

Bubbling over! The behaviour of oil futures along the yield curve*

Daniel Tsvetanov Jerry Coakley[§] Neil Kellard

Essex Finance Centre, Essex Business School, University of Essex, UK

This version: June 2015

Abstract

Using a rational bubble framework, a future spot price bubble can be shown to induce explosive behaviour in current long maturity futures prices under particular conditions. To assess this empirically, we employ a novel test of the unit root null against a mildly explosive alternative to investigate multiple bubbles in the crude oil spot and a range of futures prices along the yield curve employing monthly and weekly data from 1995 to 2013. The results indicate that the series overwhelmingly exhibit significant bubble periods ending in late 2008 even after allowing for an increase in unconditional volatility. Bubbles in the longer-dated contracts emerged as early as 2004 and are longer lasting than those in nearby and spot contracts. The bubble period was characterised by dramatic shifts in the yield curve associated with institutional spread positions that sharply increased futures prices at longer maturities. The results suggest that periods of time series disconnect between the spot and longer dated futures contracts could potentially form an input into early warning systems for macro-prudential policy.

JEL classification: G10, C15, C22

Keywords: Rational bubbles; Spot and futures prices; Bubble dating algorithm; Macro-prudential policy.

* Earlier versions of this paper have been presented at the SIRE Conference on “Finance and Commodities” at the University of St Andrews, Scotland, the 12th Annual INFINITI Conference on International Finance at Prato, Italy, and at seminars at Shangdong University and Southwestern University of Finance and Economics in China. We are grateful to the discussants and participants for their helpful comments. Tsvetanov gratefully acknowledges support from the ESRC studentship award number ES/J500045/1.

[§] Corresponding author: jcoakley@essex.ac.uk

1 Introduction

The boom in commodity prices more generally, and crude oil prices in particular, since the early 2000s and their remarkable collapse in late 2008, has sparked interest in both academic and policy circles.¹ In an influential paper, Tang and Xiong (2012) argue that increasing index investment in commodities markets since early 2000s has caused the futures prices of different commodities in the US to comove together. This was especially the case for the commodities in the two popular Goldman Sachs (GSCI) and Dow Jones (DJ-UBS) commodity indices, in both of which crude oil is a prominent component. Commodities were identified as a distinct investment category by fund managers and also began to comove positively with equity markets. Tang and Xiong argue that this reflects the financialisation of commodities markets and can explain many aspects of the recent synchronized price boom and bust of seemingly unrelated commodities.

Despite widespread agreement that institutional investor interest in commodities increased sharply in the early 2000s, there is no consensus on whether this altered the functioning of markets or contributed significantly to price behaviour in the 2004-2008 price run-up and subsequent collapse. Many attribute the boom-bust price changes to fundamental supply factors such as disruptions to production and/or to demand factors. In particular, proponents of the fundamental view have focused on the rapid growth of large emerging markets like China and India, fuelling a boom in demand for commodities and leading to the spike in commodity prices before the summer of 2008

¹ This boom and bust cycle in commodity markets also caused serious concerns among practitioners and policy makers that excessive speculation might have been the main driver of rising energy and food prices (see, for example, Masters, 2008; Soros, 2008; US Senate Permanent Subcommittee on Investigations, 2006). In response, the US Congress passed the 2010 Dodd-Frank Wall Street Reform and Consumer Protection Act that aimed, *inter alia*, at preventing excessive speculative influences on prices.

(see for example, Krugman, 2008; Hamilton, 2009; Kilian, 2009; and Irwin and Sanders, 2011). Prices subsequently fell sharply with the onset of the world recession. Others such as Masters (2008) argue that institutional investors² have contributed towards the deviation of commodity prices from their fundamental values leading to a speculative bubble that popped in late 2008.

The recent literature is more nuanced and addresses the question of whether the huge financial inflows caused by financialisation may have induced some commodity price changes. Singleton (2014) posits that informational frictions and associated speculative activity may induce prices to drift away from their fundamental values and may result in commodity price booms and busts within a rational differences of opinion framework.³ He presents new evidence that there were economically and statistically significant effects of investor flows on futures prices after controlling for a number of standard factors. The largest impacts on futures prices were from the growth in commodity fund index positions and institutional spread positions and these operated through risk or informational channels. He also established that hedge fund trading in futures spread positions impacted the shape of the term structure of oil futures prices.

This paper presents new evidence that directly relates to Singleton's findings. Its first contribution is that it formally tests the Singleton hypothesis that prices tend to drift away from their fundamental values employing recently developed bubble tests. A series of such tests is applied to examine the null hypothesis that there is no bubble in spot and futures oil prices using contracts along the yield curve with maturities of up to twenty four months. It is this novel examination of whether bubbles emerge in different

² One of the challenges of evaluating the effect of speculative behavior is the lack of precise relevant data on institutional commodity contract positions over a sufficiently long period of time.

³ Cheng and Xiong (2013) and Sockin and Xiong (2015) also stress the role of informational frictions and limits to arbitrage.

segments of the futures yield curve which is at the core of the paper. Subsequently, it applies the recent bubble dating strategy of Phillips, Shi and Yu (2013) to take into account the possibility of multiple bubbles. This strategy consistently estimates the origination and collapse dates of each bubble even when they are of different magnitudes. Our empirical analysis produces some striking results. Firstly, using monthly data for the sample period September 1995 to December 2013, the results indicate that all series exhibited extended periods of bubble behaviour that ended in late 2008. Secondly, the dating algorithm shows that the bubbles in longer-dated contracts started much earlier, in some cases as early as 2004, and thus were longer lasting than the bubble in the spot contract. The findings are qualitatively similar at the weekly data frequency. The earlier development of bubbles in distant contracts can be considered as a price-disconnect between spot and futures markets. Although our theory points to rational bubbles as a possible explanation, we remain agnostic as to the cause and note that such bubbles could be underpinned by information frictions, differences of opinion, limits to arbitrage, excess speculation, or time-varying discount factors (see Brunnermeier and Oehmke, 2012; Singleton, 2014; Phillips and Yu, 2011).

In a recent study of single bubble testing procedures, Harvey, Leybourne, Sollis, and Taylor (2015) show that changes in the unconditional variance might cause the spurious finding of a bubble in the corresponding price series. Since the volatility of crude oil futures contracts may have increased around 2004, we test for structural breaks in the unconditional variance.⁴ Evidence of breaks is found and therefore the

⁴ We are grateful to two anonymous reviewers for this suggestion.

Harvey et al. (2015) bootstrap procedure is subsequently applied.⁵ The results suggest that our original bubble findings are robust.

The second contribution is that our analysis of the crude oil futures yield curve over the 2004-2008 bubble period sheds new light on financialisation. The sharp downward slope of the yield curve in 2004 provides a clear rationale for the large investment flows into long positions. The subsequent dramatic shifts in the yield curve at longer maturities, associated with increased institutional spread positions, are consistent with sharp price increases in long-dated contracts that may have precipitated the bubble. Our Phillips, Shi and Yu (2013) test results provide more clearcut and extensive evidence of bubbles than the extant literature. Gilbert (2010) tests whether nearby futures prices exhibited bubble behaviour over the 2006 to 2008 period and strikingly reports a very short bubble in the crude oil market. Phillips and Yu (2011) and Shi and Arora (2012) find a short-lived bubble in spot oil prices over the period March to July 2008. Moreover, Phillips and Yu (2011) show a potential transmission mechanism: the bubble behaviour in commodity markets appears to have migrated from the housing market.⁶

The final contribution is that the findings may have practical policy implications. Our results raise the possibility that relevant information about potential future spot price bubbles may be gleaned from longer-dated futures prices. Bubble tests on long maturity oil futures contracts may prove useful for real-time monitoring and early warning signals for bubble formation. To provide a theoretical basis for such an investigation, this paper first outlines a rational bubble approach where it shows that,

⁵ The extant literature does not yet provide a non-stationary variance correction for the multiple bubble tests used in this paper. We therefore apply the Harvey et al. (2015) bootstrap procedure for the single bubble test as an approximation.

⁶ Lammerding et al. (2013) also present evidence of a speculative bubble in recent nearby crude oil futures prices.

under particular conditions, an expectation of a spot price bubble in the future will generate explosive behaviour in the relevant futures prices prior to the appearance of the spot price bubble.

The rest of the paper is organised as follows. Section 2 provides the theoretical background on the price-bubble relationship for storable commodities. Section 3 describes the econometric approach for testing for rational bubbles and date stamping. Section 4 presents the data and descriptive statistics. The empirical results are discussed in Section 5 whilst a final section concludes.

2 The theory of rational bubbles in commodity markets

We follow Diba and Grossman (1988b) and explain the price change of storable commodities by changes in expected future net payoffs defined as ‘fundamentals’. The current spot price of a commodity, S_t , is determined by the present value of next period's expected spot price, $E_t[S_{t+1}]$, and the marginal convenience yield, C_t :

$$S_t = \frac{E_t[S_{t+1} + C_t]}{(1 + R)} \quad (1)$$

where $E_t[.]$ denotes the expectation conditional on the information at time t and R is a time invariant interest rate. The net of storage cost marginal convenience yield measures the benefit from holding inventories per unit of commodity over the period t to $t + 1$ and is analogous to the dividend on a stock. Under the theory of storage it should satisfy the standard no arbitrage condition:

$$C_t = (1 + R)S_t - F_{1,t} \quad (2)$$

where $F_{1,t}$ denotes the futures price at time t for delivery of a commodity at $t + 1$. If S_t and C_t are both integrated processes, the series are cointegrated with specific cointegrating vector $[1, -1/R]$ (see Campbell and Shiller, 1987). Solving (2) for $F_{1,t}$ and re-arranging, it follows that the difference between contemporaneous futures and spot prices is the interest forgone in storing the commodity over the period t to $t + 1$, less the marginal convenience yield.

Solving the difference Equation (1) forward and applying the law of iterated expectations, yields:

$$S_t = E_t \left[\sum_{\tau=0}^{\infty} \frac{1}{(1+R)^{\tau+1}} C_{t+\tau} \right] + \lim_{\tau \rightarrow \infty} E_t \left[\frac{1}{(1+R)^{\tau}} S_{t+\tau} \right]. \quad (3)$$

Imposing the transversality condition, $\lim_{\tau \rightarrow \infty} E_t[(1+R)^{-\tau} S_{t+\tau}] = 0$, eliminates the last term in (3). It follows that the price S_t collapses to the discounted sum of expected future payoffs, i.e. the fundamental value which will be denoted by S_t^f . However, if the transversality condition does not hold, there are infinitely many solutions to (3) that take the form:

$$S_t = S_t^f + B_t \quad (4)$$

where B_t is a bubble component that has to satisfy:

$$B_t = (1+R)^{-1} E_t [B_{t+1}]. \quad (5)$$

In other words, the bubble has to grow over time at a rate R in order for investors agree to hold the asset (see Blanchard and Watson, 1982). Diba and Grossman (1988a) argue that rational bubbles cannot be negative. If a bubble is negative, then when it erupts, it could make the price of the security negative also. In addition, if $B_t = 0$ (5) implies that

$B_{t+1} = 0$ with probability 1. It follows that the existence of bubbles would be consistent with rationality only when $B_t > 0$.

