

Forthcoming, Journal of Time Series Analysis

TESTING FOR MILD EXPLOSIVITY AND BUBBLES IN LME NON-FERROUS METALS PRICES

ISABEL FIGUEROLA-FERRETTI^a, CHRISTOPHER L. GILBERT^b
AND J. RODERICK McCRORIE^{c,*}

^a*ICADE, Universidad Pontificia de Comillas &
Research Centre for Energy Management, ESCP Europe*

^b*SAIS Bologna Center, Johns Hopkins University*

^c*School of Economics and Finance, University of St Andrews*

First version: 20 January 2014

This version: 30 November 2014

This paper applies the mildly explosive/multiple bubbles testing methodology developed by Phillips, Shi and Yu (2015a, *International Economic Review*, forthcoming) to examine the recent time series behaviour of the main six London Metal Exchange (LME) non-ferrous metals prices. We detect periods of mild explosivity in the cash and three-month futures price series in each of copper, nickel, lead, zinc and tin, but not in aluminium. We argue that convenience yield, though the formal counterpart to dividend yield in commodity markets, is not a useful basis on which to assess whether observed explosivity is indicative of bubbles (namely, departures of prices from their fundamental values). We construct other measures that provide evidence that suggests the observed explosivity in the non-ferrous metals market can be associated with tight physical markets.

Keywords: mildly explosive process; recursive regression; generalized sup ADF test; economic bubbles; commodity prices; non-ferrous metals.

JEL. C22.

*Correspondence to: Roderick McCrorie, School of Economics and Finance, University of St Andrews, St Andrews, KY16 9AR, UK. E-mail: mccrorie@st-andrews.ac.uk

1. INTRODUCTION

This paper tests for types of non-stationary behaviour in London Metal Exchange (LME) metals prices with a view towards characterizing their time series properties. Through a baseline test of a null of unit-root non-stationarity, it relates to a subject area of a significant part of John Nankervis's work. Perhaps John's best known contribution in this area was through the *Econometrica* paper by DeJong, Nankervis, Savin and Whiteman (1992), which sits alongside a significant amount of other work published in many of the principal econometrics and statistics journals. Here, our concern is directed towards mildly explosive alternatives, which constitutes a similar approach but with interest centred on an autoregressive parameter possibly taking a value on the other (right) side of unity. As such, the paper has connections to work being undertaken by John's former colleagues in the Finance Group at the Essex Business School and, in particular, to the paper by Coakley, Kellard and Tsvetanov (2013) which uses the same test as here in an investigation of WTI crude oil prices. This work uses the multiple bubbles dating technology recently proposed by Phillips, Shi and Yu (2015a, 2015b, PSYa,b) which, as discussed below, provides a framework in which periods of mild explosivity can be related to economic bubbles.

In a companion to this paper, Figuerola-Ferretti, Gilbert and McCrorie (2015) have used a variant of the bubble-detection algorithm proposed by Phillips, Wu and Yu (2011, PWY) to test for bubble behaviour in commodity futures price series in the energy and agricultural sectors. Following through on preliminary work for UNCTAD by Gilbert (2010), this work was primarily focussed on testing for single bubbles around the time of the recent 2007-08 financial crisis. Homm and Breitung (2012) established in simulation work that when the data contain a single bubble the PWY (2011) procedure is particularly effective as a mildly explosive/bubble-detection algorithm in comparison with other, potentially rival approaches.

Within the metals sector, non-ferrous metals form a distinct class through the common demand factors they face (see, e.g., Scherer and He, 2008) and the inelastic supply that can emerge from rigid capacity constraints faced by producers which at times leave them unable to respond to increases in demand. From an economic point of view, inelastic supply combined with fluctuating demand constitute a *prima facie* reason why we might expect the non-ferrous metals market to be prone to volatile and perhaps bubble-like behaviour. Using the original PWY test applied to representative

commodities in the energy, agricultural, non-ferrous and precious metals sectors, Gilbert (2010) reported evidence for multiple bubbles in copper. While Homm and Breitung (2012) showed the original PWY test has some efficacy in the presence of two bubbles, PSY (2015b) showed that their recently proposed single and multiple bubbles test dominates both the original and an improved sequential version of the PWY test. We therefore see the PSY test as the more appropriate for analyzing non-ferrous metals prices.

The purpose of this paper is to use the PSY bubbles-detection algorithm to examine the official cash prices and three-month futures prices of the six major LME non-ferrous metals. Our first aim in applying the PSY test is to gain evidence to characterize the differences in mildly explosive/bubble behavior within the non-ferrous metals class. The recent literature on commodity market bubbles has focused on energy and food prices, both of which are important to consumers. Non-ferrous metals have received relatively little attention, perhaps because metals prices affect consumer prices only indirectly through the prices of manufactured goods. However, non-ferrous metals markets benefit from the availability of high quality data and, as Gilbert (2010) showed, there is evidence that non-ferrous metals prices may have been subject to multiple periods of explosive prices during the initial decade of this century. They therefore provide a good test bed for econometric procedures which set out to detect multiple bubbles.

Our second aim is to assess whether the observed mildly explosive periods can be explained by the behaviour of supply and demand fundamentals that economic theory suggests are the main price drivers. This links to the commodities literature where an important recent issue has been whether price movements are based on fundamentals, or whether the key role has instead been played by increased speculation resulting from financialization, which is the process through which large numbers of financial actors, specifically investment banks, hedge funds and index investors, have become involved in commodity futures markets. Through the use of proxy variables to represent fundamentals, the PSY algorithm, when set in the context of a standard asset pricing equation, offers the basis of testing whether observed periods of mild explosivity represent departures from a commodity's fundamental value or simply reflect changes in the fundamentals themselves. As noted by PSY (2015a) and others, this type of question has underpinned governmental and regulatory discussions on constructing appropriate policy responses to the recent global financial crisis.

Section 2 introduces the PWY-PSY bubbles-testing methodology, outlining the tests that follow and giving an overview of the literature. Section 3 introduces the main LME non-ferrous metals (aluminium, copper, nickel, lead, zinc and tin) and describes their raw time series properties over the chosen sample period. In Section 4, the PSY multiple bubbles test is applied to the cash price and three-month futures price series of each metal. Some interpretation of our results is offered in Section 5. Section 6 concludes.

2. THE PSY BUBBLE-TESTING METHODOLOGY

PWY (2011), Phillips and Yu (2011, PY) and PSY (2015a, 2015b) have recently developed a statistical testing methodology based upon a reduced-form autoregressive (AR) model and an invariance principle for mildly explosive processes established under various conditions by Phillips and Magdalinos (2007a, 2007b, PM). The tests look for evidence of temporary regime shifts of mild explosivity that are embedded in data evolving as a stochastic trend. Chosen proxy variables thought to represent economic fundamentals allow the results, when set in the context of a rational asset pricing model, to be assessed in terms of whether or not periods of mild explosivity are consistent with departures from each metal's fundamental value. In addition to PSY's application to the S&P500 index and Coakley et al's (2013) to crude oil prices, the PSY procedure has already been applied to the Hong Kong residential property market (Yiu, Yu and Lu, 2013), food commodity markets (Etienne, Irwin and Garcia, 2014, 2015) and precious metals prices (Figuerola-Ferretti and McCrorie, 2015).

The PSY test is prototypically based on a null specification involving a random walk process and a local-to-zero intercept term:

$$x_t = dT^{-\eta} + \theta x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}(0, \sigma^2), \quad \theta = 1, \quad t = 1, \dots, T; \quad (1)$$

where $x_0 = O_p(1)$, d and σ ($0 < \sigma < \infty$) are unknown parameters and η is a localizing coefficient that controls the magnitude of the drift as $T \rightarrow \infty$ which, in principle, could be estimated on the basis of data (see PSY, 2015b). When $\eta > \frac{1}{2}$, the drift is small compared with the random walk component and, under the given conditions, the partial sums of ε_t satisfy the functional central limit theorem

$$T^{-1/2} \sum_{t=1}^{\lfloor T \cdot \cdot \rfloor} \varepsilon_t \Rightarrow \sigma W(\cdot), \quad (2)$$

where $\lfloor \cdot \rfloor$ is the floor function (giving the integer part of the argument) and W is standard Brownian motion. The prototypical null set-up can be generalized in various directions, e.g. to allow for martingale difference sequence errors or some weak dependence using the results of PM (2007b).

PWY (2011) consider a single-bubble data generation process (DGP) under the alternative hypothesis but, following the approach of PSY (2015a, b) for the reasons discussed above, we are interested in single and multiple bubbles alternatives. The DGP under the alternative is taken to exhibit K bubble episodes in the sample period, represented in terms of sample fraction intervals $B_i = [\tau_{i,e}, \tau_{i,f}]$ ($i = 1, 2, \dots, K$), within periods of prevailing martingale-type behavior in the intervals $N_0 = [1, \tau_{1,e}]$, $N_j = [\tau_{j-1,f}, \tau_{j,e}]$ ($j = 1, 2, \dots, K-1$) and $N_K = [\tau_{K,f}, T]$, as follows:

$$x_t = (x_{t-1} + \varepsilon_t) 1(t \in N_0) + (\delta_T x_{t-1} + \varepsilon_t) 1(t \in B_i) \\ + \sum_{i=1}^K \left(\sum_{l=\tau+1}^t \varepsilon_l + x_{\tau_{i,f}}^* \right) 1(t \in N_i) \quad (t = 1, \dots, T) \quad (3)$$

$$\delta_T = 1 + c/T^\alpha, \quad c > 0, \quad \alpha \in (0, 1). \quad (4)$$

Under the conditions on c and α , the autoregressive parameter δ_T is greater than unity and the model has what PM (2007) called a mildly-integrated root (here, on the explosive side of unity). In (3), bubble implosion is modelled for each i by $x_{\tau_{i,f}}^* = x_{\tau_{i,e}} + x_i^*$, where $x_i^* = O_p(1)$, reflecting an aspect of the PWY/PSY approach whereby the process is assumed to collapse abruptly to the value of the last pre-bubble observation plus an $O_p(1)$ perturbation, representing a re-intialization of the process from which it resumes its trend. In a recent working paper, Phillips and Shi (2014) have discussed alternative and possibly more realistic bubble collapse scenarios in which the collapse can be “sudden”, “disturbing” or “smooth” in line with recent literature (e.g., Huang, Zheng and Chia, 2010). Harvey, Leybourne, Sollis and Taylor (2014) have also modelled bubble implosion that is less abrupt than in PSY.

