and futures markets. Lombardy and Van Robays (2011) provide a structural model along the lines of the Kilian and Murphy (2014) model but with incomplete pass-through. Acharya et al. (2013) propose a model of commodities in which the interaction of commodity producers averse to price fluctuations and capital-constrained speculators induces a link between a financial friction in the futures market and spot prices. Sockin and Xiong (2015) and Basak and Pavlova (2016) similarly provide models of financialization that impact upon both commodity futures and spot prices. In a paper that has had a significant influence on the oil price literature, Singleton (2014) demonstrates that money flows associated with index

[20] Hamilton (2009) also points out that if the short-run price elasticity of gasoline demand is zero, there is scope for speculation to drive up the real price of oil without crude oil inventories being affected. This condition is not testable within the PSY mild explosivity framework. We therefore simply report that in their structural model, Kilian and Murphy (2014) related the gasoline demand elasticity to the short-run elasticity of oil demand and provided a posterior median estimate of the latter of -0.26, and showed there was an 84% probability that this value was below -0.09 (*ibid*., Table II). In a different model, Baumeister and Peersman (2013) reported evidence to suggest the short-run elasticity of oil demand was time varying but that the boundary of their 95% posterior credible set was less than zero throughout their sample (*ibid*., Figure 4). Taken at face value, both sets of results suggest that the case of a zero price elasticity of demand is an unlikely explanation for any run-up in crude oil prices seen before their high point in July 2008.

investors help predict changes in oil futures prices, although recent work by Irwin and Sanders (2014) and Hamilton and Wu (2015) may mollify this conclusion.

The above motivates using the PSY strategy to assess whether the evidence supports there having been financial speculation in the oil futures market during one or both of the detected mildly explosive episodes, either from different sources or in excess of that implied by normal backwardation or the classical theory of storage. Recent work by Robe and Wallen (2016) has tied measures of oil price volatility and economy-wide financial conditions as captured by the equity VIX to the oil futures market. Indeed, Cheng et al. (2015) propose a transmission mechanism from financial traders to futures markets whereby a "convective risk flow" from speculators, who sell in response to rises in risk as prices fall, to hedgers, who operate on the other side of the market, reallocates risk from groups less able to bear the risk to groups which can. Empirical evidence, where in their framework the VIX acts as a proxy for the risk appetite of financial traders and funding constraints, corroborates commodity index traders having an impact on commodity futures prices, and thereby a role for financialization. Accordingly, we shall assess whether there is evidence that the VIX had a role to play in recent oil price movements, in particular during the two detected mildly explosive episodes. Oil has its own measure of volatility, the CBOE Crude Oil Volatility Index (OVX), which we shall use to establish a benchmark.

5.4. OVX

The OVX was launched by CBOE as an oil-related volatility index on 3 June 2008 during the unfolding of the GFC. Data are available at the CBOE official website. It measures oil market uncertainty through options taken on crude oil prices, specifically the market's expectation of the 30-day volatility of WTI nominal prices applying the standard CBOE volatility index methodology (discussed briefly in Section 5.5 below) to options on the United States Oil Fund.[21] Implied volatility indices are generally derived from *option prices*, and reflect market expectations on the *future volatility* over the lifetime of the option. Such measures, therefore, allow us the possibility of

[21] The United States Oil Fund (USO) is an exchange-traded security designed to track changes in crude oil prices. By holding near-term futures contracts and cash, the performance of the Fund is intended to reflect, as closely as possible, the spot price of WTI crude oil, less USO trading expenses.

incorporating market-based expectations in a way that traditional fundamental variables do not (although this aspect is less important in our approach than it is in a forecasting-based approach). For the interpretation of our results, we need to rely on some linkage between the volatility index and futures prices such as discussed in the papers cited above.

Figure 8A displays WTI benchmark future prices and the OVX.

[Figure 8A about here]

The negative relationship between WTI nominal prices and the OVX seen after the fall in the oil price in 2008 is notable right through to the end of the sample. There are spikes when volatility is high, around or just after the second (negative) mildly explosive episode detected in WTI nominal prices, and for a second time in 2015. To apply the PSY test, given the launch date, we need to alter some of the parameters chosen above. We continue to use the PSY (2015a) recommended rule-based calculation of the initial sample fraction, r_0, but use a sample only from January 2009, beyond the launch date, to ensure the chosen sample is clear of the oil price shock in 2007-08. Table 1D indicates that the GSADF statistic is significant at the 1% significance level, meaning that the null of there being no mildly explosive period is rejected; however as reported in Table 6, the second-stage of the test does not date a mildly explosive period.

[Table 6 around here]

Figure 8B shows that the BSADF sequence *does* cross the 5% critical value line at almost exactly the point of the OPEC meeting, but the crossing is not of sufficient duration for a mildly explosive episode to be declared. This may indicate the OVX, as a forward-looking variable based on expectations, adjusted once-and-for all in the aftermath of the meeting.

[Figure 8B around here]

On the basis of the ratio of WTI nominal prices to the OVX, the null of no mildly explosive episode is not rejected at the 10% significance level, as is reported in Table 1F, and so we do not proceed to the date-stamping stage of the test.

[Table 1F around here]

Taken at face value, this result is not surprising; for we would expect the OVX to act as an omnibus measure that reflects the myriad of influences on the oil price.

For precisely this reason, however, the results by themselves are uninformative in terms of identifying specific drivers. We will therefore treat the OVX results as a benchmark case against which to compare results obtained using the VIX.

5.5. VIX

The CBOE Volatility Index (VIX) is the implied volatility of Standard and Poor's S&P 500 equity index over the next 30-day period and is the premier benchmark for U.S. stock market volatility. It is the square root of the risk-neutral expectation of the variance of the S&P 500 index over this period, computed on the basis of a weighted average of prices for a range of options over the index, and is quoted as a percentage. Figure 9A shows that there are two spikes in the VIX during the sample period: the first occurs during the later stages of the 2007-08 oil price spike, *after* the mildly explosive episode detected during the run-up in prices;[22] and in mid-2011 around the time of Standard and Poor's downgrade of U.S. sovereign debt.

[Figure 9A around here]

Little upward or downward movement is seen in the VIX during the recent oil price decline and there is only a small uptick around the detected mildly explosive episode, in marked contrast to the OVX series. The GSADF statistic rejects at the 10% significance level the null of no mildly explosive episode, as reported in Table 1D, but no such episode is dated in the second stage of the test. Figure 9A shows that the BSADF sequence cuts the 5% critical value line during the 2007-08 oil price spike but not for a sufficient time period for a mildly explosive episode to be declared. The ratio of WTI prices to the VIX does not reject the null of no mildly explosive episode at the 10% significance level, as reported in Table 1F, and so we do not proceed to the date-stamping stage of the test. Inspection of the BSADF sequence in Figure 9B shows that this rejection occurs *without* the sequence cutting the 5% critical value line, which points to the need for further analysis.

[Figure 9B around here]

[22] That the spike in the VIX occurred during the *collapse-phase* of the 2007-08 spike is not unexpected but, as Table 6 indicates, no mildly explosive episode was detected in the VIX levels series during this period.

Because the ratio of the WTI nominal price series to the VIX is the same as the ratio of the WTI real prices to the VIX deflated by U.S. CPI, we consider the BSADF sequence of the ratio of the VIX to the U.S. CPI series, shown in Figure 9C, to try to gain more insight into the first-stage non-rejection. Nothing substantively different emerges compared with the VIX levels case: the BSADF sequence again cuts the 5% critical value line in late 2008 – in the collapse phase of the 2007-08 oil price spike – and there is no spike in late-2014.

