International Journal of Forecasting 41 (2025) 1260–1277

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Real-time monitoring procedures for early detection of bubbles

E.J. Whitehouse [a,*], D.I. Harvey [b], S.J. Leybourne [b]

[a] School of Economics, University of Sheffield, United Kingdom
[b] School of Economics, University of Nottingham, United Kingdom

ARTICLE INFO

Article history:
Dataset link: https://sites.google.com/site/ejwhitehouse1

Keywords:
Real-time monitoring
Bubble
Explosive autoregression
Early warning signal
Housing markets

ABSTRACT

Asset price bubbles and crashes can have severe consequences for the stability of financial and economic systems. Policymakers require timely identification of such bubbles in order to respond to their emergence. In this paper we propose new econometric procedures that improve the speed of detection for an emerging asset price bubble in real time. Our new monitoring procedures make use of alternative variance standardisations that are better able to capture the behaviour of the underlying process during a bubble phase. We derive asymptotic results to show that using these alternative variance standardisations does not increase the probability of false detection under the no-bubble (unit root) null hypothesis relative to existing procedures. However, Monte Carlo simulations demonstrate that much earlier detection becomes possible with our new procedures under the bubble (explosive autoregressive) alternative. Empirical applications to OECD housing markets and bitcoin prices show the value in terms of earlier detection of bubbles that our new procedures can achieve. In particular, we show that the United States housing bubble that preceded the global financial crisis could have been detected as early as 1999:Q1 by our new procedures.

1. Introduction

It is well known that asset price bubbles and their subsequent collapses can cause instability in economic and financial systems. Consider, for example, the global financial crisis of 2007–2008. The OECD real house price index shows that house prices in the United States grew 64% between 1995:Q1 and 2006:Q4, before falling by 13% from 2006:Q4 to 2008:Q3 (OECD, 2023). This collapse of the housing market triggered widespread financial distress, resulting in economic recessions in many countries across the world. It is now widely acknowledged that ignoring bubbles in housing markets can have severe consequences (Crowe et al., 2013). To mitigate these negative consequences, policymakers require econometric tools which can detect the emergence of bubbles as soon as they occur, allowing them to make timely policy interventions that promote financial stability. This paper provides new monitoring procedures for the emergence of asset price bubbles which allow for earlier detection than existing methods, and thus provide an early warning system to policymakers to make timely policy interventions.

Early detection of a bubble also provides a warning signal for forecasters, using either judgemental or model-based methods, prompting them to reconsider the efficacy of their forecasting techniques going forwards. For example, suppose a forecaster is using ARIMA or ARFIMA models as part of their forecasting methodology. Forecasts generated from such models using data in a bubble regime (proxied by an explosive autoregressive process, as in this paper) would be entirely inappropriate because

* Correspondence to: School of Economics, University of Sheffield, 9 Mappin Street, Sheffield, S1 4DT, United Kingdom.
E-mail address: e.whitehouse@sheffield.ac.uk (E.J. Whitehouse).

https://doi.org/10.1016/j.ijforecast.2024.12.005
0169-2070/

an explosive autoregressive process cannot be differenced (either integer or fractionally) to stationarity.

Following Phillips et al. (2011), an asset bubble can be identified by detecting the presence of explosive autoregressive behaviour in a price series and the absence of explosive autoregressive behaviour in a corresponding fundamental series. Econometric work in the area of explosive bubble detection has largely focused on historical detection, whereby explosive autoregressive regimes are identified within a sample of previously observed data. The now seminal approaches of Phillips et al. (2011) and Phillips et al. (2015) consider detecting regimes of explosive behaviour through the implementation of sub-sample right-tailed augmented Dickey–Fuller unit root tests. Further developments have included bootstrap implementations of these recursive unit root test procedures (Harvey et al., 2016; Phillips & Shi, 2020), as well as generalised least squares-based recursive unit root testing (Whitehouse, 2019). These sub-sample unit root-based approaches have been widely applied to detect asset price bubbles in a variety of markets, such as housing (Anundsen, 2019; Otero et al., 2022; Pavlidis et al., 2018), commodities (Etienne et al., 2014; Figuerola-Ferretti & McCrorie, 2016; Pastor & Ewing, 2022), and stock markets (Basse et al., 2021; Caspi & Graham, 2018; Hu & Oxley, 2018).

Whilst detecting historical episodes of bubble behaviour is important in its own right, policymakers and forecasters are likely to be more concerned with identifying bubble behaviour as it emerges. In a recent contribution, Astill et al. (2018) (hereafter, AHLST) propose a real-time monitoring procedure for the detection of explosive bubbles, in which they discriminate between the null hypothesis of a unit root (normal market conditions) and the alternative hypothesis of an explosive autoregressive root (bubble) as each new data observation becomes available. Defining the false positive rate (FPR) of the test procedure as the probability of false detection of an explosive bubble by a given point in the monitoring period, their approach is such that the FPR of their procedure under the null hypothesis of no explosive bubble can be quantified at any given monitoring horizon. This is facilitated by constructing the rolling fixed-window length statistics of Astill et al. (2017), which are based on estimators from the sub-sample regression of first differences of the data on a linear trend, and comparing the maximum of these statistics obtained across a training sample to the corresponding maximum of these statistics obtained in a monitoring period. When the null is true, simple uniform distribution arguments are used to establish the (asymptotic) FPR moving through the monitoring period.

Whilst this approach to real-time monitoring for a change from unit root to explosive behaviour is based on statistics constructed from a rolling window of data points (and comparing their maxima over training and monitoring periods), alternative approaches in the literature are based on variants of the recursive (full sample) CUSUM detectors of Chu et al. (1996). Homm and Breitung (2012) and Astill et al. (2024) base the CUSUM detector on first differences of the data. Horvath and Trapani (2024) consider a random coefficient autoregressive framework and suggest weighted CUSUM and Page-CUSUM detectors based on residuals constructed from an autoregressive estimator that uses only the training period data.

The Astill et al. (2017) test statistic itself is motivated by a Taylor series expansion of the first differences of an explosive autoregressive process, and employs a variance standardisation in its denominator. The variance standardisation used is one based on the behaviour of the process under the unit root null hypothesis, but is less appropriate under the alternative hypothesis of an explosive bubble. Under this explosive alternative, the denominator of the statistic is inflated such that the true positive rate (TPR) of the test procedure is reduced. (We define the TPR as the probability of correctly identifying the presence of an explosive bubble by a given point in the monitoring period.)

In this paper we propose the use of new variance standardisations which are more suitable under an explosive alternative hypothesis. We consider two alternative variance standardisations. The first is motivated by the autoregressive structure of the first-differenced series during an explosive bubble, whilst the second is motivated by a first-order Taylor series expansion of the first-differenced series during an explosive bubble. We develop theoretical results which show that the theoretical FPR of tests that use these new variance standardisations are asymptotically identical to the original AHLST test. However, through Monte Carlo simulations, we demonstrate that these modifications can lead to considerable improvements to the TPR under the explosive alternative. As a consequence, using our new procedures can lead to much earlier detection of an emerging bubble. We show that this improved performance can persist even when explosive bubbles occur within the training sample of data used to generate critical values for monitoring; and we demonstrate that our procedures are robust to time-varying conditional volatility in the innovations of the price series.

We first apply our new procedures to a house price-to-rent ratio in 17 OECD countries and demonstrate the substantial advantages that can be achieved in terms of early bubble detection. In particular, we note that our new procedures could have detected the US housing bubble that preceded the global financial crisis as early as 1999:Q1, which would have allowed for a timely policy response. We subsequently consider an empirical application to bitcoin prices which confirms the ability of our new procedures to monitor for an emerging bubble in financial data that are likely to exhibit time-varying conditional volatility.

In the next section we present an explosive bubble model and outline the AHLST testing framework. Section 3 provides the new test statistics and establishes asymptotic theory results relating to their behaviour under the no-bubble null hypothesis. Monte Carlo simulation results are provided in Section 4. Section 5 examines the impact on our test procedures of explosive bubbles occurring outside of the monitoring period. In Section 6 we provide empirical applications of our new tests to the housing market and to bitcoin. Section 7 concludes. Proofs are contained in an appendix.

2. Model and AHLST monitoring framework

We consider the following DGP for a time series y_t, $t = 1, \ldots, T$:

$$y_t = \mu + u_t,$$
$$u_t = \begin{cases} u_{t-1} + \varepsilon_t & t = 2, \ldots, \lfloor \tau T \rfloor \\ (1+\delta)u_{t-1} + \varepsilon_t & t = \lfloor \tau T \rfloor + 1, \ldots, T \end{cases} \quad (1)$$

with $u_1 = O_p(1)$ and $\delta \geq 0$. Here and throughout '$\lfloor . \rfloor$' denotes the integer part. We assume that the error term ε_t is a strictly stationary process with zero mean. If $\delta = 0$ then y_t admits a unit root throughout the sample period. If $\delta > 0$, then y_t exhibits explosive dynamics from $\lfloor \tau T \rfloor + 1$ providing a model of bubble behaviour.[1] The hypotheses of interest are therefore $H_0 : \delta = 0$ (unit root) and $H_1 : \delta > 0$ (unit root then explosive).

The real-time bubble monitoring framework of AHLST considers the sequential application of the Astill et al. (2017) test for an end-of-sample explosive bubble, which in turn is based upon the end-of-sample instability testing approach of Andrews (2003) and Andrews and Kim (2006). In particular, AHLST consider a training period $y_1, \ldots, y_{T^*}$, $T^* = \lfloor \lambda T \rfloor \leq \lfloor \tau T \rfloor$ for some $\lambda \in (0, 1)$, during which it is assumed that H_0 holds. Monitoring for a change from H_0 to H_1 starts at the present time period, denoted $T^\dagger$. The monitoring procedure is based on comparing the behaviour of a sub-sample statistic, denoted $A_{e,k}$, across the training and monitoring periods. Here,

$$A_{e,k} = \frac{B_{e,k}}{\sqrt{C_{e,k}}} \quad (2)$$

where

$$B_{e,k} = \sum_{t=e-k+1}^{e} (t-e+k)\Delta y_t, \; C_{e,k} = \sum_{t=e-k+1}^{e} \{(t-e+k)\Delta y_t\}^2$$

with k a finite constant denoting the length of the sub-sample over which the statistic is computed, and e the last time period used in the statistic's calculation.