The regression model usually involves transient dynamics, as in conventional (left-sided) augmented Dickey-Fuller (ADF) tests. The PSY approach implements a recursive right-tailed unit root test procedure based on an ADF-type regression model using a flexible window. Starting from a fraction r_1 and ending at a fraction r_2 of the total sample, with the window size $r_w = r_2 - r_1$, we begin by fitting a regression model of general type

$$\Delta x_t = \hat{\alpha}_{r_1, r_2} + \hat{\beta}_{r_1, r_2} x_{t-1} + \sum_{i=1}^k \hat{\psi}_{r_1, r_2}^i \Delta x_{t-i} + \varepsilon_t, \quad (5)$$

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

The earlier PWY (2011) methodology involved forming a recursive sequence of right-tailed ADF-type tests based on a *forward-expanding* sample and used the supremum of these as the basis of a test. While Homm and Breitung (2012) showed the test had some efficacy as a bubble-detection algorithm for one and two bubble alternatives, PSY (2015b) show that when the sample period contains two bubbles, the PWY procedure (when not applied sequentially) can fail to identify or consistently date-stamp the second bubble. This motivated PSY to formulate a test that covers more subsamples of the data and has more flexibility to choose a subsample that contains a bubble episode. Specifically, they formulate a *backward sup* ADF test, where the endpoint of the sample is fixed at a fraction r_2 of the total sample and the window size is allowed to expand from an initial fraction r_0 of the total sample to r_2 . The backward sup ADF statistic is defined as

$$BSADF_{r_2}(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \{ADF_{r_1}^{r_2}\}. \quad (6)$$

A test is then based on the *generalized sup* ADF (GSADF) statistic which is constructed through repeated implementation of the BSADF test procedure for each $r_2 \in [r_0, 1]$. The GSADF test statistic is defined as the supremum of the BSADF test statistics:

$$GSADF(r_0) = \sup_{r_2 \in [r_0, 1]} \{BSADF_{r_2}(r_0)\}, \quad (7)$$

PSY (2015a, Theorem 1) gives the limiting distribution of (7) under (1) with asymptotically negligible drift ($\eta > \frac{1}{2}$). Critical values obtained by numerical simulation are set against a mildly explosive alternative (for specific values of r_0). What allows the PSY procedure to be used as a bubble-detection algorithm is the non-trivial power that PSY (2015b) demonstrate the test statistics have under the single and multiple bubble alternatives (3) with various (fixed) values of K .

Date-stamping is achieved through the BSADF statistic: the origination and termination points of the first bubble, $r_{1,e}$ and $r_{1,f}$, are estimated, subject to a minimum bubble-duration condition that controls for Type I error, by

$$\hat{r}_{1,e} = \inf_{r_2 \in [r_0, 1]} \{r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T}\}, \quad (8)$$

$$\hat{r}_{1,f} = \inf_{r_2 \in [\hat{r}_e + \delta \log(t)/T, 1]} \{r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T}\}, \quad (9)$$

where $scv_{r_2}^{\beta_T}$ is the $100(1 - \beta_T)\%$ right-sided critical value of the BSADF statistic based on $\lfloor T r_2 \rfloor$ observations and δ is a tuning parameter that can be chosen, in principle, on the basis of sampling frequency. In PWY (2011), PY (2011) and PSY (2015a,b) and in all applied work we know prior to this paper with the exception of Coakley et al. (2013), the tuning parameter has been set to unity, implying a minimum bubble-duration condition of $\log(T)$ observations (corresponding to a sample fraction of $\log(T)/T$). This means that for a bubble to be declared, the BSADF statistic must have been above its critical value for at least $\lfloor T \hat{r}_{1,e} \rfloor + \lfloor \log(T) \rfloor$ observations. Conditional on a first bubble having been found and estimated to have terminated at $\hat{r}_{1,f}$, the procedure is repeated in search of a second and possibly more bubbles. PSY (2015b) show that, subject to rate conditions, this procedure provides consistent estimates of the origination and termination dates of one, two and three (and, in principle, more) bubbles.

Although we use data that is based on only one sampling frequency, we shall, on the advice of a referee, report results with different choices of the tuning parameter. PSY's default choice of unity for the tuning parameter arose in contexts of empirical applications involving monthly data. In commodity and other financial market applications, it is more natural to use daily or weekly data. Application of a criterion which is appropriate in a macroeconomic context to higher frequency data may result in the detection of only short-lived bubbles that have little policy interest – see, e.g., Etienne et al (2014, 2015) who, analysing agricultural futures daily data, reported only very short bubbles. Reporting results with alternative choices of the tuning parameter should confer some robustness on our approach.

PSY (2015a, Section 2) explain how the test procedure can be interpreted as a test for (rational) bubbles when the DGP is the standard rational asset pricing equation¹

$$P_t = \sum_{i=0}^{\infty} \left(\frac{1}{1+r_f} \right)^i E_t(D_{t+i} + U_{t+i}) + B_t, \quad (10)$$

where P_t is the (present-value) price of an asset, D_t is the payoff received from the asset, U_t represents unobservable fundamentals, and r_f is the (positive) risk-free interest rate. The quantity $(P_t - B_t)$ is called the market fundamental, with the bubble component B_t assumed to satisfy the property

$$E_t(B_{t-1}) = (1 + r_f)B_t, \quad (11)$$

which is explosive given $r_f > 0$.

The PSY procedure is a reduced form approach that does not identify the source of any mild explosivity observed in P_t . As PSY (2015a) note, when $B_t = 0$, the degree of non-stationarity in P_t in (10) is controlled by the nature of D_t and U_t , and if D_t is an I(1) process and U_t is either I(1) or I(0), P_t would be at most I(1). Under (10), the observation of mild explosivity in P_t *under such conditions on D_t and U_t* would offer evidence of bubble behaviour through a departure from fundamental value, i.e. through a non-zero bubble component. PY (2011) show that mildly explosive behavior can also arise from a time-varying rather than a constant r_f . In the case of equities, the payoff is the dividend.

The formal counterpart of the dividend yield in storable commodities markets is the convenience yield, which is the implied value of any benefits (net of insurance, deterioration and storage costs) that accrue from holding inventories of the commodity.² More precisely, convenience yield is defined as the percentage premium of the (current) cash price over a deferred (future) price less the interest rate, storage cost and the rate of deterioration, and may be interpreted as the premium stockholders will pay for immediate access to inventory of known specification and location. Miltersen and Schwartz (1998, p. 34) note that an equilibrium description of prices, inventories and convenience yield makes convenience yield endogenous. Because the interest rate and storage cost components of convenience yield typically vary relatively slowly, convenience yield is essentially the negative of the slope of the futures term structure over the relevant period (which here, given our futures market data, is three months). The cost of carry relationship linking cash and deferred prices is based on arbitrage considerations. There is no reason to suppose that a departure of price from fundamental value would upset that relationship. If that were the case, we

would expect periods of explosive prices to be common to both cash and futures prices but absent from convenience yield. If, instead, the explosion arises from the convenience yield process itself, perhaps because of a pending limitation to the supply of storage, this should be transmitted to either or both the cash and futures prices.

This contrasts with the use of dividend yield in the context of equity markets, as considered by PWY (2011), which is determined in the real economy. The dividend yield is only impacted by changes in equity prices with a significant time lag since firms announce dividends months in advance of payment and dividend policies change relatively slowly. In what follows, we have reported and interpreted results using convenience yields for the non-ferrous metals group but we have also provided the same analysis using another fundamental proxy, the stock-to-use ratio, which in her text Geman (2005, p. 144) describes (in the context of agricultural commodities) as “a key number in technical analysis rules of trading commodity markets and options pricing as well”. Any excess of consumption over production implies a rundown of stocks (inventories).

The PSY procedure dates mildly explosive periods within the sample period which we can declare as “bubbles” when such periods represent departures from “fundamental value”, as evidenced by the behaviour of the chosen fundamental proxies. One final question in implementing the approach is whether to use real or nominal prices and there are different conventions in the Economics and Finance literatures. On the recommendation of a referee, we chose the former and report results for real series. We deflate the nominal prices using the U.S. Producer Price Index (PPI), giving the price of each metal in terms of a basket of all goods at the wholesale/producer stage, with U.S. usage weights being taken because the metals prices are denominated in U.S. dollars. The results reported in Section 4 change little if nominal prices are used.

3. NON-FERROUS METALS PRICES SINCE 2000

Metals fall into four broad groups: precious metals, ferrous metals, non-ferrous metals and minor metals. In this paper we look at the six major non-ferrous metals (aluminium, copper, nickel, lead, zinc and tin), all of which are traded on the London Metal Exchange (LME). We do not consider ferrous metals (iron ore and steel). Until around 2010, iron ore and steel prices were based on annual contracts negotiated

between producers and consumers. The LME steel billet contract was only launched in 2008 and trades very low volumes while futures trading in iron ore began only in 2010. Precious metals, which are closely linked to monetary sector assets, form a separate, asset class that has been considered by Figuerola-Ferretti and McCrorie (2015). Prices for minor metals are quoted by a number of price reporting services but these markets are thin and, when supply is abundant, prices change at infrequent intervals.

Aluminium and copper are the two most important non-ferrous metals by value, both being extensively used across the entire range of industry and construction, in particular in electrical applications. Nickel is the third most important, its main use being as an input to the production of stainless and special steels. Lead and zinc, which have lower value-to-weight ratios, exhibit more complicated price behaviour arising in part from their being joint products. Many lead-zinc mines produce the two metals in proportions determined by geological factors with the consequence that either one or the other is often in excess supply. Tin is the least important member of the group by value. The tin industry was in decline for much of the period since 2000 owing to the virtual disappearance of tin-plating, formally its principal end use. It has, however, recently found new application in the microelectronics sector, resulting in a resurgence of demand.