[Figure 9C around here]

Running the PSY test on the ratio of WTI nominal prices to the S&P 500 index itself indicates that the null of no mildly explosive episode is rejected at the 1% significance level and, as reported in Table 7, the same mildly explosive episodes are detected as using WTI nominal prices alone even given the evidence reported by Silvennoinen and Thorp (2013) that the two variables have become increasingly correlated under financialization.

[Table 7 around here]

This offers prima facie evidence to suggest there were factors beyond stock prices driving both departures from random walk behavior. At face value, the failure of the ratio of WTI prices to the VIX to detect the departures seen in the WTI series offers corroborative evidence consistent with Cheng et al. (2015) and Robe and Wallen (2016) who tie WTI futures prices to the VIX. The supporting evidence, however, suggests otherwise. Firstly, the rise in the VIX occurs *after* the high point of crude oil prices in 2008 and therefore after the dated mildly explosive episode; there is no spike in the VIX in late-2014 in contrast to that seen in the oil-specific OVX; and the results are the same for both nominal and real prices. On the basis of the PSY testing strategy, it is therefore difficult to support the VIX having played a role in driving the level of oil prices around 2007-08 or during the post-2014 oil price decline. Inspection of the BSADF sequence in Figure 9B suggests that the VIX series is considerably more volatile than the oil price series, and this volatility translates into the ratio of prices. It may be that the PSY test, which derives its statistical power from its concentration of focus on the autoregressive parameter, and relies on a minimum duration condition in the date-stamping stage, is not conducive towards assessing the actual role of the VIX. Market expectations of oil price volatility should be higher during periods of financial stress and so we would expect the VIX to be a good predictor of the *volatility* in oil prices, but not necessarily their levels as is tested for here.

5.6. Non-commercial positions and speculation

The evidence presented above strongly points to the mildly explosive episode detected in the run-up in oil prices during the 2007-08 spike as having been driven by a demand-side fundamental: a mildly explosive episode in the BDI leading indicator proxy for global economic activity preceded the same in both WTI and Brent benchmark oil prices in mid-2008. Alongside this, on the basis of a counterfactual from Kilian and Lee's (2014) approach, which embodied speculation within the classical theory of storage, above-ground inventories were seen not to have been decisive. Sockin and Xiong (2015) argue, however, that the period coinciding with our detected mildly explosive episode was characterized by a global economy whose developed economies had begun to show signs of weakness and that, even if growth in the emerging economies remained strong, a price rise of 40% between January and July 2008 was not justified by global economic activity alone.

Using the PSY test applied to Commitments of Traders (CoT) position data published by the U.S. Commodity Futures Trading Commission (CFTC), we can assess whether and to what extent there is a contribution from financialization that may have amplified the effects of this rising demand. This data breaks down the overall open interest between the positions of commercials ("hedgers") and non-commericials ("speculators").[23] Büyükşahin and Harris (2011), Büyükşahin and Robe (2014) and others employ Working's *T*-index as a measure of "excess speculation": it calculates the amount of speculation in excess of what is minimally necessary to meet short and long hedging demand. Write long and short commercial (hedge) positions as H_L and H_S and long and short non-commercial (speculative) positions as S_L and S_S respectively. Working's *T*-index is then defined by

$$T = 1 + \frac{1(H_S \geq H_L) \cdot S_s + 1(H_S < H_L) \cdot S_L}{H_L + H_S} \tag{8}$$

[23] Commercial positions are associated with producers and consumers of the commodity. Non-commercial positions reflect the activities of financial traders for investment purposes.

If T = 1, the level of non-commercial activity is just sufficient to be available as counterparties for the commercial imbalance. Any excess over unity implies that speculators are acting as counterparties for each other. Here, we provide a *direct test* for excess speculation by applying the PSY test to a time series of Working's T-index.

The GSADF statistic as reported in Table 1G shows that the null of no mildly explosive period is rejected at the 1% significance level.

[Table 1G around here]

The BSADF sequence as shown in Figure 10 indicates than at no point prior to the high point of nominal and real oil prices in 2008 did the sequence cut the 5% critical value line.

[Figure 10 around here]

There is therefore no evidence provided by the PSY test to suggest that excess speculation played a role either in the run-up in oil prices from 2003 or during the 2007-08 spike seen during the GFC. The index rises from late-2014 onward but no mildly explosive episode coinciding with the detected periods in nominal and real prices is reported in Table 8.

[Table 8 around here]

Such is the rise in the index that an episode is dated from end-April 2015 to the beginning of January 2016. This may indicate that there have been changes in the oil market that have occurred during the recent oil price decline, although any such consideration is outside the scope of the current paper. The main result corroborates the conclusion of Sanders and Irwin (2014) and Hamilton and Wu (2015) that excess speculation through financialization did not drive either the 2007-08 oil price spike or the recent post-June price decline.

6. Conclusion

This paper provides an analysis of crude oil prices over the last decade through the lens of the recently proposed PSY (2015a) mildly explosive/bubbles technology. The mere observation that a time series is upward or downward trending over a period, such as is informally observed in the so-called commodities boom and bust in 2007-08, or in oil price behavior between 2003-08 or post-June 2014, does not constitute prima facie evidence for the rejection of random walk or martingale behavior. Following the pioneering paper by PWY (2011), we use the PSY test robustly to detect departures from random walk behavior in the direction of mildly explosive alternatives as the basis of a strategy to assess the continuing controversies over recent oil price drivers, notably during the 2007-08 oil price spike seen during the GFC and the recent post-June 2014 decline. In particular, we consider the role of traditional demand- and supply-side fundamentals and the impact of financialization and possible excess speculation as drivers of oil prices. Our reduced-form strategy creates an evidence base against which the implications of structural models of the oil market can be assessed.

Our specific conclusions are as follows:

1. A statistically significant rise beyond that consistent with random walk behavior is detected in both Brent and WTI real and nominal benchmark crude oil futures prices between mid-May and mid-July 2008, in the midst of the GFC. A similar, statistically significant fall is seen in prices, beginning in late-November or early-December 2014 and lasting for two or so months.
2. The results are robust to a potentially changing dollar exchange rate, meaning that the depreciating dollar numeraire in early-2008 and the appreciating dollar in 2014 are not decisive in the two, detected departures from random walk behavior.
3. All of the results are robust to possible non-stationary volatility in the shocks driving the oil prices, except possibly those pertaining to the first episode when WTI real prices are used.
4. The behavior of other commodities was not decisive in explaining oil prices during the 2007-08 spike or the post-June 2014 decline.
5. Evidence is provided for the rise in oil prices in 2007-08 being driven by a demand-side fundamental: a mildly explosive episode was detected in the BDI leading indicator proxy for global economic activity that immediately preceded the same in both WTI and Brent benchmark oil prices in mid-2008. While the BDI rose during the 2003-08 oil price rise, only prior to the 2007-08 oil price spike did it rise to the extent that a null of random walk behavior could be rejected.
6. On the basis of a counterfactual obtained from Kilian and Lee (2014), whose method embodied speculation within the classical theory of storage, above-ground

inventories were found not to have been decisive in the run-up in prices during the 2007-08 spike or during the post-2014 price decline.