The term $B_{e,k}$ is motivated by a first-order Taylor series expansion of the first differences of an explosive process. Specifically, over the explosive regime $t > \lfloor \tau T \rfloor$, Δy_t can be written as

$$\Delta y_t = \delta(1+\delta)^{t-\lfloor \tau T \rfloor - 1} u_{\lfloor \tau T \rfloor} + \sum_{j=0}^{t-\lfloor \tau T \rfloor - 1} (1+\delta)^j \Delta \varepsilon_{t-j}.$$

The stochastic behaviour of Δy_t is dominated by the first of the right-hand-side terms. Approximating $(1+\delta)^{t-\lfloor \tau T \rfloor - 1}$ using a first-order Taylor series expansion around $\delta = 0$ gives

$$(1+\delta)^{t-\lfloor \tau T \rfloor - 1} \approx 1 + (t - \lfloor \tau T \rfloor - 1)\delta$$

and the following approximation is obtained:

$$\Delta y_t = \delta(1-\delta)u_{\lfloor \tau T \rfloor} + \delta^2 u_{\lfloor \tau T \rfloor}(t - \lfloor \tau T \rfloor) + e_t \quad (3)$$

where e_t contains the higher-order terms in the Taylor series expansion and other lower-order terms. The term $B_{e.k}$ is then the sub-sample analogue of the numerator component of the OLS coefficient estimator in a regression of Δy_t on $(t - \lfloor \tau T \rfloor)$. Standardising $B_{e.k}$ by $\sqrt{C_{e.k}}$ is a White-type studentisation intended to provide a degree of robustness to unconditional heteroskedasticity (although this formally lies outside the assumptions of the DGP).

To employ $A_{e,k}$ in a bubble monitoring context, suppose that we intend to start monitoring at the current time period, $t = T^\dagger$. We have an initial training sample of data $t = 1, \ldots, T^*$, where $T^* = T^\dagger - k$. A set of training-sample statistics is first produced by computing the $A_{e,k}$ statistic over rolling sub-samples of length k within this training sample. The maximum training-sample statistic, $A^*_{\max} = \max_{e \in [k+1, T^*]} A_{e,k}$, forms the critical value for monitoring. At time $t = T^\dagger$, we compute the first monitoring statistic using data from $t = T^\dagger - k + 1, \ldots, T^\dagger = T^* + 1, \ldots, T^* + k$, with subsequent monitoring statistics computed as each new observation occurs. The null hypothesis is rejected in favour of an explosive bubble at the first point where the monitoring statistic exceeds the training-sample critical value. We can write this decision rule as

Reject H_0 at time e if $A_{e,k} > A^*_{\max}$

for an arbitrary point in the monitoring period, $t = e$. We refer to this bubble detection procedure as $A_{MAX}(k)$.

AHLST show that, under H_0, for an arbitrary point in the monitoring period T',

$$\lim_{T\to\infty} P\left(\max_{e \in [T^*+k, T']} A_{e,k} > \max_{e \in [k+1, T^*]} A_{e,k} \right) = \alpha \quad (4)$$

where

$$\alpha = \lim_{T\to\infty} \left(\frac{T' - T^* - k + 1}{T' - 2k + 1} \right) = \lim_{T\to\infty} \left(\frac{T' - T^*}{T'} \right). \quad (5)$$

The key to establishing the results (4) and (5) lies in recognising that the $A_{e,k}$ form a strictly stationary sequence across both the training and monitoring periods. Large sample results connected to the uniform distribution property of the *location* of the maximum value of $A_{e,k}$ are then appealed to (no limit theory applies to $A_{e,k}$ itself).

For given values of T^* and k, the approximate FPR of $A_{MAX}(k)$ at monitoring point T' is

$$\alpha \approx \frac{T' - T^* - k + 1}{T' - 2k + 1}. \quad (6)$$

We can also rearrange (6) to identify the monitoring time period T' at which the FPR of the procedure will (approximately) reach the level α, allowing determination of how far monitoring into the future can be done whilst maintaining a chosen FPR:

$$T' \approx \frac{T^* + k - 1 - \alpha(2k-1)}{1-\alpha}.$$

We can contrast the AHLST approach, which uses the fixed-sample $A_{e,k}$ as a detector, to those of Chu et al. (1996), Homm and Breitung (2012), Astill et al. (2024), and Horvath and Trapani (2024). Their null limit theory applies directly to a full-sample CUSUM detector and

[1] Since our focus here is on rapid detection in the explosive regime, we do not concern ourselves with behaviour after the bubble has terminated (e.g. collapse and/or return to unit root behaviour). Instead we simply consider a model where the explosive regime runs to the end of the sample.

associated boundary function, with the property that for large T^*, the probability the detector exceeds the boundary function for some $T' - T^* \geqslant 1$ is less than or equal to a user-chosen FPR. The AHLST approach is therefore more informative about the FPR of the procedure than the CUSUM-based approaches in the sense that no bound is involved in (4), and the (approximate) FPR can be calculated from (6) at each monitoring point T', with the procedure not potentially conservative, unlike the CUSUM-based methods. The cost of this extra information is that the FPR cannot be determined independently of the value of the ratio T^*/T'. That is, we are not free to choose α without consideration of T^*/T'. In this sense, the AHLST approach may be better suited to relatively short-range monitoring, since the FPR grows with T'. Long-range monitoring is more feasible with the CUSUM approaches in terms of having the flexibility to bound FPR to typical values (say to 0.05 or 0.10) over a long monitoring horizon, but because the CUSUM detectors are divergent, the TPR can be low. Finally, we note that α is informative when the monitoring point T' is the same order as the training-sample size T^*. If the number of monitoring periods $T' - T^*$ grows at a faster rate in T than the training period length T^*, then $\alpha = 1$; if $T' - T^*$ grows at a slower rate than T^* then $\alpha = 0$. CUSUM-based techniques are studied by Horvath and Trapani (2024) for these different cases.

3. Modified statistics

The White-type studentisation in $A_{e,k}$ of (2) based on $C_{e.k} = \sum_{t=e-k+1}^{e}\{(t-e+k)\Delta y_t\}^2$ is a feasible implementation of an ideal, but infeasible, standardisation that would involve the unobserved errors ε_t, i.e. a feasible proxy for $\sum_{t=e-k+1}^{e}\{(t-e+k)\varepsilon_t\}^2$. This approach essentially imposes the null hypothesis, since under H_0, $\Delta y_t = \varepsilon_t$. However, under the alternative H_1, $\Delta y_t = \Delta u_t = \delta u_{t-1} + \varepsilon_t \neq \varepsilon_t$ for $t > \lfloor \tau T \rfloor$, suggesting that the null-based variance standardisation is inappropriate under the alternative. When u_t is explosive, the denominator of $A_{e,k}$ becomes inflated, relative to using the unobserved ε_t, with obvious negative implications for the TPR of the procedure. In the context of monitoring, this then negatively impacts upon the speed at which a bubble will be detected. Moreover, the structure of $A_{e,k}$ is such that $\left|A_{e,k}\right| \leq k^{1/2}$ by virtue of the Cauchy–Schwarz inequality:

$$\left(\sum_{t=e-k+1}^{e}(t-e+k)\Delta y_t.1\right)^2$$

$$\leq \sum_{t=e-k+1}^{e}\{(t-e+k)\Delta y_t\}^2 \sum_{t=e-k+1}^{e} 1^2$$

$$= k\sum_{t=e-k+1}^{e}\{(t-e+k)\Delta y_t\}^2$$

from which it is clear that

$$\frac{\left(\sum_{t=e-k+1}^{e}(t-e+k)\Delta y_t\right)^2}{\sum_{t=e-k+1}^{e}\{(t-e+k)\Delta y_t\}^2} \leq k$$

i.e. $\left|A_{e,k}\right| \leq k^{1/2}$. This implicit constraint could also have a detrimental effect on the TPR. In view of these two potential issues affecting the procedure's TPR, we consider replacing Δy_t in the variance standardisation by other quantities which better approximate ε_t under the alternative and do not result in the $k^{1/2}$ form of bound associated with the original $A_{e,k}$ statistic.

The first modification that we consider is motivated by the autoregressive structure of Δy_t for $t > \lfloor \tau T \rfloor$. That is,

$$\Delta y_t = -\delta\mu + \delta y_{t-1} + \varepsilon_t.$$

We then replace Δy_t in the denominator of $A_{e,k}$ (for all e) with the residuals $\hat{\varepsilon}_t$ obtained from an OLS regression of Δy_t on a constant and y_{t-1} over the sub-sample $t = e-k+1, \ldots, e$. The modified statistic is then

$$A_{e,k}^{AR} = \frac{\sum_{t=e-k+1}^{e}(t-e+k)\Delta y_t}{\sqrt{\sum_{t=e-k+1}^{e}\{(t-e+k)\hat{\varepsilon}_t\}^2}}.$$

Our second alternative modification is motivated by the Taylor series approximation of Δy_t over the explosive period given in (3). Accordingly, we replace Δy_t in the denominator of $A_{e,k}$ with OLS residuals $\hat{e}_t$ obtained from a regression of Δy_t on a constant and time trend over the sub-sample $t = e-k+1, \ldots, e$:

$$A_{e,k}^{TR} = \frac{\sum_{t=e-k+1}^{e}(t-e+k)\Delta y_t}{\sqrt{\sum_{t=e-k+1}^{e}\{(t-e+k)\hat{e}_t\}^2}}.$$

In what follows, we denote the bubble detection procedure that replaces $A_{e,k}$ in $A_{MAX}(k)$ with $A_{e,k}^{AR}$ as $A_{MAX}^{AR}(k)$, and the procedure that replaces $A_{e,k}$ with $A_{e,k}^{TR}$ as $A_{MAX}^{TR}(k)$. In the following theorem, we establish the theoretical FPRs of the $A_{MAX}^{AR}(k)$ and $A_{MAX}^{TR}(k)$ monitoring procedures as $T \to \infty$.

Theorem 1. *Under H_0 and assuming that $\{\varepsilon_t\}$ satisfies the mixing conditions of Ferreira and Scotto (2002, p. 476), then as $T \to \infty$,*

$$\lim_{T\to\infty} P\left(\max_{e\in[T^*+k,T']} A_{e,k}^{AR} > \max_{e\in[k+1,T^*]} A_{e,k}^{AR}\right) = \alpha$$

$$\lim_{T\to\infty} P\left(\max_{e\in[T^*+k,T']} A_{e,k}^{TR} > \max_{e\in[k+1,T^*]} A_{e,k}^{TR}\right) = \alpha$$

where α is as defined in (5).

Theorem 1 shows that the theoretical FPRs associated with the new procedures $A_{MAX}^{AR}(k)$ and $A_{MAX}^{TR}(k)$ are asymptotically identical to the theoretical FPR of $A_{MAX}(k)$ discussed in the previous section.

4. Finite sample simulations

To examine the finite sample performance of our two new procedures, $A_{MAX}^{AR}(k)$ and $A_{MAX}^{TR}(k)$, relative to the original $A_{MAX}(k)$ procedure, we conduct a Monte Carlo simulation exercise. Data are generated by (1), setting $\mu = 0$ (without loss of generality) and $\varepsilon_t \sim NIID(0, 1)$. To ensure that we generate only positive explosive regimes, in line with typical bubble behaviour, we set $u_1 = 100$. We consider sub-sample window sizes of $k = \{5, 10, 15\}$. The beginning of the monitoring period is set to $T^\dagger = 200$ such that the training sample end date is $T^* = 200 -$

k throughout. Simulations are conducted using 10,000 replications in GAUSS 20.