In this paper, we analyze weekly official data for cash (current) and three-month (futures) LME aluminium, copper, lead, nickel, tin and zinc prices from January 2000 to December 2013 (731 observations before lag creation).⁴ The upper panel of Table 1 reports descriptive statistics for both cash and three-month prices, and also for convenience yields discussed in Section 5 below where convenience yields are defined by equation (12). Cash prices are uniformly slightly more volatile than the three-month prices and application of the Jarque-Bera test strongly suggests all the price distributions are non-normal. The lower panel of Table 1 reports descriptive statistics for two copper-specific speculative variables which we also discuss in Section 5.

[Table 1 around here]

The LME has been the dominant world non-ferrous metals futures market throughout the period under consideration. There are two other important metals futures markets: COMEX, a subdivision of the New York Mercantile Exchange (NYMEX), and the Shanghai Futures Exchange (SHFE). COMEX trades aluminium

and copper contracts. Its copper prices are closely arbitrated with LME prices but there is very little volume in its aluminium contract. The SHFE trades aluminium, copper, lead and zinc, for which it sets reference prices for intra-Chinese commerce. However, the financial and administrative costs associated with importing metal into China, particularly those associated with access to hard currency, can result in substantial differences between world prices and internal Chinese prices.

The sample encompasses the end of the period of low metals demand and the emergence of a run-up in metals prices that culminated around the 2008 financial crisis. It covers a number of different demand and supply regimes which, given our later emphasis on fundamentals, we shall describe in this section. While we associate the raw price movements with supply and demand changes that economic theory predicts explains them, we are careful not to assert causality.

Figure 1 charts the three-month futures price series for the period 2000-13. Non-ferrous metals prices had been subdued in the 1990's, a period associated with low sectoral profitability and low levels of investment in new mine capacity. Prices continued to be weak in the initial years of the new decade and in a number of cases fell even farther over 2001-02. The closing months of 2003 saw the start of a period of renewed GDP growth in the OECD in conjunction with rapid industrial growth in Asia, particularly China. An approximate measure of global industrial production can be formed by weighting the IMF "Advanced Economies" industrial production index with those for the four BRICs (Brazil, China, India and the Russian Federation).⁵ On this measure, global industrial production grew at an average rate of 5.1% over the five-year period 2003-07 as against 2.4% over the previous five years (1998-2002). Standard economic theory predicts that such increased production will lead to an increase in demand for metals.

[Figure 1 about here]

Metals intensities, which measure the metal content of industrial production, are generally seen in the literature as being determined by engineering considerations and are little affected by prices over the short to medium term (see, e.g., Radetzki and Tilton, 1990). Metals demand is therefore inelastic, meaning that the demand response to a change in price is less than proportionate. A new mine will take ten years to construct and major extension of an existing mine will take around five years. Supply is therefore also inelastic over the short to medium term. The combination of inelastic supply with the increase in demand over the period 2003-07 should, according to

standard economic theory, push metals prices higher. Copper and nickel prices moved higher from the end of 2003. Lead showed a similar pattern. Zinc prices were also stronger after 2003, but most notably so in those periods in which lead prices were relatively low.

In sharp contrast with the prices of other five non-ferrous metals, aluminium prices languished at relatively low levels throughout the 2004-08 price boom. The demand side drivers for aluminium are the same as those for copper and nickel and indeed some correlation is observed between aluminium price movements and those in copper and nickel. The difference in the price patterns can be related to the supply side. The last decade saw a very substantial expansion of Chinese aluminium smelting capacity with the consequence that, over the sample period, Chinese trade in aluminium remained in approximate supply and demand balance, recording an average deficit of 0.9% over 1998-2002, an average surplus of 1.7% over 2003-07 and almost exact balance over 2008-11.⁶ This contrasts with copper, where Chinese consumption of refined copper exceeded production by an average of 3.9% of world refined consumption over 1998-2002, rising to 6.6% in 2003-07 and 15.0% in 2008-11. Similarly, Chinese nickel consumption exceeded domestic production by 1.0% in 1998-2002, 6.8% for 2003-07 and 14.8% for 2008-11.

Taking the stock-to-use ratio measure of the market fundamental, over the LME warehouse stocks of copper averaged 2.1 weeks consumption over the five years 1998-2002 but this fell to an average of 0.8 weeks consumption over 2003-07. Conversely, in aluminium, the stock-consumption ratio rose from 1.5 weeks to 1.8 weeks of 2002 consumption over the same period.

Non-ferrous metals prices were already substantially lower than their 2006-07 peaks by the time of the 2008 Lehman Brothers bankruptcy, widely seen as the pivotal point of the recent financial crisis (see PY, 2011). Prices recovered from mid-2009. From 2010 to the end of the sample, the non-ferrous group effectively split into two with copper, lead and tin prices remaining strong while aluminium, nickel and zinc suffered from excess production and weak prices.

Cash and three-month prices are linked by a cost of carry relationship and so typically move closely together. According to this relationship, the three-month price will exceed the cash price by the warehousing and interest cost less the convenience yield which reflects the option value of immediate access to the metal. When stocks are plentiful, convenience yield is near zero and the market is said to be at “full

carry". Since warehousing costs and interest rates vary only slowly, the two prices generally move in step. However, when stocks are limited, convenience yield can be both high and variable resulting in a backwardation market, in which the cash price is at a premium to the three-month price. A temporary shortage will have greater impact on the cash price than on the three-month price making. This is reflected in higher cash price volatilities – see Table 1, Column 1. In our empirical work, we therefore consider both cash and three-month prices.

To maintain compatibility with the rational bubbles model, we follow PWY and PSY in analyzing price changes and not price returns.⁷ Prices are measured in U.S. dollars per ton and are deflated by the U.S. PPI (all items) by interpolation of published monthly data onto a weekly basis. We calculate the BSADF statistics with up to five lags (the maximum lag length reported in the Table 1 tests) of the dependent variable. The lag length is chosen in each recursive sub-sample to minimize the BIC.

The PSY procedure requires the choice of a smallest sample width fraction r_0 to initialize the computation of the GSADF and BSADF test statistics, whose statistical properties then become a function of this choice. We follow PSY (2015a) and take $r_0 = 0.01 + 1.8/\sqrt{T}$, calibrating critical value generation and bubble identification procedures conditional upon this value. In our sample, $T = 729$, implying $\lfloor r_0 T \rfloor = 56$. When we come to bubble identification, we also follow PSY in setting a minimum bubble-duration length of $\delta \log(T)$, where δ is the tuning parameter discussed above. There is a danger that choosing too short a criterion may generate bubble-type phenomena ("froth") which, while of concern to market participants, will be of little policy interest. The standard case where the tuning parameter is unity implies a minimum bubble-duration length in terms of weeks of $\ln(729) \approx 7$. On the recommendation of a referee, we have explored other minimum bubble lengths, and have specifically chosen $\delta = 2$ (corresponding to 13 weeks, or a quarter) and $\delta = 4$ (corresponding to 26 weeks, or half a year). Critical values are based on 5,000 replications (the same number as in PSY) over the subsample running from $\lfloor r_0 T \rfloor$ to the full span of the data T .

The imposition of a minimum bubble-duration condition is not innocuous with respect to test size. In PSY, critical values are generated without imposing the

condition. This means that estimates that date-stamp the origination and collapse of bubbles, which do incorporate the minimum bubble-duration condition, are constructed on a basis that is different from the test statistic used to generate critical values. PWY (2015a) offered some sampling evidence with various values of r_0 and T , and the rule they subsequently proposed for choosing r_0 as a function of the sample size T can be interpreted as a device to control for size distortion. Nevertheless, the rule is *ad hoc* and in our case, where we use various values of the tuning parameter, we might expect some size distortion that is not reflected in PSY's experiments; for an increase in δ results in fewer BSADF statistic excesses qualifying as bubbles. In response to a referee, we have reported only pure PSY estimates and have therefore not attempted to correct for size distortion that is specific to our problem. We leave the general issue of size distortion in the PSY methodology to future work.

4. TEST RESULTS

The GSADF statistic tests in the direction of at least one episode of mildly explosive behaviour in the sample. Table 2 compares the twelve GSADF statistics (two contracts for each of the six metals) with critical values generated for the given sample size and rule-based value of r_0 . The test statistics reject the hypothesis of no explosive periods for all the metals with the exception of aluminium, irrespective of the significance level. We conclude that there is no evidence for mildly explosive behaviour in LME aluminium prices but corroboratory evidence for periods of mild explosivity in the prices of the other five metals.

[Table 2 around here]

The strength of the PSY procedure relative to earlier procedures for testing for explosivity is its robustness to the presence of multiple bubbles. Table 3 lists the number of periods of explosive prices for the cash and three-month prices. In the first block of numbers, which takes all excesses of the backward SADF statistic over its critical value into account, a large number of bubbles is recorded. Lower test sizes are associated with higher critical values, and so in general fewer bubble periods are found. However, the number of separate bubble periods identified can rise as test size is decreased if a short-lived dip in the backward SADF sequence causes a long bubble period to be split into two shorter periods. This happens for the copper cash price.

[Table 3 around here]

As discussed above, PWY and PSY impose a minimum length condition which any backward SADF excess must meet. Holding the critical values constant, an increase in the minimum bubble length reduces the number of identified mildly explosive periods. We take minimum bubble lengths of $\delta \log(T)$ for values of $\delta = 1, 2$ and 4 , giving 7, 13 and 26 weeks respectively. In Table 3, we also report (for interest) the number of mildly explosive periods the PSY procedure detects when no minimum length condition is imposed. This corresponds to the case where critical value generation and the detection of mildly explosive periods takes place on the same basis. Imposition of the minimum length condition eliminates all the “froth” identified in the first block of the table and, except in the cases of tin (three or four bubbles) and aluminium (no bubbles), results in detection of either one or two detected bubble periods for each metal.

Table 3 also notes the number of periods in which the backward SADF excess is associated with a decline in prices. Imposition of the minimum length criterion eliminates most of these “mildly imploding” prices. In PSY (2015a,b), test power is computed under alternative hypotheses of the form of (3), where bubbles collapse instantaneously. As noted in Section 2, there is ongoing work that examines different models of bubble implosion and it is possible the PSY procedure might have some efficacy in such contexts. Following Yiu, Yu and Lu (2013), who observed this same phenomenon in their application of the PSY testing methodology to the Hong Kong residential property market, we also report the small number of detected instances of the same.

