Abstract

Despite the successful resolution to accommodate various forms of heteroskedasticity and to limit the size of test, the accuracy of real time detection by Astill et al. (2018) suffers greatly when crash is present in the training period, as the test strongly relies on the training period statistic. Therefore, the paper introduces the novel methods to overcome the shortcoming of the real time monitoring procedure: 1) crash data generating process, 2) historic crash dating and 3) enhanced monitoring procedure excluding crash data from the training period. Via simulation and real data application, the enhanced monitoring is proven that when a crash series is present in the training period data, it outperforms that of Astill et al. (2018).

Dating for start and end period of historical crash and its application to real time bubble and crash detection

Name : Taejun Kim

Student Number : 20388559

Supervisor : Stephen J. Leybourne and David I. Harvey

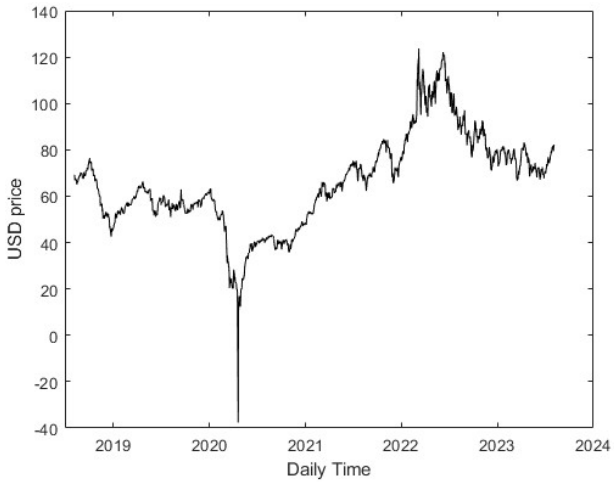
The Dissertation is presented in part fulfilment of the requirement for the completion of an MRes in the School of Economics, University of Nottingham. The work is the sole responsibility of the candidate.

Contents

1 Introduction

Asset price bubble has been continuously highlighted in various financial economic studies. Following Diba & Grossman (1988), attempts to quantify the bubble behaviour through data generating processes and to detect the change point from unit root to bubble behaviour, have been extensively introduced. While most of the studies deal with the positive explosive behaviour, there has been less emphasis on crash series, negative explosive behaviour. However, Fantazzini (2016) proposes a method employing a generalized supremum augmented dickey fuller test by Phillips et al. (2015) to study the negative explosive behaviour of oil price between 2020 and 2021 and to address a presence of negative explosive behaviour.

Figure 1: Historic Daily Oil Pirce



While the real time detection is of concern, Whitehouse et al. (2023) propose a real time monitoring method for a crash conditional on detecting bubble before identifying it. Considering the crash in price series on the figure 1 as it is not necessarily followed by the bubble process, detection methodology for an unconditionally occurring crash is desirable. Moreover, while the method by Whitehouse et al. (2023) relies on the statistic

by Astill et al. (2017) and the real time monitoring procedure by Harvey et al. (2021), an assumption that there is no bubble and/or no crash in the training period is needed. If a significant proportion of data in the training period contains a crash and a recovery from it, it can severely distort the accuracy of the real time monitoring method.

Therefore, the paper has its novelty in 1) introducing a new data generating process that characterizes a crash and a recovery series before a bubble or a crash, 2) proposing a method to accurately date the start of a crash and the end of recovery from the training period, and 3) introducing a real time monitoring process with enhanced accuracy by removing the historical crash date from the training period. Through simulations and real world application, the paper discovers that the longer the crash and recovery regimes and the steeper the magnitudes of them are, the more distorted the accuracy is.

The paper consists of 7 sections. Section 1 is the introduction. Section 2 discusses the chronical evolution of econometric methods regarding bubble and crash detection. Section 3 contains the novel data generating process. Section 4 introduces the econometric methods to be employed. Section 5 conducts the Monte Carlo simulations for robustness check of the methodology. Section 6 undertakes the real data application of the methodology. Section 7 concludes the paper with the overall remarks.

2 Literature Review

In research seminal, Diba & Grossman (1988) introduce the conventional left-tailed unit root tests, the Dickey-Fuller (DF) test (Dickey & Fuller, 1979), as a means to detect the asset price bubble. The test is applied to both levels and first differences of data. While explosive process cannot be differenced to stationarity, the rejection of H_0 : the first differenced series is stationary, signals the explosive behaviour. However, Evans (1991) argues that in case of periodically collapsing bubbles, the procedure by Diba & Grossman

(1988) has insufficient power.

Alternatively, Phillips et al. (2011) introduces a forward recursive right-tailed DF test (PWY test). The procedure uses the supremum of a set of the test applied to the stock price and dividend indices in levels as a means to compare with the critical value. The procedure signals a bubble in a price index when a stock price index rejects H_0 : the series are a unit root process by PWY test while a corresponding dividend index fails to reject H_0.

While the methods discussed above are to detect the historical bubble that already occurred in the data, detecting the bubble at the end of a sample is of interest. Phillips et al. (2015) come up with a generalized version of the PWY with a minimum sample size (PSY test) and introduce a backward recursive right-tailed DF test for its dating purpose. This procedure performs with better power than that of the PWY test when detecting the end of sample bubble.

While the backward procedure has a drawback on its asymptotic validity as it assumes the end of sample bubble is a none vanishing fraction of the sample, Astill et al. (2017) propose the test procedure based on the end of sample co-integration breakdown test by Andrews (2003) and Andrews & Kim (2006), and for the asymptotic validity, the procedure assumes a finite length of end of sample bubble regime. Further assuming first difference stationarity in the sub-sample data, the Astill et al. (2017) procedure presents greater power than that of Phillips et al. (2015) for the end of sample bubble detection. Moreover, the test can accommodate shortly lived sub-sample period bubbles and finite variance breaks as they are asymptotically negligible when constructing the critical value.

Greater interests lie on real time detection of a bubble. However, sequential application of the above mentioned test is not size controlled as the overall false-positive rate (FPR), the probability of falsely rejecting H_0, is not specifiable. Therefore, it creates well-known multiplicity problem, meaning FPR becomes inflated as monitoring horizon grows.

Homm & Breitung (2012) devise a cumulative sum of squares (CUSUM) based method as a solution, and under certain conditions, the procedure manages to control FPR with an expanding monitoring horizon. However, their asymptotic critical value is highly conservative. Instead of the conservative critical value, they recommend to employ finite-sample critical values simulated from a Gaussian random walk, but asymptotically, FPR does not rely on the normality assumption. Therefore, this assumption would affect the procedure with finite sample sizes. Moreover, the procedure is based on the assumption that the driving shocks are unconditionally homoscedastic and serially uncorrelated.

Harvey et al. (2016) argue under time-varying volatility, the CUSUM procedure generates inflated empirical FPR. While volatility clustering is commonly shown on the financial time series (Engle, 1982), Astill et al. (2018) adopts the methods of Astill et al. (2017) and Harvey et al. (2021) to solve the FPR problem. Based on the uniformity argument of Harvey et al. (2021), they propose a procedure to limit the empirical FPR by restricting the monitoring end point. Their procedure resultantly outperforms the CUSUM procedure under heteroskedasticity and non-stationary volatility.

However, as the Astill et al. (2018) procedure relies on the Astill et al. (2017) statistic and, more importantly, the MAX procedure by Harvey et al. (2021), a significantly large period of bubble or crash and a following crash or recovery from them in the training period can affect its true positive rate (TPR), the probability of correctly detecting bubble from the start of monitoring period, as those data inflate the training period statistic. Therefore, it is desirable to identify the bubble or crash in the training period.

While the backward reclusive DF method by Phillips et al. (2015) can be used for historic bubble dating, Harvey et al. (2017) shows that conditioning on the bubble detection, their procedure has greater accuracy than that of Phillips et al. (2015).

3 Data Generating Process (DGP)

3.1 Bubble DGP

$$y_t = y_0 + u_t \tag{1}$$

$$u_t = \begin{cases} u_{t-1} + \varepsilon_t & t = 2, ..., \lfloor \tau_1 T \rfloor \\ (1+\delta_1)u_{t-1} + \varepsilon_t & t = \lfloor \tau_1 T \rfloor + 1, ..., \lfloor \tau_2 T \rfloor \\ (1-\delta_2)u_{t-1} + \varepsilon_t & t = \lfloor \tau_2 T \rfloor + 1, ..., \lfloor \tau_3 T \rfloor \\ u_{t-1} + \varepsilon_t & t = \lfloor \tau_3 T \rfloor + 1, ..., \lfloor \tau_4 T \rfloor \end{cases} \tag{2}$$

The basic frame of DGP process follows that of Harvey et al. (2017). $0 \leqslant \tau_1 \leqslant \tau_2 \leqslant \tau_3 \leqslant \tau_4 \leqslant 1$ while $\lfloor \cdot \rfloor$ denotes the floor function. $\delta_1 \geqslant 0$ and $\delta_2 \geqslant 0$, and $y_0 = O_p(1)$. Importantly, the sequence $\{\varepsilon_t\}$ satisfies the assumption that the stochastic process is such that $\varepsilon_t = C(L)\eta_t$ and $C(L) := \sum_{j=0}^{\infty} C_j L^j$ while $C(1)^2 > 0$ and $\sum_{j=0}^{\infty} i|C_i| < \infty$, and the sequence $\{\eta_t\}$ is $\eta_t \sim IID(0,1)$ with $E(\eta_t^4) < \infty$. This assumption accommodates some cases of heteroskedasticity in DGP and to be able to conduct correct inference. Moreover, Harvey et al. (2017) imposes $(1+\delta_1)^{(\tau_2-\tau_1)}(1-\delta_2)^{(\tau_3-\tau_2)} \geqslant 1$, as a restriction. As mean reversion feature is the driving force to characterize the exponential crash, if left unrestricted, the process will flatten out and behave like a zero mean stationary process, thus by the restriction, the process can maintain the exponential crash as its dominant feature even after the crash ends.