Fig. 1 displays the cumulative rejection frequencies of the three procedures at T' (i.e. the proportion of replications where we detect H_1 at any point in the monitoring period up to and including T'). First, considering the empirical FPRs of the monitoring procedures obtained under the no-bubble null hypothesis of $\delta = 0$, it is clear from Figs. 1(a)–1(c) that the empirical FPR of the new $A^{AR}_{MAX}(k)$ and $A^{TR}_{MAX}(k)$ procedures track the theoretical FPR given by (6) very closely, as expected from Theorem 1. We can therefore confirm that these new procedures do not suffer from increased false detection of a bubble under H_0 relative to $A_{MAX}(k)$.

We now turn our attention to the empirical rejection frequencies obtained under the alternative hypothesis of an explosive bubble, with $\delta > 0$. We set $\lfloor \tau T \rfloor = 1.1T^{\dagger}$ such that the first observation of the explosive bubble regime occurs at $\lfloor \tau T \rfloor + 1 = 221$. Given that the focus of this paper is on rapid detection of an explosive bubble, and that we confirmed above that the new procedures have very similar false rejection probabilities under H_0 to the original $A_{MAX}(k)$ procedure, in Figs. 1(d)–1(o) we now display empirical rejection frequencies for $T' = 218$ onwards, to concentrate attention on the observations immediately before and after $\lfloor \tau T \rfloor$. We consider explosive bubble magnitude settings of $\delta = \{0.01, 0.02, 0.03, 0.04\}$.

In Figs. 1(d)–1(f), with $\delta = 0.01$, the procedures offer similar TPRs, especially when $k = 5$, with the new $A^{AR}_{MAX}(k)$ and $A^{TR}_{MAX}(k)$ procedures performing marginally better than the original $A_{MAX}(k)$ in the case of $k = 10$ and $k = 15$. As δ increases, the relative advantage of our new procedures over $A_{MAX}(k)$ is found to be more pronounced. These TPR differences can be substantial at monitoring horizons in the immediate region of the explosive bubble start date. For example, in Fig. 1(h) with $k = 10$ and $\delta = 0.02$, at $T' = 224$ (three observations after the bubble start date), $A^{TR}_{MAX}(k)$ has a TPR of 0.487 relative to only 0.244 for $A_{MAX}(k)$. In Fig. 1(k) with $\delta = 0.03$, the TPR of $A_{MAX}(k)$ increases only to 0.271, whereas the TPR of $A^{TR}_{MAX}(k)$ increases to 0.696, resulting in a 0.425 TPR differential. Finally, considering $\delta = 0.04$ in Fig. 1(n), the TPR of $A^{TR}_{MAX}(k)$ is now 0.824 relative to 0.294 for $A_{MAX}(k)$, resulting in a very large TPR difference of 0.530. The new $A^{AR}_{MAX}(k)$ procedure also demonstrates substantial gains over the original $A_{MAX}(k)$ procedure, with TPRs similar to $A^{TR}_{MAX}(k)$, although the latter generally displays slightly higher TPR levels. For example, for the settings discussed above, $A^{AR}_{MAX}(k)$ has TPR gains over the original $A_{MAX}(k)$ procedure of 0.156, 0.278, and 0.359 for $\delta = \{0.02, 0.03, 0.04\}$, respectively. In a small number of cases, particularly for $k = 5$, as monitoring moves further away from the explosive bubble start date, $A^{AR}_{MAX}(k)$ marginally outperforms $A^{TR}_{MAX}(k)$.

To further demonstrate the earlier detection capabilities of the new procedures, Fig. 2 displays histograms of the detection dates for $A_{MAX}(k)$, $A^{AR}_{MAX}(k)$, and $A^{TR}_{MAX}(k)$.[2] We report results for $k = 10$. First, we note that for all procedures, increasing the explosive magnitude δ decreases the delay in detection, as would be expected. Second, we observe that for a given setting of δ, $A^{AR}_{MAX}(k)$ and $A^{TR}_{MAX}(k)$ achieve a higher frequency of detection at dates shortly after $\lfloor \tau T \rfloor + 1$, thus reducing the delay in detection and confirming the results discussed previously.

The results in this section are in line with what would be expected from Section 3, where we outlined the limitations of the $A_{MAX}(k)$ statistic under H_1 relative to our proposed modifications. The large TPR differences between the new $A^{AR}_{MAX}(k)$ and $A^{TR}_{MAX}(k)$ procedures and the original $A_{MAX}(k)$ procedure, especially shortly after the explosive regime begins, translate into more rapid detection of a bubble. This is clearly of critical importance in the context of real-time monitoring. Comparing our two new procedures to each other, $A^{TR}_{MAX}(k)$ generally offers a marginally higher TPR than $A^{AR}_{MAX}(k)$, with a small number of exceptions occurring when $k = 5$ and the monitoring horizon is longer.

Finally, it is well known that financial time series data, which may exhibit explosive bubble behaviour, are often subject to time-varying conditional volatility. AHLST demonstrate that the $A_{MAX}(k)$ procedure is robust to conditional heteroskedasticity in the price series innovations, with the finite sample empirical FPR of $A_{MAX}(k)$ closely matching the theoretical FPR under these circumstances. To confirm that this appealing feature also applies to our new procedures, we also consider ε_t in (1) to be a conditionally heteroskedastic $GARCH(1, 1)$ process instead of $\varepsilon_t \sim NIID(0, 1)$ as before. Specifically, we consider $\varepsilon_t = \sqrt{h_t}\eta_t$, with $h_t = 0.1 + 0.1z^2_{t-1} + 0.8h_{t-1}$ and $\eta_t \sim NIID(0, 1)$. We set $h_0 = \varepsilon_t = 0$. Table 1 gives the empirical FPR of the procedures, again using $k = 10$, under this $GARCH(1, 1)$ specification, along with the corresponding results for $NIID(0, 1)$ innovations. The two sets of results are extremely similar, confirming the robustness of the new procedures to this feature, and it is clear that the empirical FPR values are close to the corresponding theoretical FPR for each value of T', as was observed for the $NIID(0, 1)$ innovations in Figs. 1(a)–1(c).

5. Training-sample bubbles

The theoretical FPR given by (6) relies on the assumption that the unit root null hypothesis, H_0, holds in the training period, $y_1, \ldots, y_{T^*}$. As AHLST note, in practice it is possible that this assumption will be violated. One option available to practitioners is to pre-test the training period data for the presence of an explosive regime through the application of a bubble detection procedure such as the Phillips et al. (2015) test. However, there is no guarantee that all explosive regimes in the training sample would be detected, especially if these are of low magnitude or short duration. It is also the case that, in practice, it may not be possible to find a sufficiently long period of training-sample data that are free from explosive behaviour to allow for monitoring. AHLST provide Monte Carlo simulation results to demonstrate that the empirical FPR of the $A_{MAX}(k)$ procedure is reduced when an explosive bubble occurs in the training sample, due to larger

[2] The detection date is the date at which the explosive regime is detected, for each Monte Carlo replication where detection occurs within the monitoring period.

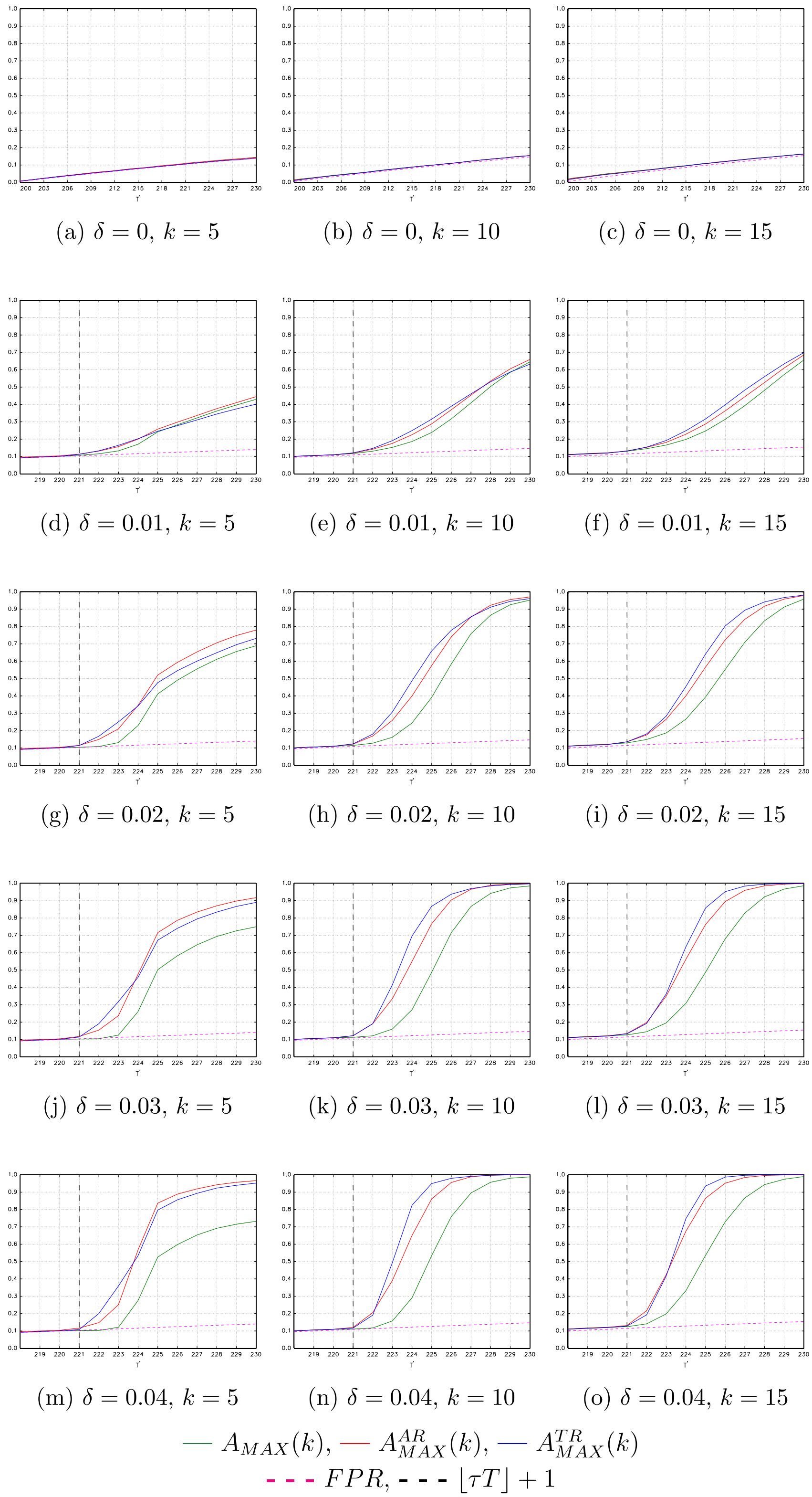


Fig. 1. Empirical rejection frequencies of $A_{MAX}(k)$, $A^{AR}_{MAX}(k)$, and $A^{TR}_{MAX}(k)$.

