Testing for Co-explosive Behaviour in Financial Time Series

ANDRIA C. EVRIPIDOU,† DAVID I. HARVEY,† STEPHEN J. LEYBOURNE†
and ROBERT SOLLIS‡

†*School of Economics, University of Nottingham, UK (e-mail: dave.harvey@nottingham.ac.uk)*
‡*Newcastle University Business School, UK*

Abstract

This article proposes a test to determine if two price series that each contain an explosive autoregressive regime consistent with the presence of a bubble are related in the sense that a linear combination of them is integrated of order zero. We refer to such a phenomenon as 'co-explosive behaviour', and propose a test based on a stationarity testing framework. The test allows the explosive episode in one series to lead (or lag) that in the other by a number of time periods. We establish the asymptotic properties of the test statistic and propose a wild bootstrap procedure for obtaining critical values that are robust to heteroskedasticity. Simulations show that the proposed test has good finite sample size and power performance. An empirical application to detect whether co-explosive behaviour exists among a set of precious and non-ferrous metals is presented.

I.   Introduction

Detecting the presence of asset price bubbles has long been of interest to applied economists. Following the global financial crisis in 2008, which was preceded by suspected price bubbles in housing, commodity, and stock markets, there has been a renewed interest in econometric methods for detecting bubbles, and several new tests for bubbles have been proposed. The tests developed by Phillips, Wu and Yu (2011) (PWY) and Phillips, Shi and Yu (2015) (PSY) adopt an autoregressive framework and propose tests for bubbles based on whether a price series contains an explosive autoregressive component, while the corresponding fundamentals series does not. These tests have proved to be extremely popular in empirical research. In empirical applications, PWY and PSY apply their tests to detect US stock market bubbles, and the tests have also been found to be useful for detecting bubbles in commodity prices (Gutierrez, 2012; Figuerola-Ferretti, Gilbert and McCrorie, 2015; Harvey *et al.*, 2016), real estate prices (e.g. Pavlidis *et al.,* 2016; Deng *et al.*, 2017), and bond markets (e.g. Phillips and Yu, 2011; van Lamoen, Mattheussens and Droes, 2017; Huston and Spencer, 2018). It is widely recognized that these types of bubble-detection tests can provide very useful information for central banks, financial regulators, and investors.

JEL Classification numbers: C22, C12, G14.

In contrast to the burgeoning literature on detecting bubbles in individual price series via econometric methods for detecting explosive autoregressive behaviour, a much smaller literature exists on modelling the relationship *between* price series that contain bubbles. However, it is important for risk managers, portfolio managers, and monetary policymakers to understand whether bubbles in different markets are related. A particularly important issue is whether bubbles that appear in one market are likely to migrate to another market, since if bubbles migrate between markets, their systemic risk to the financial sector as a whole will be greater than if they do not. Phillips and Yu (2011) propose a method for detecting bubble migration between two price series, based on dating bubble regimes in the individual series, obtaining two sets of recursively estimated autoregressive parameters, and then testing for a relationship between the recursive estimates using regression-based methods. Other papers that have considered the relationship between different price series that contain bubbles have tended to focus primarily on comparing the timeline of when explosive autoregressive behaviour is detected in the individual series, using the bubble-dating methodology proposed by PWY and PSY, see, for example, Shi *et al.* (2016) and Pavlidis *et al.* (2016) for housing market applications.

In this article, we focus on the explosive autoregressive characteristics of bubble series, and our approach differs from the aforementioned research in that, rather than comparing the results from simply detecting explosive autoregressive regimes in each series individually, we develop a regression-based test to determine whether two series with explosive regimes are related. More specifically, in the context of two price series that each contain one or more regimes of explosive autoregressive behaviour consistent with the presence of bubbles, our test can be used to determine if a linear combination of the series is integrated of order zero, $I(0)$, a property we refer to as 'co-explosive behaviour'.[1] Our work is in the same vein as Engsted and Nielsen (2012) who consider co-explosive behaviour in the context of a bivariate vector autoregression (where the explosive behaviour is present for the full sample), testing the null of no co-explosivity against the alternative of co-explosive behaviour. The test we develop in this article employs a variant of the stationarity test of Kwiatkowski *et al.* (1992) (KPSS) to test the null of co-explosive behaviour, allowing for the explosive behaviour in the series to possibly be present in sub-sample regimes only. The test is straightforward to compute and allows an explosive episode in one series to lag or lead that in the other by a number of time periods, thereby allowing detection of contemporaneous or dynamic correlation between series that contain explosive regimes, and providing information on the nature of potential explosive regime migration from one price series to another. In the context of processes with mildly explosive regimes, we show that the asymptotic null distribution of the test statistic does not depend on the precise properties of the regressor series, but will depend on the pattern of any heteroskedasticity present in the regression model errors – a relevant consideration when any application involves financial data. To overcome this problem, we adopt a simple wild bootstrap procedure that reproduces that same pattern of heteroskedasticity in the bootstrap samples, thereby allowing asymptotic size-controlled

---

[1]In what follows, a generic $I(0)$ series ϵ_t say, is one which permits a representation $\epsilon_t = c(L)\varepsilon_t$ where L is the lag operator and $c(L) = \sum_{k=0}^{\infty} c_k L^k$ with $\sum_{k=0}^{\infty} k|c_k| < \infty$ and $c(1) \neq 0$. Here ε_t is a (possibly heteroskedastic) martingale difference sequence, cf. Cavaliere and Taylor (2009).

inference to be conducted. Monte Carlo simulation exercises show that the bootstrap procedure appears to control size well in the presence of heteroskedasticity. We also demonstrate the power performance of our bootstrap test to reject under an alternative of no co-explosivity, and examine the ability of the test to identify the timing of explosive regime migration when co-explosivity is present.

As an empirical application, we employ our test to investigate whether co-explosive behaviour exists among a set of precious and non-ferrous metals. Several recent empirical studies have detected periods of explosive autoregressive behaviour in metals prices consistent with the presence of bubbles – see Figuerola-Ferretti *et al.* (2015), Harvey *et al.* (2016) and Figuerola-Ferretti and McCrorie (2016). However, there has been very little previous empirical research on the impact of possible bubbles on the long-run relationship between metals prices. While Escribano and Granger (1998) and Baur and Tran (2014) do consider this issue, they focus only on the relationship between Gold and Silver prices. Furthermore, both studies employ orthodox cointegration methods and attempt to capture the impact of a bubble on the long-run relationship between the series using dummy variables to allow for level shifts in the long-run model. Importantly, neither of these studies explicitly allow for the presence of explosive autoregressive behaviour in the regression models and subsequent hypothesis tests. Our co-explosive testing methodology does not suffer from this limitation, and therefore can provide robust insight into whether co-movements in explosive metal prices can be detected. Note that for commodity prices the underlying fundamental (equivalent to the dividend for stocks) is an unobserved 'convenience yield'; see Pindyck (1993). Figuerola-Ferretti *et al.* (2015) have, however, cast doubt on the efficacy of using an imputed convenience yield as a fundamentals proxy variable to support running a bubble test on the ratio of the raw prices and this imputed value. In our empirical exercise, we simply report results using the price series alone, recognizing that findings of explosive autoregressive behaviour do not necessarily imply the presence of bubbles when only the price series are examined.

The article is organized as follows. In section II we outline the range of explosive models we consider for an individual series. Our model for a co-explosive relationship between two individual series is presented in section III. Our test for co-explosive behaviour and its asymptotic properties are explained in section IV; this section also discusses the wild bootstrap algorithm and confirms its large sample validity. In section V, the finite sample size and power properties of the co-explosive test procedure are considered. Section VI discusses how to conduct the test when the lag/lead timing is unknown. Our application to metals prices is discussed in section VII, while section VIII concludes. Proofs of our main results and additional simulation and empirical results are collected in Appendix S1. In what follows, we adopt the following notation: $\lfloor \cdot \rfloor$ denotes the integer part, $\mathbf{1}(\cdot)$ the indicator function, $\overset{w}{\longrightarrow}$ weak convergence, $\overset{p}{\longrightarrow}$ convergence in probability, and $\overset{w}{\longrightarrow}_p$ weak convergence in probability.

II.   The explosive autoregressive model

We consider the following DGP for a generic series x_t, $t = 1, \ldots, T$, which is similar to that used in Harvey, Leybourne and Sollis (2017),

$$x_t = \mu_x + u_{x,t}$$

$$u_{x,t} = \begin{cases} u_{x,t-1} + \epsilon_{x,t}, & t = 2, \ldots, \lfloor \tau_{x,1} T \rfloor, \\ (1 + \delta_{x,1})u_{x,t-1} + \epsilon_{x,t}, & t = \lfloor \tau_{x,1} T \rfloor + 1, \ldots, \lfloor \tau_{x,2} T \rfloor, \\ (1 - \delta_{x,2})u_{xt-1} + \epsilon_{x,t}, & t = \lfloor \tau_{x,2} T \rfloor + 1, \ldots, \lfloor \tau_{x,3} T \rfloor, \\ u_{x,t-1} + \epsilon_{x,t}, & t = \lfloor \tau_{x,3} T \rfloor + 1, \ldots, T \end{cases} \tag{1}$$

where μ_x is a constant term, $\tau_{x,1}$, $\tau_{x,2}$ and $\tau_{x,3}$ are sample fractions, and $\delta_{x,1} > 0$ and $\delta_{x,2} > 0$. We assume that $u_{x,1} = o_p(T^{1/2})$ and that $\epsilon_{x,t}$ is an error process with zero mean and is integrated of order zero, $I(0)$. As such, the model (1) imposes autoregressive unit root, or $I(1)$, behaviour, on x_t up to time $\lfloor \tau_{x,1} T \rfloor$, after which x_t is an explosive autoregressive process until time $\lfloor \tau_{x,2} T \rfloor$. This is followed by a stationary (mean-reverting) regime until time $\lfloor \tau_{x,3} T \rfloor$. The series then returns to $I(1)$ behaviour for the final regime.