DGP in (1) and (2) presents a unit root behaviour during the period, t, from 1 to $\lfloor \tau_1 T \rfloor$. This behaviour changes to an explosive process from $\lfloor \tau_1 T \rfloor + 1$ to $\lfloor \tau_2 T \rfloor$. Then stationary collapsing regime takes place from $\lfloor \tau_2 T \rfloor + 1$ to $\lfloor \tau_3 T \rfloor$, and a unit root behaviour starts

from $\lfloor \tau_3 T \rfloor + 1$ and lasts until $\lfloor \tau_4 T \rfloor$. An example DGP is plotted using GAUSS 19.1 in the figure 2-(a) with $T = 250$, $y_0 = 100$, $\tau_1 = 0.2$, $\tau_2 = 0.4$, $\tau_3 = 0.45$, $\tau_4 = 0.8$, $\delta_1 = 0.01$, $\delta_2 = 0.02$, $\varepsilon_t \sim IIDN(0,1)$ and $seed = 3$

$$y_t^i = -y_t \;\; while \;\; t = 1, ..., \lfloor \tau_4 T \rfloor \tag{3}$$

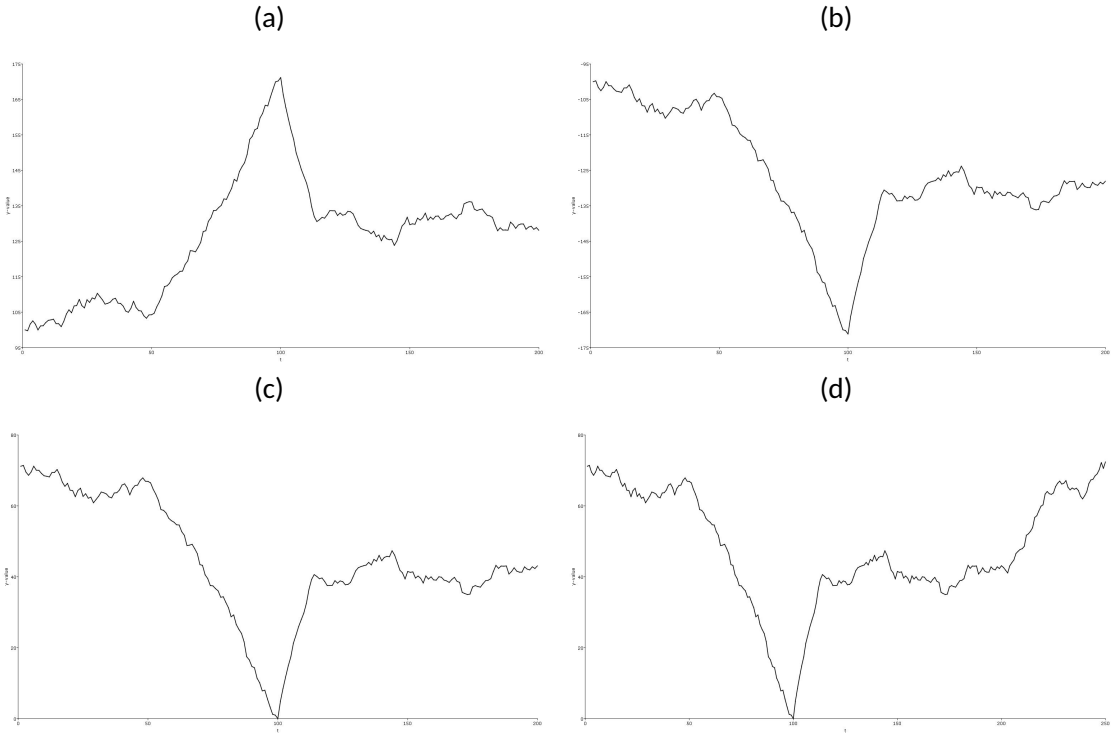
$$\mu = |min(y_t^i)| \;\; while \;\; t = 1, ..., \lfloor \tau_4 T \rfloor \tag{4}$$

$$y_t^b = \begin{cases} y_{t-1}^i + \varepsilon_t + \mu & t = 1, ..., \lfloor \tau_1 T \rfloor \\ (1+\delta_1) y_{t-1}^i + \varepsilon_t + \mu & t = \lfloor \tau_1 T \rfloor + 1, ..., \lfloor \tau_2 T \rfloor \\ (1-\delta_2) y_{t-1}^i + \varepsilon_t + \mu & t = \lfloor \tau_2 T \rfloor + 1, ..., \lfloor \tau_3 T \rfloor \\ y_{t-1}^i + \varepsilon_t + \mu & t = \lfloor \tau_3 T \rfloor + 1, ..., \lfloor \tau_4 T \rfloor \\ (1+\delta_3) y_{t-1} + \varepsilon_t & t = \lfloor \tau_4 T \rfloor + 1, ..., T \end{cases} \tag{5}$$

To characterize a crash, the paper introduces two steps. In (3), it is to inverse the DGP created via (1) and (2), then it is shown on the figure 2-(b). To make the the value of y_t realistic, in (4), it computes the absolute value of the minimum of inverted data and adds the value to the inverted process as in (5), and the visualization is on the figure 2-(c).

Lastly, to add a bubble, from $\lfloor \tau_4 T \rfloor$ to the end of sample period, T, an explosive behaviour is generated as in (5). With $\delta_3 = 0.01$, the combined data are plotted in the figure 2-(d). The plot demonstrates the similar behaviour as in the figure 1. The notable features of the behaviour are i) the duration of crash, $\lfloor \tau_2 T \rfloor - \lfloor \tau_1 T \rfloor$, is longer than that of recovery, $\lfloor \tau_3 T \rfloor - \lfloor \tau_2 T \rfloor$, and ii) the steep of recovery, δ_2, is sharper than that of crash, δ_1.

Figure 2: Unit root - Crash - Recovery - Unit root - Bubble (UCRUB) DGP



3.2 Crash DGP

$$y_t = y_0 + u_t \tag{6}$$

$$u_t = \begin{cases} u_{t-1} + \varepsilon_t & t = 2, ..., \lfloor \tau_1 T \rfloor \\ (1 + \delta_1)u_{t-1} + \varepsilon_t & t = \lfloor \tau_1 T \rfloor + 1, ..., \lfloor \tau_2 T \rfloor \\ (1 - \delta_2)u_{t-1} + \varepsilon_t & t = \lfloor \tau_2 T \rfloor + 1, ..., \lfloor \tau_3 T \rfloor \\ u_{t-1} + \varepsilon_t & t = \lfloor \tau_3 T \rfloor + 1, ..., \lfloor \tau_4 T \rfloor \\ (1 + \delta_3)u_{t-1} + \varepsilon_t + \mu_2 & t = \lfloor \tau_4 T \rfloor + 1, ..., T \end{cases} \tag{7}$$

Similar procedure is adopted for the crash DGP. (6) and (7) generate a unit root behaviour during the period, t, from 1 to $\lfloor \tau_1 T \rfloor$. From $\lfloor \tau_1 T \rfloor + 1$ to $\lfloor \tau_2 T \rfloor$, it manifests an

explosive behaviour. Then from $\lfloor \tau_2 T \rfloor + 1$ to $\lfloor \tau_3 T \rfloor$, a stationary collapse follows. From $\lfloor \tau_3 T \rfloor + 1$ to $\lfloor \tau_4 T \rfloor$, it comes back to a unit root process, then from $\lfloor \tau_4 T \rfloor + 1$ to the end of sample, again it behaves in an explosive manner. An example DGP is plotted using GAUSS 19.1 in the figure 3-(a) with $y_0 = 100$, $\tau_1 = 0.2$, $\tau_2 = 0.4$, $\tau_3 = 0.45$, $\tau_4 = 0.8$, $T = 250$, $\delta_1 = 0.01$, $\delta_2 = 0.02$, $\delta_3 = 0.01$ and seed of 3.

$$y_t^i = -y_t \;\; while \;\; t = 1, ..., T \tag{8}$$

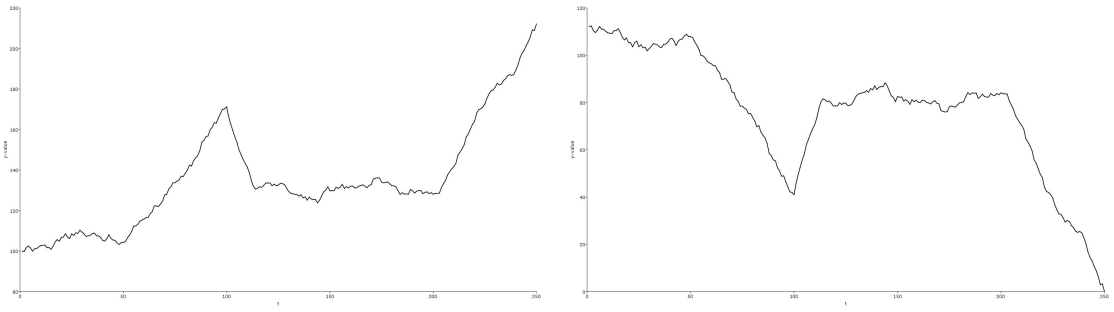
$$\mu = |min(y_t^i)| \;\; while \;\; t = 1, ..., T \tag{9}$$

$$y_t^c = \begin{cases} y_{t-1}^i + \varepsilon_t + \mu & t = 1, ..., \lfloor \tau_1 T \rfloor \\ (1 + \delta_1) y_{t-1}^i + \varepsilon_t + \mu & t = \lfloor \tau_1 T \rfloor + 1, ..., \lfloor \tau_2 T \rfloor \\ (1 - \delta_2) y_{t-1}^i + \varepsilon_t + \mu & t = \lfloor \tau_2 T \rfloor + 1, ..., \lfloor \tau_3 T \rfloor \\ y_{t-1}^i + \varepsilon_t + \mu & t = \lfloor \tau_3 T \rfloor + 1, ..., \lfloor \tau_4 T \rfloor \\ (1 + \delta_3) y_{t-1}^i + \varepsilon_t + \mu & t = \lfloor \tau_4 T \rfloor + 1, ..., T \end{cases} \tag{10}$$

To characterize the unit root-crash-recovery-unit root-crash (UCRUC) process, data are inverted for the entire time period as in (8), and to make the value realistic, the absolute value of minimum of the inverted data is added to the inverted series as in (10). The plot accommodating all discussed modification is on the figure 3-(b).