7. A declining BDI in 2014 indicates that part of the oil price decline in the latter half of the year can be explained by lower economic activity. A lack of a mildly explosive episode in BDI levels in 2014, however, indicates that demand factors were not decisive in explaining the later, detected shift to a mildly explosive regime. The evidence therefore indicates that the drivers of the 2007-08 spike and post-June 2014 price decline were different.
8. On the basis of a counterfactual provided by Kilian (2016), we found that U.S. fracking was not decisive in the rejection of random walk behavior in oil prices in late-2014.
9. The beginning of a detected (negative) mildly explosive episode during the post-June 2014 price decline is date-stamped to around or in the immediate aftermath of an OPEC meeting in late-November 2014 at which OPEC production strategy notably changed.
10. In spite of recent work tying crude oil futures to the equity VIX, we find no hard evidence to suggest that the VIX influenced oil price levels during the sample period.
11. On the basis of Working's *T*-index constructed using CFTC position data, there is no evidence to suggest that excess speculation was a contributory factor to either the 2007-08 oil price spike or the post-2014 oil price decline.

The argument outlining how the above evidence, obtained on the basis of formal statistical testing, brings clarity to many of the recent controversies in the oil price literature is made in the main body of this paper.

Acknowledgements

We thank Ramon Bermejo for assistance in data collection, and Christopher Gilbert for use of his GAUSS programs to compute wild-bootstrap-adjusted PSY critical values. We also thank participants at the Energy, Commodities and Finance (ECOMFIN) Conference, ESSEC Business School, June 2016, and the 4th Rimini Workshop in Time Series Econometrics, June 2016, for helpful comments. Figuerola-Ferretti thanks the Spanish Ministry of Education and Science for support under grants MICINN ECO2010-19357, ECO2012-36559 and ECO2013-46395, and McCrorie, The Carnegie Trust for the Universities of Scotland under grant no. 31935.

Bibliography

Acharya, V.V., Lochstoer, L.A., Ramadorai, T., 2013. Limits to arbitrage and hedging: evidence from commodity markets. *Journal of Financial Economics* 109(2), 441-465.

Alquist, R., Kilian, L., 2010. What do we learn from the price of crude oil futures? *Journal of Applied Econometrics* 25, 539-573.

Alquist, R., Kilian, L., Vigfusson, R.J., 2013. Forecasting the price of oil. In. G. Elliott and A. Timmermann (eds) *Handbook of Economic Forecasting*, 2, North-Holland, pp. 427-507.

Arezki, R., Blanchard, O.J., 2014. The 2014 oil price slump: seven key questions. *IMFDirect – The IMF Blog*, December 22.

Baffes, J., Ayhan Kose, M., Ohnsorge, F., Stocker, M., 2015. The great plunge in oil prices: causes, consequences and policy responses. Policy Research Note, World Bank Group. Available at: http://documents.worldbank.org/curated/en/726831468180852545.

Basak, S., Pavlova, A., 2016. A model of financialization of commodities. Forthcoming, *Journal of Finance*. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2201600.

Baumeister, C., Kilian, L., 2016. Understanding the decline in the price of oil since June 2014. *Journal of the Association of Environmental and Resource Economists* 3(1), 131-158.

Baumeister, C., Peersman, G., 2013. The role of time-varying price elasticities in accounting for volatility changes in the crude oil market. *Journal of Applied Econometrics* 28, 1087-1109.

Büyükşahin, B., Harris, J.H., 2011. Do speculators drive crude oil prices? *The Energy Journal* 32(2), 167-201.

Büyükşahin, B., Lee, T.K., Moser, J.T., Robe, M.A., 2013. Physical markets, paper markets and the WTI-Brent spread. *Energy Journal* 34(3), 129-151.

Büyükşahin, B., Robe, M.A., 2014. Speculators, commodities and cross-market linkages. *Journal of International Money and Finance* 42, 38-70.

Caballero, R.J., Fahri, E., Gourinchas, P.-O, 2008. Financial crash, commodity prices and global imbalances. *Brookings Papers on Economic Activity*, Fall, 1-55.

Caspi, I., Katzke, N., Gupta, R., 2016. Date stamping historical periods of oil price explosivity: 1876-2014. Forthcoming, *Energy Economics*. (DOI: http://dx.doi.org/10.1016/j.eneco.2015.03.029)

Cheng, I.-H., Xiong, W., 2014. Financialization of commodity markets. *Annual Review of Financial Economics* 6(1), 419-441.

Cheng, I.-H., Kirilenko, A., Xiong, W., 2015. Convective risk flows in commodity futures markets. *Review of Finance* 19, 1733-1781.

Deaton, A.S., Laroque, G., 1992. On the behaviour of commodity prices. *Review of Economic Studies* 59, 1-23.

Deaton, A.S., Laroque, G., 1996. Competitive storage and commodity price dynamics. *Journal of Political Economy* 104, 896-923.

EIA, 2016. What drives crude oil prices? An analysis of 7 factors that influence oil markets, with chart data updated monthly and quarterly. Posted 12 July 2016 at: www.eia.gov/finance/markets/reports_presentations/eia_what_drives_crude_oil_prices.pdf.

Etienne, X.L., Irwin, S.H., Garcia, P., 2015. Price explosiveness, speculation and grain futures prices. *American Journal of Agricultural Economics* 97, 65-87.
Fantazzini, D., 2016. The oil price crash 2014/15: Was there a (negative) financial bubble? *Energy Policy* 96, 383-396.
Fattouh, B., Kilian, L., Mahadeva, L., 2013. The role of speculation in oil markets: what have we learned so far? *Energy Journal* 34(3), 7-33.
Figuerola-Ferretti, I., Gilbert, C.L., and McCrorie, J.R., 2015. Testing for mild explosivity and bubbles in LME non-ferrous metals prices. *Journal of Time Series Analysis* 36, 763-782.
Figuerola-Ferretti, I., McCrorie, J.R., 2016. The shine of precious metals around the Global Financial Crisis. *Journal of Empirical Finance* 38, Part B, 717-738.
Greenaway-McGrevy, R., Phillips, P.C.B., 2016. Hot property in New Zealand: empirical evidence of housing bubbles in the metropolitan centres. *New Zealand Economic Papers,* 50(1), 88-113.
Hamilton, J.D., 2008. Oil and the macroeconomy. In Durlauf, S.N. & L.E. Blume (eds.) *The New Palgrave Dictionary of Economics*, second edition.
Hamilton, J.D., 2009. Causes and consequences of the oil shock of 2007-08 (and discussion). *Brookings Papers on Economic Activity*, Spring, 215-283.
Hamilton, J.D., Wu, J.C., 2015. Effects of index-fund investing on commodity futures prices. *International Economic Review* 56(1), 187-205.
Harvey, D.I., Leybourne, S.J., Sollis, R., Taylor, A.M.R., 2016. Tests for explosive financial bubbles in the presence of non-stationary volatility. *Journal of Empirical Finance* 38, Part B, 548-574.
Kilian, L., 2009. Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. *American Economic Review* 99, 1053-1069.
Kilian, L., 2016. The impact of the fracking boom on Arab oil producers. Preprint available at: www-personal.umich.edu/~lkilian/erf_egypt.pdf.
Kilian, L., Hicks, B., 2013. Did unexpectedly strong economic growth cause the oil price shock of 2003-08? *Journal of Forecasting* 32, 385-394.
Kilian, L., Lee, T.K., 2014. Quantifying the speculative component in the real price of oil: the role of global oil inventories. *Journal of International Money and Finance* 42, 71-87.
Kilian, L., Murphy, D.P., 2014. The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics* 29(3), 454-478.
Lombardi, M.J., Van Robays, I., 2011. Do financial investors destabilize the oil price? Working Paper No. 1346, European Central Bank.
Morana, C., 2013. Oil price dynamics, macro-finance interactions and the role of financial speculation. *Journal of Banking and Finance* 37, 206-226.
Phillips, P.C.B., Magdalinos, T., 2007. Limit theory for moderate deviations from unity. *Journal of Econometrics* 136 (1), 115-130.
Phillips, P.C.B., Shi, S.-P., 2014. Financial bubble implosion. Cowles Foundation Discussion Paper No. 1967.
Phillips, P.C.B., Shi, S.-P., Yu, J., 2013. Testing for multiple bubbles: historical episodes of exuberance and collapse in the S&P 500. Cowles Foundation Discussion Paper No. 1914, Yale University.
Phillips, P.C.B., Shi, S.-P., Yu, J., 2015a. Testing for multiple bubbles: historical episodes of exuberance and collapse in the S&P 500. *International Economic Review* 56 (4), 1043-1077. **[PSYa]**