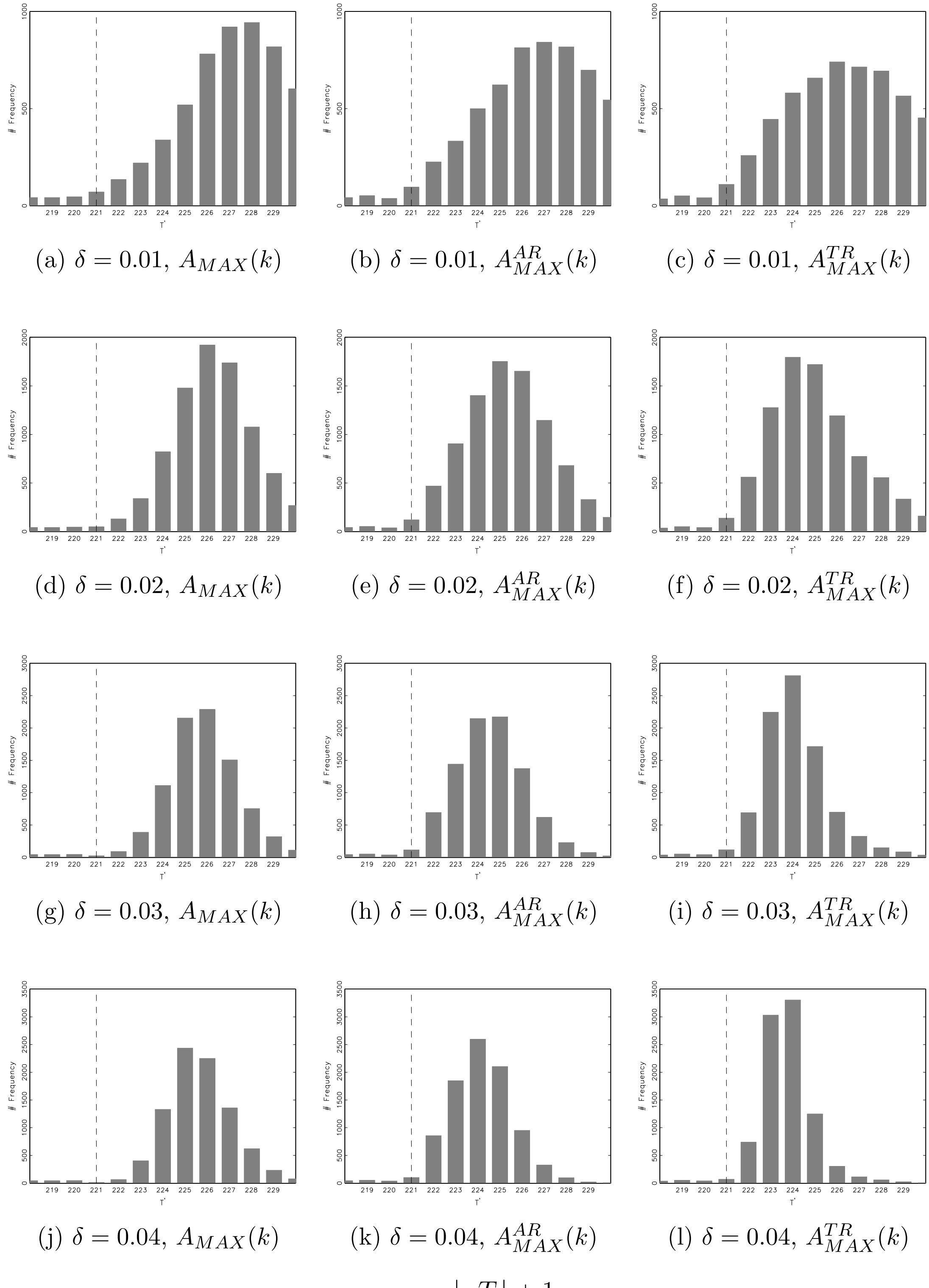


Fig. 2. Histograms of $A_{MAX}(k)$, $A^{AR}_{MAX}(k)$, and $A^{TR}_{MAX}(k)$ detection dates.

values of $A^*_{\max}$. However, under the alternative hypothesis, whilst the empirical TPR of the $A_{MAX}(k)$ procedure is reduced relative to the case where there is no bubble within the training sample, these reductions are modest,

Table 1
Empirical false positive rates of $A_{MAX}(k)$, $A_{MAX}^{AR}(k)$, and $A_{MAX}^{TR}(k)$ under *NIID* and *GARCH*(1, 1) errors, $k = 10$.

		NIID errors			*GARCH*(1, 1) errors		
T'	α	$A_{MAX}(k)$	$A_{MAX}^{AR}(k)$	$A_{MAX}^{TR}(k)$	$A_{MAX}(k)$	$A_{MAX}^{AR}(k)$	$A_{MAX}^{TR}(k)$
200.000	0.006	0.015	0.013	0.010	0.014	0.013	0.010
201.000	0.011	0.021	0.019	0.017	0.020	0.018	0.017
202.000	0.016	0.025	0.023	0.022	0.025	0.023	0.022
203.000	0.022	0.030	0.029	0.028	0.031	0.029	0.027
204.000	0.027	0.035	0.034	0.034	0.036	0.034	0.034
205.000	0.032	0.041	0.039	0.039	0.041	0.039	0.038
206.000	0.037	0.045	0.043	0.043	0.045	0.043	0.042
207.000	0.043	0.051	0.048	0.048	0.051	0.048	0.047
208.000	0.048	0.054	0.052	0.052	0.054	0.052	0.052
209.000	0.053	0.058	0.057	0.057	0.059	0.058	0.057
210.000	0.058	0.064	0.062	0.062	0.066	0.063	0.062
211.000	0.063	0.070	0.068	0.068	0.071	0.069	0.068
212.000	0.067	0.075	0.073	0.073	0.076	0.073	0.074
213.000	0.072	0.079	0.078	0.077	0.081	0.078	0.078
214.000	0.077	0.084	0.083	0.083	0.086	0.082	0.083
215.000	0.082	0.089	0.088	0.087	0.090	0.088	0.088
216.000	0.086	0.092	0.093	0.092	0.094	0.093	0.092
217.000	0.091	0.097	0.097	0.097	0.099	0.096	0.097
218.000	0.095	0.101	0.101	0.101	0.103	0.101	0.100
219.000	0.100	0.105	0.106	0.106	0.107	0.105	0.105
220.000	0.104	0.110	0.110	0.110	0.112	0.109	0.109
221.000	0.109	0.115	0.114	0.115	0.117	0.113	0.112
222.000	0.113	0.119	0.120	0.120	0.121	0.119	0.119
223.000	0.118	0.125	0.125	0.125	0.127	0.124	0.124
224.000	0.122	0.130	0.130	0.130	0.132	0.129	0.129
225.000	0.126	0.133	0.134	0.134	0.134	0.132	0.133
226.000	0.130	0.138	0.138	0.138	0.139	0.137	0.137
227.000	0.135	0.141	0.142	0.142	0.143	0.141	0.141
228.000	0.139	0.147	0.148	0.146	0.149	0.146	0.146
229.000	0.143	0.151	0.151	0.150	0.152	0.149	0.150
230.000	0.147	0.154	0.155	0.154	0.155	0.153	0.154

demonstrating the ability of the procedure to monitor for an explosive bubble even when the training-sample H_0 assumption is violated.

To investigate the impact of a training-sample bubble on our modified procedures, $A_{MAX}^{AR}(k)$ and $A_{MAX}^{TR}(k)$, we extend the DGP in (1) and consider the following DGP for a time series y_t, $t = 1, \ldots, T$:

$$y_t = \mu + u_t + x_t \tag{7}$$

$$u_t = \begin{cases} u_{t-1} + \varepsilon_t & t = 2, \ldots, \lfloor \tau_1^* T \rfloor \\ (1 + \delta_1^*) u_{t-1} + \varepsilon_t & t = \lfloor \tau_1^* T \rfloor + 1, \ldots, \lfloor \tau_2^* T \rfloor \\ \varepsilon_t & t = \lfloor \tau_2^* T \rfloor + 1 \\ u_{t-1} + \varepsilon_t & t = \lfloor \tau_2^* T \rfloor + 2, \ldots, \lfloor \tau T \rfloor \\ (1 + \delta) u_{t-1} + \varepsilon_t & t = \lfloor \tau T \rfloor + 1, \ldots, T \end{cases} \tag{8}$$

$$x_t = (u_{\lfloor \tau_2^* T \rfloor} - u_1) \mathbb{I}(t > \lfloor \tau_2^* T \rfloor + 1) \tag{9}$$

with $u_1 = O_p(1)$ and $\delta \geq 0$. As before, we assume that the error term ε_t is a strictly stationary process with zero mean. We consider $\lfloor \tau_1^* T \rfloor < \lfloor \tau_2^* T \rfloor < T^* < \lfloor \tau T \rfloor$, $\delta_1^* > 0$, and $\delta \geq 0$, such that the DGP admits an explosive regime that reverts back to a unit root process within the training sample.[3] The inclusion of x_t in (7) and the indicator function in (9) prevent the magnitude of the training sample explosive regime from entering the dynamics of the monitoring period explosive regime (cf. Harvey et al., 2020). Without this correction, setting $\delta_1^* = \delta$ (for example) would lead to a more pronounced explosive regime for the monitoring period bubble than the training period bubble.

We conduct a Monte Carlo simulation exercise with data generated by (7)–(9), setting $\mu = 0$, $\varepsilon_t \sim NIID(0, 1)$ and $u_1 = 100$ as before. Fig. 3 displays empirical rejection frequencies of $A_{MAX}(k)$, $A_{MAX}^{AR}(k)$, and $A_{MAX}^{TR}(k)$ when a training-sample explosive bubble exists. Here, we present results only for $k = 10$, as the results are qualitatively similar for other settings of k. We set $T^\dagger = 200$ as before, and consider $\lfloor \tau_1^* T \rfloor = 0.5T^\dagger = 100$ and $\lfloor \tau_2^* T \rfloor = 105$, such that the training-sample explosive regime lasts five observations.

Figs. 3(a)–3(d) consider the null hypothesis, $\delta = 0$, where no bubble is present in the monitoring period, allowing us to examine the impact of a training-sample bubble on the empirical FPRs of the three procedures. We set the magnitude of the training-sample explosive regime to $\delta_1^* = \{0.01, 0.02, 0.03, 0.04\}$ in Figs. 3(a)–3(d) respectively. Examining these figures, it is clear that the presence of an explosive regime within the training sample reduces the empirical FPR of all procedures as expected, such that all procedures have empirical FPRs below the theoretical FPR. This reduction is more pronounced the higher the value of δ_1^*, as would be expected. An ordering of the test procedures emerges, with $A_{MAX}^{TR}(k)$ generally demonstrating the lowest empirical FPR and

3 In unreported simulations, we also considered a DGP that permits an explosive regime followed by a stationary crash regime (cf. Harvey et al., 2017) within the training sample. The results were qualitatively similar to the no-crash case presented here.

