Center for
Quantitative
Economics

New evidence for explosive behavior of commodity prices

Jeanne Diesteldorf, Sarah Meyer and Jan Voelzke †

50/2016

† Department of Economics, University of Münster, Germany

Working Paper

New evidence for explosive behavior of commodity prices

Jeanne Diesteldorf · **Sarah Meyer** ·
Jan Voelzke

July 14, 2016

Abstract Over the past decade, the academic literature has engaged in a lively discussion about speculative bubbles in commodity markets. A number of papers have empirically investigated explosive behavior, albeit employing different econometric approaches that examine univariate time series for the existence of bubbles. Largely, these individual tests have been applied on a set of univariate time series such as stock indices or commodity prices, ignoring the multiple testing nature of the problem.

Departing from there, we examine explosive behavior in commodity futures markets, by employing a panel-data set containing the ten most liquid agricultural futures contracts traded in the US to conduct the Generalised Sub-ADF test by Phillips et al. (2015). We aggregate individual test-results using a stagewise rejective multiple test procedure. Overall, our results show evidence for several periods of explosive behavior in the markets for wheat, cattle, cocoa, coffee and cotton over the past 35 years. Our paper is the first to employ this approach on a panel data set, thereby solving econometric shortcomings of previously published work in the extant literature.

Keywords Speculative bubbles · Derivatives · Commodity markets · Panel-test · GSADF-test

J. Diesteldorf
University of Muenster
Department of Economics
E-mail: jeanne.diesteldorf@wiwi.uni-muenster.de

S. Meyer
University of Muenster
Department of Economics
E-mail: sarah.meyer@wiwi.uni-muenster.de

J. Voelzke
University of Muenster
Department of Economics
E-mail: mail@janvoelzke.de

1 Introduction

Over the past decade, the academic literature as well as the public debate have engaged in a lively discussion on the increasing financialization of commodity markets. Between 2006 and 2011, commodity prices worldwide more than doubled exhibiting soaring price spikes in 2007 and 2011, price crashes between mid 2007 and 2008, and a continuing decline since 2013.

Several fundamentals such as growth in emerging markets, weather shocks, declining inventory or consumption growth are put forward as price increasing factors. However, it has not been proven that these factors are responsible for the price spikes and crashes. In addition, it is difficult to quantify their impact due to the lack of valid data availability.

The public opinion, i.e. politicians, regulators and parts of the media, argue that low interest rates as well as the attractiveness of raw materials regarding portfolio diversification supported the increasing financialization of commodity markets, albeit feeding speculative bubbles in agricultural prices.

In consequence, a number of papers in the academic literature have empirically investigated explosive behavior in commodity markets. Gutierrez (2013) employs the sup-ADF test to examine the existence of price bubbles between 2007 and 2008 in the US agricultural commodity markets for wheat, corn, soybeans and rice. He uses a bootstrap methodology to compute the finite sample distributions of recently proposed tests. Monte-Carlo simulations confirm that the bootstrap methodology works well, and allow the author to identify explosive processes and collapsing bubbles for wheat, corn and rough rice. There was less evidence of exuberance in soybean prices.

Areal et al. (2014) apply the generalized version of the sup-ADF test to monthly price indices of food, beverages, agricultural raw material, cereals, dairy, meat, oils and sugar, as well as to 28 agricultural commodity prices between 1980 and 2012. These tests identify that price bubbles occurred for some commodities within food markets. However, the authors underline the methodological issue that the detection of relatively few bubbles may be caused by the data frequency. By employing daily prices, they might have detected bubbles that remain disguised under lower frequency data.

Etienne et al. (2014) estimate the sup-ADF test using daily data from individual futures contracts instead of a continuously calculated time series. They investigate if speculative bubbles exist in 12 US agricultural futures markets between 1970 and 2011 and conclude that all 12 markets experienced multiple periods of price explosiveness. However, the authors report those bubble episodes to represent a very small portion between 1.5 and 2% of total price behavior during the 42-year period. Moreover, most bubbles are found to be short-lived with 80-90% lasting fewer than 10 days. This finding is qualitatively confirmed by Etienne et al. (2015).

Paulson et al. (2013) apply a regime switching regression model to test for the presence of periodically collapsing bubbles in the US markets for wheat, corn, soybeans, rice, sugar and cotton. The authors examine daily settlement prices referring to the first and second nearby futures contracts, respectively, between 1989 and 2011. Their results yield no evidence for the presence of bubbles other than in the market for soybeans.

Adämmer & Bohl (2015) employ the momentum threshold autoregressive (MTAR) approach to test for speculative bubbles in US corn, soybean and wheat prices. They approximate fundamental values of the commodities under scrutiny by using real crude oil prices and real exchange rates. Their empirical results support the hypothesis that speculative bubbles are present in wheat prices between 2003 and 2013. For corn and soybeans, their findings are inconclusive.