Taken together, Equations (1) and (2) imply that the futures price is an unbiased predictor of the future spot price

$$F_{1,t} = E_t[S_{t+1}] \quad (6)$$

Given the decomposition of the spot price in (4), it follows that the contemporaneous futures price with maturity $t + T$ embodies information about the expected value of the bubble component over the period t to $t + T$:

$$F_{T,t} = E_t[S_{t+T}^f] + E_t[B_{t+T}] \quad (7)$$

In line with (5), the bubble is expected to grow exponentially at rate R , i.e. $E_t[B_{t+T}]$ contains the root $(1 + R)^T$ that is greater than unity. If the bubble erupts at some future time $t + j$, for $1 < j < T$, it will induce explosive behaviour in the price series of futures contracts with maturity greater than $t + j$.

To test for rational bubbles in the stock market, Diba and Grossman (1988b) motivate the use of stationarity tests. However, Evans (1991) suggests that this approach would not efficiently detect periods of explosive behaviour if bubbles collapse periodically. Consistent with the process in (5), he considers bubbles described by:

$$B_{t+1} = [(1 + R)B_t I\{B_t \leq \alpha\} + [\varphi + \pi^{-1}(1 + R)\theta_{t+1}(B_{t+1} - (1 + R)^{-1}\varphi)I\{B_t > \alpha\}]u_{t+1}] \quad (8)$$

where $0 < \varphi < (1 + R)\alpha$, u_{t+1} is a positive *iid* variable with $E_t[u_{t+1}] = 1$, and $I\{\cdot\}$ is an indicator function that assumes a value of 1 when the condition in the braces is true and 0 otherwise. θ_{t+1} is an *iid* Bernoulli process and the probability of $\theta_{t+1} = 0$ is $(1 - \pi)$

and $\theta_{t+1} = 1$ is π , where $0 < \pi < 1$. Such a bubble would start to grow at a rate $(1 + R)\pi^{-1}$ once it exceeds some threshold level α , but with a probability $(1 - \pi)$ the bubble will collapse to an expected mean level φ . Since a bubble never collapses to zero, it will start growing again without violating the non-negativity constraint given in Diba and Grossman (1988a).

The non-linear bubble process in (8) causes the data series to exhibit global characteristics similar to a stationary process. As a result, conventional unit root tests applied to the full sample would lack power, failing to adequately test the null hypothesis of no bubbles. Phillips et al. (2011) suggest the application of a unit root test in a recursive window framework to overcome this drawback. Their approach allows the test statistics to be time dependent and therefore is able to detect explosive behaviour in time series even when bubbles are periodically collapsing.

3 Testing for mildly explosive behaviour

3.1 Bubble tests

Suppose we observe the sequence $\{Y_t\}_{t=1}^T$ and estimate the following autoregression:

$$\Delta Y_t = \mu_Y + \delta_r Y_{t-1} + \sum_{i=1}^k \Delta Y_{t-i} + u_t \quad \text{for } t = 1, 2, \dots, [rT] \quad \text{and } r \in [r_0, 1] \quad (9)$$

where u_t is white noise, r is a fraction of the total sample and $[x]$ denotes the integer part of x . The recursive window of the regression expands forward by one observation at a time from some initial sample $[r_0 T]$. Consistent with Phillips et al. (2011), tests adopted in our empirical analysis are performed under the null that the time series contains a unit root at every t :

$$Y_t = Y_{t-1} + u_t \quad \text{for} \quad t = 1, 2, \dots, T. \quad (10)$$

Under the alternative hypothesis, Phillips et al. (2011) specify a data generating process where the series starts as a unit root but switches to a regime of mildly explosive behaviour (δ is greater than unity but still in its vicinity) at date $\lfloor r_e T \rfloor$ until $\lfloor r_f T \rfloor$. At date $\lfloor r_f T \rfloor$ the series returns to a unit root regime. The model is defined as:

$$Y_t = Y_{t-1} I\{t < \lfloor r_e T \rfloor\} + \delta_n Y_{t-1} I\{\lfloor r_e T \rfloor \leq t \leq \lfloor r_f T \rfloor\} \\ + \left(\sum_{j=\lfloor r_f T \rfloor+1}^t u_j + Y_{\lfloor r_f T \rfloor}^* \right) I\{t > \lfloor r_f T \rfloor\} + u_t I\{t \leq \lfloor r_f T \rfloor\} \quad (11)$$

where $\delta_n = 1 + cn^{-\eta}$ with $c > 0$ and $\eta \in (0,1)$, and $Y_{\lfloor r_f T \rfloor}^* = Y_{\lfloor r_e T \rfloor} + O_p(1)$.

The restrictions on the parameter c and values of η over the specified open interval yield the mildly explosive process discussed in Phillips and Magdalinos (2007a, 2007b). The boundary as $\eta \rightarrow 1$ includes the local to unity case where defining a bubble period is not possible (see Phillips and Yu, 2011).

Phillips et al. (2011) suggest the application of the augmented Dickey-Fuller t -statistics to the recursive autoregression in (9) to test the null hypothesis of a unit root or no bubbles. The test statistic is given as:

$$SADF(r_0) = \sup_{r \in [r_0, 1]} \{ADF_0^r\} \quad (12)$$

where ADF_0^r is the ADF statistic from (9) evaluated between $t = 1$ and $t = \lfloor rT \rfloor$.

Homm and Breitung (2012) propose a similar procedure, but under the assumption that a break in the autoregressive coefficient occurs at observation $\lfloor rT \rfloor$. The model is written as:

$$\Delta Y_t = \begin{cases} \mu_Y + u_t & \text{if } t \leq \lfloor rT \rfloor \\ \mu_Y + \delta_r Y_{t-1} + u_t & \text{if } t > \lfloor rT \rfloor \end{cases} \quad (13)$$

The null hypothesis $H_0: \delta_r = 0$ is then tested against the alternative $H_1: \delta > 0$ by using the supremum of a sequence of backward recursive Chow tests. Specifically, the test statistic is the following:

$$SDFC(r_0) = \sup_{r \in [0, 1-r_0]} \{DFC_r^1\} \quad \text{and} \quad DFC_r^1 = \frac{\sum_{t=\lfloor rT \rfloor+1}^T \Delta Y_t^* Y_{t-1}^*}{\hat{\sigma}_r \sqrt{\sum_{t=\lfloor rT \rfloor+1}^T (Y_{t-1}^*)^2}} \quad (14)$$

where

$$\hat{\sigma}_r^2 = (T-2)^{-1} \sum_{t=2}^T (\Delta Y_t - \mu_Y - \hat{\delta}_r Y_{t-1} I\{t > \lfloor rT \rfloor\})^2,$$

$$\Delta Y_t^* = \Delta Y_t - \frac{1}{T} \sum_{j=\lfloor rT \rfloor+1}^T \Delta Y_j \quad \text{and} \quad Y_{t-1}^* = Y_{t-1} - \frac{1}{T} \sum_{j=\lfloor rT \rfloor+1}^T Y_{j-1}$$

and $\hat{\delta}_r$ is the least squares estimator of δ from Equation (13).

Homm and Breitung (2012) also motivate the use of Buseti-Taylor statistics on the assumption that the series has a unit root up to observation $\lfloor rT \rfloor$ after which it switches to a regime of explosive behaviour. Using a random walk model to forecast the final value Y_T from the periods $Y_{\lfloor rT \rfloor}, Y_{\lfloor rT \rfloor+1}, \dots, Y_{T-1}$ should result in a large sum of squared forecast errors. The modified version of the statistic is given by:

$$SBT(r_0) = \sup_{r \in [0, 1-r_0]} \{BT_r^1\} \quad (15)$$

where

$$BT_r^1 = \frac{T-1}{(T-\lfloor rT \rfloor)^2 \sum_{t=1}^T \hat{\epsilon}_{0,t}^2} \sum_{t=\lfloor rT \rfloor+1}^T \hat{\epsilon}_{1,t,1,t}^2,$$

and $\hat{\epsilon}_{0,t}$ are the OLS residuals from the regression of ΔY_t on an intercept, $t = 1, \dots, T$, whilst $\hat{\epsilon}_{1,t}$ are the OLS residuals from the regression of $(Y_t - Y_{t-1})$ on an intercept, $t = \lfloor rT \rfloor + 1, \dots, T$. Evidence against the null hypothesis of no bubble in all of the above tests is obtained by comparing the sup statistics with the corresponding right-sided critical values from the limit distribution. However, the above procedures do not facilitate the identification of the explosive period. In doing so, it is desirable to set some minimum duration for this bubble period to successfully discriminate between bubbles and short-lived blips.

3.2 The bubble dating algorithm

Techniques that can help identify bubble periods are useful as a real-time monitoring procedures and early warning signals for bubble formation. They also overcome some weaknesses of the suggested tests above. For example, the $SDFC(r_0)$ and $SBT(r_0)$ procedures assume that the time series switches to mildly explosive behaviour at some date over the interval $[0, 1 - r_0]$. Homm and Breitung (2012) using extensive Monte Carlo simulations find that, when there is a one-time change in the time series behaviour, even if this change is random, both tests have an advantage over the $SADF(r_0)$ test in terms of power. But, if there is an additional break signalling a return to a unit root process, this advantage disappears.