Figure 3: Unit root - Crash - Recovery - Unit root - Crash (UCRUC) DGP

(a) (b)



4 Econometric Method

The paper considers broadly two econometric methodologies: Bayesian Information Criteria (BIC) model selection method proposed by Harvey et al. (2017) and Real time bubble and crash detection by Astill et al. (2018) and Whitehouse et al. (2023).

4.1 BIC Crash Dating

Under an assumption that bubble is present in data, the DGP in (1) and (2) can produce 4 different cases based on the value of δ_1, δ_2 and τ_i where $i = 1, 2, 3, 4$. However, unlike Harvey et al. (2017), the aim for the study is to correctly specify the start period of crash and the end period of recovery, thus the paper considers only one case of DGP that UCRUB(or UCRUC) is characterized whereas $\delta_1 > 0$, $\delta_2 > 0$ and $0 < \tau_1 < \tau_2 < \tau_3 < \tau_4 < 1$. Most importantly the inverted series from crash to bubble are considered for the implementation of the methodology. Moreover, the paper sets τ_4 as the end of sample period for the bubble dating, thus regardless of the behaviour the series has after $\lfloor \tau_4 T \rfloor$, it yields the similar results between UCRUB and UCRUC.

Assuming that DGP is the true data generating process, the paper employs two models, one dummy model, $Model\ 1$, fitted based on Unit root-Crash-Unit root (UCU) case

and the true model, $Model\ 2$, representing UCRU, to fit with the data while $D_t(a,b) = \mathbb{1}(\lfloor aT \rfloor < t \leqslant \lfloor bT \rfloor)$ with $\mathbb{1}$ being an indicator function whereas $\hat{y}_{01}$ and $\hat{y}_{02}$ are imposed to prevent $\hat{\varepsilon}_{1t}$ and $\hat{\varepsilon}_{2t}$ from varying from the series mean, y_0.

$$Model\ 1\text{: } \Delta y_t = \hat{y}_{01} D_t(\tau_1, \tau_2) + \hat{\delta}_1 D_t(\tau_1, \tau_2) y_{t-1} + \hat{\varepsilon}_{1t}$$
$$Model\ 2\text{: } \Delta y_t = \hat{y}_{01} D_t(\tau_1, \tau_2) + \hat{y}_{02} D_t(\tau_2, \tau_3) + \hat{\delta}_1 D_t(\tau_1, \tau_2) y_{t-1} + \hat{\delta}_2 D_t(\tau_2, \tau_3) y_{t-1} + \hat{\varepsilon}_{2t}$$

These fitted models serve to select the regime change points along with all possible dates, τ_1, τ_2 and τ_3, that minimize sum of squared residual, $SSR_i(\cdot) = \sum_{t=2}^{T} \hat{\varepsilon}_{it}^2$ with $i = \{1, 2\}$. While estimating the regime change estimators, there are two restrictions imposed: $y_{\lfloor \tau_2 T \rfloor} > y_{\lfloor \tau_1 T \rfloor}$ and $y_{\lfloor \tau_2 T \rfloor} > y_{\lfloor \tau_3 T \rfloor}$. These conditions makes sure the periods between $\lfloor \tau_1 T \rfloor$ and $\lfloor \tau_2 T \rfloor$ be an explosive regime while the periods between $\lfloor \tau_2 T \rfloor$ and $\lfloor \tau_3 T \rfloor$ are a stationary collapse regime. The consistency property of these estimators are proven on Theorem 1 of Harvey et al. (2017).

$$Model\ 1\text{: } (\hat{\tau}_1, \hat{\tau}_2) = arg \min_{0<\tau_1<\tau_2<1, y_{\lfloor \tau_2 T \rfloor} > y_{\lfloor \tau_1 T \rfloor}} SSR_1(\tau_1, \tau_2)$$
$$Model\ 2\text{: } (\hat{\tau}_1, \hat{\tau}_2, \hat{\tau}_3) = arg \min_{0<\tau_1<\tau_2<\tau_3<1, y_{\lfloor \tau_2 T \rfloor} > y_{\lfloor \tau_1 T \rfloor}, y_{\lfloor \tau_2 T \rfloor} > y_{\lfloor \tau_3 T \rfloor}} SSR_2(\tau_1, \tau_2, \tau_3)$$

To obtain the efficient change points, Harvey et al. (2017) adopt BIC model selection method. The method is to compute the BIC values based on the estimated change points of each model and to select a model with change points that minimizes the BIC value between the two.

$$i_{opt} = arg \min_{i \in \{1,2\}} BIC_i$$
$$BIC_1 = T \ln\{T^{-1} SSR_1(\hat{\tau}_1, \hat{\tau}_2)\} + (2+2)\ln(T)$$
$$BIC_2 = T \ln\{T^{-1} SSR_2(\hat{\tau}_1, \hat{\tau}_2, \hat{\tau}_3)\} + (4+3)\ln(T)$$

Moreover, following Harvey et al. (2017), on the dating algorithm, except for the final

regime, the paper imposes the minimum required length to each regime to be specified as a changed regime from one another. In particular, if the proportion of crash in the sample is $s_c = O(T)$, then the proportion of recovery in the sample is $s_r < s_c$. This restriction allows gradually recovery rather than an instantaneous recovery and prevents the instantaneous recovery from signaling a different model. Harvey et al. (2017) argue that this feature is empirically more appealing as economic agents does not react in unison thus the price adjustment takes place gradually.

4.2 Real Time Monitoring

Astill et al. (2018) propose sequential application of Astill et al. (2017) test for an end of sample bubble in the Harvey et al. (2021) framework. The test statistic is a variant of that of Andrews (2003) and Andrews & Kim (2006) for end of sample co-integration breakdown detection. The test is based on the assumption that no bubble exists in its training period, up to T^*, but as shortly-lived bubble or finite variance breaks are asymptotically negligible when computing the critical value, some degree of abnormalities in the training period can be accommodated. Astill et al. (2017) further implement studentization and white-type correction to the test statistic to robustify for unconditional heteroskedasticity, and this takes a form of (11).

$$S_{e,m} = \frac{\sum_{t=e-m+1}^{e}(t-e+m)\Delta y_t}{\sqrt{\sum_{t=e-m+1}^{e}\{(t-e+m)\Delta y_t\}^2}} \tag{11}$$

m refers to the user-chosen finite-length of window, and e is the most recent time period of y_t. The test procedure is as follows: **1**. Calculating training period statistics up until the end of training period, T^*, such that $S_{e,m}$ for $e = m+1, \dots T^*$. **2**. based on the training period statistics, calculating a critical value, cv_π, given significance level, π, such that $cv_\pi = S_{\lfloor(1-\pi)(T^*-m)\rfloor}$ where $S_j, j = 1, \dots, T^* - m$. **3**. Comparing cv_π with the first

statistic out of the training period, $S_{T^*+m,m}$, and if $cv_\pi < S_{T^*+m,m}$, reject H_0 of bubble is not detected at the end of sample.

While a sequential application of Astill et al. (2017) is of interest, false positive rate (FPR), the probability of falsely detecting a bubble from the start of monitoring, increases co-linearly with the increase in the monitoring frame. Therefore, Astill et al. (2018) adopt the real time monitoring procedure proposed by Harvey et al. (2021) to restrict FPR by predetermining the monitoring frame. The procedure consists of three test, MAX, SEQ, and UNI.

MAX employs the maximum training period statistic, $S^*_{max} = max_{e\in[m+1,T^*]}S_{e,m}$, as a means to compare with the statistic calculated after the monitoring start period, $T^\dagger = T^* + m$. If $S^*_{max} < S_{T^\dagger,m}$, reject H_0 of bubble is not detected during the monitoring period and terminate the monitoring. If not, monitoring continues from $T^* + m + 1$ to $T^{'}$ whereas $T^{'}$ is a period when $S^*_{max} < S_{T^{'},m}$ must hold.

Alternatively, SEQ makes use of the critical value computed during the training period, and instead of simply using the value as a means to compare, it uses the length of sequential period where the statistic in the training period exceeds the critical value. On application, $R_e = \mathbb{1}(S_{e,m} > cv_\pi)$ is used to compute the length of a sequence of contiguous exceedances, such that $R(L,U) = (U - L + 1)\prod_{e=L}^{U} R_e$ where $U \geqslant L$. Then, defining the longest contiguous sequence of exceedances in both the training period, p^*, and the monitoring period, $p^{'}$, such that

$$p^* = \max_{L,U\in[m+1,T^*]} R(L,U) \quad p^{'} = \max_{L,U\in[T^*+m,T]} R(L,U)$$

If $p^* < p^{'}$, reject H_0 of bubble is not detected during the monitoring period.