[Table 4 around here]

We now turn to bubble identification (or “date stamping”) and focus on the estimates which impose a minimum bubble length. Table 4 summarizes the periods for which mildly explosive price developments are identified using the 95% critical value for the backward SADF statistic for both the deflated cash and three-month prices. The table excludes “mildly implosive” episodes as discussed above.

Figures 2-7 plot the BSADF test statistic sequences and the associated 5% critical value sequence for the three-month prices. The figures also show (in faint) the time path of the three-month price series. During periods of rapidly rising prices, the BSADF statistic and the price tend to rise together but this co-movement stops if the

price falls back even if only temporarily. (Charts for the corresponding cash prices are similar but are omitted to conserve space.)

[Figures 2-7 around here]

The majority of the periods of mildly explosive prices are common to cash and futures prices although in many instances, the start and end dates differ by a few weeks. This is important in relation to the discussion of convenience yield in Section 2 and we revert to this issue in Section 5 below. Which of the two markets moves first will depend on the origin of the movement. Speculators will trade futures prices and so a speculatively-induced bubble is likely to be seen first in futures prices, as in the lead market in 2007. Physical traders will buy or sell cash so a fundamentals-based movement may show up first in the cash market, as in copper in 2003.

There is considerable commonality in the estimated mildly explosive periods across the five metals. Two major episodes stand out. The final months of 2003 saw explosive growth emerging almost simultaneously in four markets – copper, nickel, lead and tin. In all cases, the upward price explosion faltered in the early months of 2004. The second major common episode initiated in the final months of 2005 (copper and zinc) or mid-2006 (nickel). In this case, explosivity persisted for longer. The pattern is seen most clearly in the estimates which employ the 13 week minimum bubble length. It is somewhat more confused in the estimates using a 7 week minimum bubble length where the higher critical values force breaks in the periods seen as continuous using the 13 week criterion. The period of explosive prices lasted 13 months in copper, ending in November 2006, for 11 months in nickel (July 2006 to June 2007) and for 15 months in zinc (November 2005 to January 2007). Explosive price periods are detected later in 2007 for lead and tin with the tin bubble extending through to the summer of 2008

5. INTERPRETATION

Our secondary purpose, beyond characterizing the time series properties of the cash and three-month prices of the main non-ferrous metals, is to ask whether the price movements that are observed can be related to the supply and demand of fundamentals, as economic theory predicts. In equity markets, the market fundamental may be represented by the dividends on the basket of shares that make up the market index, and share prices should be linearly related to this fundamental (Gordon, 1959). PWY (2011) relate NASDAQ stock prices to the NASDAQ dividend yield and show

evidence of mild explosivity in the former but not the latter. They interpret this as evidence that the price explosion observed in the late 1990s cannot be explained by the market fundamental, and instead point towards a financial market epiphenomenon. From a formal standpoint, convenience yield is the commodity market analogue of dividend yield in the equities market. This is measured as the percentage premium of the cash over the deferred price net of interest and warehousing and loadout costs. The standard formula used in the finance literature follows by writing the cash (spot) price at date t as S_t , the three-month (future) price as F_t , to give the convenience yield c_t as

$$c_t = \frac{(1 + r_t)^{0.25} S_t - (F_t - w_t)}{S_t}, \quad (12)$$

where r_t is the three-month rate of interest expressed at an annual rate and w_t is the level of warehousing costs (paid after three months) – see Gospodinov and Ng (2013, p. 209). We take three-month dollar London Interbank Offered Rate (LIBOR) as the rate of interest.

In the analysis of equity market bubbles, the approach has been to divide the equity price index by the dividend yield, as in PWY. If the equity price series shows explosive behaviour absent from the behaviour of the ratio of equity prices to the dividend yield, one can conclude that explosive equity prices derive from the fundamental. If, instead, the ratio remains explosive, this would indicate equity prices have departed from the fundamental and can be inferred to be a bubble phenomenon. This procedure is problematic when applied to commodity prices. When inventories become high, the futures structure moves to “full carry” and convenience yield approaches zero. The ratio of commodity prices to convenience yield is therefore unbounded. Furthermore, if warehousing costs are ignored or incorrectly measured, convenience yield as calculated by equation (12) can become negative. Figuerola-Ferretti and Gonzalo (2010), who model convenience yields using LME non-ferrous metals prices, noted that measured convenience yields are often negative. Negative convenience yields make the ratio between the commodity price and the measured convenience yield uninterpretable.

Warehousing costs differ across warehouses. The LME regulates maximum rental rates for exchange-registered warehouses but warehouse operators often give substantial discounts on these rates. We follow Gospodinov and Ng (2013) and measure convenience yield by equation (12) setting $w_t = 0$. Estimated convenience

yield does indeed become negative over substantial periods of time for all five of the metals we are examining. Since the ratio approach is unavailable, we perform the more limited exercise of testing to see whether convenience yield is mildly explosive. The final column of Table 2 reports the GSADF statistics for the six convenience yield series. The statistics for three of the six metals (copper, nickel and lead) reject the hypothesis of no periods of explosive growth at the 1% level while those for two metals (aluminium and zinc) fail to reject even at the 10% level. However, in all six cases, excesses of the backward recursive ADF statistic over the corresponding critical value are transient and no bubbles are identified using the 7 week minimum bubble cut-off criterion. The implication is that the periods of explosive prices we have identified are common to the cash and futures price but absent from the convenience yield. The implication is that the explosive property arises out of the price and not the convenience yield process, at least as measured by equation (12). This suggests either that explosive prices had the effect of leading to a departure of prices from fundamental values or alternatively that convenience yield is not a satisfactory measure of fundamental value in metals futures markets.

Commodity market analysts typically refer to the stock-to-use ratio and the market balance as the market fundamentals. The stock-to-use ratio, which measures the ratio of the carryover from the previous crop year to the current year's consumption, is the most widely used measure of market fundamentals in the agricultural economics literature – see Bobenrieth et al (2013) who relate food price spikes to low stock-to-use ratios. The measure has also occasionally been used in metals markets – see IMF (2011).

The stocks-to-use measure works well in agricultural markets. The measure is typically employed on annual data using end-crop year stocks. The current year's harvest is revealed towards the end of the crop year so that the end-crop year stock-to-use ratio provides a scaled measure of availability for the new crop year. The measure is less suited to metals industries where production is continuous. As an alternative but related measure of market fundamentals, we construct the consumption-supply ratio (*CSR*), defined as the ratio of consumption of the metal in the quarter in question to production in the same quarter plus the stock level at the end of the previous quarter. This combines the metals balance used by metals industry practitioners with the agricultural stock-to-use measure used in agriculture. Because short term production and consumption elasticities are very low, this ratio can be effectively

taken as predetermined in the same way as the stock-to-use ratio in agriculture. Like the stock-to-use ratio, the consumption-supply ratio is scale-free.

In principle, the ratio satisfies $CSR < 1$ by construction, with a value close to unity indicating a tight market. In practice, the ratio may exceed unity because stocks are incompletely measured. This is true in our sample where our stock variable relates solely to exchange stocks (stocks in LME-registered warehouses). Two-year averages of the *CSR* are shown in Table 5.

[Table 6 around here]

We relate seasonally-adjusted quarterly values to the number of weeks in each quarter in which we have identified mildly explosive prices at the 95% significance level.⁸ The results are broadly similar irrespective of the minimum length hurdle imposed when we identify periods of explosive prices. For brevity, we focus below on the base case of a 7 week minimum bubble length.

We regress both the cash and three-month versions of this variable, for each of the five metals for which we have identified explosive price periods, on the consumption-supply. These are doubly truncated variables including a preponderance of zero values (49 or 50, depending on the metal, out of a total of 54 observations). To avoid least squares bias, we estimate using a double threshold Tobit model where the upper threshold of 13 weeks corresponds to a quarter in which prices are explosive in every single week.

The upper panel of Table 6 reports the estimated coefficients and *t*-statistics from equations. There is statistical significance at the 95% level in only a single case: three months tin. However, the coefficients and *t*-statistics are similar across metals. This suggests taking a pooled approach. The central panel of Table 7 reports the corresponding estimates from fixed effects, pooled and random effects Tobit models. There is little difference among these three sets of estimates and the coefficient of the consumption-supply ratio is significant in each case.⁹

[Table 7 around here]

We further experimented by adding convenience yield, averaged over the quarter, to the panel Tobit regressions (results not reported). The *t*-statistics associated with the consumption-supply ratio were little changed while the convenience yield coefficient was never statistically significant. These results give some support for the view that periods of explosive prices can be related to the fundamentals of physical

supply and demand, and that convenience yield is not in general a good measure of this fundamental.

The claim that periods of mildly explosive behaviour in the non-ferrous metals markets were associated with market fundamentals does not rule out a role for speculation which may have exacerbated or otherwise amplified explosive movements arising out of market tightness. We can only provide a direct test of this hypothesis for the copper market where we are able to take advantage of the Commitments of Traders position data published by the U.S. Commodity Futures Trading Commission (CFTC) for the COMEX copper futures market. We look specifically at the net level of non-commercial positions on the COMEX market, since these are widely interpreted as speculative positions. In addition, following Sanders, Irwin and Merrin (2010), we use Working's (1960) T as a measure of "excess speculation". Write long and short commercial ("hedge") positions as H_L and H_S and long and short non-commercial ("speculative") positions as S_L and S_S respectively. The net non-commercial position is $S_N = S_L - S_S$. Working's T index is then defined by

$$T = 1 + \frac{1(H_S \geq H_L) \cdot S_S + 1(H_S < H_L) \cdot S_L}{H_L + H_S} \quad (13)$$

If $T = 1$, the level of non-commercial activity is just sufficient to be available as counterparties for the commercial imbalance. Any excess over unity implies that speculators are acting as counterparties for each other.

We proceed as with the convenience yield exercise and look for periods of explosivity in these two measures of speculative activity. The GSADF statistic of 0.065 fails to reject the hypothesis of no periods of mild explosivity for the net non-commercial position variable. Instead, this hypothesis is rejected for the Working index at the 10% level where the GSADF is 2.125. However, the backward recursive sup ADF procedure fails to identify any explosive periods using the minimum bubble length criterion of seven weeks.