The technique described next is robust to multiple breaks in the series. Under the data generating process in (11), the series switches from unit root to mildly explosive behaviour for a period of time and then returns to a unit root. The break dates are defined as $\lfloor r_e T \rfloor$ and $\lfloor r_f T \rfloor$, respectively. Phillips et al. (2011) provide consistent estimators of the break points as the first and last observation at which the ADF_0^r statistic is significant at some level β_T . However, Phillips et al. (2013) consider a more

general case where the *ADF* statistics are still computed for the interval $r_2 \in [r_0, 1]$ but now for each recursion the starting point varies over the feasible range $r_1 \in [0, r - r_0]$. They show that this procedure is more efficient when multiple bubbles are present in the data. The generalized test statistic is:

$$GSADF(r_0) = \sup_{\substack{r_2 \in [r_0, 1] \\ r_1 \in [0, r_2 - r_0]}} \{ADF_{r_1}^{r_2}\} \quad (16)$$

To estimate the fraction points for the origination and collapse points of the bubble, Phillips et al. (2013) recommend using a backward sup *ADF* test to improve dating accuracy. In essence, this is an *SADF*(r_0) test on a backward increasing sample in which the endpoint $[r_2 T]$ is fixed and the initial observation changes over the region $t = 1$ and $t = T[r_2 - r_0]$. The backward *SADF*(r_0) test can therefore be written:

$$BSADF_{r_2}(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \{ADF_{r_1}^{r_2}\} \quad (17)$$

and the origination and collapse points:

$$\hat{r}_e = \inf_{r_2 \geq r_0} \{r_2 : BSADF_{r_2}(r_0) > cv_{\beta_T}(r_2)\} \text{ and} \quad (18)$$

$$\hat{r}_f = \inf_{r_2 \geq \hat{r}_e + \gamma \frac{\ln(T)}{T}} \{r_2 : BSADF_{r_2}(r_0) < cv_{\beta_T}(r_2)\}$$

where cv_{β_T} are the right-side $100\beta_T\%$ critical values of the *SADF*(r_0) statistics for $[Tr_2]$ observations. It should be noted that $\gamma \frac{\ln(T)}{T}$ is the minimum duration necessary for a part of the series to qualify as a bubble period. The parameter is chosen based on the sampling frequency so that periods shorter than $\left\lfloor \gamma \frac{\ln(T)}{T} \right\rfloor$ observations are considered insignificant (see Phillips and Yu, 2011). Finally, it can be seen that the date stamping methodology above, based on the backward sup *ADF* test, corresponds to the *GSADF*(r_0) as:

$$GSADF(r_0) = \sup_{r_2 \in [r_0, 1]} \{BSADF_{r_2}(r_0)\} \quad (19)$$

4 Data and sample characteristics

Daily WTI crude oil prices for the spot and futures contracts on NYMEX were downloaded from DataStream for the period September 1995 to December 2013. The starting date of the sample was dictated by the availability of data for longer-dated contracts from 15 to 24 months. In fact, data were collected for futures contracts with a range of maturities along the yield curve including 1, 3, 6, 9, 12, 15, 18, 21 and 24 months. It should be noted that NYMEX crude oil futures usually expire on the third business day prior to the 25th calendar day of the month preceding the delivery month.

For the empirical analysis, a monthly series of 220 observations for each spot and associated nine futures contracts was constructed using closing daily prices on the last business day of each month. As an aid to explanation, Table 1 summarises the relationship between the selected price of contract j (i.e., $F_{j,t}$) and its expiration and delivery month, for the first observation of the sample at $t = 1$.

[Table 1 around here]

Remembering that our sample begins in September 1995, Table 1 shows that on the last business day of that month, the contract that expires the following month provides the price for observation $(F_{1,1})$. Likewise, on that same day, the contract that expires in three months provides the price for observation $(F_{3,1})$. The first observation for contract maturities 6 to 24 months is analogously collected. The data are sampled monthly, and therefore as we move forward to the last business day of October, we collect the prices $(F_{1,2})$, $(F_{3,2})$ and so on. In this manner, our constant maturity series are constructed.

Table 2 provides descriptive statistics of log returns for both spot and futures contracts over the full 1995-2013 sample in Panel A and three non-overlapping subsamples.

[Table 2 around here]

The first subsample extends from 1995 up to the end of 2003 and the results are displayed in Panel B. The second subsample results cover the January 2004 to June 2008 price run-up period and are in Panel C. The start point was chosen to highlight the effect of increased investment flows and, in particular, institutional investor flows into commodities since 2004 as highlighted, *inter alios*, by Tang and Xiong (2012). The final subsample commences in July 2008 and extends up to the end of 2013 and so covers the financial crisis and economic downturn period.

Table 2 reveals some interesting patterns. First, the full sample period and the pre-2004 period are strikingly similar, especially in terms of their mean log return and standard deviation. Log returns exhibit positive skewness that increases along the yield curve for both periods although it is more pronounced at longer maturities for the full sample period. Both periods exhibit mild excess kurtosis that is always negative for the full sample but mainly negative for the pre-2004 subsample.

Second, the 2004-2008 subsample results in Panel C exhibit sharp changes as compared with those for the earlier subsample in Panel B. The mean monthly log return is now in the 2.7% to 3% range which is some 4.5 to 6 times that in the earlier period. Given the search for yield prior to the sub-prime crisis, the increased post-2004 investment in passive CIFs (Commodity Index Funds) is thus no surprise. More strikingly, the mean standard deviation is actually lower at the short end (up to 3 months) of the yield curve post-2004 as compared with the pre-2004 period despite the post-2004 mean return being 4.5 times higher. The log return standard deviation decreases monotonically along the yield curve from 0.078 for the nearby 1-month contract to just 0.059 for the 24-month contract which made trading in distant contracts even more attractive since returns increased modestly along the yield curve. This was

an important factor in the increased popularity of managed money spreads highlighted, by Singleton (2014), where investors go long a distant contract and short a nearby one. The final difference between the pre-and post-2004 subsamples is that all contracts in Panel C are negatively skewed but those in Panel B are correspondingly positively skewed whilst excess kurtosis is similarly mild in both periods.

Overall, spot and futures oil price dynamics changed significantly after 2004. Prices changes remained in a very narrow range during 1995-2003 in comparison to the 2004-2013 period when the maximum price for all contracts was close to \$140 per barrel. Prices reached an all-time high in 2008 followed by a sharp collapse. The above factors point to distinct investor behaviour in the post-2004 period. For example, Büyüksahin et al. (2009) show that the growth of large net positions in long-dated contracts by hedge funds and other investors dates from 2004 and 2005.

To further investigate the relationship between spot and futures prices, Figure 1 plots the sample average of the futures-spot differential (i.e. the basis) over three periods: the pre-2004 period in panel A, the January 2004 - June 2008 price run-up period in Panel B, and the post-bubble period July 2008 – December 2013.

[Figure 1 around here]

Figure 1 can be viewed as depicting the average term structure or yield curve for futures prices. In line with the literature, the basis is measured as:

$$\frac{F_{j,t} - S_t}{S_t} \quad (20)$$

The sharply downward sloping futures yield curve in Panel A implies that the crude oil futures market was backwardated prior to 2004 and the average basis at all maturities is always negative. A trader with a long futures position on average would realize a positive return from rolling her position forward into the cheaper (next) nearby

contract. Panel B shows that the futures market was still mainly backwardated in the price run-up period from 2004 to August 2008 but the slope is now much flatter (see below for more details).

Panel C shows that the futures yield curve switched to contango with a positive slope in the period commencing late 2008. By contrast with the earlier periods, the average basis at all maturities is always positive. Contracts further out along the curve now cost more and so rolling a nearby long position to the next contract would on average be loss making. This is consistent with recent studies such as Acharya, Lochstoer, and Ramadorai (2012) and Cheng and Xiong (2014) that stress that investors' risk-bearing capacity and thus their appetite for risk sharing vary over time. In particular, their reduced risk appetite during crises may lead them to unwind their positions rather than take on extra risk.

5 Empirical results

5.1 Bubble test results

As noted earlier, the power of the $SDFC(r_0)$ and $SBT(r_0)$ tests deteriorate if the bubble bursts within the sample period. Hogg and Breitung (2012) suggest that successive observations following the explosive period are excluded from the sample to overcome this problem. The spectacular downturn in the price of oil since June 2008 bears a close resemblance to a collapsing bubble. Therefore, the $SDFC(r_0)$ and $SBT(r_0)$ statistics are estimated over the period September 1995 to June 2008. The $SADF(r_0)$ and $GSADF(r_0)$ statistics are obtained from the whole sample. The number of lags k in Equation (9) is determined by the Bayesian information criterion (BIC) with a maximum lag 12. Phillips et al. (2013) showed that the dynamic lag length selection results in a satisfactory size

for the $SADF(r_0)$ test but positive size distortion for the $GSADF(r_0)$ test. Thus, in line with their empirical analysis, we set the lag order k to zero for the $GSADF(r_0)$ test to lower the probability of a type 1 error.

The recursive regressions were run with an initial window size of 44 observations (20% of the total sample) due to the sample size. The test results are given in Panel A of Table 3 whilst Panel B provides various right-sided critical values.

[Table 3 around here]

An analysis of the results in the Table 3 leads to several conclusions. First, all statistics for the spot price series readily reject the null hypothesis of no bubbles at the 1% level. Second, all statistics for all of the futures price series also provide evidence against the null hypothesis of no bubble at the 1% level. This result is novel and is one of the original findings of this study. Finally, the evidence supporting bubbles becomes stronger as the maturity of futures contracts increases. These patterns are consistent with the positive sample excess kurtosis and negative skewness that we observe in Table 2. Bubbles will contribute towards positive price changes during the boom phase and large negative price changes during the bust phase. Such a positive feedback mechanism will cause the mass of the probability distribution to be more concentrated on the right with fatter tails than a normally distributed variable (see, *inter alios*, Blanchard and Watson, 1982; Camerer, 1989).

The support for bubbles in the spot and nearby futures is generally consistent with Shi and Arora (2012) although the timing differs, as we shall see in the next subsection. They use the convenience yield implied by the nearest and second nearest contracts in three different regime-switching models and report a short-lived bubble in late 2008 and early 2009. Our bubble timing is more consistent with Phillips and Yu (2011) who find a short bubble episode between March and July 2008 for spot oil

prices. Furthermore, their empirical findings offer some evidence in support of the sequential hypothesis of Caballero et al. (2008) who relate the oil bubble to global imbalances, the subprime crisis and volatile asset prices.

5.2 Date stamping the bubbles

Whilst Table 3 provides strong evidence of bubbles, they need to be date stamped. Figure 2 presents the results from the Phillips et al. (2013) *GSADF* procedure on bubble origination and collapse dates at the 5% level. We follow the literature in imposing a minimum six month bubble duration period on the grounds that shorter episodes can be regarded as just blips.

[Figure 2 around here]

There are some striking patterns in the results. First, they indicate a bubble in spot prices and the nearby futures (contract 1) from February 2008 to July 2008. The identical results for the spot and nearby contracts are no surprise as otherwise there would have been arbitrage opportunities. Second, the bubble duration increases along the yield curve. For example, it ranges from just 7 months (February to August 2008) for contract 3 to some four and a half years for contract 24. Contracts 6 through to 18 exhibit evidence of multiple bubbles but there is a continuous bubble for contracts 21 and 24 from April 2004 to August 2008.