In terms of modelling possible bubble behaviour, (1) allows for a number of different specifications (models) simply by changing $\tau_{x,1}$, $\tau_{x,2}$ and $\tau_{x,3}$. If $0 < \tau_{x,1} < \tau_{x,2} = 1$ the explosive regime, possibly arising from a bubble, emerges at (sample proportion) $\tau_{x,1}$ and runs up to the sample end (Model 1); while if $0 < \tau_{x,1} < \tau_{x,2} = \tau_{x,3} < 1$ the explosive regime terminates at $\tau_{x,2}$ and is unit root to the sample end (Model 2). If $0 < \tau_{x,1} < \tau_{x,2} < \tau_{x,3} = 1$ the explosive regime terminates with an offsetting stationary collapse regime, proxying a possible bubble crash, that continues to the sample end (Model 3); while if $0 < \tau_{x,1} < \tau_{x,2} < \tau_{x,3} < 1$, the collapse is followed by a unit root period to the sample end (Model 4). Here $\delta_{x,1}$ and $\tau_{x,2} - \tau_{x,1}$ together control the magnitude and duration of the explosive period, while $\delta_{x,2}$ and $\tau_{x,3} - \tau_{x,2}$ determine the speed and duration of the collapse (should one occur as in Models 3 and 4). For simplicity, we refer to an explosive regime, possibly followed by a stationary collapse regime, as an 'explosive episode'.

In what follows the explosive autoregressive component $\delta_{x,1}$ will be parameterized as $\delta_{x,1} = \delta_{x,1,T} = c_{x,1} T^{-\alpha_{x,1}}$ with $\alpha_{x,1} \in (0, 1)$, such that the process is mildly explosive (cf. PWY). For the stationary collapse autoregressive component, we correspondingly assume $\delta_{x,2} = \delta_{x,2,T} = c_{x,2} T^{-\alpha_{x,2}}$ with $\alpha_{x,2} \in (0, 1)$.

III. The co-explosive model

Suppose y_t and x_t are two observed series, with x_t containing an explosive episode that has been generated from one of Models 1–4. Also, let z_t define a latent, or unobserved, process that is also generated from one of Models 1–4, where x is replaced by z in (1), and $\delta_{z,1,T}$, $\delta_{z,2,T}$, $c_{z,1}$, $c_{z,1}$, $\alpha_{z,1}$ and $\alpha_{z,2}$ are the corresponding model parameters with $\alpha_{z,1}$, $\alpha_{z,2} \in (0, 1)$. We may then consider a DGP for y_t of the following form:

$$y_t = \mu_y + \beta_x x_{t-i} + \beta_z z_t + \epsilon_{y,t}, \tag{2}$$

where $\epsilon_{y,t}$ is a mean zero $I(0)$ error term. We permit correlation to exist between $\epsilon_{y,t}$, $\epsilon_{x,t}$ and $\epsilon_{z,t}$. Here, y_t is a process containing explosive dynamics, either driven by x_{t-i} if $\beta_x > 0$ and $\beta_z = 0$, or driven by z_t if $\beta_z > 0$ and $\beta_x = 0$ (or both x_{t-i} and z_t if $\beta_x > 0$ and $\beta_z > 0$). In (2), if $\beta_z = 0$ and $\beta_x > 0$, we can then consider that y_t and x_{t-i} are 'co-explosive', in the sense that a linear combination of these processes, $y_t - \mu_y - \beta_x x_{t-i}$, is $I(0)$. This

implies that the linear combination is stationary across all sub-regimes involved in y_t and x_{t-i}, so that 'co-explosivity' also implies cointegration in the $I(1)$ regimes of y_t and x_{t-i} as well as a stationary linear combination in the explosive regimes. The co-explosivity is contemporaneous if $i = 0$. If $i > 0$, the co-movement occurs with a lag of i time periods, with the explosive episode in x_t preceding that in y_t; in this circumstance we can therefore think of the explosive regime in x_t migrating into y_t after i time periods. If $i < 0$, then a co-explosive relationship again exists between y_t and x_{t-i}, but now the explosive episode in y_t leads the corresponding episode in x_t by i periods. If $\beta_z > 0$ and $\beta_x = 0$, then y_t contains an explosive episode driven by the unobserved process z_t; the observed process x_t is then irrelevant as a co-explosive variate for y_t. In this article, we focus on the concept of positive co-explosivity $\beta_x > 0$. However, if significant periods of negative explosivity are detected in a price series, then our approach could also be used to analyse negative co-explosivity with other relevant series.[2],[3]

In terms of testing for co-explosive behaviour between the observed series, the hypotheses of interest may be stated as

$$H_0 : \beta_x > 0, \beta_z = 0,$$
$$H_1 : \beta_x = 0, \beta_z > 0.$$

Under the null hypothesis H_0, y_t and x_t are co-explosive, with the linear combination $y_t - \mu_y - \beta_x x_{t-i}$ being $I(0)$; under the alternative hypothesis H_1, the processes are not co-explosive but 'differ' by a process that retains explosive dynamics. In the case where $\beta_x > 0$ and $\beta_z > 0$, x_t alone is not sufficient to explain the explosive behaviour observed in y_t since it depends on the additional unobserved component z_t. Such a scenario is not consistent with the null hypothesis since there is no linear combination of y_t and observed x_t that is $I(0)$, hence we implicitly treat such a case as part of the alternative of no co-explosivity.[4]

We allow for heteroskedasticity in $\epsilon_{y,t}$ via the following assumption:

Assumption 1. Let $\epsilon_{y,t} = \sigma_t v_t$ where $v_t \sim IID(0,1)$ with $E|v_t|^r < K < \infty$ for some $r \geq 4$. The volatility term σ_t satisfies $\sigma_t = \omega(t/T)$, where $\omega(\cdot)$ is non-stochastic and strictly positive.

Assumption 1 implies that the error variance is non-stochastic, bounded and displays a countable number of jumps. A detailed discussion of the class of variance processes allowed is given in Cavaliere and Taylor (2007); this includes variance processes displaying (possibly) multiple one-time volatility shifts (which need not be located at the same point in the sample as the putative regimes associated with explosive behaviour), polynomially (possibly piecewise) trending volatility and smooth transition variance

---

[2]For example, Figuerola-Ferretti, McCrorie and Paraskevopoulos (2019) find statistically significant evidence of negative mild-explosivity in crude oil prices. It is possible that such behaviour may be in a co-exploding relationship with related credit derivatives. A further investigation of this point lies outside the scope of this article, but would be an interesting topic for future research.

[3]Taking a different approach, Phillips and Yu (2011) consider a negative relationship between the autoregressive coefficients in the y_t and x_t processes in their model of bubble migration.

[4]Note that we exclude the possibility $\beta_x = \beta_z = 0$ from our analysis, since in such a case y_t would be $I(0)$ and our focus is on y_t (and x_t) series containing an explosive regime.

breaks, among others. The conventional homoskedasticity assumption, that $\sigma_t = \sigma$ for all tt, is also permitted, since here $\omega(s) = \sigma$ for all s. Assumption 1 requires that the volatility process is non-stochastic and that v_t is an *IID* sequence. These restrictions are placed in order to simplify our analysis but can be weakened, without affecting the main results of the article, to allow for cases where $\omega(\cdot)$ is stochastic and independent of v_t and where v_t is a martingale difference sequence satisfying certain moment conditions; see Cavaliere and Taylor (2009) for further details. While Assumption 1 requires the $\epsilon_{y,t}$ to be serially independent for the purposes of transparency in our theory, this assumption can be relaxed to permit serially correlated $\epsilon_{y,t}$ so long as an appropriate modification is made to the variance estimator in the test statistic we propose – see section IV.

In what follows we make use of the *variance profile* of the process:

$$\eta(s) = \left(\int_0^1 \omega(h)^2 dh \right)^{-1} \int_0^s \omega(h)^2 dh.$$

This variance profile satisfies $\eta(s) = s$ under homoskedasticity while it deviates from s in the presence of heteroskedasticity. Notice also that the quantity $\overline{\omega}^2 = \int_0^1 \omega(h)^2 dh$ is equal to the limit of $T^{-1} \sum_{t=1}^T \sigma_t^2$, and is the (asymptotic) average error variance. We will also make use of the invariance principle from Theorem 1(i) of Cavaliere and Taylor (2007), which establishes that

$$T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} \epsilon_{y,t} \xrightarrow{w} \overline{\omega} W^\eta(r),$$

where the process $W^\eta(r) = \int_0^r dW(\eta(s))$, with $W(r)$ denoting a standard Brownian motion on $[0, 1]$, is known as a *variance-transformed* Brownian motion.

IV. Testing for co-explosivity

To test H_0 against H_1 we consider the KPSS-type statistic:

$$S = \hat{\sigma}_y^{-2} (T - |i|)^{-2} \sum_{t=i\mathbf{1}(i>0)+1}^{T+i\mathbf{1}(i<0)} \left(\sum_{s=i\mathbf{1}(i>0)+1}^t \hat{e}_{y,s} \right)^2, \tag{3}$$

where $\hat{e}_{y,t} = y_t - \hat{\mu}_y - \hat{\beta}_x x_{t-i}$ is the OLS residual from a regression of y_t on a constant and x_{t-i}, and $\hat{\sigma}_y^2 = (T - |i|)^{-1} \sum_{t=i\mathbf{1}(i>0)+1}^{T+i\mathbf{1}(i<0)} \hat{e}_{y,t}^2$. Statistics of this general form were previously suggested in the context of cointegration testing where x_t is $I(1)$, with the aim to distinguish between the model errors (i.e. a linear combination of y_t and x_t) being $I(0)$ or $I(1)$ processes; see Leybourne and McCabe (1993) and Shin (1994). However, we would a priori expect this form of statistic to also be suitable in the current co-explosive testing context where x_t contains an explosive episode, with S having the ability to distinguish between the residuals $\hat{e}_{y,t}$ being $I(0)$ or containing an unobserved explosive component. At this stage we assume i is known, so that the test is being implemented with the true lag/lead.

Asymptotic behaviour of *S*

We next establish the large sample behaviour of *S*.

   Theorem 1.   For Models 1–4, under Assumption 1,

   (i) Under H_0,

$$S \xrightarrow{w} \int_0^1 V^\eta(r)^2 dr$$

      where

$$V^\eta(r) = W^\eta(r) - rW^\eta(1).$$

   (ii) Under H_1,

$$S = O_p(T^{2\alpha_{z,1}-1}).$$

   The result of Theorem 1(i) shows that the limit null distribution of *S* is dependent on the pattern of heteroskedasticity present in $\epsilon_{y,t}$. This obviously makes tabulation of limit critical values infeasible. In the homoskedastic case where $W^\eta(r) = \int_0^r dW(s) = W(r)$, the limit critical values coincide with those of Table 1 of KPSS (the demeaned case). Otherwise, these critical values will not be appropriate due to the impact of the heteroskedasticity. Theorem 1(ii) implies that, provided $\alpha_{z,1} \in (1/2, 1)$, comparison of *S* with any finite critical values will result in a consistent test under the alternative since $\lim_{T\to\infty} \Pr(S > c) = 1$ for any finite *c*. Although *S* does not diverge when $\alpha_{z,1} \in (0, 1/2]$, in the next section we find that comparison of *S* with wild bootstrap critical values can still result in a consistent test for the full range of $\alpha_{z,1} \in (0, 1)$.