Phillips, P.C.B., Shi, S.-P., Yu, J., 2015b. Testing for multiple bubbles: limit theory of real time detectors. *International Economic Review* 56 (4), 1079-1133. **[PSYb]**

Phillips, P.C.B., Wu, Y., Yu, J., 2011. Explosive behavior in the 1990s Nasdaq: when did exuberance escalate asset values? *International Economic Review* 52 (1), 210-226. **[PWY]**

Phillips, P.C.B., Yu, J., 2011. Dating the timeline of financial bubbles during the subprime crisis. *Quantitative Economics* 2 (3), 455-491. **[PY]**

Robe, M.A., Wallen, J., 2016. Fundamentals, derivatives market information and oil price volatility. *Journal of Futures Markets* 36(4), 317-344.

Sanders, D.R., Irwin, S.H., 2014. Energy futures prices and commodity index investment: new evidence from firm-level position data. *Energy Economics* 46, S57-S68.

Shi, S., Arora, V., 2012. An application of models of speculative behavior to oil prices. *Economics Letters* 115, 469-472.

Silvennoinen, A., Thorp, S., 2013. Financialization, crisis and commodity correlation dynamics. *International Financial Markets, Institutions and Money* 24, 42-65.

Singleton, K.J., 2014. Investor flows and the 2008 boom/bust in oil prices. *Management Science* 60(2), 300-318.

Sockin, M., Xiong, W., 2015. Informational frictions and commodity markets. *Journal of Finance* LXX (5), 2063-2098.

Tsvetanov, D., Coakley, J., Kellard, N., 2016. Bubbling over! The behavior of oil futures along the yield curve. *Journal of Empirical Finance* 38, Part B, 516-533.

Working, H. (1960). Speculation on hedging markets. *Stanford University Food Research Institute Studies* 1, 185-220.

Appendix A: Tables

Table 1A
GSADF Test Statistics
ICE Brent Crude nominal front-month futures
2003 - 2016 (weekly data)

	Brent	**Brent/CPI**	**Brent/SDR**
GSADF	4.382	4.349	4.377

This table reports the GSADF statistic for ICE Brent Crude nominal front-month futures prices and the GSADF statistics for this series deflated by a U.S. CPI series interpolated weekly by and an SDR factor to control for the changing dollar numeraire. The initial window for recursive estimation is 47 weeks. The ADF lag is chosen to minimize the BIC over every subsample with maximum lag length set at 5 weeks. Standard PSY Critical values: 2.069 (10%), 2.282 (5%) and 2.664 (1%).

Table 1B
GSADF Test Statistics
NYMEX WTI Crude nominal front-month futures
2003-2016 (weekly data)

	WTI	**WTI/CPI**	**WTI/SDR**
GSADF	4.935	4.898	4.848

This table reports the GSADF statistics for NYMEX WTI Crude nominal front-month futures month prices and the GSADF statistics for this series adjusted by U.S. CPI and by an SDR factor to control for the changing dollar numeraire. Standard PSY Critical values: 2.069 (10%), 2.282 (5%) and 2.664 (1%).

Table 1C
GSADF Test Statistics
LME Copper and Zinc nominal three-month futures
& Bloomberg Commodity Index
2003-2016 (weekly data)

	Cu	**Zi**	**BCI**
GSADF	7.963	6.528	2.187

This table reports GSADF statistics for LME copper and zinc nominal three-month futures prices and the Bloomberg Commodity Index. Standard PSY Critical values: 2.069 (10%), 2.282 (5%) and 2.664 (1%).

Table 1D
GSADF Test Statistic
Fundamental measures (in levels)
2003-2016 (weekly data)

	BDI	Stocks	VIX	OVX *
GSADF	4.315	1.882	3.238	3.056

This table reports GSADF statistics for the Baltic Dry Index (BDI); U.S. end-of-week stocks (EIA data); and the VIX and OVX volatility index measures. Standard PSY Critical values: 2.069 (10%), 2.282 (5%) and 2.664 (1%).

*Note that for the OVX measure, the sample runs from 2009 to 2016 and the initial value window, also chosen using PSY's rule, is 35 weeks.

Table 1E
GSADF Test Statistics
WTI Crude Nominal Front-Month Prices deflated by fundamental proxy variables and financial variables
2003-2016 (weekly data)

GSADF	WTI/Baltic Dry Index	WTI/inventories	WTI/S&P500	WTI/HY
	4.850	4.930	4.703	2.101

This table reports GSADF statistics for WTI front month NYMEX futures weighted by (a) the Baltic Dry Index; (b) U.S. end-of-week (WTI) crude oil stocks (Source: EIA); (c) S&P 500 index; (d) Barclays US Corporate High Yield average. Standard PSY Critical values: 2.069 (10%), 2.282 (5%) and 2.664 (1%).

Table 1F
GSADF Test Statistics
Nominal Crude Oil Prices deflated by Volatility Measures
2003-2016 VIX deflated (weekly data)
2007-2016 VOX deflated (weekly data)

	WTI/OVX	WTI/VIX	Brent/VIX
GSADF	0.992	0.525	1.072

This table reports GSADF statistics for WTI Crude nominal front-month prices deflated by the OVX index and GSADF statistics for both WTI Crude and Brent Crude nominal front-month futures prices deflated by the CBOE VIX index. Standard PSY Critical values: 2.069 (10%), 2.282 (5%) and 2.664 (1%).

Table 1G
GSADF Test Statistic
Working's *T* index
2003-2016 (CFTC weekly data, reported Tuesdays)

	CFTC
GSADF	2.823

This table reports the GSADF statistic for Working's T index calculated using data on Commercial Long, Commercial Short, Non-Commerical Long and Non-Commercial Short positions. The initial window for recursive estimation is 47 weeks. The ADF lag is chosen to minimize the BIC over every subsample with the maximum lag length set at 5 weeks. Standard PSY Critical values: 2.069 (10%), 2.282 (5%) and 2.664 (1%).

Table 2A: ICE Brent Crude nominal front-month futures
2003-2016 (weekly data)

Estimated start and end dates for periods of mildly explosive price behavior

Nominal		**CPI deflated**		**SDR deflated**	
Start	**End**	**Start**	**End**	**Start**	**End**
16-5-2008	18-7-2008	16-5-2008	18-7-2008	16-05-2008	25-07-2008*
05-12-2014	06-02-2015	05-12-2014	13-02-2015	5-12-2014	06-2-2015

This table reports mildly explosive periods in ICE Brent crude weekly nominal front-month futures and the same deflated by an interpolated U.S. CPI series and by the currency value of the SDR basket using the PSY procedure with 5% size.
*This period is only detected at the 10% level.