These two methods are subject to the uniformity argument proposed by Harvey et al. (2021). The argument serves to limit the increase of FPR by restricting the monitoring

frame. Under H_0 of bubble is not detected during the monitoring period, the theoretical FPR takes a form of (12), the limit ratio of the number of $S_{e,m}$ computed in the monitoring period to the number of $S_{e,m}$ computed across the whole period.

$$FPR = \lim_{T^*,T' \to \infty} \frac{T' - T^* - m + 1}{T' - 2m + 1} \tag{12}$$

(12) allows to approximate FPR, and this enables to calculate an end point of monitoring for a given FPR level. By setting a desirable FPR level, a desirable monitoring end point can be set as in (13).

$$FPR \approx \frac{T' - T^* - m + 1}{T' - 2m + 1} \Rightarrow T' \approx \frac{T^* + m - 1 - FPR(2m-1)}{1 - FPR} \tag{13}$$

If either MAX or SEQ rejects H_0 of bubble is not detected during the monitoring period, it triggers UNI, a union of rejections where UNI rejects H_0. FPR of UNI is not subject to the uniformity argument unlike MAX or SEQ, and is bigger than that of the two. However, the procedure is still based on the same Astill et al. (2017) statistics, it does not differ much from the other two procedures.

Additionally, the paper introduces the real time crash detection tests, MIN and SEQ_c, which are crash analogous of Astill et al. (2018), inspired by Whitehouse et al. (2023). Whitehouse et al. (2023) present the minimum statistic to monitor a crash conditioning that MAX rejects H_0 of bubble is not detected during the monitoring period. Inspired by Whitehouse et al. (2023), the paper introduces MIN procedure, and it adopts the minimum statistic computed during the training period such that $S^*_{min} = min_{e \in [m+1,T^*]} S_{e,m}$, as a means to compare with the statistic calculated after the monitoring start period. The procedure rejects H_0 of crash is not detected during the monitoring period when $S^*_{min} > S_{T',m}$. In case of not rejecting, it updates T' in the same manner as MAX.

As opposed to the SEQ, SEQ_c constructs the critical value from the lower percentile such that $cv_{1-\pi}$. This critical value works differently for the hypothesis testing. While everything else is analogous with SEQ, the maximum sequence of contiguous undershooting of statistic in the training period is considered as a means to compare with that of the monitoring period such that

$$p_c^* = \max_{L,U \in [m+1, T^*]} R^c(L,U) \quad p_c^{'} = \max_{L,U \in [T^*+m, T^{'}]} R^c(L,U)$$

where $R^c(L,U) = (U-L+1)\prod_{e=L}^{U} R_e^c$, $U \geqslant L$ and $R_e^c = \mathbb{1}(S_{e,m} < cv_{1-\pi})$. Like SEQ, the procedure also rejects H_0 of crash is not detected during the monitoring period when $p_c^* < p_c^{'}$.

For the UNI, it works the same as Astill et al. (2018), and both MIN and SEQ_c have the exactly same uniformity argument property.

4.3 Enhanced Real time Monitoring

While MAX and MIN procedure heavily relies on Astill et al. (2017) statistic, the existence of crash series in the training period can undermine the accuracy of MAX and MIN due to the construction of the statistic, (11), where it is summation of Δy_t. Therefore, the recovery periods from $\lfloor \tau_2 T \rfloor$ to $\lfloor \tau_3 T \rfloor$ inflate S_{max}^* whereas the crash periods from $\lfloor \tau_1 T \rfloor$ to $\lfloor \tau_2 T \rfloor$ lower down S_{min}^*. Then, this results in decreasing accuracy of both procedures. Therefore, the paper suggests to apply BIC procedure first to the training period, then to conduct the real time detection afterward.

5 Simulation

All Monte Carlo simulations are conducted via GAUSS 19.1 with 2000 replication.

5.1 BIC Crash Dating

To test for dating accuracy and efficient model selection of BIC for the crash series generated by DGP UCRU, $Model\ 1$ is employed as a dummy model to test for the distinguishing power of the BIC when model is miss-specified. While the BIC method can only be conducted on the condition that the crash is existing in the sample, prior to applying the BIC method, the paper implements the generalized supremum augumented dicky fuller test by Phillips et al. (2015) (PSY) to test for the existence of a crash within the sample. Moreover, the exercise compares the accuracy estimated dates by PSY test and BIC procedure.

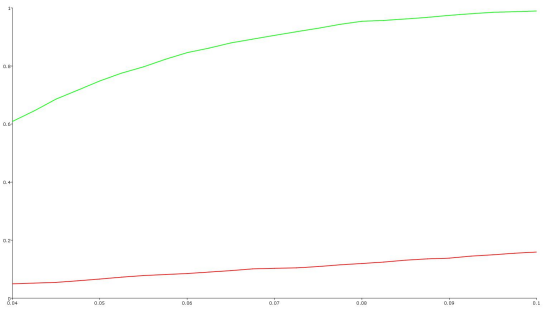
The algorithm generates DGP UCRU based on $T = 200$, $\tau_1 = 0.3$, $\tau_2 = 0.5$, $\tau_3 = 0.55$, $y_0 = 100$, $\varepsilon_t \sim IIDN(0, 1)$ and $seed = 3$. Moreover, various values of δ_1 and δ_2 are considered such that $\delta_2 = \{0.04, 0.0425, ..., 0.1\}$ whereas $\delta_1 = \delta_2/2$. These are to characterize the realistic features that the duration of recovery is shorter and the steep for recovery is sharper than those of a crash as observed in the figure 1. Moreover, duration restrictions for BIC are imposed such that $s_c = 0.1$ and $s_r = 0.04$ to bolster the above discussed empirical features. Before applying PSY and BIC procedure, the generated series are inverted.

In simulation, through all δ_2 values, priory applied PSY test had power of 1 and BIC $Model\ 2$ is constantly chosen. These results imply that the prerequisite condition of BIC application is satisfied and the algorithm selects the more efficient model between the two. However, while correctly specifying the model is not necessary for measuring the dating accuracy, with regard of $Model\ 1$ (UCU) case, only $\lfloor \tau_1 T \rfloor$ and $\lfloor \tau_2 T \rfloor$ are plotted for the accuracy measure.

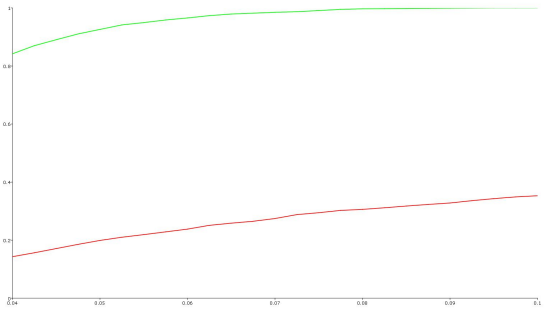
The figure 4 shows the crash start dating accuracy between PSY and BIC procedures while Y-axis refers to accuracy ranging from 0 to 1 and X-axis is the range of δ_2. For the accuracy measure, four cases are considered: **(a)** probability of accurately estimating at the true date, $\lfloor \tau_1 T \rfloor$, **(b)** probability of accurately estimating in $\lfloor \tau_1 T \rfloor \pm 1$, **(c)** probability of accurately estimating in $\lfloor \tau_1 T \rfloor \pm 5$, and **(d)** probability of accurately estimating in $\lfloor \tau_1 T \rfloor \pm 10$. In all four cases, for both PSY and BIC, accuracy rises as δ_2 increases. However, on (a) and (b), the accuracy difference between the two procedures is larger than 0.6 consistently throughout δ_2 values.

Figure 4: Crash Start Dating Accuracy

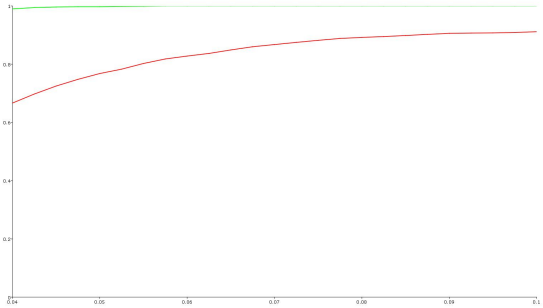
(a) Accuracy at $\lfloor \tau_1 T \rfloor$



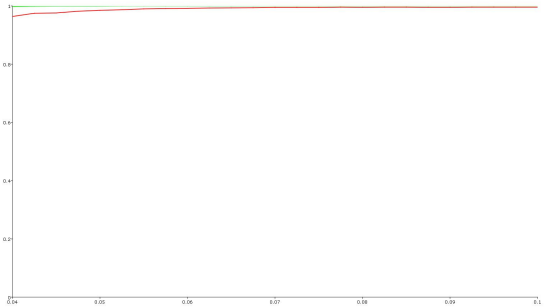
(b) Accuracy in $\lfloor \tau_1 T \rfloor \pm 1$



(c) Accuracy in $\lfloor \tau_1 T \rfloor \pm 5$



(d) Accuracy in $\lfloor \tau_1 T \rfloor \pm 10$



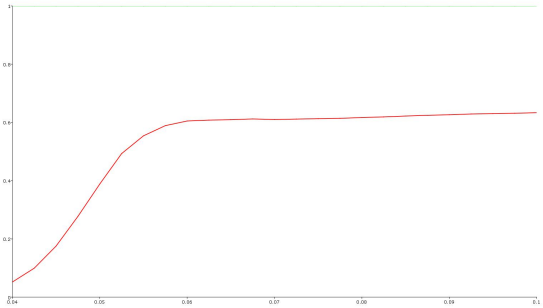
BIC accuracy : ——, PSY accuracy : ——,

Dating accuracy for recovery start date (crash end date) is plotted on the figure 5. For the accuracy measure, $\lfloor \tau_2 T \rfloor$ analogous cases of the figure 4 are considered: **(a)** probability of accurately estimating at the true date, $\lfloor \tau_2 T \rfloor$, **(b)** probability of accurately estimating in $\lfloor \tau_2 T \rfloor \pm 1$, **(c)** probability of accurately estimating in $\lfloor \tau_2 T \rfloor \pm 5$, and **(d)** probability of

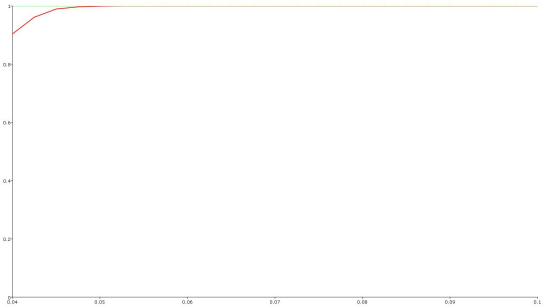
accurately estimating in $\lfloor \tau_2 T \rfloor \pm 10$. In all four cases, BIC shows the perfect accuracy in specifying the regime change point whereas PSY also shows powerful accuracy in (b), (c) and (d), but it performs poorly at identifying the transition period at the very date (a).