   Remark 1.   Notice that the limit null distribution does not depend in any way on the process for the regressor x_t. This arises since, as is shown in the proof of Theorem 1(i), we find that the partial sum process of $\hat{e}_{y,t}$ is such that

$$T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} \hat{e}_{y,t} = T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} (\epsilon_{y,t} - \overline{\epsilon}_y) + O_p(T^{(\alpha_{x,1}-1)/2}) \qquad (4)$$

and, under our assumption of mild-explosivity in x_t, $\alpha_{x,1} < 1$, which in turn implies that the effect of x_t is asymptotically negligible as far as *S* is concerned. The negligibility of x_t is essentially due to its stochastic order of magnitude being sufficiently large. This is not the typical situation for KPSS tests based on regression residuals. For example, the Shin (1994) KPSS test is based on regression residuals with an x_t regressor that is *I*(1); in such a case, the limit distribution of the test statistic depends on the x_t regressor because an *I*(1) order of magnitude is not large enough to induce asymptotic negligibility. The asymptotic negligibility of x_t in (4) explains why we can allow $\epsilon_{x,t}$ to be a generic *I*(0) process and also be correlated with $\epsilon_{y,t}$. An obvious implication of this result is that Theorem 1(i) would continue to hold if $\epsilon_{x,t}$ was heteroskedastic.

   Note that in Assumption 1 we have assumed that the $\epsilon_{y,t}$ are not serially dependent. If we relax this assumption to permit serial correlation in the model errors, then the results

of Theorem 1 will continue to hold provided $\hat{\sigma}_y^2$ in (3) is replaced with a suitable long-run variance estimator, such as the estimator proposed by Newey and West (1994) that uses a quadratic spectral kernel with automatic lag selection.

A wild bootstrap procedure

Since we cannot rely on the homoskedastic critical values of KPSS to guarantee large sample size robustness of S in the heteroskedastic case, we propose a bootstrap procedure in order to deliver large sample size-controlled inference. Specifically, the bootstrap procedure we employ is a wild bootstrap scheme (see, e.g. Liu, 1988; Mammen, 1993) as follows:

Algorithm

1. Generate a bootstrap sample $y_{tb} = w_t \hat{e}_{y,t}$, $t = i\mathbf{1}(i > 0) + 1, \ldots, T + i\mathbf{1}(i < 0)$, where w_t denotes an *IIDN*$(0, 1)$ sequence.

2. Calculate the bootstrap KPSS test statistic S_b using residuals $\hat{e}_{tb}$ obtained from the regression of y_{tb} on a constant and x_{t-i}. That is,

$$S_b = \hat{\sigma}_b^{-2}(T - |i|)^{-2} \sum_{t=i\mathbf{1}(i>0)+1}^{T+i\mathbf{1}(i<0)} \left(\sum_{s=i\mathbf{1}(i>0)+1}^{t} \hat{e}_{sb} \right)^2,$$

where $\hat{\sigma}_b^2 = (T - |i|)^{-1} \sum_{t=i\mathbf{1}(i>0)+1}^{T+i\mathbf{1}(i<0)} \hat{e}_{tb}^2$.

3. Repeat for $b = 1, 2, \ldots, B$ bootstrap replications and calculate the (upper tail) π-level critical value, $c_{\pi B}$ say, of the empirical CDF of S_b.

4. Reject H_0 in favour of H_1 if $S > c_{\pi B}$.

We next demonstrate the large sample behaviour of S_b.

Asymptotic behaviour of S_b

Theorem 2. For Models 1–4, under Assumption 1,
(i) Under H_0,

$$S_b \xrightarrow{w}_p \int_0^1 V^\eta(r)^2 dr.$$

(ii) Under H_1,

$$S_b = O_p(T^{\alpha_{z,1}-1}).$$

Theorem 2(i) shows that under the null H_0, the limit distribution of S_b coincides with that of S. Thus, when the number of bootstrap replications B is large, the empirical CDF of S_b is such that $\Pr(S > c_{\pi B}) = \pi$, that is, the size of S is controlled asymptotically. This robustness is obtained due to the fact that the pattern of heteroskedasticity present in the original errors $\epsilon_{y,t}$ is replicated in the bootstrap data $y_{tb} = w_t \hat{e}_{y,t}$.

Remark 2. In principle, there is no need to include the regressor x_{t-i} when constructing the bootstrap residuals $\hat{e}_{tb}$ (i.e. the $\hat{e}_{tb}$ can be obtained from the regression of y_{tb} on a

constant alone). This is because the limit null distribution of S_b does not involve x_t. However, as is apparent from (4), the 'regression effect' of x_t on S may still be significant in finite samples when $\alpha_{x,1}$ is close to 1. Consequently, excluding x_{t-i} when constructing $\hat{e}_{tb}$ may result in the finite sample distributions of S_b and S being less similar than when x_{t-i} is included. On the basis of unreported simulation evidence, we indeed found this to be the case. As such, we strongly advocate inclusion of x_{t-i} in the bootstrap regressions.

It follows from Theorems 2(ii) and 1(ii) that, for $\alpha_{z,1} \in (1/2, 1)$, S diverges to $+\infty$ while S_b converges to zero; for $\alpha_{z,1} = 1/2$, S is $O_p(1)$ with S_b again converging to zero; while for $\alpha_{z,1} \in (0, 1/2)$, S converges to zero, but S_b converges to zero at a faster rate than S. Taken together, under H_1, $S/S_b = O_p(T^{\alpha_{z,1}})$, which implies that $S/c_{\pi B} = O_p(T^{\alpha_{z,1}})$, and therefore that $\lim_{T \to \infty} \Pr(S > c_{\pi B}) = 1$. Hence the bootstrap-based test is consistent under the alternative for the full range $\alpha_{z,1} \in (0, 1)$.

Remark 3. It is important to note that our testing procedure requires knowledge of x_t, but never requires us to model its specific attributes. We do not need to know which of Models 1–4 generated it (nor z_t for that matter), only that it contains some form of explosive component. Further, the procedure will be valid in the presence of *multiple* explosive episodes in x_t (or z_t) formed from sequentially conjoining DGPs of Models 1–4; for example, a model with two explosive episodes that conjoins Models 2 and 4 ($I(1)$, explosive, $I(1)$, explosive, stationary collapse, $I(1)$). Hence the procedure can be used to determine whether a linear combination of y_t and x_t is stationary when x_t contains multiple explosive episodes.

Note that in the case where $\epsilon_{y,t}$ is permitted to be serially dependent, there is no need to use a long-run variance estimator in the construction of S_b (in contrast to S), since there is no serial dependence in the wild bootstrap sample y_{tb}.

In the next section, we investigate the size and power performance of our bootstrap procedure via Monte Carlo simulation techniques.

V. Monte Carlo simulations

In this section, we examine the finite sample properties of our wild bootstrap procedure based on S for Models 1–4. We implement the bootstrap at the nominal 0.05 level (i.e. $\pi = 0.05$) and use $B = 500$ bootstrap replications. We consider a sample size of $T = 200$. In the simulations, we generate v_t, $\epsilon_{x,t}$, and $\epsilon_{z,t}$ as $IIDN(0, 1)$ variates, independent of each other. Since the regression for $\hat{e}_{y,t}$ includes a constant, we can set $\mu_x = \mu_y = 0$ without loss of generality. Our main simulations are conducted using $\alpha_{x,1} = \alpha_{x,2} = \alpha_{z,1} = \alpha_{z,2} = 0.6$, in line with the setting adopted in PSY. In this section, we set $i = 0$ in the DGP for y_t in (2) and assume correct knowledge of this value when constructing the regression residuals $\hat{e}_{y,t}$, that is, $\hat{e}_{y,t} = y_t - \hat{\mu}_y - \hat{\beta}_x x_t$. This allows us to set $\beta_x = 0$ without loss of generality since the $\hat{e}_{y,t}$ are also invariant to β_x. Our results are based on 5,000 Monte Carlo replications throughout.

Behaviour under H_0

Our first set of simulations involves quantifying the size of the bootstrap procedure under H_0, for both homoskedastic and heteroskedastic errors. The DGPs we initially consider for x_t in (1) are as follows:

$$
\begin{aligned}
\text{Model 1:} \quad & \tau_{x,1} = \{0.2, 0.4, 0.8\} \\
\text{Model 2:} \quad & \tau_{x,1} = \{0.2, 0.4, 0.6\}; \ \tau_{x,2} = 0.8 \\
\text{Model 3:} \quad & \tau_{x,1} = \{0.2, 0.4, 0.6\}; \ \tau_{x,2} = 0.9; \ c_{x,2} = 1 \\
\text{Model 4:} \quad & \tau_{x,1} = \{0.2, 0.4, 0.6\}; \ \tau_{x,2} = 0.7; \ \tau_{x,3} = 0.8; \ c_{x,2} = 1
\end{aligned}
$$

with $c_{x,1} = \{0.3, 0.7, 1.4\}$. In the homoskedastic case we set $\sigma_t = 1$ for all t in Models 1–4. For the heteroskedastic specifications we adopt two settings for each model, given by:

$$
\begin{aligned}
\text{Model 1:} \quad & \sigma_t = 5\mathbf{1}(t \leq \lfloor \tau_{x,1}T \rfloor) + \mathbf{1}(t > \lfloor \tau_{x,1}T \rfloor); \\
& \sigma_t = \mathbf{1}(t \leq \lfloor \tau_{x,1}T \rfloor) + 5\mathbf{1}(t > \lfloor \tau_{x,1}T \rfloor) \\
\text{Models 2, 3:} \quad & \sigma_t = 5\mathbf{1}(t \leq \lfloor \tau_{x,1}T \rfloor) + \mathbf{1}(\lfloor \tau_{x,1}T \rfloor < t \leq \lfloor \tau_{x,2}T \rfloor) + 2\mathbf{1}(t > \lfloor \tau_{x,2}T \rfloor); \\
& \sigma_t = 2\mathbf{1}(t \leq \lfloor \tau_{x,1}T \rfloor) + \mathbf{1}(\lfloor \tau_{x,1}T \rfloor < t \leq \lfloor \tau_{x,2}T \rfloor) + 5\mathbf{1}(t > \lfloor \tau_{x,2}T \rfloor) \\
\text{Model 4:} \quad & \sigma_t = 5\mathbf{1}(t \leq \lfloor \tau_{x,1}T \rfloor) + \mathbf{1}(\lfloor \tau_{x,1}T \rfloor < t \leq \lfloor \tau_{x,2}T \rfloor) \\
& \quad + 2\mathbf{1}(\lfloor \tau_{x,2}T \rfloor < t \leq \lfloor \tau_{x,3}T \rfloor) + 3\mathbf{1}(t > \lfloor \tau_{x,3}T \rfloor); \\
& \sigma_t = 3\mathbf{1}(t \leq \lfloor \tau_{x,1}T \rfloor) + 2\mathbf{1}(\lfloor \tau_{x,1}T \rfloor < t \leq \lfloor \tau_{x,2}T \rfloor) \\
& \quad + \mathbf{1}(\lfloor \tau_{x,2}T \rfloor < t \leq \lfloor \tau_{x,3}T \rfloor) + 5\mathbf{1}(t > \lfloor \tau_{x,3}T \rfloor).
\end{aligned}
$$