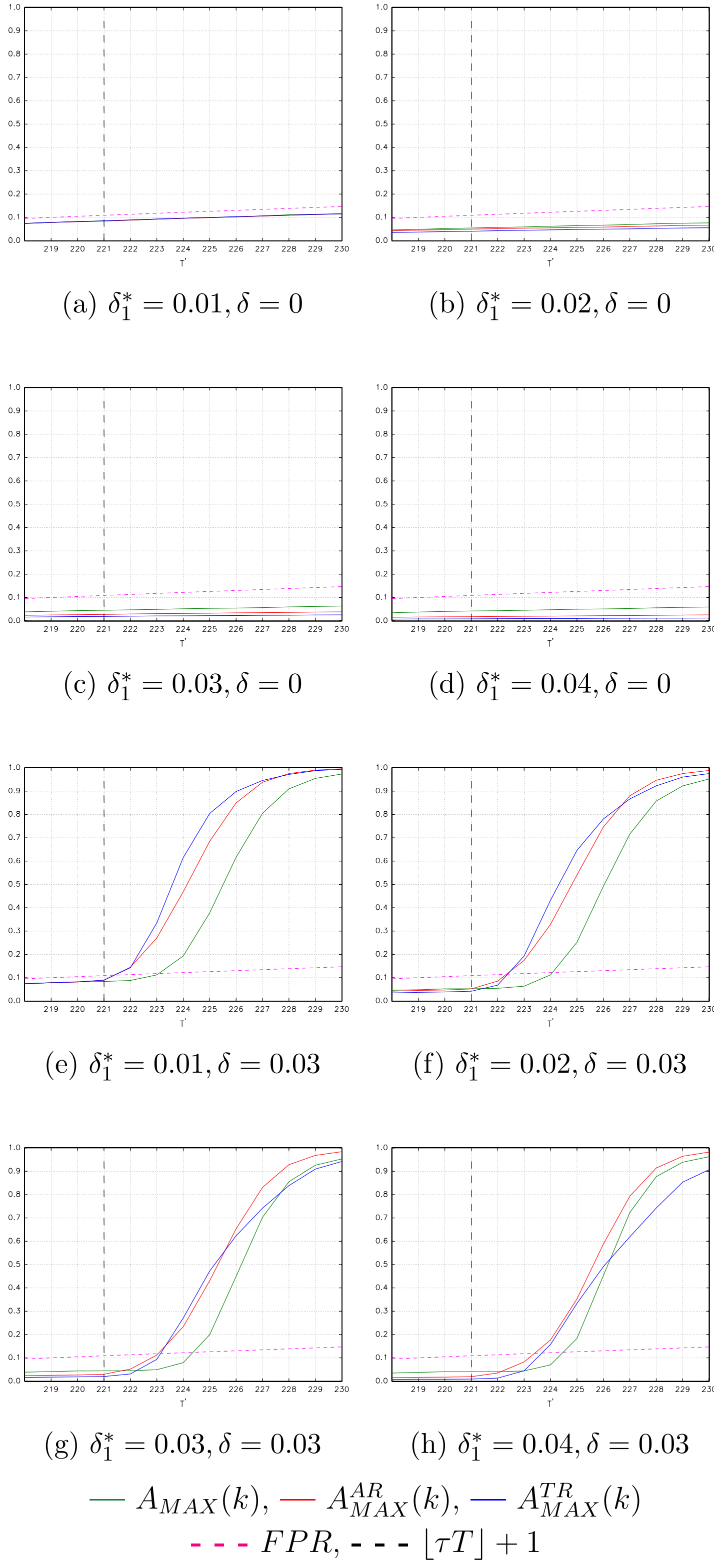


Fig. 3. Empirical rejection frequencies of $A_{MAX}(k)$, $A_{MAX}^{AR}(k)$, and $A_{MAX}^{TR}(k)$ with a training-sample bubble.

therefore offering the most conservative procedure, followed by $A^{AR}_{MAX}(k)$ and $A_{MAX}(k)$. Indeed, in the case of $\delta_1^* = 0.04$, $A^{TR}_{MAX}(k)$ has an empirical FPR close to zero.

Turning our attention now to empirical rejection frequencies obtained under the alternative hypothesis, $\delta > 0$, Figs. 3(e)–3(h) examine the empirical TPRs of the three procedures. As before, we set $\lfloor \tau T \rfloor = 1.1T^{\dagger}$ such that the start date of the explosive regime in the monitoring period is $\lfloor \tau T \rfloor + 1 = 221$. We set $\delta = 0.03$, such that these new results can be compared directly to the case without a training-sample bubble given in Fig. 1(k). First, we observe that the empirical TPR of all procedures is reduced by the presence of an explosive bubble in the monitoring period, in line with the reduced empirical FPRs. For example, at $T' = 226$, $A_{MAX}(k)$ has an empirical TPR of 0.716 in the case of $\delta_1^* = 0$, 0.616 with $\delta_1^* = 0.01$, 0.492 with $\delta_1^* = 0.02$, 0.449 with $\delta_1^* = 0.03$, and 0.453 with $\delta_1^* = 0.04$. $A^{AR}_{MAX}(k)$ has empirical TPRs of 0.904, 0.850, 0.747, 0.653, and 0.588, respectively, for the same settings, whilst $A^{TR}_{MAX}(k)$ has empirical TPRs of 0.937, 0.898, 0.781, 0.624, and 0.491. Nevertheless, it is pleasing to see that all procedures maintain reasonable levels of power to detect a monitoring period bubble, even with equivalent or greater magnitude bubbles occurring within the training period. We note that the new $A^{AR}_{MAX}(k)$ and $A^{TR}_{MAX}(k)$ procedures continue to outperform the original $A_{MAX}(k)$ procedure close to the explosive bubble start date, demonstrating their superior performance for early detection of a bubble. We do note, however, that for the high magnitude training-sample bubble $\delta_1^* = 0.04$, the empirical TPR of $A_{MAX}(k)$ overtakes that of $A^{TR}_{MAX}(k)$ further away from the bubble start date. In this case, $A^{AR}_{MAX}(k)$ still retains a small advantage over $A_{MAX}(k)$.

6. Empirical applications

To demonstrate how the superior performance demonstrated by our procedures in Section 4 can translate into earlier detection of an explosive bubble in practice, we now consider two empirical applications. In the first, we apply our procedures to the housing market in 17 OECD countries. In the second, we apply our procedures to bitcoin prices.

6.1. Housing market bubbles

It is now acknowledged that the rise and subsequent collapse of a bubble in the United States housing market was the catalyst for the global financial crisis of 2007–2008. The economic distress following the collapse of this so-called subprime bubble – named for the increase in mortgage lending to subprime borrowers which helped fuel rising house prices – indicates the widespread consequences that bubbles in housing markets can cause. As such, economists now pay significant attention to housing market dynamics.

Whilst economists are now agreed that the US housing market was subject to a bubble, there was a lack of consensus at the time. In an article published in 2002, Baker (2002) noted that over the previous seven years (1995–2002) US house prices had increased by 30% above the inflation rate and that this rise in prices was not supported by corresponding rises in fundamentals (such as rent) and was therefore due to a bubble. In contrast, then Federal Reserve Chairman Alan Greenspan remarked that as housing has high transaction costs and limited arbitrage opportunities, it was not reasonable to compare the housing market to the bubble and crash behaviour that was seen in stock markets (Monetary Policy and the Economic Outlook, 2002). The real-time monitoring procedures considered in this paper can provide empirical evidence of the emergence of asset price bubbles in situations such as this, where their presence is contested.

In previous work, Harvey et al. (2020) use a BIC-based bubble date-stamping procedure to examine the timing of bubble and crash regimes in the housing markets of 20 OECD countries from 1975–2018. Whilst patterns of behaviour differ somewhat across countries, the paper identifies three broad periods of explosivity that are common to many of these countries: the 1980s, the mid-2000s, and end-of-sample explosivity that was on-going as of 2018. It is therefore clear that the United States was not unique in demonstrating explosive behaviour in the housing market prior to the global financial crisis.

To confirm that explosive behaviour detected in a price series is the result of an asset bubble, we need to exclude the possibility that prices are being driven by explosive behaviour in the asset's fundamental value. If both the price series and the corresponding fundamental value of an asset are exhibiting the same growth behaviour, then prices can be considered justified by the fundamental value, and this would therefore not be a sign of a bubble. In the context of house prices, rent is commonly used as a proxy for the fundamental value of housing, given that rent is the return a homeowner can realise whilst holding the asset. In a pseudo-real-time monitoring exercise, Whitehouse et al. (2023) monitor for a bubble and crash in US house prices by applying the $A_{MAX}(k)$ procedure, and a subsequent crash detection procedure, to a house price-to-rent ratio. In doing so, they detect the emergence of a bubble in US house prices in 2000:Q1.

To examine the performance of our new test procedures, in this paper we apply $A^{AR}_{MAX}(k)$ and $A^{TR}_{MAX}(k)$, in addition to the original $A_{MAX}(k)$ procedure, to the logarithms of house price-to-rent ratios of 17 OECD countries[4] using data from the OECD (OECD, 2023). For all countries, monitoring begins in 1997:Q1. Our training sample begins in 1972:Q3 ($T' = 99$) for all countries except Belgium, where data are available only from 1976:Q2 ($T' = 84$).[5] We note that there is a wealth of empirical evidence (see, *inter alia*, (Holly et al., 2010; Malpezzi, 1999; Meen, 2002)) that real house prices in the US and UK exhibit unit root behaviour during the time period of our chosen training sample, lending validity to our unit root null hypothesis. We set $k = 10$ for all countries. Monitoring by each test procedure continues either until a bubble is detected or until 2010:Q1 if no bubble is detected by that procedure.

4 Australia, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Spain, Sweden, the United Kingdom, and the United States.

5 We focus on these 17 OECD countries only, as limited data availability for all other countries would leave us with $T' < 50$.

Table 2
House price-to-rent ratio: $A_{MAX}(k)$, $A^{AR}_{MAX}(k)$, and $A^{TR}_{MAX}(k)$ bubble detection dates.

Country	$A_{MAX}(k)$	$A^{AR}_{MAX}(k)$	$A^{TR}_{MAX}(k)$
Australia	**1999:Q4**	2000:Q1	2000:Q1
Belgium	2004:Q2	**1999:Q3**	**1999:Q3**
Canada	**2000:Q4**	**2000:Q4**	2001:Q1
Denmark	**1997:Q1**	**1997:Q1**	**1997:Q1**
Finland	**1998:Q1**	–	–
France	2006:Q2	**2004:Q1**	2005:Q2
Germany	–	–	–
Ireland	1999:Q4	**1998:Q4**	**1998:Q4**
Italy	–	2001:Q2	**2000:Q2**
Japan	–	–	–
Netherlands	1997:Q3	**1997:Q1**	**1997:Q1**
New Zealand	**1997:Q1**	**1997:Q1**	2003:Q1
Norway	–	**1997:Q3**	**1997:Q3**
Spain	–	**2004:Q1**	**2004:Q1**
Sweden	1999:Q4	**1999:Q3**	1999:Q4
United Kingdom	2003:Q4	**2000:Q1**	**2000:Q1**
United States	2000:Q1	**1999:Q1**	1999:Q3

Bold indicates the procedure(s) with the earliest detection date.