We summarise in Figure 3 the bubble periods at the 10% significance level.

[Figure 3 around here]

In contrast to the results in Figure 2, this now shows a continuous bubble for the 2004-2008 period for contracts with maturities of 12 months and beyond. Again, the origination of the bubbles is in early 2004 whereas the collapse date for contracts 18 to 24 shifts to September 2008. The earlier appearance of bubbles in longer maturity

contracts is interesting from a macro-prudential perspective. The evidence in this paper suggests that central banks could usefully employ bubble tests on longer maturity crude oil futures prices as part of an early warning system for detecting overheating in commodity and financial markets.

5.3 Robustness checks

5.3.1 Breaks in unconditional volatility

The results so far suggest that longer-dated futures contracts have exhibited bubble behaviour since early 2004, which also marks the beginning of the financialisation of commodity futures markets. However, the descriptive statistics in Table 2 indicate that after 2004 there was an increase in the unconditional price volatility, especially in the case of longer-dated futures contracts. Given the work of Harvey et al. (2015), this raises the question of whether the documented bubble behaviour in crude oil spot and futures prices is spurious due to a break in the unconditional variance. Thus we test for possible (multiple) breaks in the unconditional volatility of crude oil spot and futures prices.

Following Rapach et al. (2008) and Vivian and Wohar (2012), we employ a modified version of the cumulative sum of squares statistic suggested by Inclàn and Tiao (1994). Let $\Delta Y_t^* = \Delta Y_t - \frac{1}{T} \sum_{j=1}^T \Delta Y_j$, so that the statistic is given by:

$$MIT = \sup_k |T^{-0.5} G_k| \quad (21)$$

where

$$G_k = \lambda^{-0.5} \left[C_k - \frac{k}{T} C_T \right],$$

$$C_k = \sum_{t=1}^k (\Delta Y_t^*)^2,$$

$$\hat{\lambda} = \hat{\gamma}_0 + 2 \sum_{l=1}^m [1 - l(m+1)^{-1}] \hat{\gamma}_l$$

$$\hat{\gamma}_l = T^{-1} \sum_{t=l+1}^T [(\Delta Y_t^*)^2 - \hat{\sigma}^2][(\Delta Y_{t-l}^*)^2 - \hat{\sigma}^2],$$

$$\hat{\sigma}^2 = T^{-1} C_T,$$

and the lag truncation parameter is selected as in Newey and West (1994). Because of the non-parametric adjustment, the *MIT* statistic has been shown to possess good size properties even when ΔY_t^* is characterised by some form of temporal dependencies like autocorrelation or autoregressive conditional heteroscedasticity (see Sansò et al., 2004). To allow for multiple breaks in the variance, we apply the iterative algorithm described in Steps 0-3 in Inclàn and Tiao (1994, p.916) and employ the *MIT* statistic.

The break dates are depicted in Figure 4 along with the full-sample graph for each price series.

[Figure 4 around here]

Interestingly, the test results indicate that the spot and nearby contract (contract 1) exhibited two structural breaks (i.e., one in April 2004 and another in September 2007), whereas there is only one break in July 2005 for all other contracts. Whilst the break date in the longer dated contracts does not coincide exactly with the origination date of the explosive periods, there is still a possibility that the bubble test results reported above are misleading.

Currently in the literature we do not have a non-stationary variance correction for the multiple bubble tests used in this paper. Therefore, as an approximation, we apply

the bootstrap algorithm suggested by Harvey et al. (2015, p.11) to the single bubble SADF(r_0) statistic defined in (12).⁷ The results are summarised in Table 4.

[Table 4 around here]

Compared to the non-bootstrapped results in Table 3, unsurprisingly in Table 4, we find slightly weaker evidence of bubbles in the shorter-dated contracts even if the rejection is marginal at the 5% significance level in most cases. Nevertheless, all series exhibited bubble behaviour at the 10% significance level, whereas for longer-dated contracts (i.e. contracts 21 and 24) the test statistic is significant at the 5% level. Therefore, even allowing for non-stationary variance, the robustness results in this section support our baseline findings of bubbles in oil price series.

5.3.2 Weekly frequency

Finally, we examine the effect of data frequency on the bubble periods that we have identified. Using the same logic as with the monthly data, we create weekly series by taking Tuesday's settlement price for each maturity.⁸ In particular, we roll over contracts on the last business day of the month prior to nearby expiration to create continuous series with a reasonably constant maturity. The bubble test results are reported in Table 5.

[Table 5 around here]

Consistent with the results in Table 3, each test statistic is significant at the 1% level and the evidence for bubbles becomes even stronger as the maturity of the contracts

⁷ Note that Harvey et al. (2015) focus only on the SADF bubble test and offer no correction for the bubble date-stamping strategy. It is possible to extend their methodology to the bubble dating algorithm described in section 3.2 but we leave it as an area for future research since the asymptotic properties of the bootstrapped real time detectors have yet to be defined.

⁸ Monday's settlement price is used for those weeks when the last business day of the month is a Monday.

increases. The bubble periods identified by the date-stamping strategy are depicted in Figures 5 and 6 for the 5% and 10% levels of significance, respectively.

[Figures 5 and 6 around here]

Again, the results are consistent with the bubble periods depicted in Figures 2 and 3. Overall, the finding that longer-dated futures contracts may serve as an early warning system for bubble detection in the spot market appears robust at the weekly frequency also. For completeness, we also apply the unconditional variance break test and the Harvey et al. (2015) test to our weekly series. The results of the former can be seen in Table 6 and the latter in Table 7.

[Table 6 around here]

[Table 7 around here]

The results in Table 6 suggest that more breaks are found in weekly, as opposed to monthly data. Given the additional noise in weekly data, it is perhaps unsurprising that the results in Table 7 provide less evidence for bubbles than monthly counterparts. However, it is striking that 8 of the 10 series still suggest a bubble at the 10% level.

5.4 Discussion

The *GSADF* data stamping results indicate strong evidence of bubbles commencing in early 2004 for longer maturity contracts. This bubble origination date is significantly earlier than that for spot prices and contracts with shorter maturity. Contracts 12, 15, 18, 21 and 24 show evidence of multiple and/or continuous bubbles starting in early 2004. The upshot is that there is a clear indication that futures contracts with maturity above six months have been traded at prices considerably higher than their fundamental level since approximately early 2004. Interestingly, the year 2004 marks the increase of investment flows into commodity derivatives market that many

researchers believe changed oil futures price behaviour (see, e.g., Sockin and Xiong, 2015, Tang and Xiong, 2012, and Singleton, 2014).

Can we shed any further light on the appearance of bubbles from 2004? Figure 7 depicts the average basis or carry of all contracts at different maturities *for each calendar year*, along the yield curve, from January 2004 to June 2008.⁹

[Figure 7 around here]

It is immediately clear that the average yield curve 2004-2008 in Figure 1 Panel B hides a number of dramatic details and shifts over time. The first is that the yield curve is far more sharply backwardated in 2004 as compared with the pre-2004 period. This made passive commodity index investment seem even more attractive around 2004. The second is that the yield curve shifts dramatically in each subsequent year. It shifts upward and flattens in 2005¹⁰ whilst it moves into contango in 2006 before becoming almost horizontal in 2007 and mildly backwardated in 2008. Third, the shifts are most dramatic at longer maturities and probably reflect the massive flows into managed money commodity spread positions in those years. In that sense, the bubble duration evidence from the *GSADF* detector for the longer maturity contracts is consistent with the financialisation of commodities hypothesis.

Büyüksahin and Robe (2014) and Büyüksahin et al. (2009) utilise a unique CFTC Large Trader Reporting System (LTRS) position dataset, which allows them to investigate the exact type of investor and their changing investment patterns over the last decade. They argue that positions of hedge funds and non-registered participants in long-dated futures have increased significantly since the early 2000s. However, hedge

⁹ The period from January to June was employed for 2008.

¹⁰ The mild contango at shorter maturities in 2005 may reflect hedge funds front-running the Goldman roll (see Mou, 2011).

funds are generally recognised as sophisticated institutional investors. Their trading activities should correct for any price deviations from fundamental levels and contribute towards market efficiency, preventing the occurrence of bubbles (see Fama, 1965). Contrary to the efficient market prospective, Abreu and Brunnermeier (2003) consider an economy where rational agents would deliberately allow bubbles to persist. Hedge funds may have been aware of the mispricing of long-dated futures but could have opted to benefit from it in the short run rather than correct it.

Alternatively, Allen and Gorton (1993) provide a model in which a bubble arises in rational expectations equilibrium because of institutional investors' agency problem. Portfolio managers' payoffs have the form of a call option which will induce them to speculate on the future asset price path. In this context, our results are consistent with the existence of excess speculation in futures markets but can offer no definitive evidence on causality.

The contrasting results on bubbles for the spot and nearby contracts and those for the longer-dated futures contracts point to a potential disconnect between prices and fundamentals in the crude oil market. One possibility is that this might be driven by the type of rational bubble model outlined earlier. Another alternative is, that it might suggest a violation of market efficiency in the 2004-2008 period. Typically, market efficiency would imply cointegration between spot and longer-dated futures contracts which require both series to have a similar order of integration. Our results contrast with those of Büyükkahin et al. (2009). Their ADF test results for the period from July 2000 to August 2008 showed that nearby, 1- and 2-year futures prices were all $I(1)$ and their results indicated the prices were cointegrated with one cointegrating vector. One possible explanation is the following. A well-known limitation of Dickey-Fuller type tests is that when applied to the full sample and the data are described by the

generating process in Equation (11), with a coefficient close to but greater than one, the unit root null hypothesis cannot be properly assessed. Evans (1991) showed that periodically collapsing bubble processes behave like an $I(1)$ process and full-sample unit root tests have low power.

A final possibility is that the increase in crude oil spot and futures prices might be due to changes in expected fundamentals.¹¹ This type of rationale is given by Balke and Wohar (2001) for the stock price run-up in the 1990s. They show that after World War II, price-dividend and price-earnings ratios were quite persistent and subject to structural breaks. In this scenario, fundamentals can have a significant effect on price changes and even make price series appear explosive whilst adjusting towards the new equilibrium. They conclude that the 1990s stock price run-up can be rationalised by a combination of increasing expected real dividend (earnings) growth and declining expected future discount rates. Analogously, the crude oil futures price run up 2004-2008 may have reflected changes in future expected demand and supply fundamentals.