For brevity, we adopt a shorthand notation for these heteroskedasticity settings, with $\sigma_t = [\sigma_1; \sigma_2; \sigma_3; \dots]$ denoting the values of σ_t in each of the sequential regimes defined above, for example, the first Model 1 specification is denoted $\sigma_t = [5; 1]$ and the last Model 4 specification $\sigma_t = [3; 2; 1; 5]$. Our specifications introduce heteroskedasticity via volatility shifts whose number and timings are coincident with various autoregressive regimes present in the DGP x_t. This is done largely to avoid introducing further (essentially arbitrary) break timings into the simulation DGP and, while not a requirement, would seem a plausible restriction to impose. For each model, the second volatility pattern simply reverses the volatility magnitudes chosen in the first. To gauge the effectiveness of the wild bootstrap procedure in correcting size under heteroskedasticity, in addition to the bootstrap based on $y_{tb} = w_t \hat{e}_{y,t}$ (the wild bootstrap), we also compute a bootstrap based on $y_{tb} = w_t$ (an *IID* bootstrap), denoting these by W and I, respectively.

Table 1a shows the size of the bootstrap procedures under Model 1. In the benchmark homoskedastic case, $\sigma_t = 1$, both I and W have rejection frequencies close to the nominal 0.05 level throughout. In the first heteroskedastic case, $\sigma_t = [5; 1]$, I is seen to over-reject the null for the smaller values of $\tau_{x,1}$, while in the second heteroskedastic case, $\sigma_t = [1; 5]$, it under-rejects for small $\tau_{x,1}$ and over-rejects for the larger $\tau_{x,1}$. In contrast, W controls

TABLE 1

Finite sample rejection rates of bootstrap S under H_0

(a) Model 1

| | $\sigma_t = 1$ | | | | | | $\sigma_t = [5; 1]$ | | | | | | $\sigma_t = [1; 5]$ | | | | | |
| | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | |
$\tau_{x,1}$	I	W	I	W	I	W	I	W	I	W	I	W	I	W	I	W	I	W
0.2	0.049	0.047	0.047	0.048	0.050	0.052	0.123	0.051	0.157	0.053	0.162	0.053	0.037	0.049	0.027	0.047	0.027	0.049
0.4	0.049	0.048	0.049	0.049	0.050	0.051	0.080	0.050	0.085	0.055	0.089	0.059	0.051	0.051	0.040	0.049	0.038	0.051
0.8	0.051	0.053	0.048	0.050	0.048	0.050	0.046	0.049	0.045	0.048	0.050	0.049	0.093	0.062	0.080	0.059	0.079	0.058

(b) Model 2, $\tau_{x,2} = 0.8$

| | $\sigma_t = 1$ | | | | | | $\sigma_t = [5; 1; 2]$ | | | | | | $\sigma_t = [2; 1; 5]$ | | | | | |
| | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | |
$\tau_{x,1}$	I	W	I	W	I	W	I	W	I	W	I	W	I	W	I	W	I	W
0.2	0.051	0.050	0.052	0.053	0.051	0.052	0.103	0.049	0.137	0.054	0.152	0.053	0.087	0.054	0.024	0.051	0.007	0.045
0.4	0.052	0.051	0.052	0.051	0.051	0.051	0.075	0.050	0.079	0.051	0.080	0.054	0.076	0.054	0.025	0.053	0.006	0.046
0.6	0.052	0.052	0.048	0.050	0.048	0.050	0.060	0.052	0.071	0.053	0.072	0.057	0.075	0.055	0.040	0.052	0.014	0.049

(Continued)

TABLE 1

(Continued)

(c) Model 3, $\tau_{x,2} = 0.9$, $c_{x,2} = 1$

| | $\sigma_t = 1$ | | | | | | $\sigma_t = [5; 1; 2]$ | | | | | | $\sigma_t = [2; 1; 5]$ | | | | | |
| | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | |
$\tau_{x,1}$	I	W	I	W	I	W	I	W	I	W	I	W	I	W	I	W	I	W
0.2	0.051	0.050	0.052	0.051	0.049	0.052	0.107	0.050	0.148	0.054	0.162	0.054	0.195	0.062	0.082	0.053	0.030	0.047
0.4	0.054	0.050	0.050	0.049	0.049	0.052	0.072	0.052	0.078	0.053	0.086	0.056	0.161	0.060	0.061	0.053	0.024	0.046
0.6	0.054	0.055	0.047	0.047	0.049	0.049	0.058	0.051	0.067	0.054	0.066	0.059	0.133	0.060	0.069	0.054	0.025	0.050

(d) Model 4, $\tau_{x,2} = 0.7$, $\tau_{x,3} = 0.8$, $c_{x,2} = 1$

| | $\sigma_t = 1$ | | | | | | $\sigma_t = [5; 1; 2; 3]$ | | | | | | $\sigma_t = [3; 2; 1; 5]$ | | | | | |
| | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | | $c_{x,1} = 0.3$ | | $c_{x,1} = 0.7$ | | $c_{x,1} = 1.4$ | |
$\tau_{x,1}$	I	W	I	W	I	W	I	W	I	W	I	W	I	W	I	W	I	W
0.2	0.056	0.054	0.055	0.054	0.054	0.055	0.098	0.055	0.083	0.054	0.108	0.056	0.135	0.060	0.150	0.061	0.116	0.061
0.4	0.058	0.057	0.056	0.056	0.055	0.055	0.066	0.055	0.052	0.055	0.051	0.055	0.116	0.061	0.126	0.062	0.100	0.058
0.6	0.057	0.055	0.056	0.055	0.055	0.052	0.049	0.053	0.046	0.054	0.044	0.050	0.092	0.061	0.094	0.059	0.094	0.058

Notes: Entries in the table represent the finite sample sizes of nominal 0.05-level bootstrap S tests, with I and W denoting the test implemented using an *IID* bootstrap and a wild bootstrap, respectively. The sample size is $T = 200$ and Models $1-4$ are as defined in section II, with $\tau_{x,1}$, $\tau_{x,2}$ and $\tau_{x,3}$ denoting the start of the explosive regime, the end of the explosive regime, and the end of the stationary collapse regime in x_t, respectively. The explosive and stationary collapse offset parameters are given by $\delta_{x,1} = c_{x,1}T^{-0.6}$ and $\delta_{x,2} = c_{x,2}T^{-0.6}$, respectively. The volatility magnitudes are given by σ_t, with the full σ_t specification given for each model in section V.

the rejection rate well throughout. For Model 2, in Table 1b we observe a similar pattern of results regarding I and W for $\sigma_t = 1$ and $\sigma_t = [5; 1; 2]$ as seen in Model 1, although I is generally over-sized for all $\tau_{x,1}$ when $\sigma_t = [5; 1; 2]$. For $\sigma_t = [2; 1; 5]$, while W is well-behaved, under-rejection becomes even more of an issue for I, most evidently when $c_{x,1} = 1.4$. Under Model 3, Table 1c shows that I over-rejects in more cases and to a greater extent when $\sigma_t = [2; 1; 5]$. The same is true for Model 4 in Table 1d when $\sigma_t = [3; 2; 1; 5]$. Under all the heteroskedasticity patterns of Models 3 and 4, W continues to demonstrate rejection rate robustness. The results of Table 1a–d therefore highlight the efficacy of our wild bootstrap procedure in controlling size over a broad range of heteroskedastic settings. Its rejection rates lie within the range $0.045 - 0.062$ across the simulations we have considered, whereas for its *IID* counterpart, the corresponding range is $0.006 - 0.195$.

Behaviour under H_1

We now examine the ability of the bootstrap procedure to differentiate between H_0 and H_1. For brevity, we restrict attention to the homoskedastic case $\sigma_t = 1$ for all t, and only consider the wild bootstrap procedure. The DGPs for x_t we examine are a representative subset of those in the previous subsection, which use the single values $\tau_{x,1} = 0.4$ and $c_{x,1} = 0.7$, while for z_t, we generate it in a similar way to the corresponding x_t model, but allow for various values of $\tau_{z,1}$ and $c_{z,1}$. Specifically, we have:

Model 1: $\tau_{z,1} = \{0.2, 0.4, 0.8\}$; $c_{z,1} = \{0.3, 0.7, 1.4\}$

Model 2: $\tau_{x,2} = 0.8$;
$\tau_{z,1} = \{0.2, 0.4, 0.6\}$; $\tau_{z,2} = 0.8$; $c_{z,1} = \{0.3, 0.7, 1.4\}$

Model 3: $\tau_{x,2} = 0.9$; $c_{x,2} = 1$;
$\tau_{z,1} = \{0.2, 0.4, 0.6\}$; $\tau_{z,2} = 0.9$; $c_{z,2} = 1$

Model 4: $\tau_{x,2} = 0.7$; $\tau_{x,3} = 0.8$; $c_{x,2} = 1$;
$\tau_{z,1} = \{0.2, 0.4, 0.6\}$; $\tau_{z,2} = 0.7$; $\tau_{z,3} = 0.8$; $c_{z,1} = \{0.3, 0.7, 1.4\}$; $c_{z,2} = 1$

The strength of the alternative H_1 is controlled by the magnitude of $\beta_z > 0$, and we consider $\beta_z = \{0.025, 0.05, 0.1\}$.