Table 2 displays the bubble detection date of $A_{MAX}(k)$, $A^{AR}_{MAX}(k)$ and $A^{TR}_{MAX}(k)$ for each country. We note that there are only two occasions where the original $A_{MAX}(k)$ procedure detects earlier than either of the two new procedures: Australia, where $A_{MAX}(k)$ detects a bubble in 1999:Q4 and both $A^{AR}_{MAX}(k)$ and $A^{TR}_{MAX}(k)$ detect a bubble one observation later at 2000:Q1; and Finland, where $A_{MAX}(k)$ detects a bubble in 1998:Q1 and neither $A^{AR}_{MAX}(k)$ nor $A^{TR}_{MAX}(k)$ detects a bubble within the monitoring horizon.

In contrast, the new $A^{AR}_{MAX}(k)$ procedure detects a bubble earlier than the original $A_{MAX}(k)$ for nine countries (and detects a bubble at the same time as $A_{MAX}(k)$ for an additional three countries), demonstrating this new procedure's ability to provide quicker detection of an emerging asset price bubble. The difference in detection dates by these two procedures can often be very large. In the United Kingdom, for example, whilst $A_{MAX}(k)$ detects a bubble in 2003:Q4, $A^{AR}_{MAX}(k)$ detects a bubble 15 quarters earlier in 2000:Q1. In Belgium, whilst $A_{MAX}(k)$ does not detect a bubble until 2004:Q2, $A^{AR}_{MAX}(k)$ detects 19 quarters earlier in 1999:Q3. Given that the goal of real-time monitoring of bubbles is to provide timely information to policymakers, it is evident that this new procedure can help policymakers respond more rapidly to bubbles as they emerge. Finally, we note that $A^{TR}_{MAX}(k)$ detects a bubble at the same time as $A^{AR}_{MAX}(k)$ on seven occasions and quicker than $A^{AR}_{MAX}(k)$ on one occasion—namely, Italy, where $A^{TR}_{MAX}(k)$ detects in 2000:Q2, $A^{AR}_{MAX}(k)$ in 2001:Q2, and $A_{MAX}(k)$ does not detect a bubble within the monitoring horizon.

Of course, earlier signalling of a house price bubble is not an advantage unless a bubble has actually emerged. We therefore want to confirm that the bubble detection offered by our new procedures correspond to explosive upwards movements in the price-to-rent series and are not spurious detections. Fig. 4 displays the house price-to-rent ratios of the 17 OECD countries discussed, along with the bubble detection dates offered by $A_{MAX}(k)$, $A^{AR}_{MAX}(k)$, and $A^{TR}_{MAX}(k)$, and the theoretical FPR given by (6). It is clear from visual inspection of the full sample of data (which, of course, would not have been possible in real time) that the earlier detection offered by the new methods is justified. Consider the United States, for example, in Fig. 4(q). Observing the full sample of data, it is apparent that the price-to-rent ratio undergoes a period of rapid growth from the late 1990s to the mid-2000s, which is detected four quarters earlier by $A^{AR}_{MAX}(k)$ (1999:Q1) than $A_{MAX}(k)$ (2000:Q1), and two quarters earlier by $A^{TR}_{MAX}(k)$ (1999:Q3). Whilst the result of $A_{MAX}(k)$ here replicates that found by Whitehouse et al. (2023), the two new procedures are able to detect the bubble in the United States housing market even earlier than our previous paper.

To formally test for the presence of explosive bubbles in the monitoring period, we conduct the sub-sample, right-tailed unit root test procedure of Phillips et al. (2015), known as the GSADF test. This test is applied to data from 1997:Q1–2010:Q1, yielding a sample of size $T = 53$. The GSADF test results are provided in Table 3. Evidence of explosive behaviour is found in the monitoring period for 14 out of the 17 countries studied at a 0.05 significance level. This includes in two countries (Italy and Spain in Figs. 4(i) and 4(n), respectively) where the new monitoring procedures detect a bubble and the original $A_{MAX}(k)$ procedure does not.

Examining Fig. 4 also allows us to consider the dynamics of the training sample in each country. In particular, we observe that in many countries, the training sample itself contains possibly explosive regimes. To formally test for the presence of bubbles within the training sample, we can also apply the GSADF test procedure to this training sample data from $t = 1, \dots, T^*$. The GSADF test results for the training sample are also provided in Table 3. We find evidence of explosive behaviour in the training sample of eight out of 17 countries at a 0.05 significance level. Of these eight countries, $A^{AR}_{MAX}(k)$ provides an earlier detection date than $A_{MAX}(k)$ on seven occasions (the exception being Finland, where only $A_{MAX}(k)$ detects a bubble during monitoring), and $A^{TR}_{MAX}(k)$ provides an earlier detection date than $A_{MAX}(k)$ on six occasions (the additional exception being Sweden, where it detects at the same time as $A_{MAX}(k)$). From Section 5, we know that the presence of training-sample bubbles results in both reduced empirical FPRs and reduced empirical TPRs for all procedures. It is therefore reassuring firstly that the monitoring test procedures are still able to detect bubbles at all in these circumstances in the vast majority of cases, and secondly that the new tests continue to outperform the original $A_{MAX}(k)$ in terms of early detection.

Finally, we note that an additional advantage of the early detection offered by these new procedures, is that detection occurs at a point where the theoretical FPR is lower. Consider the United States, for example. $A_{MAX}(k)$ detects a bubble at 2000:Q1 with a theoretical FPR of 0.141, whereas $A^{AR}_{MAX}(k)$ detects a bubble at 1999:Q1 with a theoretical FPR of 0.102. For a practitioner concerned about detecting at a given level of significance, differences such as these may be meaningful.

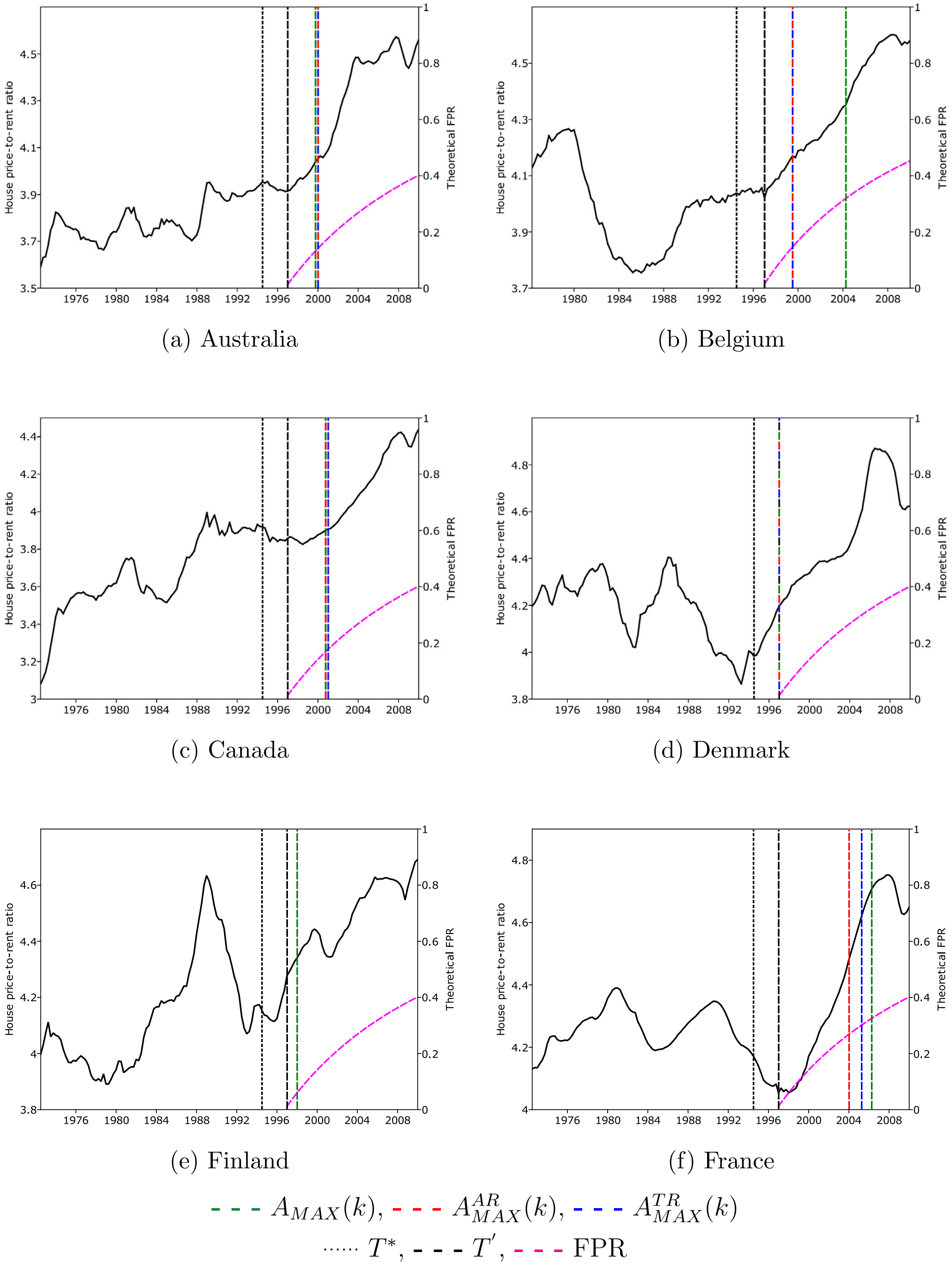


Fig. 4. House price-to-rent ratio: $A_{MAX}(k)$, $A^{AR}_{MAX}(k)$ and $A^{TR}_{MAX}(k)$ bubble monitoring.