6 Conclusions

This paper investigates the time-series properties of spot and futures crude oil prices for the presence of mildly explosive bubbles. In particular, it applies a battery of tests to monthly and weekly data for a range of futures prices along the yield curve, from the nearby contract to the 24-month contract. The sample period extends from September 1995 to December 2013 and therefore includes the period of vastly increased investment flows into commodity derivative markets since the mid 2000s. Finally, the

¹¹ We thank an anonymous referee for pointing to this explanation.

paper also employs a procedure that allows for consistent identification of bubble origination and collapse dates.

The results are novel. The Phillips et al. (2013) *GSADF* multiple bubble test results indicate a disconnect between the spot and nearby contracts on one hand and the longer-dated oil futures contracts on the other hand. The latter provide significant evidence of extensive bubble periods. The prices for 12-month and longer series exhibited bubble behaviour from early 2004 up until late 2008, coinciding with the period of increased participation of financial investors including index trackers and hedge funds. The 2004-2008 bubble period was characterised by dramatic shifts in the yield curve. These are most dramatic at longer maturities and probably reflect the massive flows into managed money commodity spread positions in those years. It is quite plausible that the popularity of managed money spread positions highlighted by Singleton (2014) and Büyüksahin and Robe (2014) contributed to the price run up at longer maturities. Although a necessary condition, our new results do not necessarily infer that excessive speculation in commodity futures markets contributed towards price deviations away from fundamental levels in the physical market. Further empirical analysis is needed to clarify in more detail the exact source of bubble like behaviour. On the methodological side, future work could also usefully examine the effect of non-stationary volatility on tests for multiple bubbles.

Finally, our results have pertinent policy implications. Since bubbles appear earlier in longer dated futures contracts, they suggest that these contracts may provide additional information, over and above that of the more examined shorter-dated alternatives. Bubble tests on long maturity crude oil futures contracts may prove useful both for real-time monitoring and in terms of providing potential early warning signals

for bubble formation. Thus evidence on bubbles in long maturity oil futures contracts could form an input into macro-prudential policy for a bubble early warning system.

References

- Abreu, D., and M.K. Brunnermeier. 2003. Bubbles and crashes, *Econometrica* 71, 173-204.
- Acharya, V.V., L.A. Lochstoer, and T. Ramadora. 2012. Limits to arbitrage and hedging: Evidence from commodity markets, *Journal of Financial Economics* 109, 441-465.
- Allen, F., and G. Gorton. 1993. Churning bubbles, *Review of Economic Studies* 60, 813-836.
- Balke, N.S, and M.E. Wohar. 2001. Explaining stock price movements: Is there a case for fundamentals? *Federal Reserve Bank of Dallas Economic Review* (Third Quarter), 22-34.
- Blanchard, O.J., and M.W. Watson. 1982. Bubbles, rational expectations, and financial markets, in P. Wachtel, ed., *Crisis in the Economic and Financial Structure*, 295-315 (Lexington Books).
- Brunnermeier, M.K., and M. Oehmke. 2012. Bubbles, financial crises, and systemic risk, in G.M Constantinides, R. Stulz, and M. Harris, eds., *Handbook of the Economics of Finance* (Elsevier Science and Technology).
- Büyükaşahin, B., and M.A. Robe. 2014. Speculators, commodities and cross-market linkages. *Journal of International Money and Finance* 42, 38-70.
- Büyükaşahin, B., M.S. Haigh, J.H. Harris, J.A. Overdahl, and M.A. Robe. 2009. Fundamentals, trader activity and derivative pricing. Paper presented at the 2009 Meeting of the European Finance Association.
- Caballero, R.J., E. Farhi, and P.-O. Gourinchas. 2008. Financial crash, commodity prices and global Imbalances, *Brookings Papers on Economic Activity*, 1-55.
- Camerer, C. 1989. Bubbles and fads in asset prices, *Journal of Economic Surveys* 3, 3-41.

- Campbell, J.Y., and R.J. Shiller. 1987. Cointegration and tests of present value models, *Journal of Political Economy* 95, 1062-1088.
- Cheng, I-H., and W. Xiong. 2013. The financialization of commodity markets, Working Paper.
- Diba, B.T., and H.I. Grossman. 1988a. The theory of rational bubbles in stock prices, *Economic Journal* 98, 746-754.
- Diba, B.T., and H.I. Grossman. 1988b. Explosive rational bubbles in stock prices, *American Economic Review* 78, 520-530.
- Evans, G.W. 1991. Pitfalls in testing for explosive bubbles in asset prices, *American Economic Review* 81, 922-930.
- Fama, E.F. 1965. Behavior of stock-market prices, *Journal of Business* 38, 34-105.
- Gilbert, C.L. 2010. Speculative influences on commodity futures prices 2006-2008. United Nations Conference on Trade and Development, Discussion Paper 197.
- Hamilton, J. 2009. Causes and consequences of the oil shock of 2007-08, *Brookings Papers on Economic Activity Spring*, 215-259.
- Harvey, D.I., S.J. Leybourne, R. Sollis, and A.M.R. Taylor. 2015. Tests for explosive financial bubbles in the presence of non-stationary volatility. Working paper.
- Homm, U., and J. Breitung. 2012. Testing for speculative bubbles in stock markets: A comparison of alternative methods, *Journal of Financial Econometrics* 10, 198-231.
- Inclàn, C. and G.C. Tiao. 1994. Use of cumulative sums of squares for retrospective detection of changes of variance, *Journal of the American Statistical Association* 89, 913-923.
- Irwin, S.H., and D.R. Sanders. 2011. Index funds, financialization, and commodity futures markets, *Applied Economic Perspectives and Policy* 33, 1-31.

- 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.
- Krugman, P. 2008. More on oil and speculation, *New York Times*. Available online: <http://krugman.blogs.nytimes.com/2008/05/13/more-on-oil-and-speculation/>
- Lammerding, S.H., P. Stephan., M. Trede., and B. Wilfling. 2013. Speculative bubbles in recent oil price dynamics: Evidence from a Bayesian Markov-switching state-space approach, *Energy Economics* 36, 491–502.
- Masters, M.W. 2008. Testimony before the US Senate Committee of Homeland Security and Government Affairs, Washington (DC), May 20.
- Mou, Y, 2011. Limits to arbitrage and commodity index investment: Front-running the Goldman Roll. Available at SSRN: <http://ssrn.com/abstract=1716841>.
- Newey, W.K., and K. D. West. 1994. Automatic Lag Selection in Covariance Matrix Estimation, *Review of Economic Studies* 61, 631-653.
- Phillips, P.C.B., and J. Yu. 2011. Dating the timeline of financial bubbles during the subprime crisis, *Quantitative Economics* 2, 455-491.
- Phillips, P.C.B., S.P. Shi, and J. Yu. 2013. Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. Working paper.
- Phillips, P.C.B., and T. Magdalinos. 2007a. Limit theory for moderate deviations from unit root, *Journal of Econometrics* 136, 115-130.
- Phillips, P.C.B., and T. Magdalinos. 2007b. Limit theory for moderate deviations from unity under weak dependence, in G.D.A. Phillips, and E. Tzavalis, eds., *The refinement of econometric estimation and test procedures: Finite sample and asymptotic analysis*, 123-162 (Cambridge University Press).
- Phillips, P.C.B., Y. Wu, and J. Yu. 2011. Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values, *International Economic Review* 52, 201-226.

- Rapach, D.E., J.K. Strauss, and M.E. Wohar. 2008. Forecasting stock return volatility in the presence of structural breaks, in D.E. Rapach, and M.E., Wohar, eds., *Forecasting in the presence of structural breaks and model uncertainty*, 381-416 (Emerald Group Publishing Limited).
- Sansò A., V. Arragò, and J.L. Carrion. 2004. Testing for change in the unconditional variance of financial time series, *Revista de Economia Financiera* 4, 32-53.
- Shi, S., and V. Arora. 2012. An application of a model of speculative behaviour to oil prices, *Economics letters* 115, 469-472.
- Singleton, K.J. 2014. Investor flows and the 2008 boom/bust in oil prices, *Management Science* 60, 300-318.
- Sockin, M., and W. Xiong. 2015. Informational frictions and commodity markets, Princeton University. *Journal of Finance* 70, 2063-2098.
- Soros, G. 2008. Testimony before the US Senate Commerce Committee Oversight Hearing on FTC Advanced Rulemaking on Oil Market Manipulation, Washington (DC), June 4.
- Tang, K., and W. Xiong. 2012. Index investment and financialization of commodities, *Financial Analysts Journal* 68, 54-74.
- US Senate Permanent Subcommittee on Investigations. 2006. The role of market speculation in rising oil and gas prices: A need to put the cop back on the beat, Washington (DC), June 27.
- Vivian, A., and M.E. Wohar. 2012. Commodity Volatility Breaks, *Journal of International Financial Markets, Institutions and Money* 22, 395-422.

Table 1: Description of futures contracts

A description of the relationship between the last business day of the month on which we observe the price for our data series (*price month*), the month in which the futures contract stops trading (*expiration month*), and the month in which delivery takes place (*delivery month*). NYMEX crude oil futures expire on the third business day prior to 25th calendar day of the month proceeding the delivery month. If the 25th happens to be a non-business day, the expiration day is the third business day prior to the business day proceeding the 25th. $F_{j,t}$ denotes the j^{th} contract at time t . For illustrative purposes we choose $t = 1$, which represents the price on September 29, 1995.

$F_{j,t}$	Price month	Expiration month	Delivery month
$F_{1,1}$	1995M09	1995M10	1995M11
$F_{3,1}$	1995M09	1995M12	1996M01
$F_{6,1}$	1995M09	1996M03	1996M04
$F_{9,1}$	1995M09	1996M06	1996M07
$F_{12,1}$	1995M09	1996M09	1996M10
$F_{15,1}$	1995M09	1996M12	1997M01
$F_{18,1}$	1995M09	1997M03	1997M04
$F_{21,1}$	1995M09	1997M06	1997M07
$F_{24,1}$	1995M09	1997M09	1997M10

Table 2: Descriptive statistics

This reports the main characteristics of the monthly spot and futures log return series.