In summary, our results indicate that explosive price behaviour in the non-ferrous metals markets was at least due in part to tight physical markets, particularly at the start of 2004 and through 2006. While this does not rule out a role for speculation as a generator of bubbles, it is sufficient to reject the view that the bubbles in non-ferrous metals markets during the first decade of the century were purely financial epiphenomena.

6. CONCLUSION

The PSY test that we have applied to LME metals prices entails using a reduced form approach that first and foremost tries to capture the essential time series properties in the data. A significant part of John's approach to Econometrics was in contributing methodologically to the problem of testing between stationary and non-stationary time series, in a way that helped us more powerfully characterize and distinguish between the essentially different properties of a time series in different regions of the parameter space. In this context, our paper has used the PSY recursive test to look at the extent to which it identifies bubbles in the mildly explosive region of the parameter space. We applied the test specifically to non-ferrous metals prices because we had strong *a priori* grounds to expect that such metals would be prone to mildly explosive and possibly bubble-like behaviour. Single and multiple instances of mild explosivity were detected in all the LME non-ferrous metals other than aluminium, namely copper, lead, zinc, tin and nickel, both when the test was applied with the tuning parameter set to its default value of unity and when it was increased to provide for a more stringent test.

One of the main contributions of our paper was in the use of fundamental proxy variables to examine whether the mildly explosive behaviour we found could be interpreted as representing departures from each metal's fundamental value, thereby indicating bubble behaviour. In the pioneering paper in this literature, PWY (2011) related NASDAQ stock prices to dividend yield, showing mild explosivity in the former but not the latter and, on the basis of a standard asset price equation, used this result to indicate bubble behaviour. Here, we argued that for commodity markets, the natural counterpart to dividend yield in this context, the convenience yield, was less appropriate for this purpose and we outlined alternative fundamental proxy variables that, by design, had greater efficacy to uncover whether detected mild explosivity could point towards a financial market epiphenomenon. We are using this approach to analyse other commodity prices in ongoing work.

* * *

The third author remembers an occasion when John was an integral member of the group that participated in a memorial conference for another New Zealander econometrician, Rex Bergstrom, who preceded John in having been at the University of Essex for a major part of his career. No-one at that conference could have

imagined the same circumstances would arise from John's unexpected death only seven years later. John will be remembered as someone who took an optimistic view over what Econometrics as a discipline could achieve, especially in terms of its statistical role in helping to explain the world around us. Yet his insights were derived, exactly in this context, through his sensitivity to its limitations. He was always willing to see the best in people and, in conversation, he was distinguished by the fairness and integrity with which he judged the contribution of others.

ACKNOWLEDGEMENTS

We have benefited from the GAUSS code for the PSY procedure made available on Shu-Ping Shi's website (<https://sites.google.com/site/shupingshi/home/>). We are also grateful to Vanessa Davidson (who has fond memories of being taught by John Nankervis during her Essex M.Sc.) and Paul Robinson, both of the CRU Group, and Peter Kettle of the International Tin Study Group for the quarterly data used in this analysis. We thank participants at the 21st Annual Symposium of the Society for Nonlinear Dynamics and Econometrics, University of Milano-Bicocca, March 2013, and seminar participants in the University of York, England, for comments on the paper in its preparatory stages. Figuerola-Ferretti thanks the Spanish Ministry of Education and Science for support under grants MICINN ECO2010-19357, ECO2012-36559 and ECO2013-46395, and McCrorie, The Carnegie Trust for the Universities of Scotland under grant no. 31935.

NOTES

1. The standard present-value model is the most basic model of rational asset pricing and states that the asset price, P_t , is given by the sum of current and discounted expected future payoffs (benefits) from ownership of an asset. The model explains changes in asset prices in terms of "fundamentals", namely changes in expected future payoffs and an interest rate that "discounts" expected future values to a present value. (10) is the standard present-value model with an added term to represent the bubble component, and is the basis for what are called rational bubbles (see Branch and Evans (2011) for a reasonably complete bibliography of the literature). PSY (2015a) recognize that there are other bases for economic bubbles other than models such as (10) even if they have yet to be given a proper statistical basis.
2. See, e.g., Pindyck (1993) for an explanation.
3. More precisely, convenience yield is defined as the percentage premium of the (current) cash price over a deferred (future) price less the interest rate, storage cost and the rate of deterioration. It may be interpreted as the premium stockholders will pay for immediate access to inventory of known specification and location.

4. We use settlement prices and three-month mid-prices. Prices relate to Tuesdays or, in the case that a Tuesday was a holiday, the immediately prior trading day. Data sources for the cash and three-month prices are the LME (<http://www.lme.com/en-gb/pricing-and-data/historical-data/>) and the World Bureau of Metals Statistics publication, *World Metal Statistics* (<http://www.world-bureau.com/>).
5. The weights (Advanced Economies 70.4%, Brazil 1.8%, China 15.8%, India 3.6%, Russian Federation 8.4%) are shares of world refined copper consumption in 2002. Sources: IMF, *International Financial Statistics*, and World Bureau of Metal Statistics, *World Metal Statistics*.
6. Throughout the sample period, China was a major importer of both bauxite and alumina, the raw materials from which aluminium is obtained. However, the most important value-added component in aluminium is the energy input in smelting. China has been able to use stranded electricity (generating capacity installed distant from industrial users) to fuel its aluminium smelting industry.
7. Price returns would entail using the logarithms of prices instead of the levels, and in such a context an explosive process would be reflected in a non-zero intercept, not in the autoregressive coefficient.
8. Seasonal adjustment is achieved through the use of the STAMP program – see Koopman, Harvey, Doornik and Shephard (1999).
9. The likelihood ratio tests reported in the bottom panel of Table 6 show that the fixed effects panel model cannot be rejected in favour of the single equation alternative and, in turn, that the random effects model cannot be rejected in favour of the fixed effects alternative. However, the simple pooled model, which lacks metal-specific intercepts, appears less acceptable.

REFERENCES

- Bobenrieth, E., Wright, B., and Zheng, D. (2013). Stocks-to-use ratios and prices as indicators of vulnerability to spikes in global cereal markets. *Agricultural Economics* **44** (Supp.): 43-52.
- Branch, W.A., Evans, G.W. (2011) Learning about risk and return: a simple model of bubbles and crashes. *American Economic Journal: Macroeconomics* **3**: 159-191.
- Coakley, J., Kellard, N. and Tsvetanov, D. (2013) Bubbling over! The behavior of long-dated oil futures prices. Paper presented at the SIRE Conference in Finance and Commodities, University of St Andrews, July 2013. Available at: http://www.efmaefm.org/0EFMAMEETINGS/EFMA%20ANNUAL%20MEETINGS/2014-Rome/papers/EFMA2014_0601_fullpaper.pdf
- DeJong, D.N., Nankervis, J.C., Savin, N.E. and Whiteman, C.H. (1992) Integration versus trend stationarity in time series. *Econometrica* **60**: 423-433.
- Etienne, X.L., Irwin, S.H. and Garcia, P. (2014) Bubbles in food commodity markets: four decades of evidence. *Journal of International Money and Finance* **42**: 129-155.
- (2015) Price explosiveness, speculation and grain futures prices. *American Journal of Agricultural Economics* **97**: forthcoming (doi:10.1093/ajae/aau069).

- Figuerola-Ferretti, I., Gilbert, C.L., and McCrorie, J.R. (2015) Understanding commodity futures prices: fundamentals, financialization and bubble characteristics, preprint.
- Figuerola-Ferretti, I., and Gonzalo, J. (2010). Modeling and measuring price discovery in commodity markets. *Journal of Econometrics* **158**: 95-107.
- Figuerola-Ferretti, I., McCrorie, J.R. (2015) The shine of precious metals around the global financial crisis, preprint.
- Geman, H. (2005). *Commodities and Commodity Derivatives*, Hoboken (NJ), Wiley.
- Gilbert, C.L. (2010) Speculative influence on commodity prices 2006-08. Discussion Paper 197, *United Nations Conference on Trade and Development* (UNCTAD), Geneva.
- Gordon, M.J. (1959) Dividends, earnings and stock prices. *Review of Economics and Statistics* **41**: 99-105.
- Gospodinov, N., and Ng, S. (2013). Commodity prices, convenience yields, and inflation. *Review of Economics and Statistics* **95**: 206-219.
- Harvey, D.I., Leybourne, S.J., Sollis, R., and Taylor, A.M.R. (2014) Tests for explosive financial bubbles in the presence of non-stationary volatility, preprint available at: http://www.nottingham.ac.uk/~lezdih/bubble_bootstrap.pdf
- Homm, U. and Breitung, J. (2012) Testing for speculative bubbles in stock markets: a comparison of alternative methods. *Journal of Financial Econometrics* **10**: 198-231.
- Huang, W., Zheng, H, and Chia, W.M. (2010) Financial crises and interacting heterogeneous agents. *Journal of Economic Dynamics and Control* **34**: 1105-1122.
- IMF (2011) World Economic Outlook 2011. Washington DC, International Monetary Fund.
- Jarque, C.M. and Bera, A.K. (1987) A test for normality of observations and regression residuals. *International Statistical Review* **55**: 163-172.
- Koopman, S. J., Harvey, A. C., Doornik, J. A., and N. Shephard (1999) *Structural Time Series Analysis, Modelling, and Prediction using STAMP*. London: Timberlake Consultants Press.
- Miltersen, K.R., and Schwartz, E.S. (1998). Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates. *Journal of Financial and Quantitative Analysis*, **33**, 33-59.
- Pindyck, R. (1993) The present value model of rational commodity pricing. *Economic Journal* **103**: 511-530.
- Phillips, P.C.B., Magdalinos, T. (2007a) Limit theory for moderate deviations from unity. *Journal of Econometrics* **136**: 115-130.
- (2007b) Limit theory for moderate deviations from unity under weak dependence. In Phillips, G.D.A. & E. Tzavalis (Eds.) *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis*. Cambridge University Press: Cambridge, pp. 123-162.
- Phillips, P.C.B., Shi, S.P. (2014) Financial bubble imposition. Preprint at: [file:///C:/Users/jrm210/Downloads/BubbleImplosion_August2014H%20\(1\).pdf](file:///C:/Users/jrm210/Downloads/BubbleImplosion_August2014H%20(1).pdf)
- Phillips, P.C.B., Shi, S.-P. and Yu, J. (2015a) Testing for multiple bubbles: historical episodes of exuberance and collapse in the S&P 500. *International Economic Review*, forthcoming. [PSYa]
- (2015b) Testing for multiple bubbles: limit theory of real time detectors. *International Economic Review*, forthcoming. [PSYb]