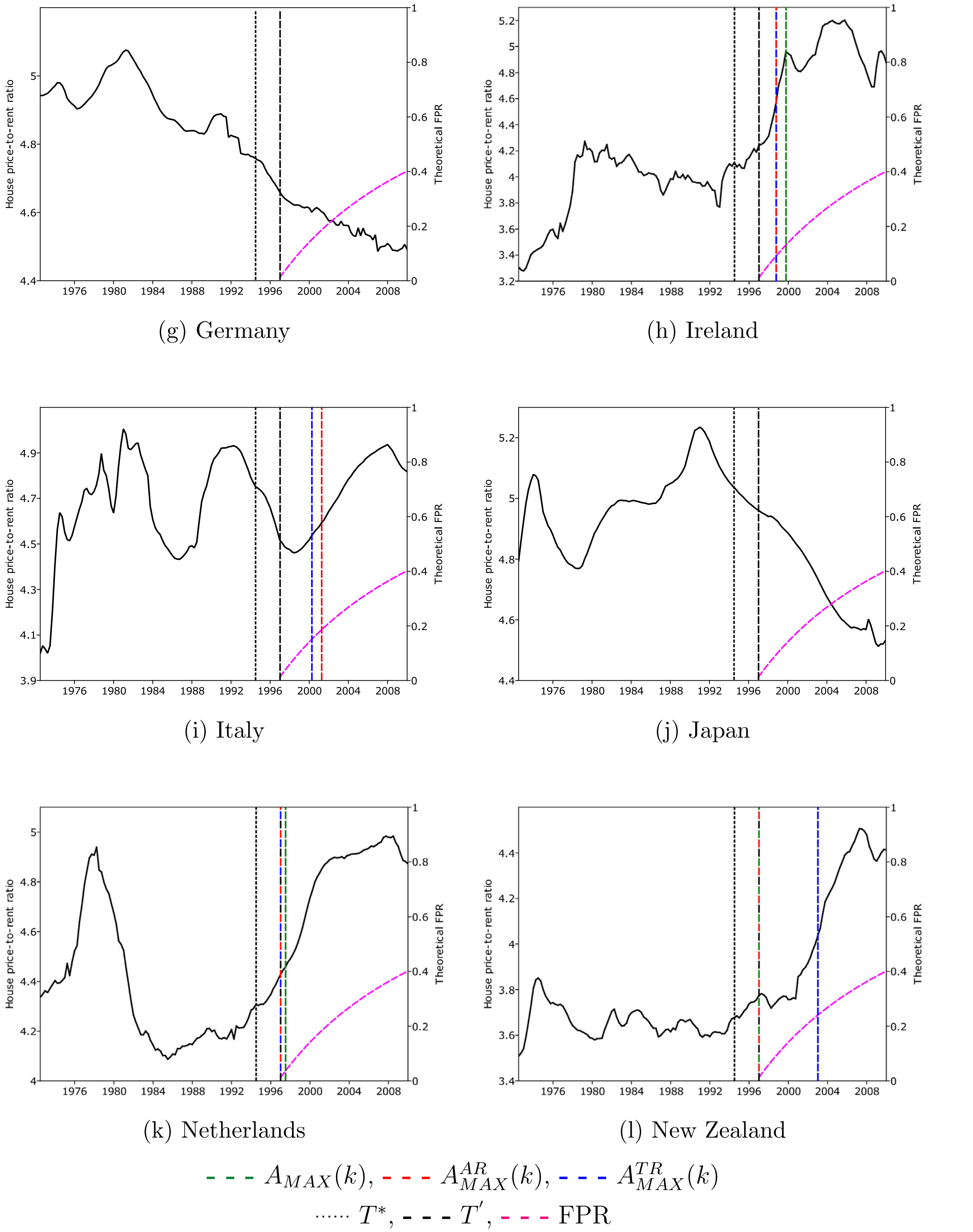
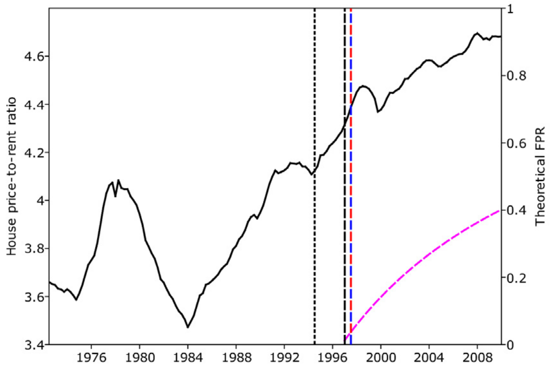


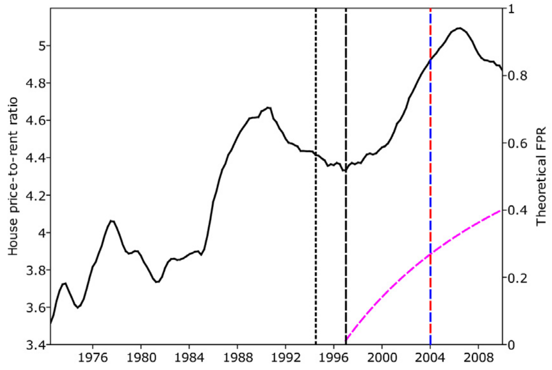
Fig. 4. (*continued*).

6.2. Bitcoin explosivity

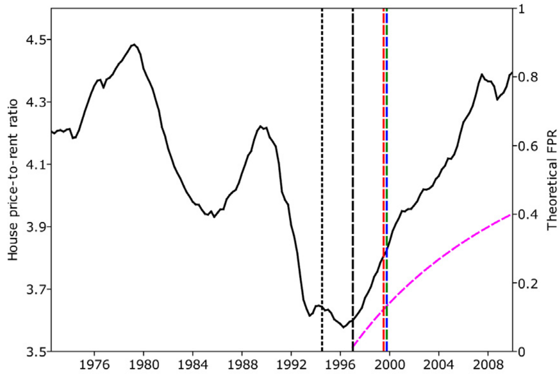
In March 2020, global instability due to the Covid-19 pandemic resulted in substantial declines to stock and bond markets worldwide. Despite suggestions that bitcoin could act as a safe-haven against these traditional asset classes, due to being independent from monetary policy and, historically, having weak correlation with traditional



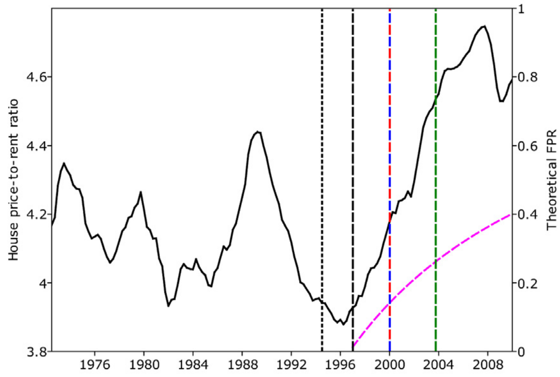
(m) Norway



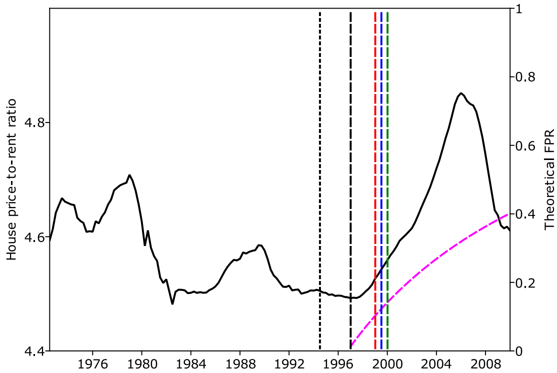
(n) Spain



(o) Sweden



(p) United Kingdom



(q) United States

- - - $A_{MAX}(k)$, - - - $A^{AR}_{MAX}(k)$, - - - $A^{TR}_{MAX}(k)$
······ T^*, - - - T', - - - FPR

Fig. 4. (*continued*).

asset price movements, bitcoin also suffered in this time period (Conlon & McGee, 2020), with prices crashing by approximately 35% between March 6th and March 29th. However, this collapse in bitcoin was short-lived, with prices recovering to their pre-Covid level by May 2020 and continuing to rise throughout the year, with 2020

Table 3
House price-to-rent ratio: GSADF tests.

Country	Training sample	Monitoring sample
Australia	1.014	2.704**
Belgium	3.342***	3.853***
Canada	1.951	4.477***
Denmark	1.334	5.018***
Finland	5.833***	9.862***
France	1.630	4.254***
Germany	0.981	0.858
Ireland	2.501**	1.589
Italy	1.128	2.434**
Japan	2.300*	4.472***
Netherlands	3.959***	3.140**
New Zealand	0.006	6.352***
Norway	2.926**	1.306
Spain	1.957	3.141**
Sweden	2.590**	2.938**
United Kingdom	2.653**	2.225**
United States	3.278***	3.721***

*, **, and *** denote rejection at the 0.10, 0.05, and 0.01 levels, respectively.

Table 4
Bitcoin: $A_{MAX}(k)$, $A^{AR}_{MAX}(k)$ and $A^{TR}_{MAX}(k)$ bubble detection dates.

$A_{MAX}(k)$	$A^{AR}_{MAX}(k)$	$A^{TR}_{MAX}(k)$
11th Oct 2020	10th Oct 2020	**9th Oct 2020**

Bold indicates the procedure(s) with the earliest detection date.

returns substantially larger than both the S&P 500 and the well-known safe-haven asset of gold. In this application, we investigate the possibility that bitcoin exhibited explosive behaviour in prices in 2020.

It is well known that financial data, and especially cryptocurrencies such as bitcoin, can exhibit time-varying conditional volatility. As demonstrated in Section 4, the procedures considered in this paper are robust to conditionally heteroskedastic innovation processes. To confirm that our procedures are able to monitor an emerging bubble in this type of financial data, we now apply the $A_{MAX}(k)$, $A^{AR}_{MAX}(k)$, and $A^{TR}_{MAX}(k)$ procedures to bitcoin prices.

As discussed above, an asset price bubble can be defined as the presence of explosive behaviour in a price series and the absence of explosive behaviour in the asset's fundamental value. For example, we may use rent as a proxy for the fundamental value of housing (as in Section 6.1), or we may use dividends as a proxy for the fundamental value of the stock market. However, in the context of cryptocurrencies, there is much less consensus within the academic literature about how to quantify the fundamental value, or indeed whether a cryptocurrency has a fundamental value at all (see e.g. Gronwald (2021)). Given the above, in what follows we test bitcoin prices directly for explosive behaviour, noting that there is disagreement in the academic literature about whether or not such explosiveness should be deemed an asset price bubble.

We test the logarithms of daily BTC–USD prices, with data obtained from Yahoo Finance. Monitoring begins on Monday August 3rd, 2020. Our training sample begins on January 1st, 2020 ($T' = 216$) and we set $k = 10$ as before. Table 4 displays the bubble detection date of $A_{MAX}(k)$, $A^{AR}_{MAX}(k)$, and $A^{TR}_{MAX}(k)$. All three procedures are able to detect a bubble in bitcoin prices, with $A^{TR}_{MAX}(k)$ detecting first on October 9th, 2020, followed by $A^{AR}_{MAX}(k)$ one day later, and $A_{MAX}(k)$ two days later. As in the case of house prices in Section 6.1, we want to confirm that this detection corresponds to the emergence of explosive behaviour and is not spurious. Fig. 5 displays the logarithm of BTC:USD prices in addition to the detection dates offered by each procedure. From visual inspection of the full sample of data (which would not be possible in a real-time monitoring scenario) it is clear that the detection dates correspond to the start of a period of rapid growth that continues into 2021. As above, we formally test for the presence of explosivity in the monitoring period (August 3rd, 2020 to June 30th, 2021; $T = 332$) using the GSADF test. The GSADF test results in Table 5 show that evidence of explosive behaviour is found at a 0.01 significance level. As in the case of several house price series that we examined, there is also evidence of explosive behaviour in the training sample, which does not prevent the detection of emerging explosive behaviour in monitoring.

Table 5
Bitcoin: GSADF tests.

Training sample	Monitoring sample
3.259***	3.403***

*** denotes rejection at the 0.01 level

Finally, we note that the two empirical applications offered in Section 6, to quarterly house prices and daily bitcoin prices, demonstrate that the proposed monitoring procedures perform well in the context of both low-frequency macroeconomic data and higher-frequency financial data.

7. Conclusion

Prompt detection of an emerging asset price bubble is crucial, as it allows policymakers to react to the event in a timely manner. In this paper we provided two new real-time monitoring test statistics for the emergence of an asset price bubble. These statistics are based on modifications of the recently developed AHLST procedure, using variance standardisations that are more appropriate under the alternative hypothesis of an explosive bubble.