Table 2a gives the powers under Model 1. For a given set of values of $\tau_{z,1}$ and $c_{z,1}$, we see that the rejection rate of the procedure is monotonically increasing with β_z, while for given values of $c_{z,1}$ and β_z, this rate monotonically decreases with $\tau_{z,1}$. Both of these findings are consistent with what intuition might suggest, since the larger is β_z, the farther H_1 lies from H_0; also, the smaller is $\tau_{z,1}$, the longer is the explosive regime. For a given $\tau_{z,1}$ and β_z, we find that the rejection rate, perhaps surprisingly, is not monotonic in the explosive parameter $c_{z,1}$. It is higher for $c_{z,1} = 0.3$ than for $c_{z,1} = 0.7$; however, it is then higher for $c_{z,1} = 1.4$ than for $c_{z,1} = 0.3$. Table 2b gives the results for Model 2. Other things equal, the results with respect to β_z and $\tau_{z,1}$ mirror the monotonic results of Table 2a. Interestingly, the non-monotonicity regarding $c_{z,1}$ is largely absent in Table 2b. For Model 3, we once more see monotonicity for β_z and $\tau_{z,1}$, but again some non-monotonicity for

TABLE 2

Finite sample rejection rates of wild bootstrap S under H_1

(a) Model 1, $\tau_{x,1} = 0.4$, $c_{x,1} = 0.7$

$\tau_{z,1}$	$\beta_z = 0.025$			$\beta_z = 0.05$			$\beta_z = 0.1$		
	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$
0.2	0.445	0.447	1.000	0.699	0.650	1.000	0.878	0.854	1.000
0.4	0.327	0.279	0.990	0.593	0.469	0.994	0.829	0.732	0.998
0.8	0.189	0.149	0.328	0.433	0.368	0.645	0.743	0.706	0.883

(b) Model 2, $\tau_{x,1} = 0.4$, $\tau_{x,2} = 0.8$, $c_{x,1} = 0.7$, $\tau_{z,2} = 0.8$

$\tau_{z,1}$	$\beta_z = 0.025$			$\beta_z = 0.05$			$\beta_z = 0.1$		
	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$
0.2	0.329	0.477	0.991	0.587	0.677	0.997	0.836	0.870	0.999
0.4	0.230	0.290	0.927	0.479	0.488	0.968	0.766	0.760	0.989
0.6	0.171	0.173	0.522	0.390	0.360	0.754	0.723	0.705	0.911

(c) Model 3, $\tau_{x,1} = 0.4$, $\tau_{x,2} = 0.9$, $c_{x,1} = 0.7$, $c_{x,2} = 1$, $\tau_{z,2} = 0.9$, $c_{z,2} = 1$

$\tau_{z,1}$	$\beta_z = 0.025$			$\beta_z = 0.05$			$\beta_z = 0.1$		
	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$
0.2	0.372	0.453	0.998	0.630	0.650	0.999	0.848	0.852	1.000
0.4	0.260	0.275	0.973	0.521	0.465	0.990	0.786	0.735	0.996
0.6	0.176	0.162	0.780	0.407	0.335	0.896	0.726	0.685	0.961

(d) Model 4, $\tau_{x,1} = 0.4$, $\tau_{x,2} = 0.7$, $\tau_{x,3} = 0.8$, $c_{x,1} = 0.7$, $c_{x,2} = 1$, $\tau_{z,2} = 0.7$, $\tau_{z,3} = 0.8$, $c_{z,2} = 1$

$\tau_{z,1}$	$\beta_z = 0.025$			$\beta_z = 0.05$			$\beta_z = 0.1$		
	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$	$c_{z,1} = 0.3$	$c_{z,1} = 0.7$	$c_{z,1} = 1.4$
0.2	0.254	0.506	0.969	0.531	0.729	0.987	0.838	0.897	0.995
0.4	0.190	0.307	0.786	0.447	0.535	0.903	0.777	0.806	0.968
0.6	0.163	0.167	0.215	0.408	0.406	0.468	0.753	0.749	0.784

Notes: Entries in the table represent the finite sample powers of nominal 0.05-level wild bootstrap S tests under homoskedasticity. The sample size is $T = 200$ and β_z is the coefficient on the unobserved component z_t. Models 1−4 are as defined in section II, with $\tau_{x,1}$, $\tau_{x,2}$ and $\tau_{x,3}$, and $\tau_{z,1}$, $\tau_{z,2}$ and $\tau_{z,3}$, denoting the start of the explosive regime, the end of the explosive regime and the end of the stationary collapse regime in x_t and z_t, respectively. The explosive and stationary collapse offset parameters are given by $\delta_{x,1} = c_{x,1}T^{-0.6}$ and $\delta_{x,2} = c_{x,2}T^{-0.6}$, and $\delta_{z,1} = c_{z,1}T^{-0.6}$ and $\delta_{z,2} = c_{z,2}T^{-0.6}$, respectively, for x_t and z_t.

$c_{z,1}$ is evident. In Table 2d, the rejection rates appear monotonic in β_z, $\tau_{z,1}$ and (apart from a single very minor exception) $c_{z,1}$. It is evident from the results of Table 2a–d then, that while the results for β_z and $\tau_{z,1}$ are in line with what we would a priori expect, those for $c_{z,1}$ imply that it plays a rather complex role in determining finite sample rejection rates.

While the results reported in Table 2 pertain to only a single sample size ($T = 200$) and a single setting for the parameters governing the degree of explosivity in the x_t and z_t series ($\alpha_{x,1} = \alpha_{x,2} = \alpha_{z,1} = \alpha_{z,2} = 0.6$), additional results reported in Appendix S1, using a range of sample sizes, demonstrate the consistency of the bootstrap procedure for different explosive parameter settings, in line with the consistency result derived in section IV.

VI. Identifying the timing of explosive regime migration

In practice, under the null of co-explosivity the value of the lag/lead parameter i is typically not known. Suppose then, we construct the regression residuals on which S and S_b are based from the fitted regression $\hat{e}_{y,t,j} = y_t - \hat{\mu}_y - \hat{\beta}_x x_{t-j}$ where j is a user-chosen value for i. To examine the effects of the possibility that $j \neq i$, we note that $\hat{e}_{y,t,j}$ can be decomposed as

$$\hat{e}_{y,t,j} = \beta_x r_{1,t,i,j} + r_{2,t,j},$$

where $r_{1,t,i,j}$ is a residual from a regression of $x_{t-i} - x_{t-j}$ on a constant and x_{t-j}, while $r_{2,t,j}$ is a residual from a regression of $\epsilon_{y,t}$ on a constant and x_{t-j}. Here, $r_{1,t,i,j}$ directly reflects the contribution to $\hat{e}_{y,t,j}$ arising from the incorrect choice of j since $r_{1,t,i,j} = 0$ for all t if and only if $j = i$. Taking the case where $j > i$ as an example then, during the explosive regime, it can be shown that

$$x_{t-i} - x_{t-j} \approx (j - i)\delta_{x,1} u_{x,t-i}.$$

A priori, then, we might expect the $\hat{e}_{y,t,j}$ to be rather larger in magnitude than when $j = i$, that is, when $\hat{e}_{y,t,j} = r_{2,t,j}$, due to the neglected explosive term $(j - i)\delta_{x,1} u_{x,t-i}$ appearing in $r_{1,t,i,j}$, consequently leading to S over-rejecting, and increasingly so with the magnitudes of $j - i$ and β_x.[5] Intuition would therefore suggest we might determine i by calculating $\hat{\sigma}_{y,j}^2 = (T - |j|)^{-1} \sum_{t=j\mathbf{1}(j>0)+1}^{T+j\mathbf{1}(j<0)} \hat{e}_{y,t,j}^2$ for a range of values of j, and estimating i by the value of j for which $\hat{\sigma}_{y,j}^2$ is minimized. That is, we can estimate i using $\hat{i} = \arg\min_{j \in J} \hat{\sigma}_{y,j}^2$ where J is the set of values of j considered and we implicitly assume that i is an element of J.

To gauge the effects of incorrect specification of i, together with the performance of the estimator $\hat{i}$, we conduct simulation exercises using null DGPs similar to those in section, again setting $i = 0$ (such that the co-explosivity is contemporaneous). We set $\beta_x = \{0.25, 0.5, 1\}$ but only now consider $c_{x,1} = 0.7$ in the homoskedastic case $\sigma_t = 1$. We allow j to take the values $j = \{-6, -2, -1, 0, 1, 2, 6\}$ in the fitted regression model, such that misspecification exists unless $j = 0$. Table 3a gives the results for Model 1. It is immediately evident that, for a given value of β_x, rejection rates increase with $|j|$,

---

[5]Notice that $r_{2,t,j}$ also depends on j, but in a much more benign way than $r_{1,t,i,j}$ as the regressand is simply $\epsilon_{y,t}$.

TABLE 3

Finite sample rejection rates of wild bootstrap S

(a) Model 1, $c_{x,1} = 0.7$

$\tau_{x,1}$	$\beta_x = 0.25$								$\beta_x = 0.5$								$\beta_x = 1$							
	j								j								j							
	−6	−2	−1	0	1	2	6	$\hat{i}$	−6	−2	−1	0	1	2	6	$\hat{i}$	−6	−2	−1	0	1	2	6	$\hat{i}$
0.2	0.228	0.060	0.050	0.060	0.053	0.050	0.296	0.052	0.493	0.083	0.045	0.053	0.049	0.087	0.583	0.053	0.690	0.127	0.037	0.053	0.042	0.141	0.729	0.053
0.4	0.252	0.061	0.047	0.052	0.050	0.067	0.378	0.051	0.552	0.086	0.045	0.052	0.050	0.116	0.672	0.052	0.742	0.136	0.036	0.052	0.050	0.191	0.809	0.052
0.8	0.532	0.116	0.069	0.056	0.062	0.114	0.675	0.055	0.823	0.251	0.093	0.056	0.089	0.263	0.937	0.056	0.928	0.442	0.152	0.056	0.144	0.519	0.980	0.056

(b) Model 2, $\tau_{x,2} = 0.8$, $c_{x,1} = 0.7$

$\tau_{x,1}$	$\beta_x = 0.25$								$\beta_x = 0.5$								$\beta_x = 1$							
	j								j								j							
	−6	−2	−1	0	1	2	6	$\hat{i}$	−6	−2	−1	0	1	2	6	$\hat{i}$	−6	−2	−1	0	1	2	6	$\hat{i}$
0.2	0.947	0.762	0.555	0.052	0.577	0.783	0.971	0.052	0.988	0.872	0.740	0.052	0.763	0.893	0.996	0.052	0.997	0.927	0.839	0.052	0.859	0.962	0.998	0.052
0.4	0.879	0.480	0.218	0.052	0.241	0.531	0.931	0.053	0.971	0.699	0.447	0.052	0.490	0.769	0.987	0.052	0.988	0.830	0.618	0.052	0.681	0.919	0.996	0.052
0.6	0.697	0.126	0.064	0.049	0.082	0.185	0.815	0.050	0.937	0.318	0.101	0.049	0.148	0.452	0.968	0.049	0.982	0.565	0.187	0.049	0.292	0.746	0.991	0.049