Table 2B: NYMEX WTI Crude nominal front-month futures
2003-2016 (weekly data)

Estimated start and end dates for periods of mildly explosive price behavior

Nominal		**CPI deflated**		**SDR deflated**	
Start	**End**	**Start**	**End**	**Start**	**End**
16-5-2008	18-7-2008	16-5-2008	18-7-2008	23-04-08	18-07-2008
05-12-2014	06-02-2015	21-11-2014	06-02-2015	21-11-14	06-2-2015

This table reports mildly explosive periods in NYMEX WTI crude oil weekly nominal front-month futures and the same deflated by an interpolated U.S. CPI series and by the currency value of the SDR basket using the PSY procedure with 5% size.

Table 2C: NYMEX WTI Crude nominal front-month futures
2003-2016 (weekly data)

Estimated start and end dates for periods of mildly explosive price behavior under wild bootstrapped critical values

WTI nominal		WTI nominal/CPI	
Start	**End**	**Start**	**End**
13-06-2008	18-07-2008	No dated 2008 period	
05-12-2014	03-01-2015	12-12-2014	6-2-2015
12-12-2014	03-01-2015*	19-12-2014	6-2-2015*

This table reports mildly explosive periods in NYMEX WTI crude weekly nominal front-month futures prices and the same deflated by an interpolated U.S. CPI series using the wild-bootstrapped-adjusted PSY procedure with 10% size.
*as detected at the 5% level.

Table 3: Other Commodities and Commodity Indices
LME Copper and Zinc three-month futures prices & Bloomberg Commodity Index
2003-2016 (weekly data)

Estimated start and end dates for periods of mildly explosive price behavior

Copper		Zinc		BCI	
Start	**End**	**Start**	**End**	**Start**	**End**
24-03-2006	16-06-2006	09-12-2005	17-02-2006	23-04-2008	18-7-2008*
		03-03-2006	16-06-2006	21-11-14	06-02-2015*

This table reports mildly explosive periods in the LME copper and zinc three-month futures prices and the Bloomberg Commodity Index (BCI) using the PSY procedure with 5% size.
*This period is only detected at the 10% level.

Table 4: Levels of fundamental proxy variables:
Baltic Dry Index and U.S. end-of-week Stocks
2003-2016 (weekly data)

Estimated start and end dates for periods of mildly explosive price behavior

BDI		U.S. stocks	
Start	**End**	**Start**	**End**
13-04-2007	01-06-2007	01-04-2005	17-06-2005*
24-08-2007	21-12-2007		

This table reports detected mildly explosive periods for the levels of a demand-side fundamental (Baltic Dry Index, BDI) and a supply-side fundamental (U.S. EIA end-of-week stocks) using the PSY test with a 5% significance level. *This period is only detected at the 10% level.

Table 5: Ratios of WTI nominal prices to fundamental proxy variables: Baltic Dry Index and U.S. EIA Stocks 2003-2016 (weekly data)

Estimated start and end dates for periods of mildly explosive price behavior

Nominal WTI/BDI		**Nominal WTI/stocks**	
Start	**End**	**Start**	**End**
10-06-2005	19-08-2005	16-05-2008	25-07-2008
03-10-2008	21-11-2008	14-11-2014	06-02-2015

This table reports detected mildly explosive periods for the ratios of WTI front-month futures prices to a demand-side fundamental (the BDI) and a supply-side fundamental (U.S. EIA end-of-week stocks) using the PSY test with a 5% significance level.

Table 6: Levels of fundamental proxy and financial variables OVX and VIX

VIX levels: Jan 2003- Apr 2016 (weekly data)
OVX levels: Jan 2009 – Apr 2016 (weekly data)

Estimated start and end dates for periods of mildly explosive price behavior

VIX	**OVX**
No mildly explosive period	No mildly explosive period

This table reports detected mildly explosive periods for the levels of the OVX and VIX measures using the PSY test with a 5% significance level.

Table 7: Ratios of WTI nominal prices to financial variables: S&P500 index
2003-2016 (weekly data)

Estimated start and end dates for periods of mildly explosive price behavior

WTI nominal/S&P500	
Start	**End**
23-04-2008	18-07-2008
21-11-2014	06-02-2015

This table reports detected mildly explosive periods for the ratios of WTI front-month futures prices to the S&P500 index using the PSY test with a 5% significance level.

Table 8: Levels of fundamental proxy variables:
CFTC 2003-2016 (weekly data, collected Tuesdays)

Estimated start and end dates for periods of mildly explosive price behavior

CFTC	
Start	**End**
21-04-2015	05-01-2016

This table reports detected mildly explosive periods for the levels of a demand-side fundamental (Baltic Dry Index, BDI) and a supply-side fundamental (U.S. EIA end-of-week stocks) using the PSY test with a 5% significance level.

Figure 1: ICE Brent and NYMEX WTI crude nominal front-month futures prices: weekly data 2003-2016

Figure 2A: Brent crude oil front-month futures nominal weekly series: BSADF sequence

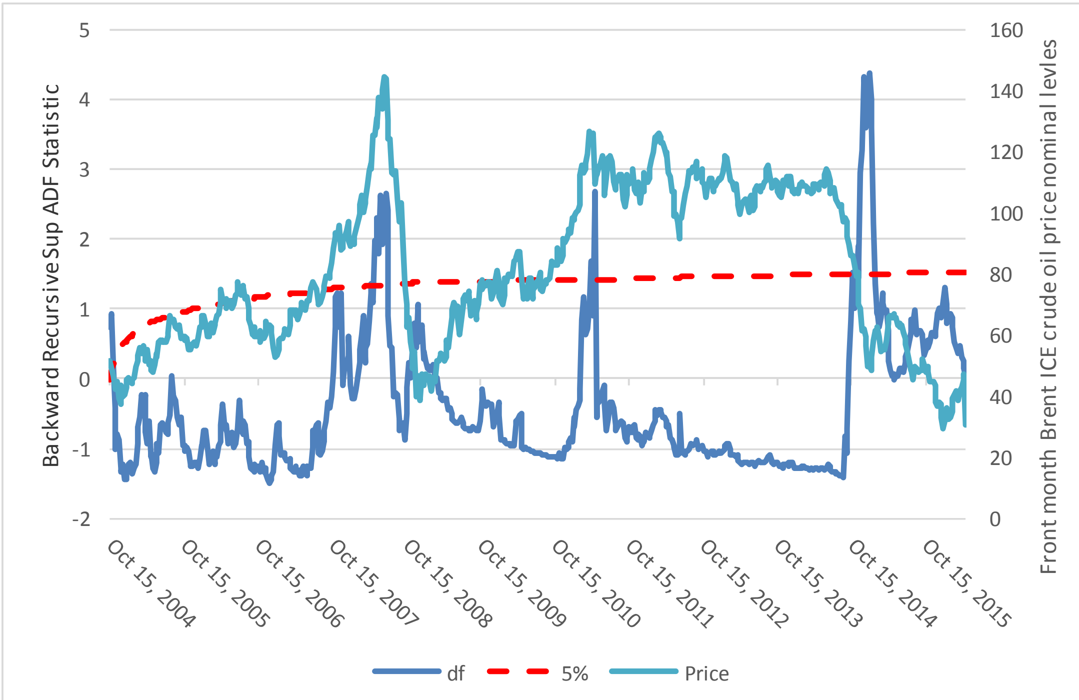

Figure 2B: WTI crude oil front-month futures nominal weekly prices: BSADF sequence

Figure 3A. WTI crude oil front-month futures nominal weekly prices: first differenced series