We showed that our new monitoring procedures maintain the same asymptotic properties as the original procedure under the null hypothesis of no explosive bubble. Using finite sample simulations, our new procedures were shown to deliver substantial advantages over the original procedure in terms of the true positive rate for detecting an explosive regime, which in turn leads to earlier detection of an emergent bubble. These advantages were shown to hold even when an explosive regime is present within the training sample of data, a situation which may well occur in practice. An empirical application to the house price-to-rent ratio in 17 OECD countries demonstrated the ability of our new test procedures to provide earlier detection of an emerging asset price bubble. In particular, we noted

Fig. 5. Bitcoin prices: $A_{MAX}(k)$, $A^{AR}_{MAX}(k)$, and $A^{TR}_{MAX}(k)$ bubble monitoring.

that the United States subprime bubble could have been identified as early as 1999:Q1 with our new procedures, whilst a bubble could have been detected in the United Kingdom's housing market in 2000:Q1. Combined with an application to bitcoin prices, we were able to demonstrate that our new procedures can detect emerging bubbles in both macroeconomic and financial data.

We also note that a practitioner who monitors for the emergence of an asset price bubble may also be interested in monitoring for the collapse of that bubble. Whitehouse et al. (2023) propose a monitoring procedure that can detect collapsing bubbles in real time, which is reliant on first detecting the emergence of a bubble using $A_{MAX}(k)$. Therefore, improving the speed of detection of a bubble using either $A^{AR}_{MAX}(k)$ or $A^{TR}_{MAX}(k)$ will improve the efficiency of the Whitehouse et al. (2023) crash monitoring procedure.

CRediT authorship contribution statement

E.J. Whitehouse: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **D.I. Harvey:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **S.J. Leybourne:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization.

Data and code availability

The data and Gauss code for this paper can be downloaded at https://sites.google.com/site/ejwhitehouse1.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Proof of Theorem 1

We first establish the result for $A^{AR}_{e,k}$. The numerator of this statistic is a measureable function of a finite number of observations on Δy_t, and under H_0, $\Delta y_t = \varepsilon_t$. In the denominator, the residuals $\hat{\varepsilon}_t$, obtained from a regression over the sub-sample $t = e-k+1, \ldots, e$, can be written as

$$\hat{\varepsilon}_t = (\Delta y_t - \overline{\Delta y}) - \hat{b}(y_{t-1} - \bar{y}_{-1})$$

where

$$\overline{\Delta y} = k^{-1} \sum_{t=e-k+1}^{e} \Delta y_t, \qquad \bar{y}_{-1} = k^{-1} \sum_{t=e-k+1}^{e} y_{t-1},$$

$$\hat{b} = \frac{\sum_{t=e-k+1}^{e} (\Delta y_t - \overline{\Delta y})(y_{t-1} - \bar{y}_{-1})}{\sum_{t=e-k+1}^{e} (y_{t-1} - \bar{y}_{-1})^2}.$$

Now,

$$\begin{aligned} y_{t-1} - \bar{y}_{-1} &= y_{t-1} - k^{-1} \sum_{s=e-k+1}^{e} y_{s-1} \\ &= (y_{t-1} - y_{e-k-1}) - k^{-1} \sum_{s=e-k+1}^{e} (y_{s-1} - y_{e-k-1}) \\ &= (u_{t-1} - u_{e-k-1}) - k^{-1} \sum_{s=e-k+1}^{e} (u_{s-1} - u_{e-k-1}) \end{aligned}$$

$$= \sum_{j=e-k}^{t-1} \varepsilon_j - k^{-1} \sum_{s=e-k+1}^{e} \sum_{j=e-k}^{s-1} \varepsilon_j.$$

Hence, $\hat{\varepsilon}_t$, $t = e-k+1, \ldots, e$, and consequently the denominator of $A^{AR}_{e,k}$, is also a measureable function of the same ε_t as in the numerator of $A^{AR}_{e,k}$. It then follows that $\{A^{AR}_{e,k}\}$ is a strictly stationary sequence and, from Lemma 2.1 of White and Domowitz (1984), it is mixing of the same size as $\{\varepsilon_t\}$. Theorem 1 assumes that the mixing conditions of $\{\varepsilon_t\}$, and hence $\{A^{AR}_{e,k}\}$, satisfy the mixing conditions of Ferreira and Scotto (2002) (see Definition on p. 476). Hence, the conditions underpinning the result in Theorem 2.1 of Ferreira and Scotto (2002) are satisfied for the sequence $\{A^{AR}_{e,k}\}$.

Theorem 2.1 of Ferreira and Scotto (2002) for the case $r = s = 1$ (in their notation) then implies that, for two disjoint subintervals $I_{T,a}$ and $I_{T,b}$ of $[1, T]$, with respective lengths a_T and b_T such that $a_T/T \to a$ and $b_T/T \to b$,

$$\lim_{T\to\infty} P\left(\max_{e\in I_{T,a}} A^{AR}_{e,k} \le \max_{e\in I_{T,b}} A^{AR}_{e,k}\right) = \frac{b}{a+b}$$

or

$$\lim_{T\to\infty} P\left(\max_{e\in I_{T,a}} A^{AR}_{e,k} > \max_{e\in I_{T,b}} A^{AR}_{e,k}\right) = \frac{a}{a+b}.$$

Setting $I_{T,a} = [T^*+k, T']$ and $I_{T,b} = [k+1, T^*]$, we obtain

$$\lim_{T\to\infty} P\left(\max_{e\in[T^*+k,T']} A^{AR}_{e,k} > \max_{e\in[k+1,T^*]} A^{AR}_{e,k}\right)$$
$$= \lim_{T\to\infty}\left(\frac{T'-T^*-k+1}{T'-2k+1}\right)$$
$$= \lim_{T\to\infty}\left(\frac{T'-T^*}{T'}\right)$$
$$= \alpha.$$

The result for $A^{TR}_{e,k}$ follows in the same way, but more directly, since the $\hat{e}_t$, $t = e-k+1, \ldots, e$, used in the denominator of $A^{TR}_{e,k}$, are obtained from a regression of Δy_t on a constant and trend (and not y_{t-1}). Hence, it is more readily apparent that $\hat{e}_t$ is a measureable function of the same ε_t as in the numerator of $A^{TR}_{e,k}$.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijforecast.2024.12.005.

Data and code availability

The data and Gauss code for this paper can be downloaded at https://sites.google.com/site/ejwhitehouse1.

References

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

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

Anundsen, A. K. (2019). Detecting imbalances in house prices: What goes up must come down? *Scandinavian Journal of Economics*, *121*, 1587–1619.

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

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

Astill, S., Harvey, D. I., Leybourne, S. J., Taylor, A., & Zu, Y. (2024). Cusum-based monitoring for explosive episodes in financial data in the presence of time-varying volatility. *Journal of Financial Econometrics*, *21*, 187–227.

Baker, D. (2002). The run-up in home prices: Is it real or is it another bubble? (p. BP200208A). Center for Economic Policy and Research, Briefing Paper.

Basse, T., Klein, T., Vigne, S. A., & Wegener, C. (2021). U.S. stock prices and the dot-com bubble: Can dividend policy rescue the efficient market hypothesis? *Journal of Corporate Finance*, *67*, Article 101892.

Caspi, I., & Graham, M. (2018). Testing for bubbles in stock markets with irregular dividend distribution. *Finance Research Letters*, *26*, 89–94.

Chu, C. S., Stinchcombe, J., & White, H. (1996). Monitoring structural change. *Econometrica*, *64*, 1045–1065.

Conlon, T., & McGee, R. (2020). Safe haven or risky hazard? Bitcoin during the Covid-19 bear market. *Finance Research Letters*, *35*, Article 101607.

Crowe, C., Dell'Ariccia, G., Igan, D., & Rabanal, P. (2013). How to deal with real estate booms: Lessons from country experiences. *Journal of Financial Stability*, *9*, 300–319.

Etienne, X. L., Irwin, S. H., & Garcia, P. (2014). Bubbles in food commodity markets: Four decades of evidence. *Journal of International Money and Finance*, *42*, 129–155.

Ferreira, H., & Scotto, M. (2002). On the asymptotic location of high values of a stationary sequence. *Statistics & Probability Letters*, *60*, 475–482.

Figuerola-Ferretti, I., & McCrorie, J. (2016). The shine of precious metals around the global financial crisis. *Journal of Empirical Finance*, *38*, 717–738.

Gronwald, M. (2021). How explosive are cryptocurrency prices? *Finance Research Letters*, *38*, Article 101603.

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., Leybourne, S., & Whitehouse, E. (2020). Date-stamping multiple bubble regimes. *Journal of Empirical Finance*, *58*, 226–246.

Holly, S., Pesaran, M. H., & Yamagata, T. (2010). A spatio-temporal model of house prices in the USA. *Journal of Econometrics*, *158*, 160–173.

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

Horvath, L., & Trapani, L. (2024). Real-time monitoring with RCA models. Discussion paper arXiv:2312.11710v1.

Hu, Y., & Oxley, L. (2018). Do 18th century 'bubbles' survive the scrutiny of 21st century time series econometrics? *Economics Letters*, *162*, 131–134.

Malpezzi, S. (1999). A simple error correction model of house prices. *Journal of Housing Economics*, *8*, 27–62.

Meen, G. (2002). The time-series behavior of house prices: A transatlantic divide? *Journal of Housing Economics*, *11*, 1–23.

Monetary Policy and the Economic Outlook (2002). Hearing before the joint economic committee. In *Congress of the United states (testimony of Alan Greenspan). 107th cong, April 17*.

OECD (2023). Housing prices (indicator). http://dx.doi.org/10.1787/63008438-en.

Otero, J., Panagiotidis, T., & Papapanagiotou, G. (2022). Testing for exuberance in house prices using data sampled at different frequencies. *Studies in Nonlinear Dynamics & Econometrics*, *26*, 675–691.

Pastor, D. J., & Ewing, B. T. (2022). Is there evidence of mild explosive behavior in Alaska North Slope crude oil prices? *Energy Economics*, *114*.

Pavlidis, E. G., Paya, I., & Peel, D. A. (2018). Using market expectations to test for speculative bubbles in the crude oil market. *Journal of Money, Credit and Banking*, *50*, 833–856.

Phillips, P. C. B., & Shi, S. (2020). Real time monitoring of asset markets: Bubbles and crises. In H. D. Vinod, & C. Rao (Eds.), *Handbook of statistics, vol. 42* (pp. 61–80). Elsevier.

Phillips, P. C. B., Shi, S., & Yu, J. (2015). Testing for multiple bubbles: historical episodes of exuberance and collapse in the S & P 500. *International Economic Review*, *56*, 1043–1078.

Phillips, P. C. B., Wu, Y., & Yu, J. (2011). Explosive behavior in the 1990s NASDAQ: When did exuberance escalate asset values? *International Economic Review*, *52*.

White, H., & Domowitz, I. (1984). Nonlinear regression with dependent observations. *Econometrica*, *52*, 143–162.

Whitehouse, E. J. (2019). Explosive asset price bubble detection with unknown bubble length and initial condition. *Oxford Bulletin of Economics and Statistics*, *81*, 20–41.

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