Statistics	Spot	Contract								
		1	3	6	9	12	15	18	21	24
Panel A: September 1995 – December 2013										
Mean	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.007	0.007
Std.	0.093	0.093	0.083	0.073	0.067	0.062	0.059	0.056	0.055	0.053
Skewness	1.421	1.470	1.903	2.452	3.016	3.421	3.731	3.725	3.605	3.351
Kurtosis	-0.453	-0.478	-0.518	-0.565	-0.654	-0.701	-0.716	-0.683	-0.623	-0.534
Panel B: September 1995 – December 2003										
Mean	0.006	0.006	0.006	0.006	0.006	0.005	0.005	0.005	0.005	0.005
Std.	0.095	0.097	0.081	0.064	0.053	0.046	0.040	0.037	0.035	0.033
Skewness	0.413	0.506	1.040	1.351	1.331	1.104	1.107	0.946	0.998	0.891
Kurtosis	-0.074	-0.086	-0.052	0.211	0.372	0.493	0.550	0.451	0.355	0.245
Panel C: January 2004 – June 2008										
Mean	0.027	0.027	0.028	0.029	0.029	0.030	0.030	0.030	0.030	0.030
Std.	0.078	0.078	0.073	0.069	0.065	0.063	0.061	0.061	0.060	0.059
Skewness	-0.992	-1.038	-1.154	-1.058	-0.953	-0.857	-0.775	-0.701	-0.597	-0.505
Kurtosis	-0.281	-0.233	-0.226	-0.218	-0.183	-0.125	-0.058	0.019	0.091	0.160
Panel D: July 2008 – December 2013										
Mean	-0.005	-0.005	-0.005	-0.006	-0.006	-0.006	-0.007	-0.007	-0.007	-0.007
Std.	0.099	0.099	0.092	0.086	0.081	0.077	0.073	0.070	0.068	0.065
Skewness	3.160	3.250	3.384	3.392	3.540	3.659	3.659	3.525	3.237	2.961
Kurtosis	-0.882	-0.992	-1.001	-1.080	-1.179	-1.247	-1.277	-1.260	-1.215	-1.170

Table 3: Testing for explosive behaviour in WTI crude oil spot and futures price series – monthly frequency

The *SDFC* and *SBT* test statistics are estimated over the sample period September 1995 to June 2008 (154 monthly observations) to enhance the power of the tests. The *SADF* and *GSADF* test statistics are estimated over the sample period September 1995 to December 2013 (220 monthly observations). All tests are calculated recursively with a fraction of the total sample $r_0 = 20\%$ (44 observations) for the initial window size. The right-sided critical values in Panel B are approximated using Monte Carlo simulations with 10,000 replications.

	<i>SADF</i>	<i>SDFC</i>	<i>SBT</i>	<i>GSADF</i>
Panel A: Test statistics				
Spot	3.675	3.863	1.341	3.675
Contract 1	3.641	3.838	1.325	3.641
Contract 3	4.120	4.210	1.475	4.120
Contract 6	4.675	4.721	1.637	4.675
Contract 9	5.052	5.102	1.724	5.052
Contract 12	4.707	5.350	1.766	5.297
Contract 15	5.241	5.515	1.787	5.461
Contract 18	5.522	5.635	1.797	5.580
Contract 21	5.807	5.750	1.810	5.807
Contract 24	5.919	5.831	1.828	5.919
Panel B: Critical values				
90%	1.041	0.709	0.788	1.584
95%	1.350	1.063	0.933	1.854
99%	1.980	1.749	1.241	2.362

Table 4: Testing for explosive behaviour in WTI crude oil spot and futures price series – bootstrapping and monthly series

The table reports the results of applying the bootstrap algorithm suggested by Harvey et al. (2015, p.11) to the $SADF(r_0)$ statistic defined in (12).

Contract	<i>p</i> -value
Spot	0.060
Contract 1	0.061
Contract 3	0.054
Contract 6	0.052
Contract 9	0.054
Contract 12	0.054
Contract 15	0.056
Contract 18	0.055
Contract 21	0.048
Contract 24	0.045

Table 5: Testing for explosive behaviour in WTI crude oil spot and futures price series - weekly frequency

This table is similar to Table 3 but we use weekly prices for the spot and futures contracts. *SDFC* and *SBT* test statistics are estimated over the sample period September 1995 to June 2008 to enhance the power of the tests. The *SADF* and *GSADF* test statistics are estimated over the full sample period September 1995 to December 2013 (954 weekly observations). All tests are calculated recursively with a fraction of the total sample $r_0 = 10\%$ for the initial window size. The right-sided critical values in Panel B are approximated using Monte Carlo simulations with 5,000 replications.

	<i>SADF</i>	<i>SDFC</i>	<i>SBT</i>	<i>GSADF</i>
Panel A: Test statistics				
Spot	3.393	3.602	2.186	2.972
Contract 1	3.266	3.832	2.566	3.266
Contract 3	3.867	4.263	3.231	3.867
Contract 6	4.534	4.723	4.060	4.534
Contract 9	4.957	4.986	4.603	4.957
Contract 12	5.228	5.239	4.920	5.228
Contract 15	5.407	5.419	5.108	5.407
Contract 18	5.522	5.535	5.192	5.522
Contract 21	5.771	5.636	5.247	5.623
Contract 24	5.924	5.691	5.228	5.678
Panel B: Critical values				
90%	1.233	1.037	0.879	1.923
95%	1.471	1.345	1.021	2.177
99%	2.042	2.025	1.364	2.679

Table 6: Volatility breaks – weekly frequency

Unconditional variance break dates are summarised below. We use a response surface as in Sansò et al. (2004) to generate the 5% critical values.

Contract	Weekly break dates (month/year)
Spot	Nov 1999, Sep 2004, Oct 2007
Contract 1	Dec 1999, Sep 2004, Oct 2007
Contract 3	Jan 2000, Sep 2004, Oct 2007
Contract 6	Jan 2000, Feb 2004, Jul 2005, Oct 2007
Contract 9	Jan 2000, Mar 2004, Jun 2005, Nov 2007
Contract 12	Jan 2000, Feb 2004, Aug 2005, Nov 2007
Contract 15	Feb 2000, Feb 2004, May 2005, Oct 2007
Contract 18	Mar 2000, Mar 2004, Jul 2005, Oct 2007
Contract 21	May 2000, Apr 2004, Jul 2005, Dec 2007
Contract 24	May 2000, Feb 2004, Jul 2005, Oct 2007

Table 7: Testing for explosive behaviour in WTI crude oil spot and futures price series – bootstrapping and weekly frequency

The table reports the results of applying the bootstrap algorithm suggested by Harvey et al. (2015, p.11) to the SADF(r_0) statistic defined in (12).

Contract	<i>p</i> -value
Spot	0.131
Contract 1	0.108
Contract 3	0.082
Contract 6	0.070
Contract 9	0.067
Contract 12	0.069
Contract 15	0.069
Contract 18	0.071
Contract 21	0.072
Contract 24	0.072

Figure 1: Futures yield curves

The figure depicts the average basis (*solid line*) of all contracts at different maturities along the yield curve over the three subsample periods – September 1995 to December 2003 (Panel A), January 2004 to June 2008 (Panel B), and July 2008 to December 2013 (Panel C).

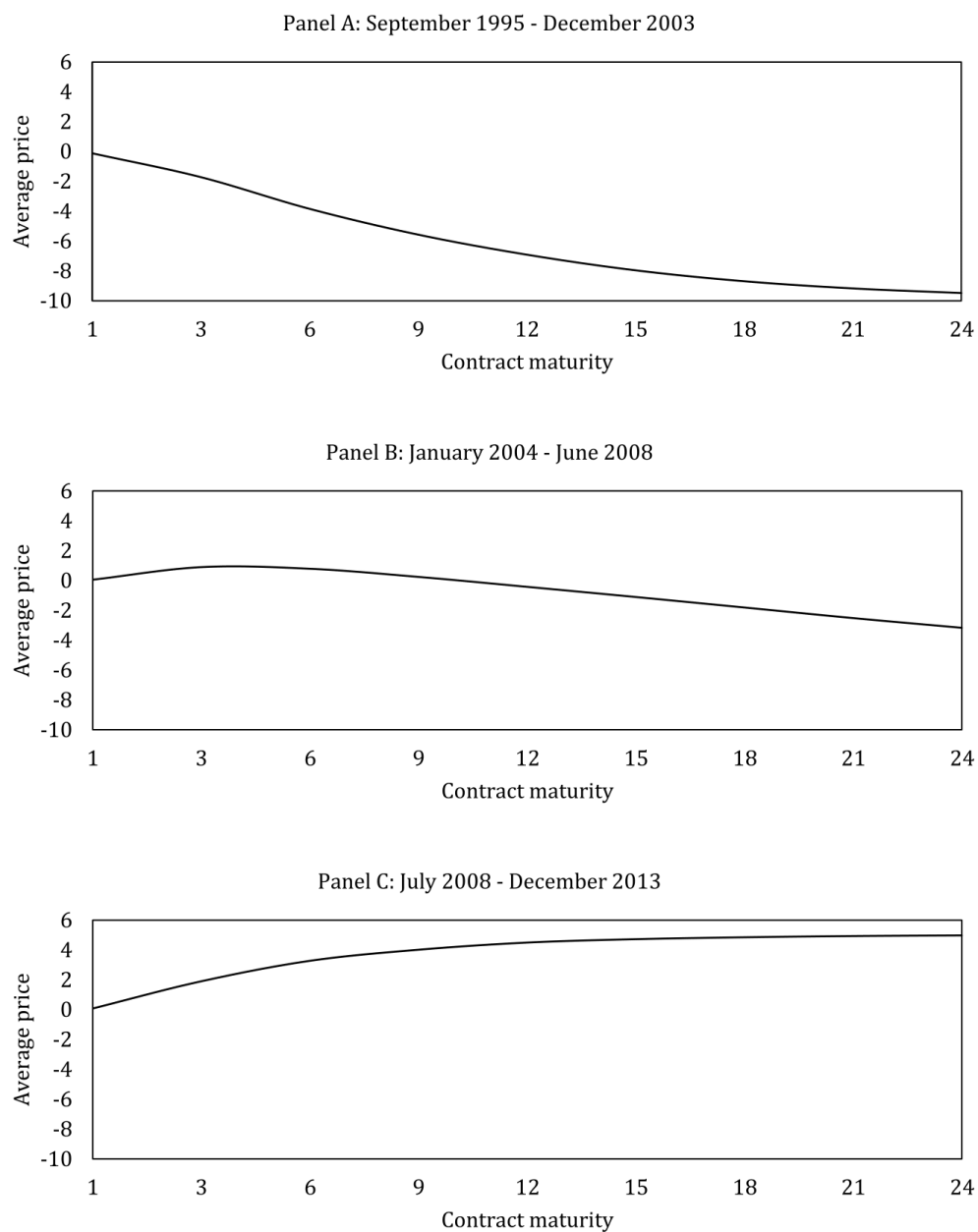


Figure 2: Bubble periods at the 5% significance level

The backward *SADF* statistic sequence (*solid line*) is plotted against the 95% critical value sequence (*dashed line*). The initial sample for the *BSADF* procedure is chosen to be 20% of the total sample (44 observations). The estimations are performed over the sample period September 1995 to December 2013. The right-sided critical 95% values are approximated using Monte Carlo simulations with 10,000 replications.