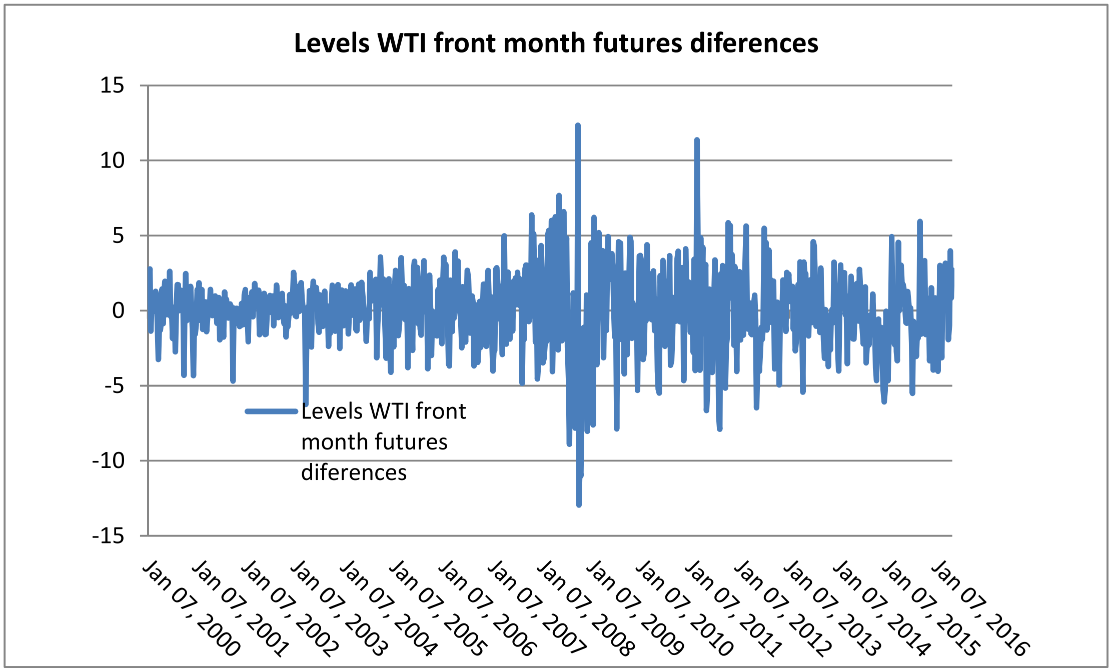

Figure 3B. Brent crude oil front-month futures nominal weekly prices: first differenced series

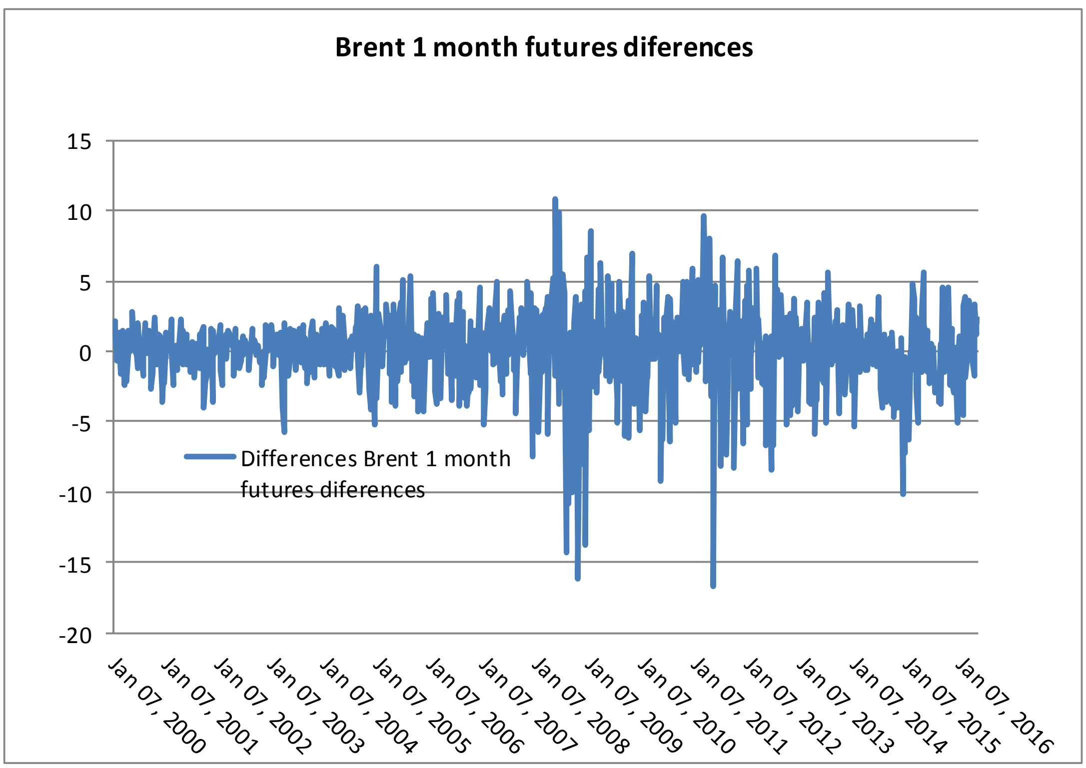

Figure 4A. BSADF sequence: WTI nominal front-month futures with wild bootstrapped critical values

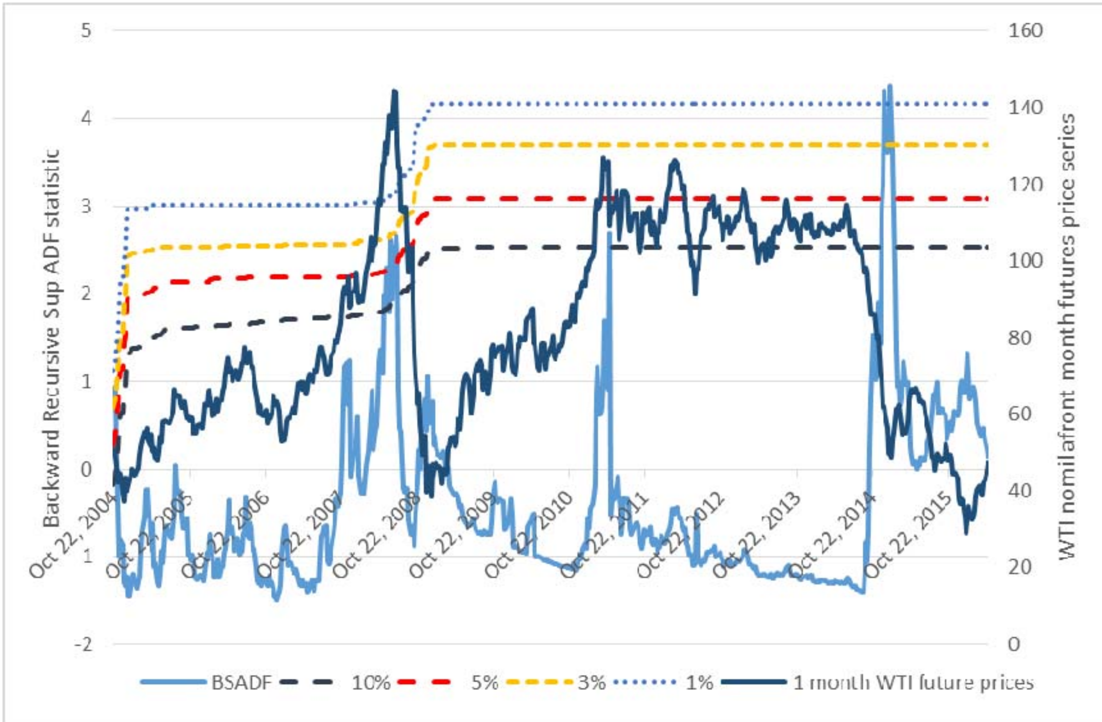

Figure 4B. BSADF sequence: Ratio of WTI nominal front-month futures to U.S. CPI with wild bootstrapped critical values