Figure 5: Recovery Start Dating Accuracy

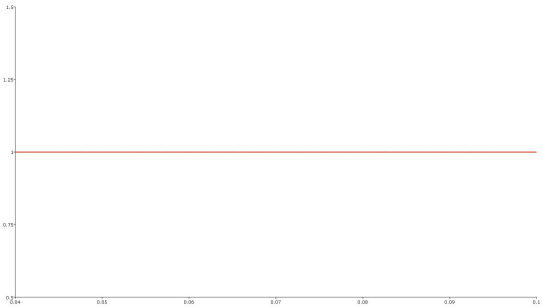
(a) Accuracy at $\lfloor \tau_2 T \rfloor$



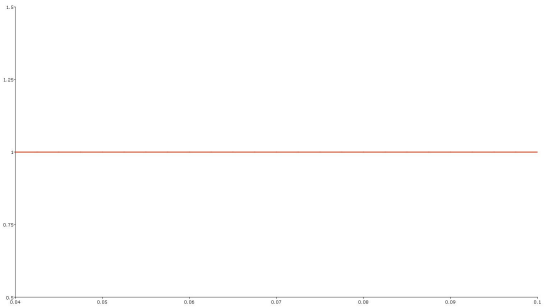
(b) Accuracy in $\lfloor \tau_2 T \rfloor \pm 1$



(c) Accuracy in $\lfloor \tau_2 T \rfloor \pm 5$



(d) Accuracy in $\lfloor \tau_2 T \rfloor \pm 10$



BIC accuracy : ———, PSY accuracy : ———,

Through the simulation, it can be concluded that under the employed DGP series, BIC procedure performs better than PSY.

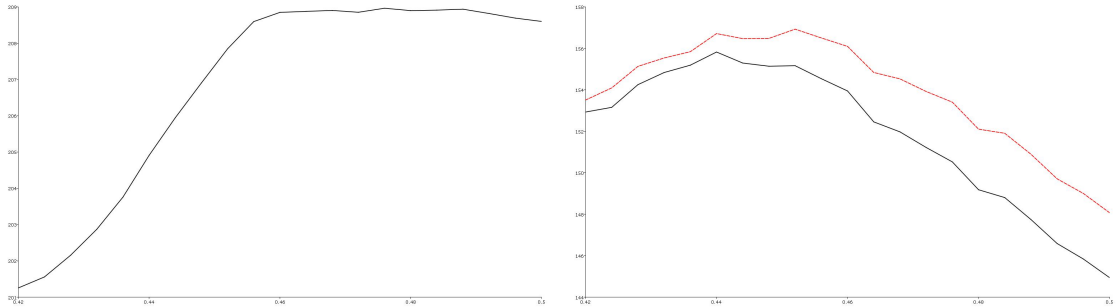
5.2 Real Time Monitoring Accuracy

In this section, two simulations for real time detection are conducted based on the proposed DGPs, UCRUB and UCRUC. The purpose of simulation is to check the influence of training period crash on monitoring accuracy and to present improvement on the detection accuracy after removing the crash series from the training period.

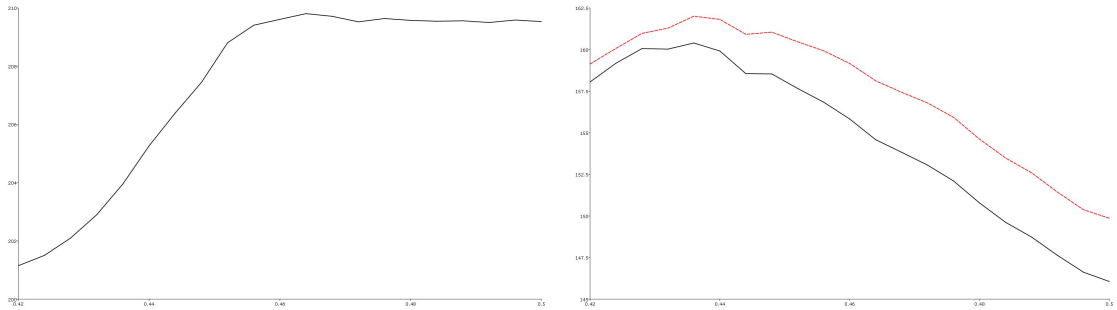
5.2.1 UCRUB

Figure 6: Real Time Bubble Detection Accuracy

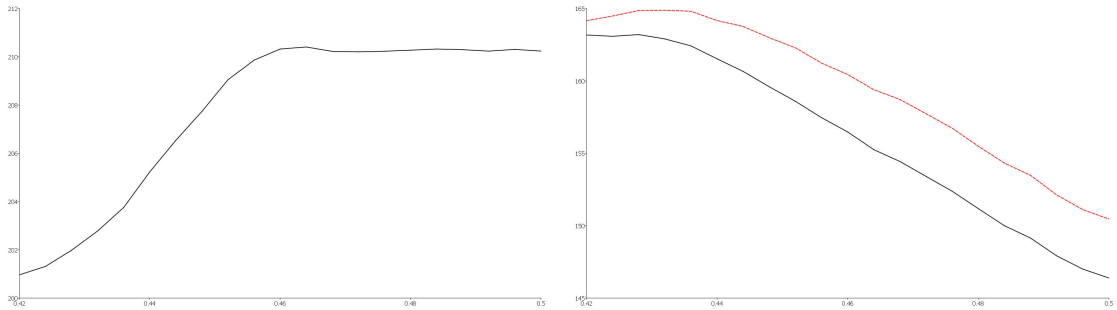
(a) Before removing crash with δ_2=0.02 (b) After removing crash with δ_2=0.02



(c) Before removing crash with δ_2=0.025 (d) After removing crash with δ_2=0.025



(e) Before removing crash with δ_2=0.03 (f) After removing crash with δ_2=0.03



Average First Rejection Date : ——, True First Rejection Date : - - - - -,

The simulation is based on the DPG UCRUB, and two dimensions that affect the accuracy of real time detection are considered for it: the length of recovery period, $\lfloor \tau_3 T \rfloor - \lfloor \tau_2 T \rfloor$, and the steep of recovery, δ_2 whereas $\tau_3 = \{0.42, 0.424, ..., 0.5\}$ and $\delta_2 = \{0.02, 0.025, 0.03\}$. In addition, $T = 250$, $y_0 = 100$, $\tau_1 = 0.3$, $\tau_2 = 0.4$, $\tau_4 = 0.81$, $\delta_1 = 0.01$, $\delta_3 = 0.03$, $\varepsilon_t \sim IIDN(0, 1)$ and $state = 290373$ are considered for DGP. For each parameter com-

bination, two thousand series are generated.

Firstly, the simulation starts with producing the real time detection given DGP. For the real time detection, $T^{\dagger} = 200$, $T^{*} = T^{\dagger} - m$ and $E = 225$, 'End of monitoring period', are set, and for each DGP series, three window length, $m = \{5, 10, 15\}$, are considered. Then, the algorithm records the earliest rejection date by UNI among the three window lengths (If not rejected, the series is not considered). Secondly, with $s_c = 0.025$ and $s_r = 0.02$, it investigates an existence of crash in the periods from 1 to $T^{\dagger}$ of each series. Once a crash is detected, the algorithm executes BIC procedure to date the start of the crash and the end of recovery. Lastly, the algorithm repeats the real time detection without the crash in training period and records the detection date and the true bubble date.

The figure 6 shows the results from the first and third steps of the algorithm. The graphs in the first column plot the average bubble detection period before removing the crash from the training period whereas those on the second column show the true bubble starting date and the bubble detection date by the procedure. For each graph, the Y-axis is the date and the X-axis is the range of τ_3. As anticipated, on the graphs in the first column, as τ_3 increases, the average detection date departs from the true bubble date, $\lfloor \tau_4 T \rfloor = 202$. Moreover, by $\tau_3 = 0.46$, it reaches the peak on all three figures, but with a higher δ_2 value, a peak is also higher. However, the graphs on the second column produce the average detection date plot trending with the true bubble date, and the gap between these two lines is less than 5 regardless of τ_3 and δ_2 values. This is a vastly contrasting result to that of the first column where the bubble detection can be delayed by 10 periods.

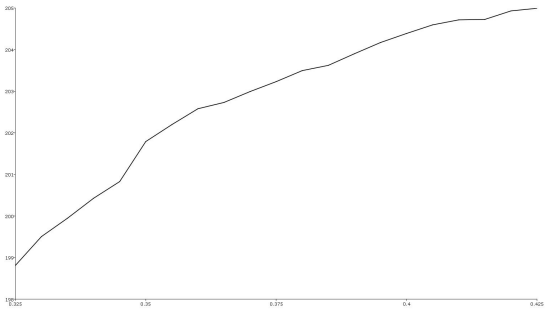
5.2.2 UCRUC

For the simulation, DPG UCRUC is considered. Unlike bubble detection, factors that affect the accuracy of real time detection for a crash are the length of crash period, $\lfloor \tau_2 T \rfloor - \lfloor \tau_1 T \rfloor$, and the steep of crash, δ_1 whereas $\tau_2 = \{0.325, 0.330, ..., 0.425\}$ and $\delta_1 = \{0.01, 0.015, 0.02\}$.