- Phillips, P.C.B., Wu, Y. and Yu, J. (2011) Explosive behavior in the 1990s Nasdaq: when did exuberance escalate asset values? *International Economic Review* **52**: 210-226. [PWY]
- Phillips, P.C.B., Yu, J. (2011) Dating the timeline of financial bubbles during the subprime crisis. *Quantitative Economics* **2**: 455-491. [PY]
- Radetzki, R., Tilton, J.E. (1990) Conceptual and methodological issues. In Tilton, J.E., ed., *World Metal Demand*. Washington D.C., Resources for the Future, pp. 13-34.
- Sanders, D.R., Irwin, S.H., and Merrin, R.P. (2010). The adequacy of speculation in agricultural futures markets: too much of a good thing? *Applied Economic Perspectives and Policy* **32**: 77-94.
- Scherer, V., He, L. (2008), The diversification benefits of commodity futures indexes: a mean-variance spanning test. In Fabozzi, F.J., Füss, R., and Kaiser, D.G., *The Handbook of Commodity Investing*, Hoboken (NJ), Wiley, pp. 241–65.
- Working, H. (1960). Speculation on hedging markets. *Food Research Institute Studies* **1**:185-220.
- Yiu, M.S., Yu, J., Lu, J. (2013) Detecting bubbles in Hong Kong residential property. *Journal of Asian Economics* **28**: 115-124.

Table 1					
Descriptive Statistics					
		Volatility	Skewness	Excess kurtosis	Normality χ_2^2
Aluminium	Cash	21.9%	0.00	3.90	23.9
	Fut. 3m	21.0%	-0.09	3.98	30.7
	c. yield	10.1%	2.35	9.54	1974
Copper	Cash	29.5%	-0.39	5.55	215.9
	Fut. 3m	29.0%	-0.40	5.54	215.4
	c. yield	12.1%	-0.06	3.83	21.3
Nickel	Cash	45.1%	-0.21	11.6	2230
	Fut. 3m	42.6%	-0.14	8.85	1043
	c. yield	14.71%	0.79	3.11	77.6
Lead	Cash	43.7%	-0.42	8.42	916.9
	Fut. 3m	42.8%	-0.44	8.70	1009
	c. yield	21.3%	0.81	3.74	97.74
Tin	Cash	37.0%	-0.37	8.93	10.86
	Fut. 3m	36.4%	-0.47	9.23	1210
	c. yield	7.2%	1.25	4.81	290.1
Zinc	Cash	38.2%	-0.16	10.19	1580
	Fut. 3m	36.4%	-0.21	9.52	1301
	c. yield	8.15%	3.78	22.48	1330
Net n.c. positions		8.88%	-0.04	5.30	160.7
Working's <i>T</i> index		15.0%	0.39	8.52	947.8

The upper panel of the table reports the descriptive statistics for the deflated weekly (Tuesday) official cash and three-month price (Fut. 3m) of the major six LME metals. Convenience yield ("c. yield") is calculated from the prices using equation (12). The lower panel reports the same statistics for the measures of speculative activity in the COMEX copper market which we employ in Section 5 of the paper. Price volatilities are the standard deviations of the weekly first differences of the (deflated) prices divided by the means of the price levels. Working's *T* index of excess speculation is defined by equation (13). The volatilities for the *T* index of "excess speculation" in the COMEX copper market are standard deviations of first differences while those net non-commercial positions are the standard deviations of the first differences of net non-commercial divided by the sum of non-commercial long and short positions. Convenience yield volatilities are calculated as the standard deviations of convenience yields levels. All volatility figures are reported on an annual basis by multiplication by $\sqrt{52}$. The skewness, excess kurtosis, normality statistics all relate to the levels series. The normality test is that given by Jarque and Bera (1987). Sample: 12 January 2000 to 31 December 2013.

Table 2			
GSADF test statistics			
	Cash	Fut. 3m	Convenience yield
Aluminium	1.913	1.454	1.822
Copper	7.339	7.651	3.905
Nickel	5.515	5.601	4.262
Lead	6.296	6.986	3.766
Tin	5.345	5.002	2.187
Zinc	8.190	7.829	-0.791

The table reports the generalized sup ADF (GSADF) statistics for the cash and three-month prices plus the convenience yield of the six major LME non-ferrous metals estimated over the sample of 729 weekly observations from January 2000 to December 2013 (two observations are lost in lag creation). Convenience yield is calculated using equation (12). The initial window for recursive estimation is 56 weeks. The ADF lag is chosen to minimize the BIC over every subsample with the maximum lag length set at 5 weeks. Critical values: 2.069 (10%), 2.282 (5%) and 2.664 (1%).

Table 3													
Number of mildly explosive periods identified													
Minimum length		1			7			13			26		
Test size		10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
Aluminium	Cash	1+2	1+2	0+1	0	0	0	0	0	0	0	0	0
	3 months	2+3	1+3	0+2	0	0	0	0	0	0	0	0	0
Copper	Cash	5+3	3+4	4+2	3	3+1	2	3	2	1	1	0	0
	3 months	5+2	3+4	3+2	3	3	2	3	3	1	1	0	0
Nickel	Cash	2+3	4+1	4+1	2	1	2	1	1	1	1	1	0
	3 months	2+2	3+1	2+1	2	2	2	2	1	1	1	1	0
Lead	Cash	2+6	2+1	3+1	2	2	2	2	2	1	1	0	0
	3 months	3+4	2+3	2+2	2	2	2	2	2	2	1	1	0
Tin	Cash	6+10	3+8	3+3	3+1	2+2	1	2+1	1	0	0	0	0
	3 months	5+11	5+6	2+3	4+1	1+2	1	2+1	1+1	1	1	0	0
Zinc	Cash	3+3	2+2	2+2	1	2+1	1+1	1	2	2	1	1	0
	3 months	2+6	2+1	2+2	3+1	2	2	2	2	2	2	2	0

The table reports the number of mildly explosive periods identified for the cash and three-month prices of the six major LME non-ferrous metals estimated over the sample of 729 weekly observations from January 2000 to December 2013 (two observations are lost in lag creation). Where two numbers are quoted, the first is the number of “positive periods” (i.e. periods in which prices rose) and the second is the number of “negative periods [of implosion]” (periods in which prices fell). Where a single number is quoted, no negative periods were found. The initial window for recursive estimation is 56 weeks. The ADF lag is chosen to minimize the BIC over every subsample with the maximum lag length set at 5 weeks.

Table 4							
Estimated start and end dates for mildly explosive periods							
		Minimum bubble length					
		7 weeks		13 weeks		26 weeks	
		Start	End	Start	End	Start	End
Copper	Cash Fut. 3m	Dec 2003	Apr 2004	Dec 2003	Apr 2004	no bubble	
	Cash Fut. 3m	Mar 2006 Jan 2006	Jun 2006	Mar 2006 Jan 2006	Jun 2006	no bubble	
	Cash Fut. 3m	Jun 2006	Aug 2006 Sep 2006	no bubble Jun 2006	Sep 2006	no bubble	
Nickel	Cash Fut. 3m	no bubble Dec 2003 Jan 2004		no bubble		no bubble	
	Cash Fut. 3m	Aug 2006 Oct 2006	Jun 2007	Aug 2006 Oct 2006	Jun 2007	Aug 2006 Oct 2006	Jun 2007
Lead	Cash Fut. 3m	Dec 2003	Mar 2004	Dec 2003	Mar 2004	no bubble	
	Cash Fut. 3m	May 2007	Nov 2007	May 2007	Nov 2007	no bubble May 2007	Nov 2007
Tin	Cash Fut. 3m	Dec 2003	Jan 2004	no bubble		no bubble	
	Cash Fut. 3m	Feb 2004	May 2004 Jun 2004	Feb 2004	Jun 2004	no bubble	
Zinc	Cash Fut. 3m	Dec 2005	Jun 2006	Dec 2005	Jun 2006	Dec 2005	Jun 2006
	Cash Fut. 3m	Jul 2006	Jan 2007	Sep 2006 Jul 2006	Jan 2007	no bubble Jul 2006	Jan 2007

This table reports the mildly explosive periods in the deflated LME prices that are identified using the PSY procedure with a 5% size. Following PSY (2015a), critical values are calculated without taking account of the minimum bubble length and are therefore the same in each column of the table. No mildly explosive periods are recorded for aluminium.

	Aluminium	Copper	Nickel	Lead	Tin	Zinc
2000-01	91.4%	85.4%	93.4%	90.0%	86.0%	89.0%
2002-03	86.5%	82.5%	93.7%	89.4%	76.9%	78.0%
2004-05	89.9%	100.1%	95.0%	99.0%	92.8%	81.2%
2006-07	91.6%	96.4%	93.8%	98.1%	91.0%	95.9%
2008-09	86.8%	91.3%	77.1%	96.2%	87.5%	86.5%
2010-11	72.1%	90.8%	75.3%	90.5%	83.0%	79.6%
2012-13	71.1%	91.8%	72.1%	89.3%	88.6%	75.9%

Defined here on quarterly data, the consumption-supply ratio is the ratio of metal consumption to production plus stocks at the end of the previous quarter. Production and consumption cover the entire world. Stocks are London Metal Exchange stocks on the final day of the quarter. The table reports figures averaged over eight quarters. Data sources: CRU Group, the International Tin Study Group and London Metal Exchange.