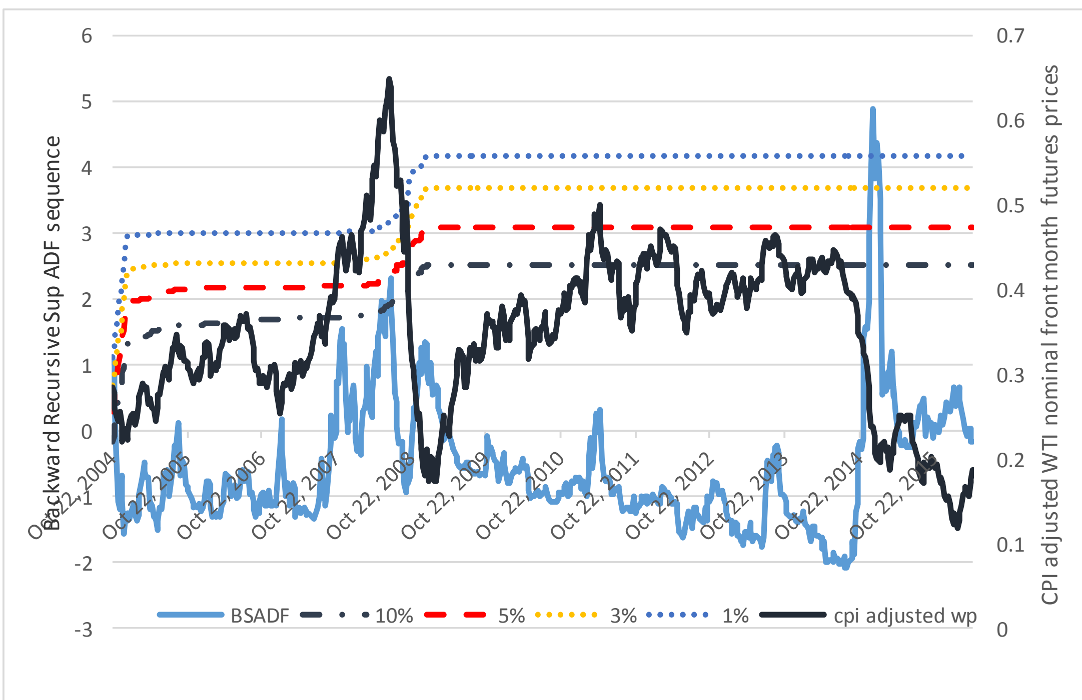

Figure 5. BSADF sequence: Bloomberg Commodity Index

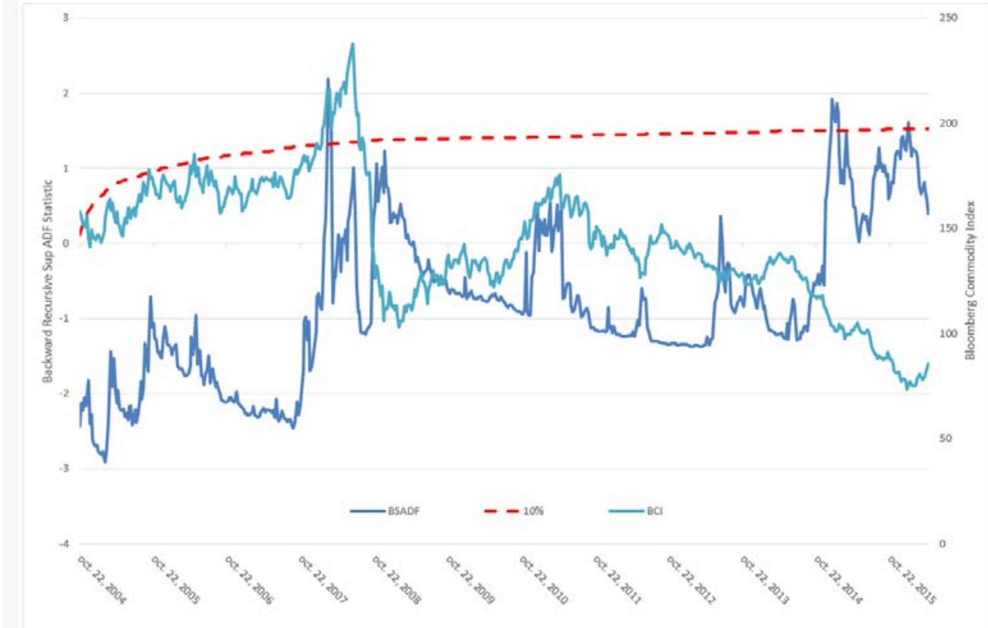

Figure 6. BSADF sequence: Baltic Dry Index

Figure 7. BSADF sequence: Ratio of WTI crude nominal prices to U.S. above-ground inventory supply data

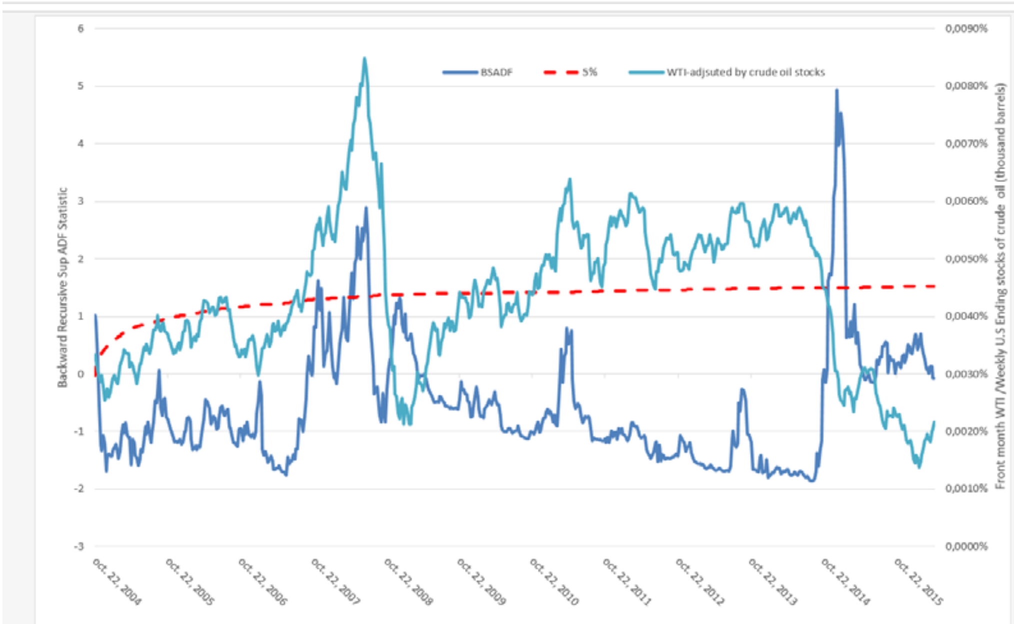

Figure 8A. WTI crude oil front month futures and the CBOE Oil Volatility (OVX) index: 2009-2016 weekly data