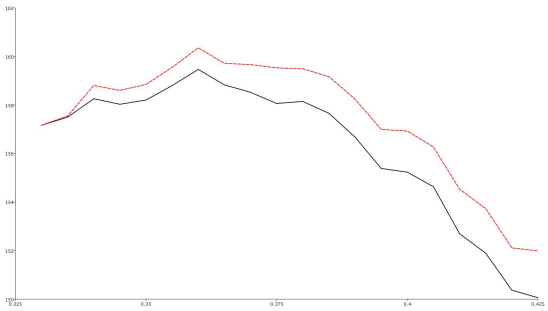
In addition, $T = 250$, $y_0 = 100$, $\tau_1 = 0.3$, $\tau_3 = \tau_2 + 0.05$, $\tau_4 = 0.81$, $\delta_2 = 0.02$, $\delta_3 = 0.03$, $\varepsilon_t \sim IIDN(0, 1)$ and $state = 290373$ are considered for DGP. Apart from the fact that the algorithm employs the crash detection procedure, the steps are analogous of the previous algorithm. On the figure 7, the simulation yields a similar result as the bubble detection case. Thus, excluding the crash before implementing the real time detection method is desirable.

Figure 7: Real Time Crash Detection Accuracy

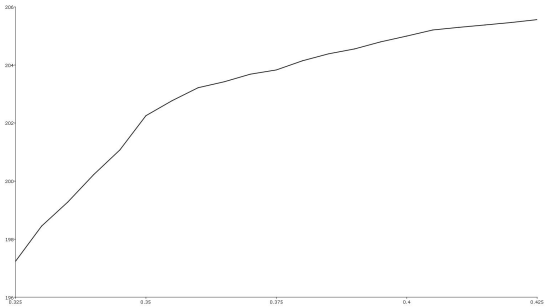
(a) Before removing crash with δ_1=0.01



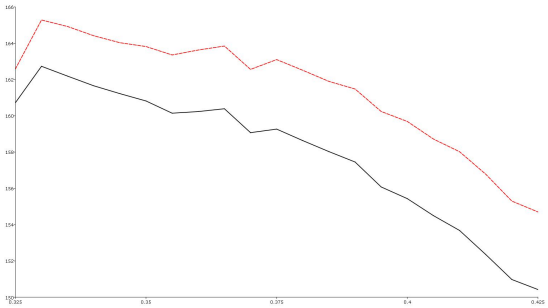
(b) After removing crash with δ_1=0.01



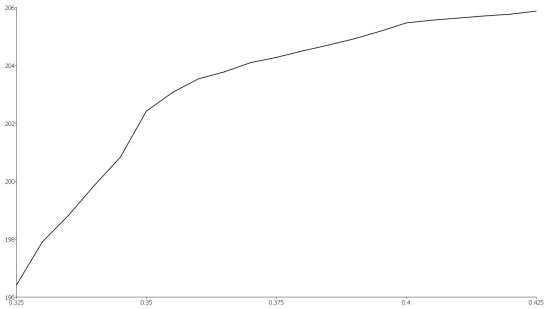
(c) Before removing crash with δ_1=0.015



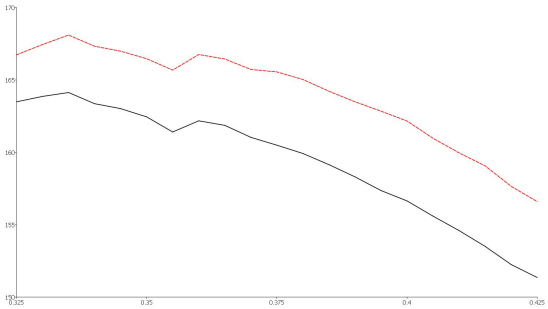
(d) After removing crash with δ_1=0.015



(e) Before removing crash with δ_1=0.02



(f) After removing crash with δ_1=0.02



Average First Rejection Date : ——, True First Rejection Date : - - - - -,

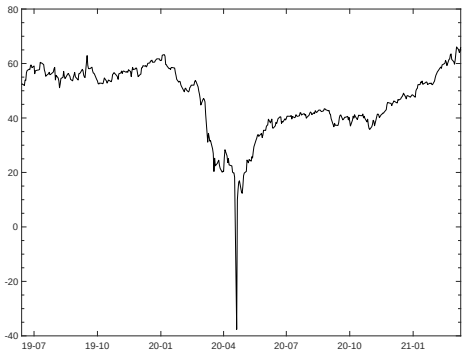
6 Application

6.1 Real Time Monitoring

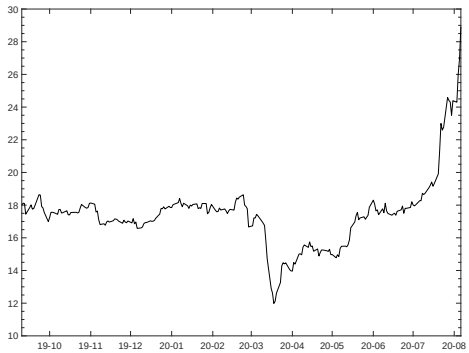
For the real data application, daily future prices of WTI, silver, copper are collected from Bloom-berg Terminal. WTI price ranges from 07-Jun-2019 to 12-Mar-2021 and Copper is from 07-Jun-2019 to 18-Dec-2020 while silver is from 04-Sep-2019 to 06-Aug-2020. The data are plotted on the figure 8. Moreover, for the analysis, monitoring starting periods are set to be 30-Oct-2020 for WTI and copper while for silver, it is 06-Jul-2020. The end of monitoring periods are set to be the end of sample periods. Three window lengths, $m = 5, 10, 15$, are used.

Figure 8: Commodity price plots

(a) WTI



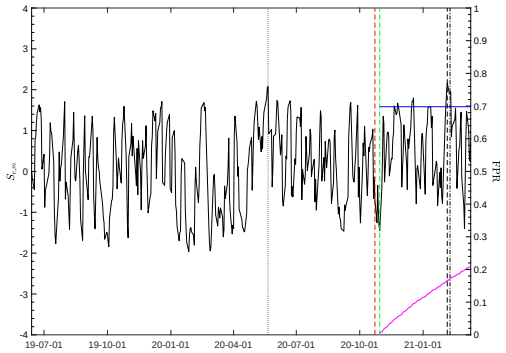
(b) Silver



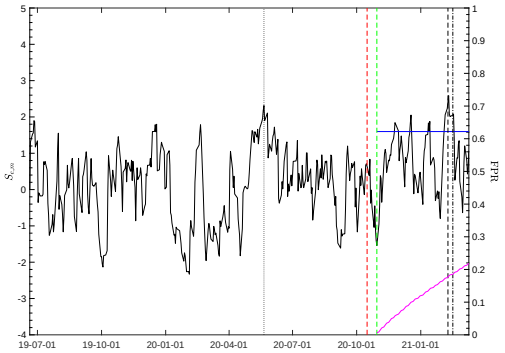
(c) Copper

Figure 9: Real Time Detection for WTI

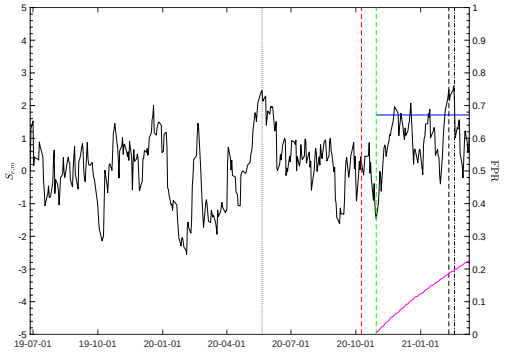
(a) m-5



(b) m-10



(c) m-15



$S_{e,m}$: ——, $max_{e\in[m+1,T^*]}S_{e,m}$: ....., SEQ first detection : - - - -,
MAX first detection : - - - - -, FPR : ——, $cv_{0.05}$: ——, T^* : - - - - -, $T^\dagger$: - - - - -

While Y-axis represents $S_{e,m}$ and X-axis indicates the sample time period, the figure 9 shows the real time detection results on WTI at each window length, m. For all adopted window length, both MAX and SEQ reject $H_0 : WTI_t$ is a unit root process, and at each window length, MAX rejects H_0 earlier than SEQ. For $m = 5$, rejection happens at 05-Feb-2021 while in case of $m = 10$ and $m = 15$, it take place 09-Feb-2021 and 10-Feb-2021 respectively. Therefore, all MAXs are corresponding to UNIs. The paper considers the first rejection among UNIs as the first bubble detection, $First$, to compare with the enhanced monitoring case.

While the real time monitoring procedure signals a bubble on WTI_t as early as 05-

Fed-2021, it fails to detect a bubble on $Silver_t$ and $Copper_t$ within the sample period. The analogous of the figure 9 for $Silver_t$ and $Copper_t$ are on Appendix, and the detection results are the table 1.