(Continued)

Bulletin

TABLE 3
(Continued)

(c) Model 3, $\tau_{x,2} = 0.9$, $c_{x,1} = 0.7$, $c_{x,2} = 1$

$\beta_x = 0.25$

$\tau_{x,1}$	j							$\hat{i}$
	-6	-2	-1	0	1	2	6	
0.2	0.979	0.895	0.780	0.050	0.777	0.885	0.982	0.050
0.4	0.950	0.741	0.503	0.050	0.502	0.742	0.954	0.050
0.6	0.873	0.385	0.129	0.051	0.144	0.377	0.876	0.051

$\beta_x = 0.5$

	j							$\hat{i}$
	-6	-2	-1	0	1	2	6	
	0.994	0.946	0.882	0.050	0.872	0.943	0.995	0.050
	0.985	0.874	0.711	0.050	0.715	0.865	0.988	0.050
	0.966	0.671	0.331	0.051	0.333	0.653	0.972	0.051

$\beta_x = 1$

	j							$\hat{i}$
	-6	-2	-1	0	1	2	6	
	0.998	0.971	0.928	0.050	0.926	0.973	0.998	0.050
	0.995	0.933	0.825	0.050	0.820	0.939	0.996	0.050
	0.992	0.824	0.555	0.051	0.535	0.820	0.991	0.051

(d) Model 4, $\tau_{x,2} = 0.7$, $\tau_{x,3} = 0.8$, $c_{x,1} = 0.7$, $c_{x,2} = 1$

$\beta_x = 0.25$

$\tau_{x,1}$	j							$\hat{i}$
	-6	-2	-1	0	1	2	6	
0.2	0.940	0.676	0.385	0.049	0.351	0.628	0.937	0.049
0.4	0.866	0.349	0.125	0.054	0.110	0.288	0.850	0.054
0.6	0.666	0.131	0.070	0.053	0.064	0.114	0.629	0.053

$\beta_x = 0.5$

	j							$\hat{i}$
	-6	-2	-1	0	1	2	6	
	0.982	0.842	0.633	0.049	0.587	0.801	0.987	0.049
	0.957	0.636	0.291	0.054	0.239	0.552	0.966	0.054
	0.905	0.308	0.102	0.053	0.093	0.259	0.908	0.053

$\beta_x = 1$

	j							$\hat{i}$
	-6	-2	-1	0	1	2	6	
	0.994	0.926	0.779	0.049	0.735	0.903	0.996	0.049
	0.984	0.818	0.506	0.054	0.421	0.769	0.990	0.054
	0.970	0.561	0.185	0.053	0.146	0.496	0.973	0.053

Notes: Entries in the table represent the finite sample sizes of nominal 0.05-level wild bootstrap S tests under homoskedasticity, for various fitted lag/lead values. The sample size is $T = 200$ and β_x is the coefficient on x_t. Models 1–4 are as defined in section II, with $\tau_{x,1}$, $\tau_{x,2}$ and $\tau_{x,3}$ denoting the start of the explosive regime, the end of the explosive regime, and the end of the stationary collapse regime in x_t, respectively. The explosive and stationary collapse offset parameters are given by $\delta_{x,1} = c_{x,1}T^{-0.6}$ and $\delta_{x,2} = c_{x,2}T^{-0.6}$, respectively. The true lead/lag i is zero; results are given for tests using different fitted lead/lag values j and for the estimated lead/lag $\hat{i}$.

that is, with the degree of estimating model misspecification (departures from $j = 0$), as expected. Rejection rates tend to be slightly higher for the positive values of j than for their negative counterparts. These rejection rates also clearly increase with β_x, again in line with expectations. For Model 2, Table 3b shows rejection rates again increasing with $|j|$. Here, however, these rejection rates are generally substantially higher than for Model 1, other things equal. This arises because the end of the explosive period (as well as the beginning) is now being mis-timed, essentially thereby amplifying the effects of misspecification. The results for Models 3 and 4 (Table 3c-d) share many more similarities with those for Model 2 than for Model 1. This is perhaps to be expected as they both also contain explosive regimes that terminate within the sample (albeit into a stationary regime rather than a random walk). At a practical level, what this implies is that our test will not be able to reliably distinguish between H_0 and H_1 being true when an incorrect choice of lag/lead parameter is employed, since both will result in the test rejecting at levels in excess of the true significance level. The results employing $\hat{\imath}$ are based on the set of values $J = \{-6, -5, \ldots, -1, 0, 1, \ldots, 6\}$. Throughout Table 3a–d, we see that the rejection rates using $\hat{\imath}$ are very close to those obtained under $j = 0$, implying that $\hat{\imath}$ is selecting the correct value of 0 with substantial regularity. These results suggest that $\hat{\imath}$ is performing very satisfactorily as an estimator of i, and on this basis, we would recommend employing it as a matter of practice to construct the residuals used in S and S_b. Of course, in any particular application, a user-specific decision will need to be made regarding the search set J and some judgement exercised.

VII. Empirical application

Data

Our dataset consists of monthly observations for the period $1993:7 - 2019:5$ on the COMEX/NYMEX prices of four precious metals: Gold, Palladium, Platinum, and Silver, and the LME prices for six non-ferrous metals: Aluminium, Copper, Lead, Nickel, Tin, and Zinc. For completeness, we compute results using both spot price data and continuous futures price data (we use the continuous price series that rolls to the nearest position contract on the first day of each month). In both cases the price data are deflated using the consumer price index (CPI) prior to testing. All of the data used were downloaded from Datastream.[6]

There are some important differences in the behaviour of precious and non-ferrous metals over the sample period we consider, reflecting their different primary roles as investment assets and manufacturing inputs. For example, driven by the increased demand from rapidly industrializing emerging markets for non-ferrous metals as manufacturing inputs (particularly Brazil, China, India, and Russia) and concerns over supply availability, non-ferrous metals prices increased very significantly in the early- and mid-2000s, more so than the precious metal prices over the same period (see Figuerola-Ferretti *et al.*, 2015,

---

[6]While several recent studies of bubbles in individual metals prices have successfully used weekly rather than monthly data (e.g. Figuerola-Ferretti *et al.*, 2015; Figuerola-Ferretti and McCrorie, 2016), we choose to use monthly data for consistency with previous studies of the relationship between metals prices that have tended to use monthly data (e.g. Escribano and Granger, 1998; Baur and Tran, 2014).

for further details on the demand and supply factors behind the risk in non-ferrous metals prices during this period). Similarly, following the 2007–08 global financial crisis and collapse in global demand and production of manufactured products, falls in non-ferrous metals prices were proportionately more extreme than for Gold and Silver. Interestingly, following the 2007–08 global financial crisis, the price of Platinum did fall by a similar proportionate amount to the non-ferrous metals, possibly reflecting its important role as a manufacturing input in the automotive industry. More generally, while all of the metals price series experienced higher volatility in the late-1980s/early-1990s and from 2005 onwards, and lower volatility in the 1990s and early 2000s, shifts in volatility are more extreme for the non-ferrous metals than for the precious metals. It should be noted that significant differences in market tightness and short-term production and consumption elasticities also exist for the different types of metals that we analyse. For example, short-term production and consumption elasticities are especially low for the case of non-ferrous metals; this can impact on the volatility of prices and even induce explosive behaviour (see Figuerola-Ferretti *et al.,* 2015, for further details).

Tests for explosive autoregression in the individual series

Prior to applying our test for co-explosive behaviour, it is important to formally test whether explosive autoregressive regimes are present in the individual series, since this is a precondition for co-explosivity (although we do not need to know the form of the explosive behaviour – see Remark 3). To do this, we use the PSY test of the unit root null hypothesis against the alternative of at least one explosive autoregressive regime. The PSY test is based on the supremum of forward and backward recursive statistics, and has good power to detect single or multiple periods of explosive autoregressive behaviour. In our empirical application, we choose the number of lagged differences in the underlying Dickey–Fuller regressions using the Bayesian information criteria (BIC), allowing for a maximum of six lags. To assess the statistical significance of the computed test statistics, we use critical values that allow for heteroskedasticity (since it is clearly a feature of our data set), obtained employing the wild bootstrap procedure proposed by Harvey *et al.* (2016) with 1999 bootstrap replications. For completeness, we also assess statistical significance using an *IID* bootstrap. The results are given in Table 4 in the form of -values.

It can be seen in Table 4 that for both the spot and futures price series, when an *IID* bootstrap is used, all of the PSY test statistics are significant at the 0.01 level, with the exception of Aluminium prices. When the wild bootstrap critical values are used, the evidence of explosive behaviour is a little weaker. However, rejections are still obtained at conventional significance for all four precious metals and for the non-ferrous metals apart from Aluminium. These results highlight that it is important to account for heteroskedasticity when testing for explosive behaviour in the prices of precious and non-ferrous metals using PSY tests, and there are fewer strong rejections when a wild bootstrap is used compared with an *IID* bootstrap. This is consistent with the results obtained by Harvey *et al.* (2016) and Figuerola-Ferretti and McCrorie (2016) in their empirical applications of the PWY/PSY tests employing wild bootstrap critical values.

TABLE 4

Application of bootstrap PSY tests for explosive behaviour to metals prices

	Spot		Futures	
	I	*W*	*I*	*W*
Precious metals				
Gold	0.000	0.035	0.000	0.028
Palladium	0.000	0.040	0.000	0.024
Platinum	0.000	0.082	0.000	0.086
Silver	0.000	0.034	0.000	0.041
Non-ferrous metals				
Aluminium	0.140	0.309	0.143	0.317
Copper	0.000	0.008	0.000	0.006
Lead	0.000	0.021	0.000	0.018
Nickel	0.000	0.030	0.000	0.023
Tin	0.000	0.062	0.000	0.072
Zinc	0.000	0.002	0.000	0.002

Note: Entries in the table represent *P*-values associated with bootstrap PSY tests, with *I* and *W* denoting the test implemented using an *IID* bootstrap and a wild bootstrap, respectively.