Table 6		
Tobit Estimates of Impact of Market Fundamentals on Bubble Incidence		
	Cash	Fut. 3m
Copper	0.15 (1.65)	0.22 (1.53)
Nickel	0.46 (1.34)	0.16 (1.56)
Lead	0.24 (1.51)	0.26 (1.51)
Tin	0.16 (1.62)	0.23 (1.93)
Zinc	0.19 (1.89)	0.30 (1.72)
Fixed effects Tobit	0.21 (3.65)	0.24 (3.60)
Pooled Tobit	0.20 (3.59)	0.23 (3.51)
Random effects Tobit	0.20 (3.45)	0.23 (3.56)
Fixed effects versus single equation	$\chi^2(4) = 3.66$ [0.4537]	$\chi^2(4) = 5.51$ [0.2389]
Pooled versus fixed effects	$\chi^2(4) = 7.74$ [0.1017]	$\chi^2(4) = 9.50$ [0.0948]
Random effects versus fixed effects	$\chi^2(3) = 7.42$ [0.0596]	$\chi^2(3) = 8.59$ [0.0353]
<p>The dependent variable is the number of weeks in the quarter in which a period of mildly explosive prices is identified in the base case of a 7 week minimum bubble threshold. The conditioning variables are the seasonally adjusted consumption-supply ratios (CSRs). The top panel of the table reports the single equation Tobit estimates with censoring at zero and 13 weeks. The middle panel reports panel Tobit estimates imposing homogeneity on the coefficients. The bottom panel reports likelihood ratio tests of the nested hypotheses. <i>t</i>-statistics are reported in (.) parentheses and tail probabilities in [.] parentheses. Sample: 2001q2-2013q4.</p>		