Table 1: Real time bubble detection and FPRs

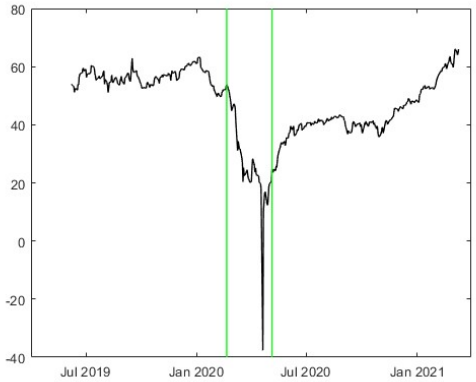
	m	SEQ	FRP_{SEQ}	MAX	FRP_{MAX}	UNI	$First$
WTI	5	09-Feb-2021	0.1707	05-Feb-2021	0.1667	05-Feb-2021	
WTI	10	16-Feb-2021	0.1855	09-Feb-2021	0.1772	09-Feb-2021	05-Feb-2021
WTI	15	18-Feb-2021	0.1969	10-Feb-2021	0.1864	10-Feb-2021	
Silver	5	N/A	N/A	N/A	N/A	N/A	
Silver	10	N/A	N/A	N/A	N/A	N/A	N/A
Silver	15	N/A	N/A	N/A	N/A	N/A	
Copper	5	N/A	N/A	N/A	N/A	N/A	
Copper	10	N/A	N/A	N/A	N/A	N/A	N/A
Copper	15	N/A	N/A	N/A	N/A	N/A	

6.2 Enhanced Real Time Monitoring

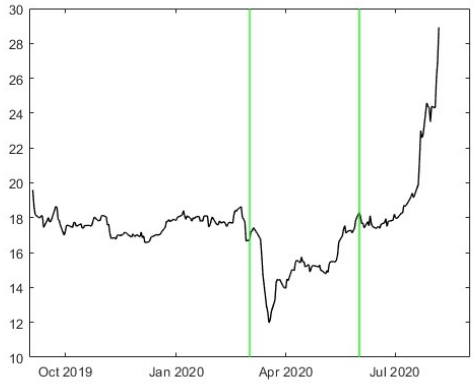
Following the real time monitoring procedure, the paper implements the BIC procedure for crash dating. In all three asset classes, crashes are detected by PSY test, and the BIC procedure is executed to the date from the beginning of sample to the monitoring start periods, to date crash start and recovery end periods within the training periods. For WTI, a crash starts at 20-Feb-2020, and recovery ends at 05-May-2020. In case of copper, a crahs begins at 28-Jan-2020, and recovery terminates at 07-Oct-2020. A bubble in silver price starts at 02-Mar-2020, and recovery ends at 01-Jun-2020. The dating results are plotted on the figure 10 where Y-axis indicates the price in USD and X-axis is the daily time period. For the enhanced real time monitoring analysis, data spanning between the crash start date and the recovery end date are removed from the sample data. When applying the BIC procedure, $s_c = 0.1$ and $s_r = 0.04$ are adopted for the crash and recovery length restrictions.

Figure 10: Crash Dating

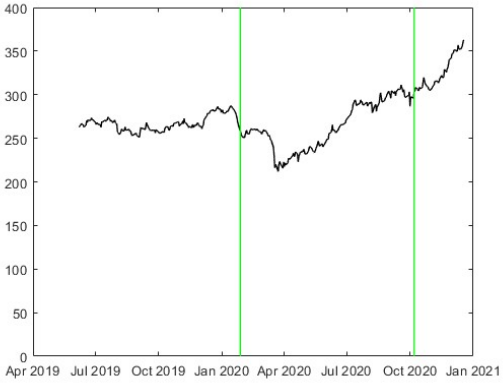
(a) WTI



(b) Silver



(c) Copper



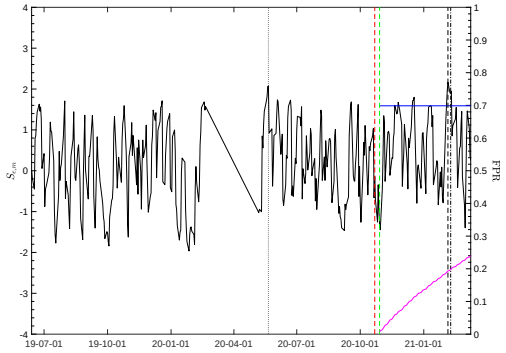
6.2.1 Enhanced WTI

The figure 11 presents the enhanced real time detection results on WTI at each window length. In all window lengths, both MAX^e and SEQ^e reject $H_0 : WTI_t$ is a unit root process, and at each window length, MAX^e rejects H_0 earlier than SEQ^e. In case of $m = 5$, rejection of both SEQ^e and MAX^e happen at the same date as the previous monitoring, but for $m = 10$, the first rejection date of SEQ^e advances by 5 day at 11-Feb-2021 whereas for $m = 15$, both SEQ^e and MAX^e advances 2 days and 3 months respectively at 16-Feb-2021 and 18-Dec-2020. As all MAX^es are corresponding to UNI^es, the paper considers the first rejection among UNI^es as the first bubble detection, $First^e$, which is

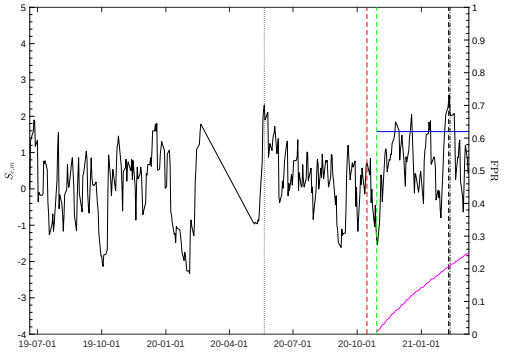
18-Dec-2020.

Figure 11: Enhanced Real Time Detection for WTI

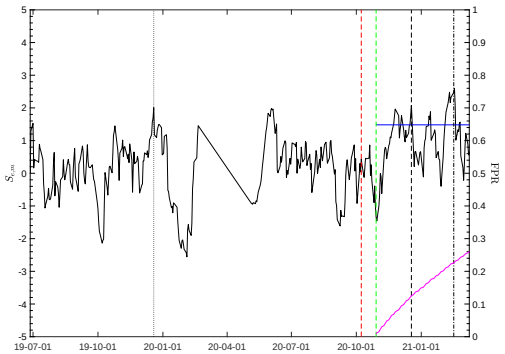
(a) m-5



(b) m-10



(c) m-15



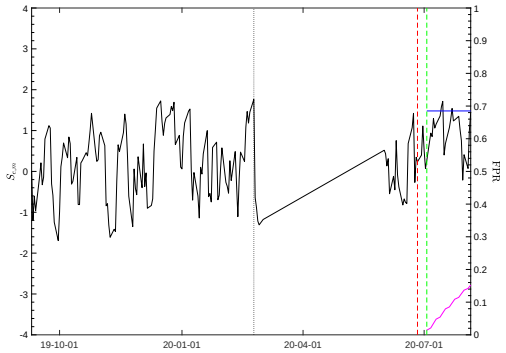
$S_{e,m}$: ——, $max_{e\in[m+1,T*]}S_{e,m}$: ....., SEQ first detection : - - - - -,
MAX first detection : - - - - -, FPR : ——, $cv_{0.05}$: ——, $T*$: - - - - -, $T^{\dagger}$: - - - - -

6.2.2 Enhanced Silver

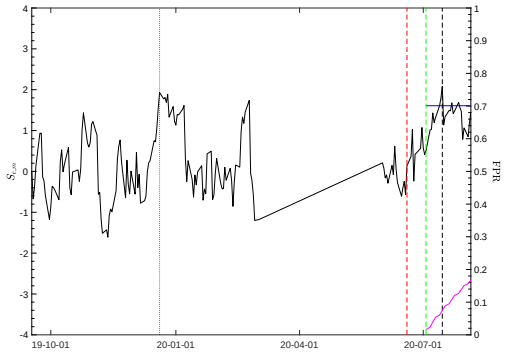
Despite the fact that there is no rejection on the previous monitoring, with $m = 10$ and $m = 15$, MAX^e firstly rejects $H_0 : silver_t$ is a unit root process at 15-Jul-2020 whereas there is still no rejection with $m = 5$, and these results are plotted on the figure 12. As both window lengths have the same rejection date, $First^e$ is 15-Jul-2020.

Figure 12: Enhanced Real Time Detection for Silver

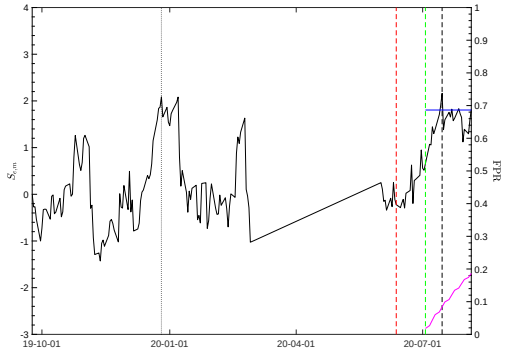
(a) m-5



(b) m-10



(c) m-15



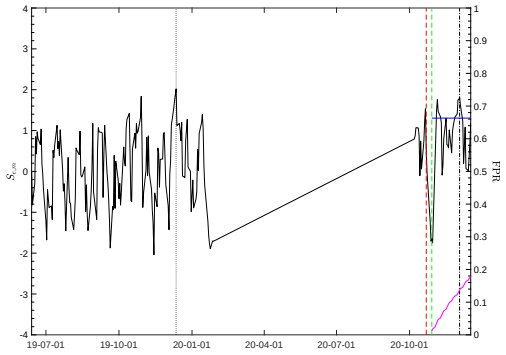
$S_{e,m}$: ——, $max_{e\in[m+1,T^*]}S_{e,m}$: ....., SEQ first detection : - - - - -,
MAX first detection : - - - - -, FPR : ——, $cv_{0.05}$: ——, T^* : - - - - -, $T^\dagger$: - - - - -

6.2.3 Enhanced Copper

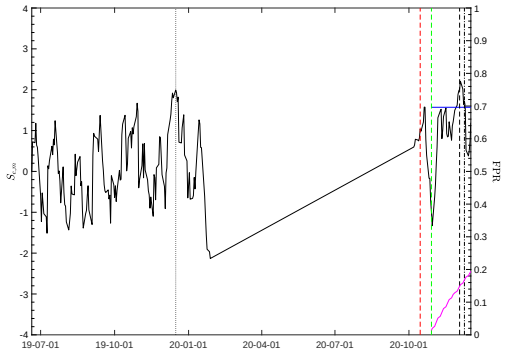
The accuracy is vastly enhanced in copper series. In contract to the fact that there is no rejection in the previous monitoring, in all window lengths, H_0 : $copper_t$ is a unit root process is rejected. For $m = 5$, SEQ^e rejects H_0 at 03-Dec-2020 while MAX^e does not reject. In case of $m = 10$, MAC^e rejects H_0 6 days earlier than SEQ^e, but for $m = 15$, SEQ^e rejects H_0 a day earlier than MAX^e. All three UNI^es are at 03-Dec-2020. Therefore, $First^e$ is at 03-Dec-2020. These results are plotted on the figure 13.