However, we still find that for all of the precious metals prices and the majority of the non-ferrous metals prices, the PSY tests suggest explosive autoregressive behaviour consistent with the possible presence of bubbles. Therefore it is possible that co-explosivity exists among these metals prices, which we investigate in the following subsection.

In Table 5 we report the periods of explosive behaviour detected using the date-stamping approach for bubbles proposed by PSY.[7] Note that this date-stamping approach does not account for possible heteroskedasticity, but we nevertheless consider the identified bubble periods as being illustrative of the behaviour of these series. We find identical or very similar results for most of the spot and futures prices. For Gold and Silver prices this approach detects explosive behaviour in 2007–08, leading up to the global financial crisis, and in the recovery and Eurozone crisis periods after the financial crisis. However, an earlier short period of explosive behaviour is also detected for Gold in 2003–04. For Platinum prices, this approach finds explosive behaviour in the period leading up to the global financial crisis, but not after the crisis. For Palladium prices, short explosive periods are detected in the late-1990s and early-2000s. PSY date-stamping for the non-ferrous metals finds explosive behaviour in the early and mid-2000s. Interestingly, for Copper (spot) and Lead (spot and futures), the third explosive period ends before the global financial crisis, and slightly earlier than the pre-crisis periods of explosive behaviour detected for Gold, Silver and Platinum. The explosive periods detected for Tin prices before the global financial crisis seem to lag those for Copper and Lead and are similar to the results for the precious metals. PSY date-stamping suggests only a single explosive period for Nickel prices in 2006–07. While these date-stamping results

---

[7] Note that PSY date-stamping is known for detecting periods of both positively and negatively explosive behaviour. Given the focus of our approach on positive co-explosive behaviour, in Table 5 we only report positively explosive periods detected by PSY date-stamping.

TABLE 5

PSY date-stamping of explosive periods in metals prices

Precious metals	Spot			Futures		
Gold	07/03–06/04	09/05–09/08	01/09–04/13	07/03–03/04	09/05–09/08	12/08–04/13
Palladium	03/98–07/98	12/98–05/99	07/99–05/01	01/98–09/98	12/98–05/01	12/10–01/11
Platinum	02/04–03/04	01/06–08/06	01/07–08/08	02/04–03/04	01/06–09/06	01/07–08/08
Silver	01/06–07/07	09/07–07/08	09/10–11/11	01/06–05/06	09/07–07/08	09/10–11/11
Non-ferrous metals						
Copper	02/04–03/04	11/04–03/05	06/05–12/06	06/05–11/06	-	-
Lead	12/03–03/04	11/04–04/05	10/06–04/08	12/03–02/04	11/04–12/04	10/06–04/08
Nickel	07/06–05/07	-	-	07/06–05/07	-	-
Tin	02/04–10/04	01/07–08/08	09/10–05/11	07/01–05/02	01/07–08/08	-
Zinc	07/97–08/97	12/04–03/05	05/05–02/07	07/97–08/97	12/04–03/05	05/05–12/06

Note: Entries in the table represent date ranges for positively explosive periods identified by the PSY date-stamping procedure, with dates in the form MM/YY.

are not directly comparable with the date-stamping results for precious and non-ferrous metals prices reported in Figuerola-Ferretti and McCrorie (2016) and Figuerola-Ferretti *et al.* (2015) because of differences in the type of data used, there are some common findings. For example, Figuerola-Ferretti and McCrorie (2016) also find evidence of explosive behaviour in precious metals prices in the period leading up to the financial crisis and after the crisis, and Figuerola-Ferretti *et al.* (2015) also find periods of explosive behaviour for non-ferrous metals prices in the early and mid-2000s.

Although the presence of explosive autoregressive behaviour in precious and non-ferrous metals prices might be consistent with the presence of speculative bubbles, it is also possible that in some cases the periods of explosive behaviour that we detect may be partly or entirely driven by improved fundamentals. However, it is very difficult to separate out with accuracy the impact of speculation and fundamentals on the explosive component of metals prices, particularly since, unlike equities, for precious and non-ferrous metals there is no single primary measure of fundamentals and in practice, proxies need to be used. Recent studies that have considered whether information on proxies for relevant fundamentals can explain explosive behaviour in precious metals prices have concluded that such proxies offer little support for the bubble hypothesis, other than for Gold prices in the run-up to the high point of the financial crisis and for Silver and Palladium prices linked to the launch of certain financial products (Figuerola-Ferretti *et al.,* 2015; Figuerola-Ferretti and McCrorie, 2016). Indeed, there is a growing consensus that the dramatic increase in non-ferrous metals prices in the early- and mid-2000s was primarily driven by the increased demand from rapidly industrializing emerging markets for these metals as manufacturing inputs (particularly Brazil, China, India, and Russia) and concerns over supply availability, rather than because of speculative behaviour and bubbles (see Figuerola-Ferretti *et al.,* 2015, for a detailed analysis of the demand and supply factors behind the risk in non-ferrous metals prices during this period).

An investigation of whether evidence of explosivity constitutes evidence of bubbles lies outside of the scope of this article, and we recognize that evidence from our test

supporting the null hypothesis of co-explosive behaviour could simply be evidence of co-explosive behaviour driven by co-explosive fundamentals. However, it is important to also recognize that our test is informative in empirical applications where data on fundamentals are not available because it looks for evidence *against* the existence of co-explosive behaviour (co-explosivity being the null hypothesis), and the absence of data on fundamentals does not affect the ability of the test to reject this null when it is false. Therefore, even when data on fundamentals are not available, a rejection from our test can be used to rule out co-bubble behaviour between pairs of series where explosive behaviour had already been detected (e.g. following an application of the PSY test), even if the co-explosive finding associated with a non-rejection from our test cannot uniquely be ascribed to co-bubbling.

Results from tests for co-explosive behaviour

 The results from applying our bootstrap test statistic S are reported in Table 6. The lag/lead parameters for constructing the residuals underpinning the S and S_b statistics are estimated using $\hat{\imath}$ from section VI with $J = \{-12, -11, \ldots, -1, 0, 1, \ldots, 11, 12\}$, thereby permitting a lag/lead of up to 1 year. To account for possible serial correlation in $\epsilon_{y,t}$ when computing S, we replace the short-run variance estimator $\hat{\sigma}_y^2$ with the long-run variance estimator proposed by Newey and West (1994), which employs a quadratic spectral kernel with automatic lag selection. The computed test statistics are given, along with the *p*-values computed from 5,000 bootstrap replications using the algorithm outlined in section IV. We computed results for all possible pairs of metals price series, excluding any pairs involving Aluminium (since this series was found to not contain explosive regimes) and for each we considered both possibilities in terms of which series is specified as y_t and which as x_t in the test for co-explosive behaviour. In order to focus on the most robust findings, we report results only for those pairs of metals for which $p \geq 0.025$ regardless of which series was specified as y_t and which as x_t ; for all other pairs of metals, co-explosive behaviour was rejected at the 0.025 level for at least one of the two possible orderings of y_t and x_t .

It can be seen from Table 6 that very similar results are obtained for both the spot and the futures prices and the same co-explosive pairs are identified irrespective of which type of data is used. Graphs of the pairs of spot prices for which our test finds evidence of co-explosive behaviour are given in Figure S1 of Appendix S1, along with the relevant co-explosive residual series. All the residual series show evidence of heteroskedasticity (that our wild bootstrap procedure accounts for when we compute the *p* -values in Table 6), with the volatility noticeably lower in the period up to 2005–06 than later in the sample.

For both the spot and the futures prices, our test reveals evidence of co-explosive behaviour for one or more pairs involving three of the four precious metal series considered: Gold, Silver, and Platinum. Interestingly, we do not find statistically significant evidence of co-explosive behaviour between any pairs of precious metals, but only for pairs involving non-ferrous metals, including Copper and Silver, Copper and Platinum, and Lead and Gold. Notice from Table 6 that for the Lead−Gold pair, $\hat{\imath} = -7$ indicating that explosive behaviour in the Lead price leads to explosive behaviour in the Gold price by 7 months.

TABLE 6

Application of wild bootstrap S tests for co-explosive behaviour to metals prices

		Spot		Futures	
y_t	x_t	$\hat{\iota}$	*p-value*	$\hat{\iota}$	*p-value*
Copper	Lead	0	0.391	0	0.391
Lead	Copper	0	0.043	0	0.041
Copper	Platinum	0	0.050	0	0.051
Platinum	Copper	0	0.071	0	0.066
Copper	Silver	0	0.095	0	0.100
Silver	Copper	0	0.073	0	0.081
Lead	Gold	-7	0.188	-7	0.192
Gold	Lead	7	0.076	7	0.076
Lead	Tin	-2	0.198	-2	0.202
Tin	Lead	2	0.127	2	0.127
Tin	Gold	0	0.346	0	0.369
Gold	Tin	0	0.036	0	0.036
Nickel	Zinc	3	0.041	3	0.040
Zinc	Nickel	-3	0.035	-3	0.036

Notes: Table rows provide results for a given pair of metals prices, stating which series is specified as y_t and which as x_t in the *S* test for co-explosive behaviour. The estimated lead/lag value is given by $\hat{\iota}$. The *p*-values are obtained from the wild bootstrap procedure. Results are reported for pairs of metals price series for which the *p*-value exceeds 0.025 regardless of which series was specified as y_t and which as x_t (pairs involving Aluminium were excluded as this series was found to not contain explosive regimes).

This result is broadly consistent with previous research on the timing of speculative behaviour in precious and non-ferrous metals prices. For example, Figuerola-Ferretti *et al.* (2015) find that the main episodes of explosive behaviour for non-ferrous metals are in the period 2003–07 and they appear to be driven by an increase in the demand for these metals from emerging economies, while Figuerola-Ferretti and McCrorie (2016) find that for precious metals the main periods of explosive behaviour tend to occur in the period leading up to the global financial crisis and in the period after the crisis (see also Zhao *et al.*, 2015, who find explosive behaviour in the Gold price over 2009–13). It can be seen in Table 6 that Gold and Lead are the only pair of precious and non-ferrous metals where we do find a large lag/lead between the explosive behaviour detected. For the Copper–Silver pair, and the Copper–Platinum pair, $\hat{\iota} = 0$, indicating a contemporaneous co-explosive relationship. We also find statistically significant evidence of co-explosive behaviour for three pairs of non-ferrous metals prices. A contemporaneous co-explosive relationship is found for Copper and Lead, while explosive behaviour in the Lead price is found to lead explosive behaviour in the Tin price by 2 months. Nickel and Zinc are also found to be co-explosive, with explosive behaviour in Nickel lagging that in Zinc by 3 months.