Figure 1: Deflated LME three month prices, 2000-13

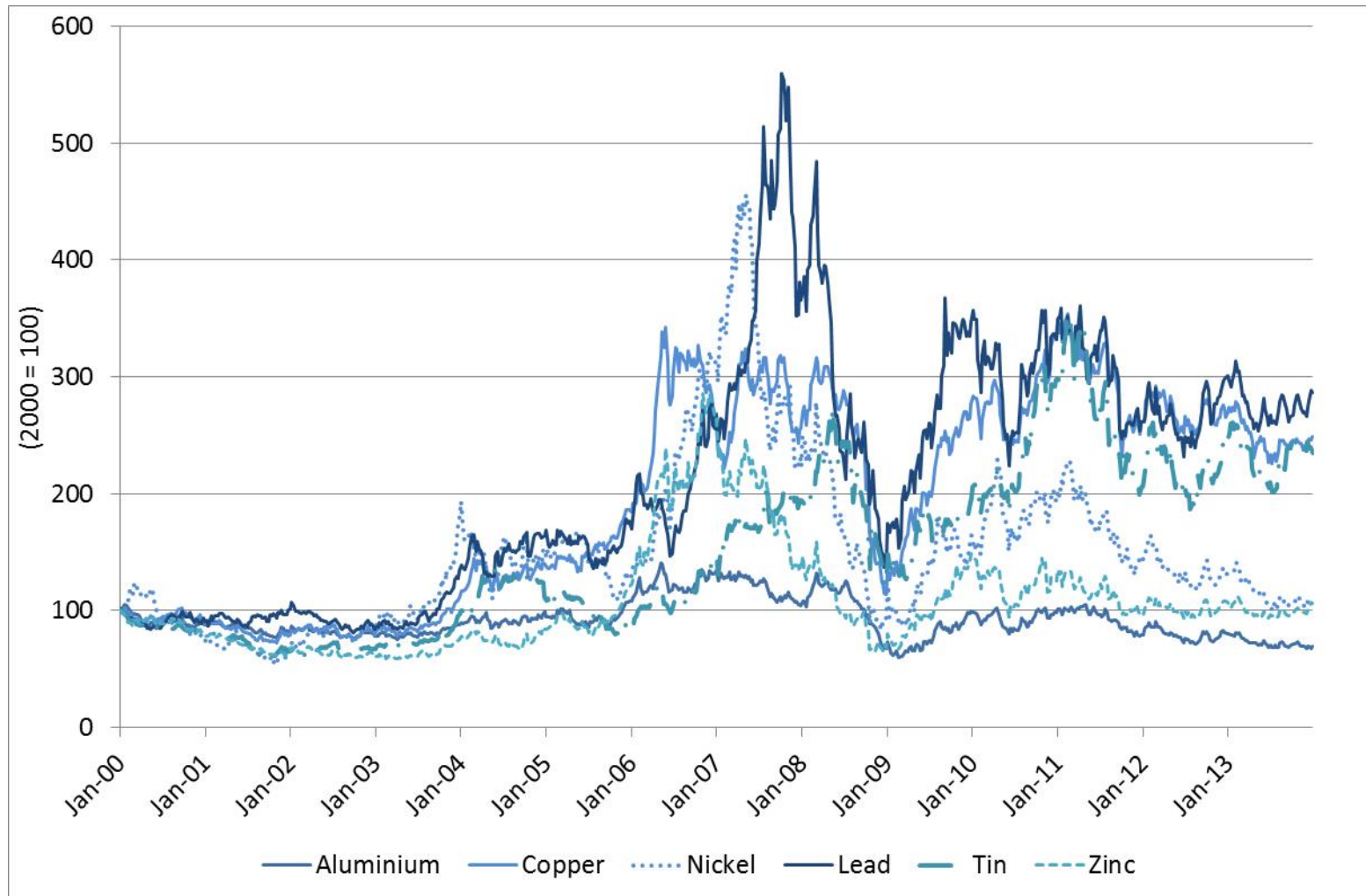


Figure 2: Deflated three months aluminium BSADF sequence

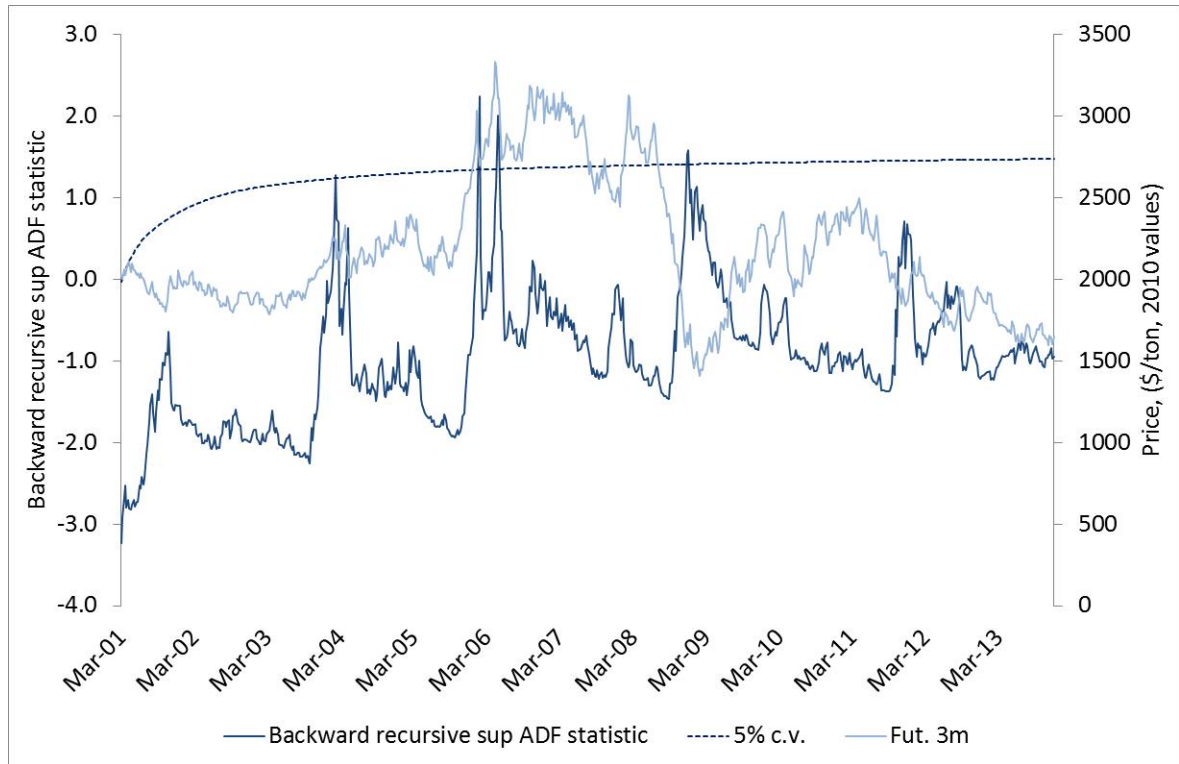


Figure 3: Deflated three months copper BSADF sequence

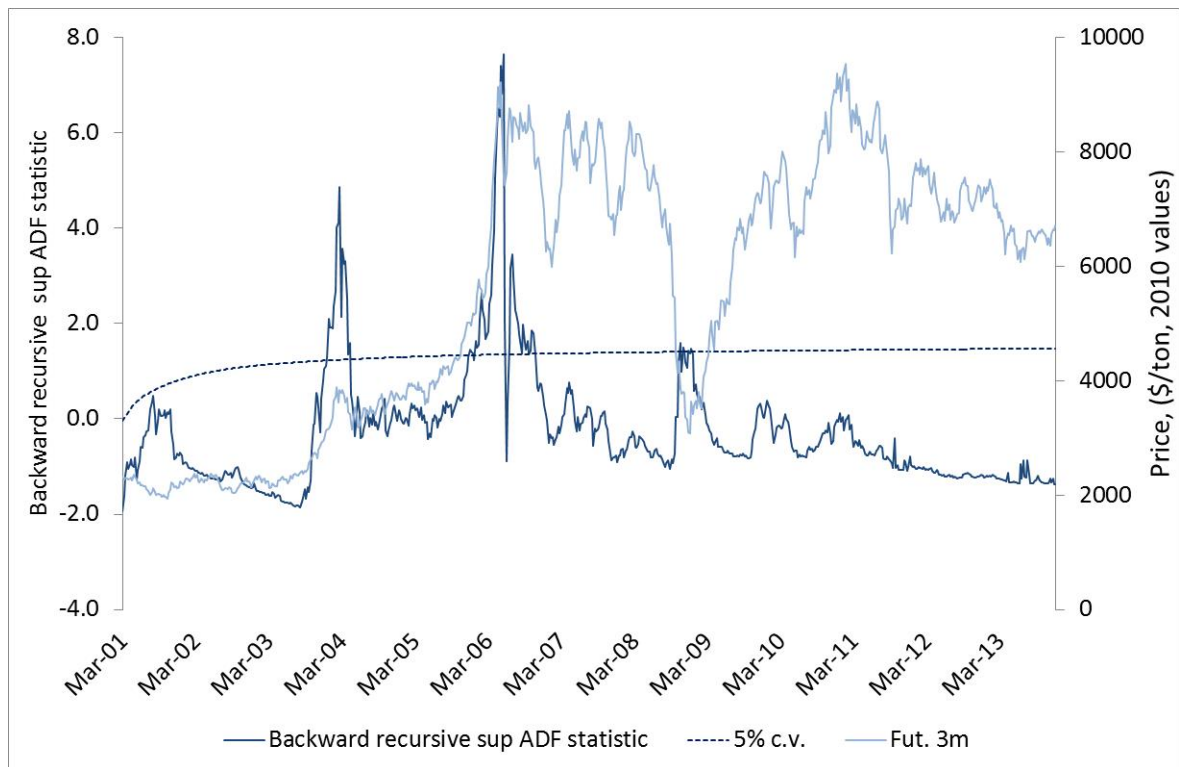


Figure 4: Deflated three months nickel BSADF sequence

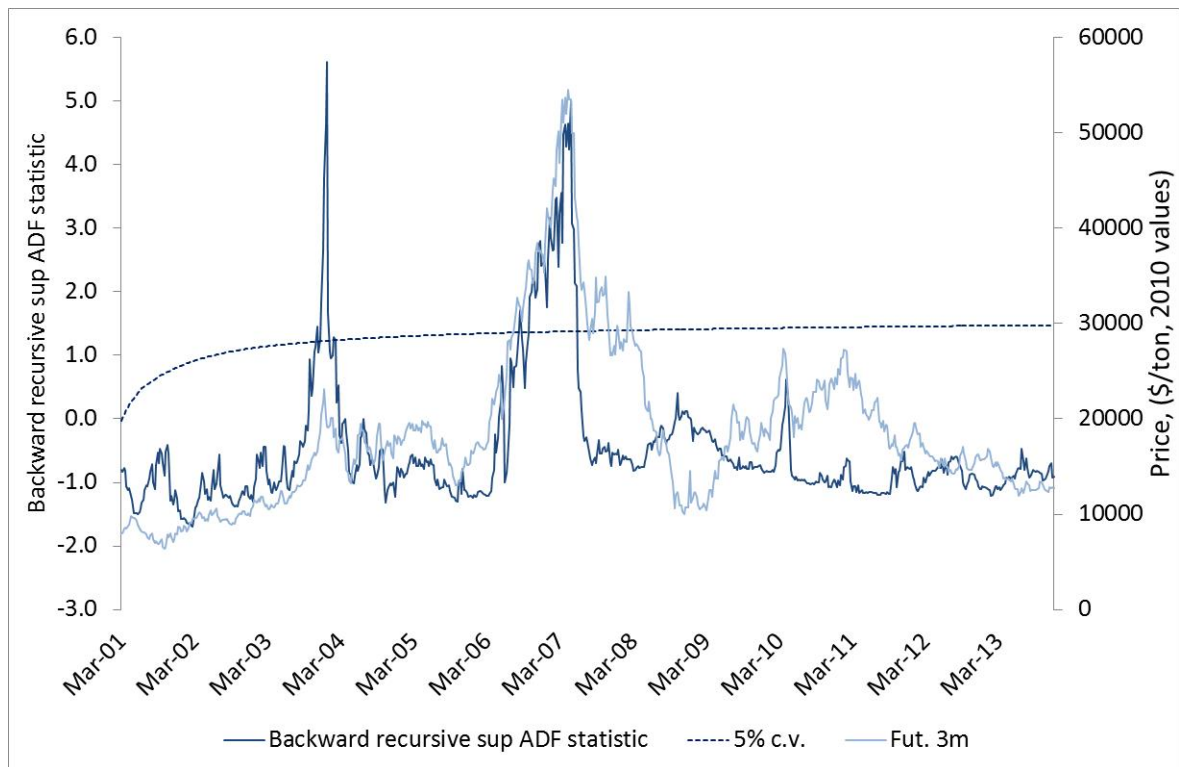


Figure 5: Deflated three months lead BSADF sequence

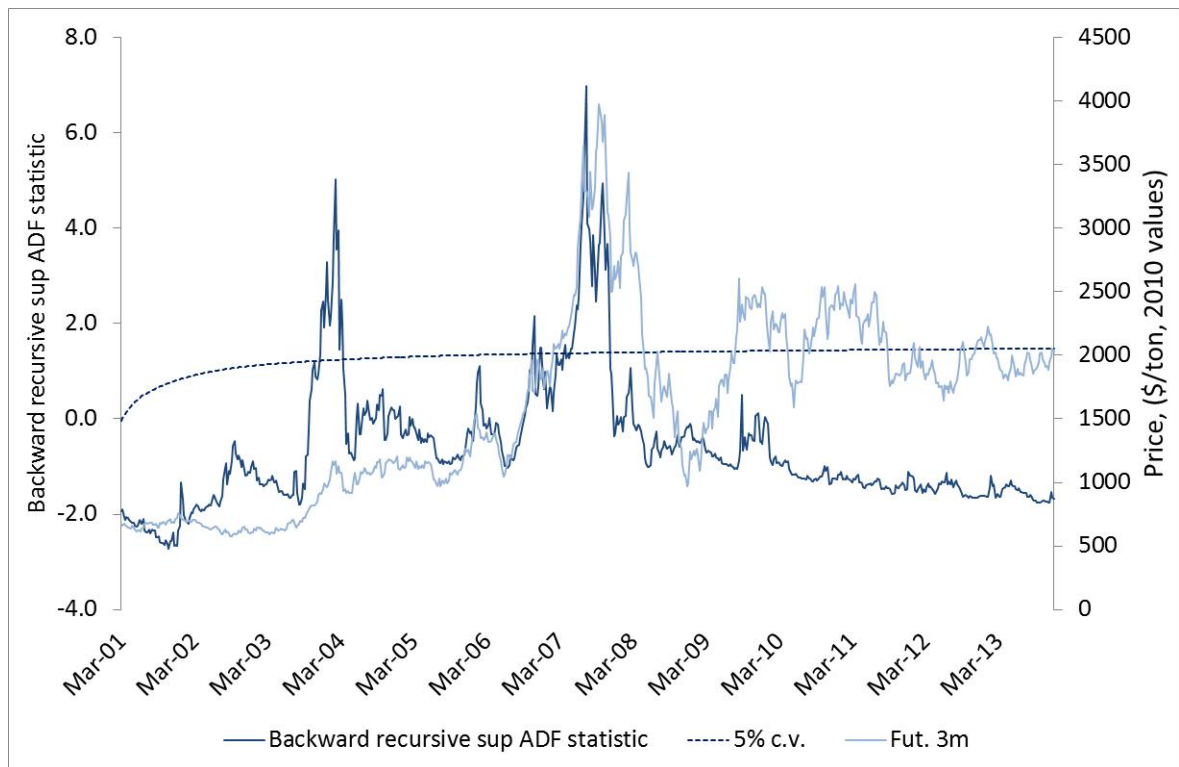


Figure 6: Deflated three months tin BSADF sequence

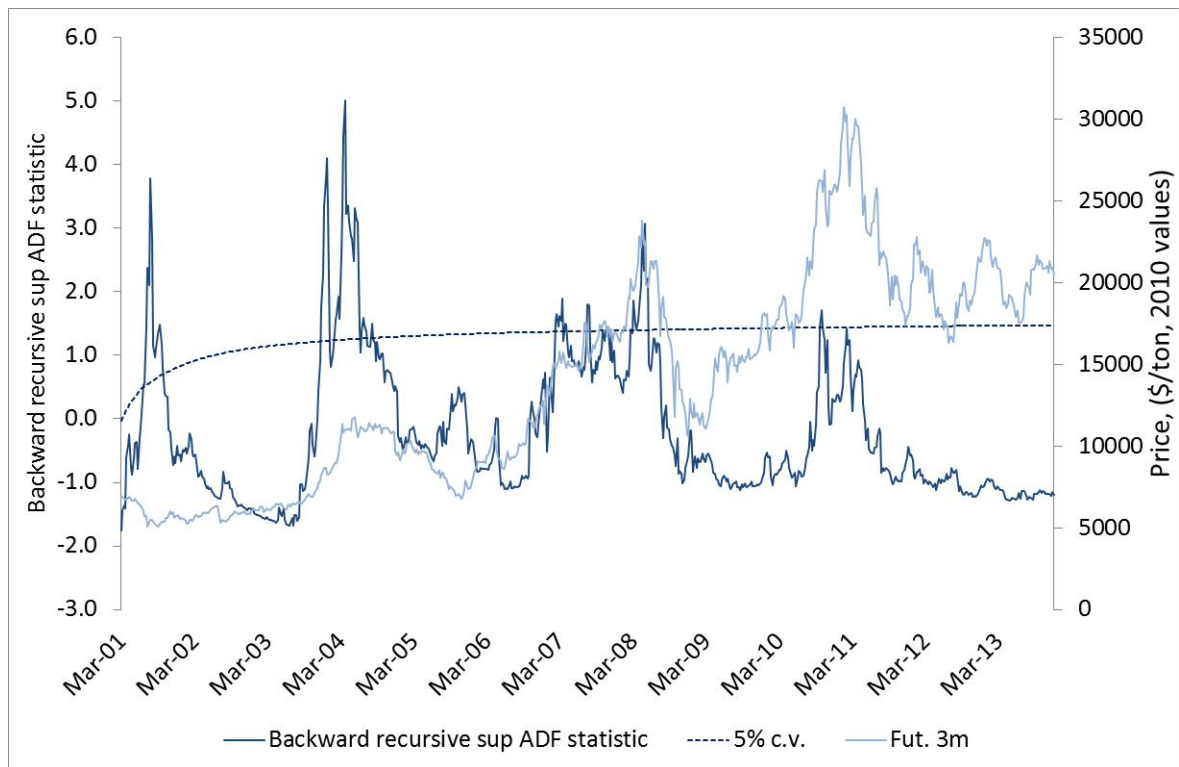


Figure 7: Deflated three months zinc BSADF sequence