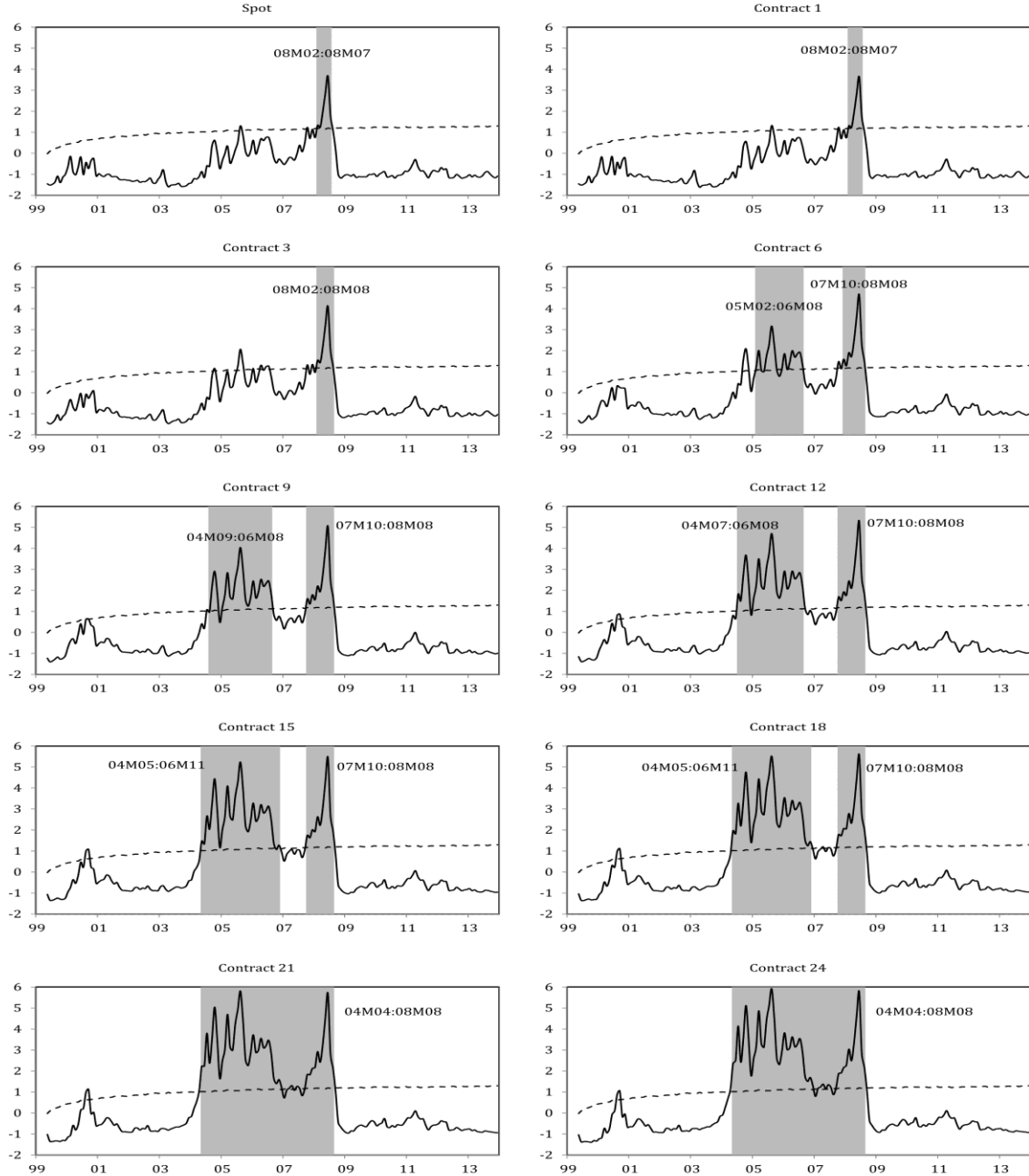


Figure 3: Bubble periods at the 10% significance level

The backward *SADF* statistic sequence (*solid line*) is plotted against the 90% critical value sequence (*dashed line*). The initial sample for the *BSADF* procedure is chosen to be 20% of the total sample (44 observations). The estimations are performed over the sample period September 1995 to December 2013. The right-sided critical 90% values are approximated using Monte Carlo simulations with 10,000 replications.

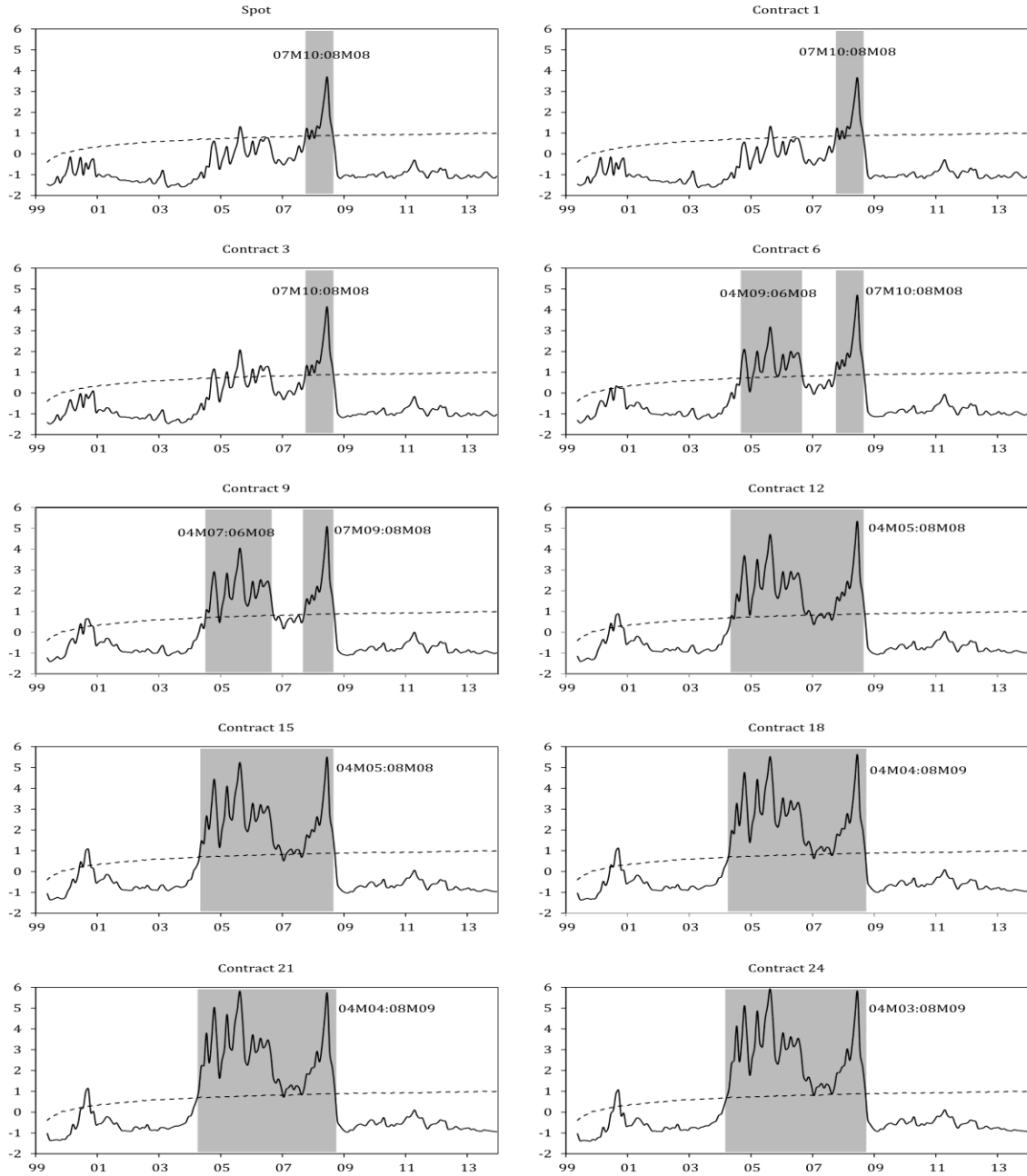


Figure 4: Volatility breaks

Unconditional volatility break dates are summarised along with the price series of each contract. The vertical line represents the structural break point in the unconditional variance of each corresponding series. We use a response surface as in Sansò et al. (2004) to generate the 5% critical values.