On comparing the computed lag values given in Table 6 and the start and end dates of the explosive regimes suggested by the PSY date-stamping in Table 5, we do not find an exact correspondence (perhaps not surprisingly given the different methodologies involved), but we do observe some common patterns across the two sets of results. For example, for both the spot and futures prices, while the first shorter explosive period detected by PSY date-stamping for Gold precedes the first period detected for Lead, the

positions of the second and third periods are broadly consistent with our finding of a negative (positive) lag for the Lead–Gold (Gold–Lead) pair. Similarly, the explosive periods detected by PSY date-stamping for Lead and Tin are broadly consistent with our finding of a negative (positive) lag for the Lead–Tin (Tin–Lead) pair. Although our test detects co-explosive behaviour for multiple pairs of precious and non-ferrous metals, we recognize that it does not reveal conclusive information on whether this is due to the migration of speculative behaviour in metals markets. A more detailed analysis of fundamentals and trading volumes for the relevant pairs of metals would be required to understand the extent to which our results represent evidence of speculative behaviour spreading between the different types of metals.

Our finding of co-explosive links between some of the precious metals and non-ferrous metals prices is interesting because while there has been considerable recent research on the relationship between non-ferrous metals prices (e.g. Ciner, Lucey and Yarovaya, 2020), and between precious metals prices (e.g. Kucher and McCoskey, 2017), there has been relatively little recent research on the relationship between precious and non-ferrous metals. A comprehensive empirical analysis of co-movements and trends in a large sample of precious and non-ferrous metals prices by Rossen (2015) does find evidence that certain precious and non-ferrous metals prices are strongly correlated, including Copper, Zinc, and the main precious metals. However, in this research, correlation is measured using the Pearson correlation coefficient, which is not designed for processes with shifts between regimes of $I(1)$ and explosive autoregressive behaviour. In contrast, our procedure is designed for this situation and delivers robust inference.

It is also interesting that we do not find statistically significant evidence of co-explosive behaviour between Gold and Silver prices, since both Escribano and Granger (1998) and Baur and Tran (2014) do find some evidence of cointegration between Gold and Silver spot prices using monthly data and orthodox cointegration techniques. However, our results, which are computed using a more recent sample of data that begins in the 1990s, are not inconsistent with those in Escribano and Granger (1998) and Baur and Tran (2014), as both of these studies find a weakening of the relationship between Gold and Silver prices in the 1990s. Other empirical studies using conventional cointegration tests have also found evidence that questions the existence of cointegration between Gold and Silver prices, particularly since the 1990s. For example, using daily data for the 1990s and Johansen (1991) trace and maximum eigenvalue test statistics, Ciner (2001) finds no evidence of cointegration between Gold and Silver futures prices, while using monthly spot price data for 1970:1–2015:5, Pierdzioch, Risse and Rohloff (2015) find that Gold and Silver prices are cointegrated only occasionally and that for long periods there is no evidence of a long-run relationship. Escribano and Granger (1998) and Baur and Tran (2014) also investigate the short-run relationship between Gold and Silver prices using linear and nonlinear error correction models (ECMs). Baur and Tran (2014) find that short-run causality runs from the Gold price to the Silver price but not in the opposite direction. However, their analysis does not explicitly account for the fact that the Gold and Silver price series contain regimes of explosive autoregressive behaviour. When the Gold and Silver prices are explosive autoregressive processes in levels, first-differences will also be non-stationary since an explosive series cannot be differenced to stationarity – see

Diba and Grossman (1988) for further details. Therefore, standard *t*-tests computed using ECMs as in Baur and Tran (2014) could lead to erroneous conclusions. Escribano and Granger (1998) also do not explicitly recognize the potentially explosive nature of the raw price data, which could impact on the standard errors obtained for regressions employing such series.

VIII.   Summary

In this article, we have proposed a test procedure for the detection of co-explosive behaviour involving two price series that each contain regimes of explosive autoregressive behaviour. Our test statistic is built on a variation of the stationarity test of KPSS. We establish the asymptotic properties of our test statistic and show that the null distribution is dependent on the pattern of heteroskedasticity in the model errors. To obtain appropriately size-controlled critical values for use in empirical applications, we propose a wild bootstrap procedure and find that, in addition to its asymptotic size control, the procedure has good finite sample size and power performance. The new test for co-explosive behaviour constitutes a simple, robust, and effective procedure that provides a valuable addition to the set of methods for analysing series with bubble components, allowing an analysis of co-movements between pairs of series. An empirical application to detect co-explosivity between the prices of precious and non-ferrous metals is included, which illustrates the practical value of the proposed test. We find convincing evidence from PSY tests that many of the metals prices in our sample include regimes of explosive autoregressive behaviour. Interestingly, we only uncover statistically significant evidence of co-explosive behaviour for pairs involving non-ferrous metals. While previous empirical research using orthodox statistical methods has found evidence suggesting that some precious and non-ferrous metals prices are strongly correlated, our methods are significantly more robust because they permit episodes of explosivity and heteroskedasticity.

References

Baur, D. G. and Tran, D. T. (2014). 'The long-run relationship of gold and silver and the influence of bubbles and financial crises', *Empirical Economics*, Vol. 47, pp. 1525–1541.

Cavaliere, G. and Taylor, A. M. R. (2007). 'Testing for unit roots in time series models with non-stationary volatility', *Journal of Econometrics*, Vol. 140, pp. 919–947.

Cavaliere, G. and Taylor, A. M. R. (2009). 'Heteroskedastic time series with a unit root', *Econometric Theory*, Vol. 25, pp. 1228–1270.

Ciner, C. (2001). 'On the long-run relationship between gold and silver prices: a note', *Global Finance Journal*, Vol. 12, pp. 299–303.

Ciner, C., Lucey, B. and Yarovaya, L. (2020). 'Spillovers, integration and causality in LME non-ferrous metal markets', *Journal of Commodity Markets*, Vol. 17, Article 100079.

Deng, Y., Girardin, E., Joyeux, R. and Shi, S. (2017). 'Did bubbles migrate from the stock to the housing market in China between 2005 and 2010?' *Pacific Economic Review*, Vol. 22, pp. 276–292.

Diba, B. T. and Grossman, H. I. (1988). 'Explosive rational bubbles in stock prices?' *American Economic Review*, Vol. 78, pp. 520–530.

Engsted, T. and Nielsen, B. (2012). 'Testing for rational bubbles in a coexplosive vector autoregression', *Econometrics Journal*, Vol. 15, pp. 226–254.

Escribano, A. and Granger, C. W. J. (1998). 'Investigating the relationship between gold and silver prices', *Journal of Forecasting*, Vol. 17, pp. 81–107.

Figuerola-Ferretti, I., Gilbert, C. L. and McCrorie, J. R. (2015). 'Testing for mild explosivity and bubbles in LME non-ferrous metals prices', *Journal of Time Series Analysis*, Vol. 36, pp. 763–782.

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

Figuerola-Ferretti, I., McCrorie, J. R. and Paraskevopoulos, I. (2019). 'Mild explosivity in recent crude oil prices', *Energy Economics*, Vol. 87, 104387.

Gutierrez, L. (2012). 'Speculative bubbles in agricultural commodity markets', *European Review of Agricultural Economics*, Vol. 40, pp. 217–238.

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

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

Huston, J. H. and Spencer, R. W. (2018). 'Quantitative easing and asset bubbles', *Applied Economics Letters*, Vol. 25, pp. 369–374.

Johansen, S. (1991). 'Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models', *Econometrica*, Vol. 59, pp. 1551–1580.

Kucher, O. and McCoskey, S. (2017). 'The long-run relationship between precious metal prices and the business cycle', *Quarterly Review of Economics and Finance*, Vol. 65, pp. 263–275.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992). 'Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?' *Journal of Econometrics*, Vol. 54, pp. 159–178.

Leybourne, S. J. and McCabe, B. P. M. (1993). 'A simple test for cointegration', *Oxford Bulletin of Economics and Statistics*, Vol. 12, pp. 157–166.

Liu, R. Y. (1988). 'Bootstrap procedures under some non-I.I.D. models', *Annals of Statistics*, Vol. 16, pp. 1696–1708.

Mammen, E. (1993). 'Bootstrap and wild bootstrap for high dimensional linear models', *Annals of Statistics*, Vol. 21, pp. 255–285.

Newey, W. K. and West, K. D. (1994). 'Automatic lag selection in covariance matrix estimation', *Review of Economic Studies*, Vol. 61, pp. 631–654.

Pavlidis, E., Yusupova, A., Paya, I., Peel, D., Martinez-Garcia, E., Mack, A. and Grossman, V. (2016). 'Episodes of exuberance in housing markets: in search of the smoking gun', *Journal of Real Estate Finance and Economics*, Vol. 53, pp. 419–449.

Phillips, P. C. B., Wu, Y. and Yu, J. (2011). 'Explosive behavior in the 1990s Nasdaq: when did exuberance escalate stock values?' *International Economic Review*, Vol. 52, pp. 201–226.

Phillips, P. C. B. and Yu, J. (2011). 'Dating the timeline of financial bubbles during the subprime crisis', *Quantitative Economics*, Vol. 2, pp. 455–491.

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

Pierdzioch, C., Risse, M. and Rohloff, S. (2015). 'Cointegration of the prices of gold and silver: RALS-based evidence', *Finance Research Letters*, Vol. 15, pp. 133–137.

Pindyck, R. S. (1993). 'The present value model of rational commodity pricing', *Economic Journal*, Vol. 103, pp. 511–530.

Rossen, A. (2015). 'What are metals prices like? Co-movement, price cycles and long-run trends', *Resources Policy*, Vol. 45, pp. 255–276.

Shi, S.-P., Valadkhani, A., Smyth, R. and Vahid, F. (2016). 'Dating the timeline of house price bubbles in Australian capital cities', *Economic Record*, Vol. 92, pp. 590–605.

Shin, Y. (1994). 'A residual based test for cointegration', *Econometric Theory*, Vol. 10, pp. 91–115.

van Lamoen R., Mattheussens, S. and Droes, M. (2017). *Quantitative Easing and Exuberance in Government Bond Markets: Evidence from the ECB's Expanded Asset Purchase Program*, DNB Working Paper, No. 548/March 2017.

Zhao, Y., Chang, H.-L., Su, C.-W. and Nian, R. (2015). 'Gold bubbles: when are they most likely to occur?' *Japan and the World Economy*, Vol. 34-35, pp. 17–23.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Supplementary Appendix.