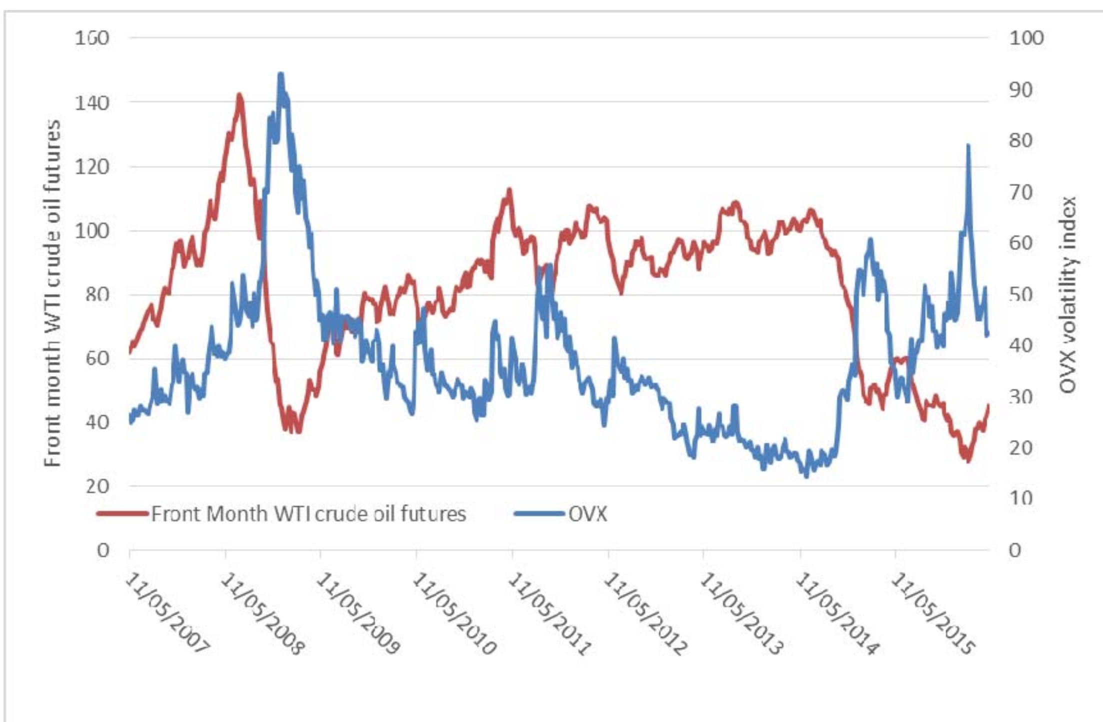

Figure 8B: BSADF sequence: OVX levels

Figure 9A: BSADF sequence: VIX levels

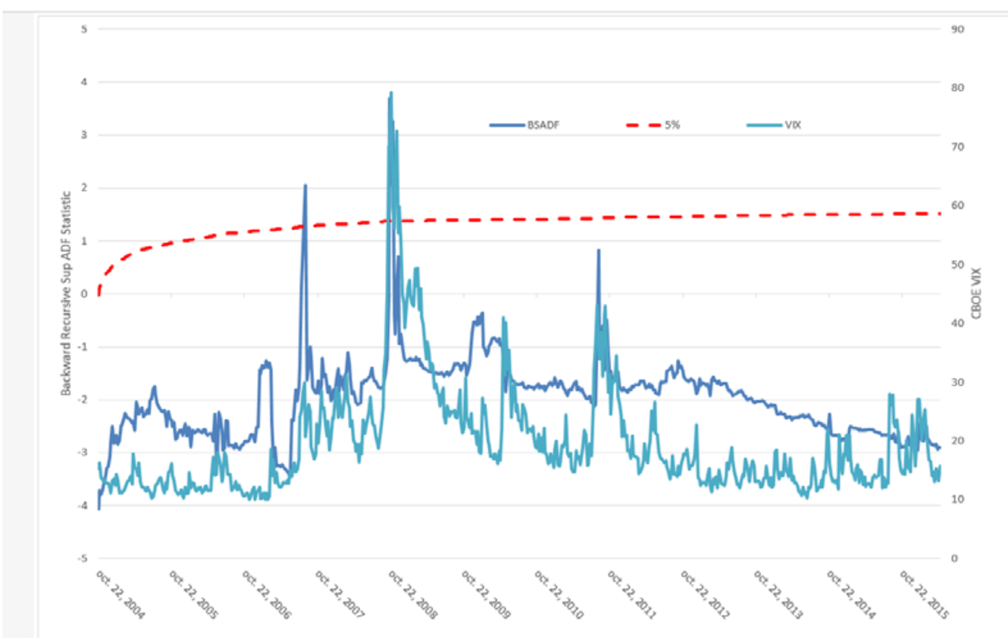

Figure 9B: BSADF sequence: Ratio of WTI front-month futures nominal price to VIX

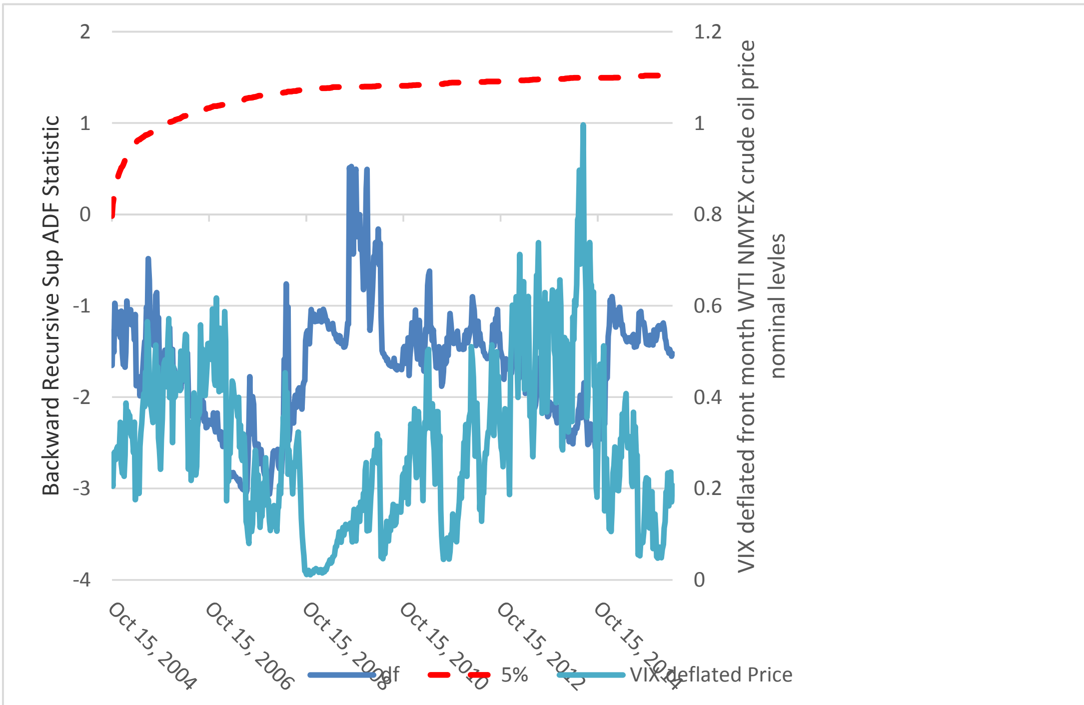

Figure 9C: BSADF sequence: VIX levels deflated by U.S. CPI

Figure 10. BSADF sequence: Working *T*-index (WTI CFTC position data)