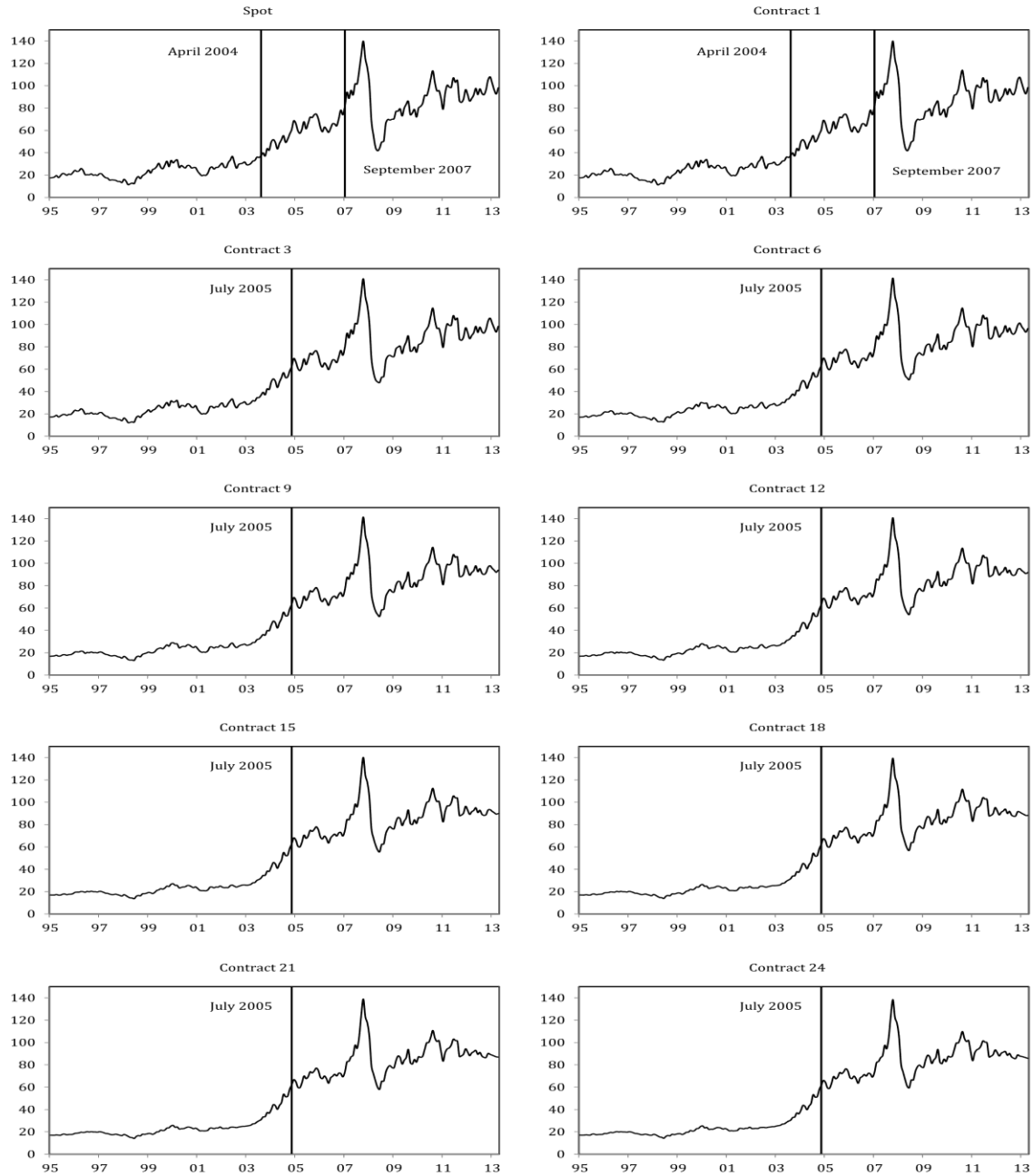


Figure 5: Bubble periods at the 5% significance level - weekly frequency

The backward *SADF* statistic sequence (*solid line*) is plotted against the 95% critical value sequence (*dashed line*). The initial sample for the *BSADF* procedure is chosen to be 20% of the total sample (96 observations). The estimations are performed over the sample period September 1995 to December 2013. The right-sided critical 95% values are approximated using Monte Carlo simulations with 10,000 replications.

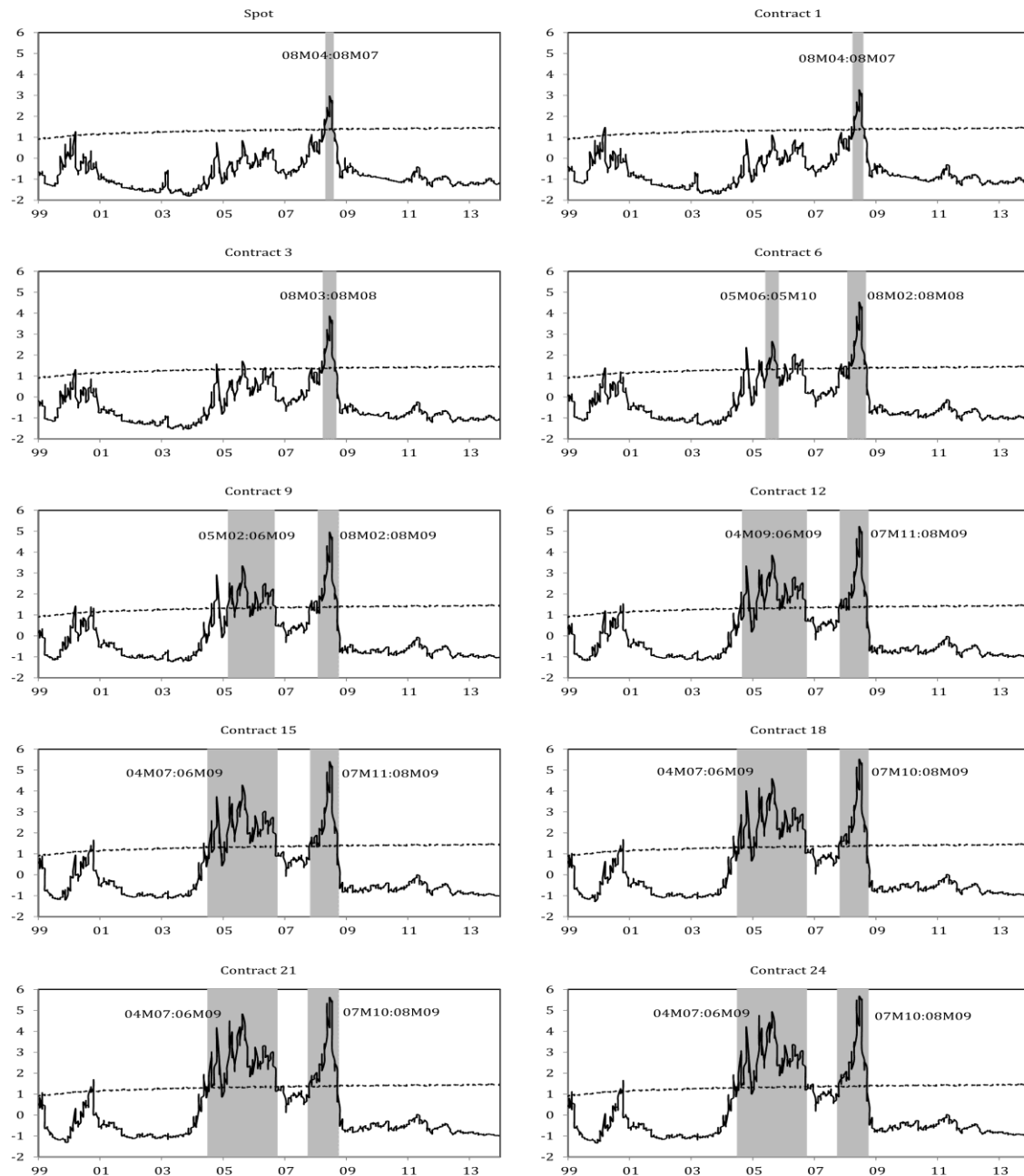


Figure 6: Bubble periods at the 10% significance level - weekly frequency

The backward *SADF* statistic sequence (*solid line*) is plotted against the 90% critical value sequence (*dashed line*). The initial sample for the *BSADF* procedure is chosen to be 20% of the total sample (96 observations). The estimations are performed over the sample period September 1995 to December 2013. The right-sided critical 90% values are approximated using Monte Carlo simulations with 10,000 replications.

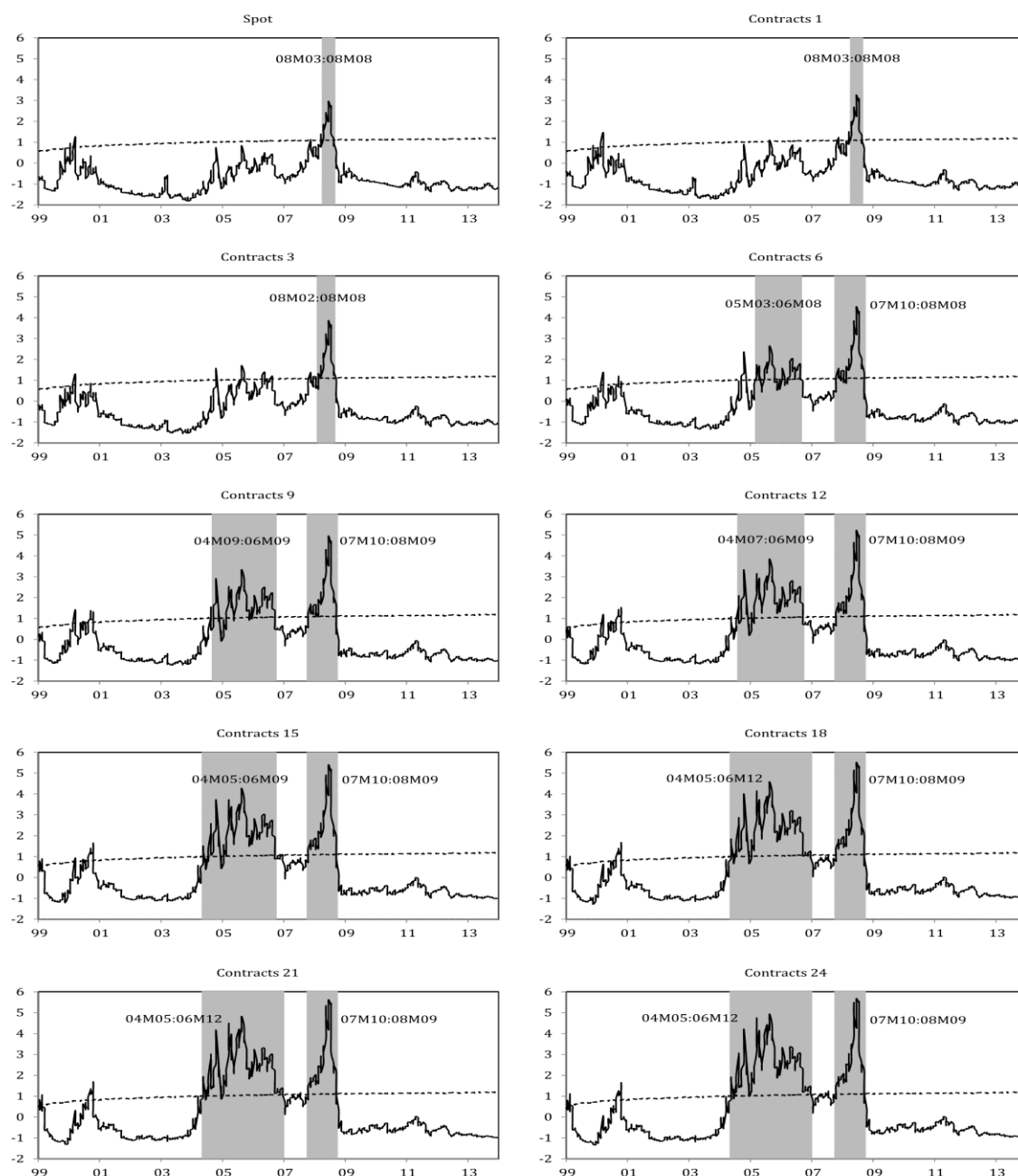


Figure 7: Annual futures yield curves 2004-2008

This depicts the average basis of all contracts at different maturities for each year along the yield curve from January 2004 (*solid line*) to June 2008.