Figure 13: Enhanced Real Time Detection for Copper

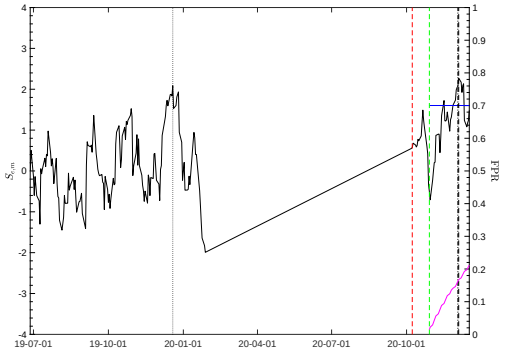
(a) m-5



(b) m-10



(c) m-15



$S_{e,m}$: ——, $max_{e\in[m+1,T^*]}S_{e,m}$: ....., SEQ first detection : - - - -,
MAX first detection : - - - - -, FPR : ——, $cv_{0.05}$: ——, T^* : - - - - -, $T^\dagger$: - - - - -

6.2.4 Remark

For the case of WTI, comparing to the previous monitoring results, not only for MAX^e but also for SEQ^e, accuracy is improved by rejecting H_0 earlier. More importantly, the biggest improvement is from MAX^e with $m = 15$ where the rejection advances by 3 months at 18-Dec-2020. While for $silver$ there is no rejection in the previous monitoring, under the enhanced monitoring, with $m = 10$ and $m = 15$, MAX^e now rejects H_0. In case of $copper$, the accuracy vastly enhanced as in all window lengths, H_0 is rejected in contrast to the previous results. In all asset indices, $First^e$ happens at a desired time

period. Validity of this result is supported by visual inspection on the data. Therefore, the paper concludes that the newly proposed procedure improves the real time monitoring accuracy.

Table 2: Enhanced real time bubble detection and FPRs

	m	SEQ^e	FRP^e_{SEQ}	MAX^e	FRP^e_{MAX}	UNI^e	$First^e$
WTI	5	09-Feb-2021	0.1978	05-Feb-2021	0.1933	05-Feb-2021	
WTI	10	11-Feb-2021	0.2110	09-Feb-2021	0.2064	09-Feb-2021	18-Dec-2020
WTI	15	16-Feb-2021	0.2252	18-Dec-2020	0.1254	18-Dec-2020	
Silver	5	N/A	N/A	N/A	N/A	N/A	
Silver	10	N/A	N/A	15-Jul-2020	0.0746	15-Jul-2020	15-Jul-2020
Silver	15	N/A	N/A	15-Jul-2020	0.0840	15-Jul-2020	
Copper	5	03-Dec-2020	0.1368	N/A	N/A	03-Dec-2020	
Copper	10	09-Dec-2020	0.1676	03-Dec-2020	0.1486	03-Dec-2020	03-Dec-2020
Copper	15	03-Dec-2020	0.1625	04-Dec-2020	0.1677	03-Dec-2020	

7 Conclusion

While the real time monitoring procedure by Astill et al. (2018) successfully performs under the cases of various heteroskedasticities, the underlying assumption that there is no training period crash, can affect the TPR of the procedure due to the construction of statistic and the feature of MAX procedure. Therefore, the paper proposes novel methods : 1) proposing a data generating process to create an unconditional crash and recovery from it, 2) introducing a BIC model selection to date historic crash, and 3) Conducting the enhanced real time monitoring excluding the training period crash and recovery date.

The paper conducts the Monte Carlo simulation with 2000 replications for the accuracy of historic crash dating via BIC modeling and real time monitoring for the cases where a bubble or a crash experiences a significant length of a crash and recovery from it in the training period and where those data are removed from the training period. The simulation results find that the longer the recovery length and the higher the magnitude of it

are, the more distorted accuracy of bubble detection is, conversely, the longer the crash length and the higher the magnitude of it are, the worse the accuracy of crash detection is.

Simulation results gain further validity through real world application employing daily WTI, silver and copper data. Simple application of Astill et al. (2018) fails to detect the bubble for silver and copper while for WTI, it detects as early as 05-Feb-2021. However, through the enhanced real time monitoring procedure, bubbles are detected at 15-Jul-2020 and 03-Dec-2020 respectively for silver and copper whereas the detection date for WTI advances to 18-Dec-2020.

However, this procedure has its shortcoming that using the proposed BIC method, it can only specify one crash series within a sample, thus a methodology to specify multiple bubble in the training period is necessary. Moreover, the procedure is still subject to the FPR inflation as it cannot control the FPR but only restrict. Therefore, the paper suggests the direction of future research to control the FPR. Furthermore, defining an unconditional crash and recovery from it together with a bubble and its crash allows to determine supremum and infimum of a random walk region. This region can be interpreted as a temporary fundamental of certain financial time series and be helpful to forecast the magnitude of a crash or recovery regime followed by a bubble and a crash respectively.

References

Andrews, D. (2003). End-of-sample instability tests. *Econometrica*, *71*(6), 1661–1694.

Andrews, D., & Kim, J.-Y. (2006). Tests for cointegration breakdown over a short time period. *Journal of Business & Economic Statistics*, *24*(4), 379–394.

Astill, S., Harvey, D. I., Leybourne, S. J., Sollis, R., & Robert Taylor, A. (2018). Real-time monitoring for explosive financial bubbles. *Journal of Time Series Analysis*, *39*(6), 863–891.

Astill, S., Harvey, D. I., Leybourne, S. J., & Taylor, A. R. (2017). Tests for an end-of-sample bubble in financial time series. *Econometric Reviews*, *36*(6-9), 651–666.

Diba, B. T., & Grossman, H. I. (1988). Explosive rational bubbles in stock prices? *The American Economic Review*, *78*(3), 520–530.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, *74*(366a), 427–431.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, 987–1007.

Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review*, *81*(4), 922–930.

Fantazzini, D. (2016). The oil price crash in 2014/15: Was there a (negative) financial bubble? *Energy Policy*, *96*, 383–396.

Harvey, D. I., Leybourne, S. J., & Sollis, R. (2017). Improving the accuracy of asset price bubble start and end date estimators. *Journal of Empirical Finance*, *40*, 121–138.

Harvey, D. I., Leybourne, S. J., Sollis, R., & Taylor, A. R. (2016). Tests for explosive financial bubbles in the presence of non-stationary volatility. *Journal of Empirical Finance*, *38*, 548–574.

Harvey, D. I., Leybourne, S. J., Sollis, R., & Taylor, A. R. (2021). Real-time detection of regimes of predictability in the us equity premium. *Journal of Applied Econometrics*, *36*(1), 45–70.

Homm, U., & Breitung, J. (2012). Testing for speculative bubbles in stock markets: a comparison of alternative methods. *Journal of Financial Econometrics*, *10*(1), 198–231.

Phillips, P. C., Shi, S., & Yu, J. (2015). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the s&p 500. *International economic review*, *56*(4), 1043–1078.

Phillips, P. C., Wu, Y., & Yu, J. (2011). Explosive behavior in the 1990s nasdaq: When did exuberance escalate asset values? *International economic review*, *52*(1), 201–226.

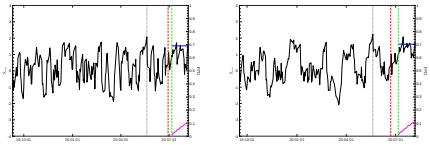
Whitehouse, E. J., Harvey, D. I., & Leybourne, S. J. (2023). Real-time monitoring of bubbles and crashes. *Oxford Bulletin of Economics and Statistics*, *85*(3), 482–513.

8 Appendix

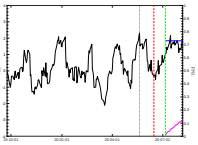
8.1 Real Time Detection for Silver

Figure 14: Real Time Detection for Silver

(a) m-5 (b) m-10



(c) m-15

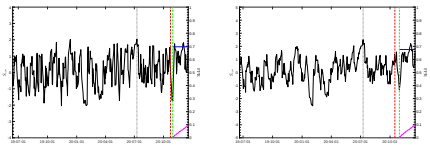


$S_{e,m}$: ——, $max_{e\in[m+1,T^*]}S_{e,m}$: ....., FPR : ——, $cv_{0.05}$: ——, T^* : - - - - -, $T^\dagger$: - - - - -

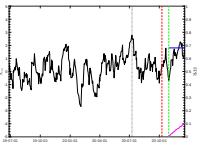
8.2 Real Time Detection for Copper

Figure 15: Real Time Detection for Copper

(a) m-5 (b) m-10



(c) m-15



$S_{e,m}$: ——, $max_{e\in[m+1,T^*]}S_{e,m}$: ....., FPR : ——, $cv_{0.05}$: ——, T^* : - - - - -, $T^\dagger$: - - - - -