Apparently, there prevails a broad discussion in the academic literature on the topic of speculative bubbles in price time series, interdepending on the theoretical debate of the question of whether an econometric model needs a bubble component to explain empirically observed behavior of price time series. Gürkaynak (2008) gives a vast review on the topic.

As outlined above, the empirical studies on explosive behavior employ different econometric approaches to investigate several time series for the existence of speculative bubbles. Often, they apply the same test on several individual time series, interpreting the results and significance statement as if they were single tests. This entails a statistical mistake: Testing more than one time series, such as considering an entire panel data set, is a multiple test procedure. The resulting, somewhat implicit research issue is therefore no longer a test of each individual time series but a test of the overall hypothesis of explosiveness in one of the time series or in a particular sub-group. In consequence, ignoring this option of multiple testing leads to a misinterpretation of the significance level and may thereby result in erroneous conclusions. Controlling the multiple levels can be achieved by the classical Bonferroni scheme or one of the developments in multiple test procedures as those by Holm (1979), Rüger (1978) or Simes (1986). Especially, the latter improves the power drastically in comparison to the classical Bonferroni scheme all the while still allowing for a broad class of dependence structures. In this investigation, we will rely on the structure and hypotheses of Simes (1986), with all its assumptions and implications, but also follow Hommel (1988) and use his advancement, which allows to build and test arbitrary subset-hypotheses.

Our paper conducts a panel test approach jointly investigating explosive behavior in ten different agricultural commodity price time series. In doing so, we are able to circumvent this inherent mistake outlined above and employ a procedure that is statistically correct.

We aim at detecting market exuberance in commodity futures prices, albeit remaining outside of the discussion about the general necessity of a specific bubble model. We use the Generalized Sup-Augmented Dicky Fuller test (GSADF) by Phillips et al. (2015). Already, the Sup-ADF test (Phillips et al. (2011)) proofed to perform satisfactorily well.[1] Nevertheless it has reduced power for long(er) time series, for those with multiple bubbles and/ or if em-

[1] Cp. e.g. Homm & Breitung (2012).

ployed on rapidly changing markets. In contrast, the GSADF-test overcomes these drawbacks. Moreover, it possesses the advantage to "detect market exuberance arising from a variety of sources, including mildly explosive behavior that may be induced by changing fundamentals such as a time-varying discount factor".[2]

The testing procedure can also be used to time-stamp phases of explosive behavior, but, in the first place, it delivers a classical test-statistic for the entire time series. This very value can then be compared to the corresponding critical value or a p-value can be calculated by Monte-Carlo methods, since the asymptotic distribution of the test-statistic under the Null hypothesis is known.

Our paper proceeds as follows. Section 2 discusses our method and introduces the data used for estimation. In Section 3, we provide the empirical results and compare them to the existing literature. Finally, Section 4 concludes.

2 Method and Data

We examine explosive behavior in US commodity futures prices. $p_{i,t}$ denotes the price of commodity i at time t. For each commodity, we test the hypothesis of explosive behavior. Hence, the underlying reduced form empirical model can be stated as:

$$p_{i,t} = \mu_i + \delta_i p_{i,t-1} + \sum_{k=1}^{p} \phi_k \Delta p_{i,t-k} + \varepsilon_{i,t}, \tag{1}$$

with $\varepsilon_{i,t} \overset{iid}{\sim} N(0, \sigma_i^2)$. The Null hypothesis can be stated as $H_{i,0} : \delta_i = 1$ against $H_{i,1} : \delta_i > 1$. The GSADF test procedure essentially calculates the supremum over all ADF test statistics for all possible starting points and interval lengths that lie within the time period examined.[3]

For the estimation of the model parameters, the calculation of the test statistics, the critical values resp. the individual p-values and the graphics, we used the Rtadf package for Eviews by Caspi et al. (2014). For the initial window, we used a fraction of 8% of the observations. This is close to the suggested value of approximately 0.1 that would result in taking the proposed formula $0.01 + 1.8\sqrt{T}$ of Phillips et al. (2015). However, decreasing this value mildly allows to detect earlier explosive episodes. The lag length p is chosen by AIC.[4] For the calculation of the p-values, we simulate 20 000 realizations

[2] Cp. Phillips et al. (2015), p. 1045.

[3] For a detailed description see Phillips et al. (2015).

[4] Further parameters $d = 1$ and $\eta = 1$ are set in accordance to the literature.

of the test statistic under the null hypothesis.[5]

Simply rejecting the null hypotheses for all the time series whose test statistics exceed some level-α critical value may lead to an excessive number of rejections. This approach ignores the multiple testing nature of the problem and does therefore not control the multiple level α. To overcome this problem, we follow Meyer & Trede (2016) and propose to adopt Simes' classical intersection test as a panel aggregation method to control the Familywise Error Rate (FWER), i.e. the probability of falsely rejecting at least one individual null hypothesis, at level α. Accordingly, we set up the global null hypothesis H_0 that all individual null hypotheses $H_{i,0},\ i = 1, ..., N$ are true. This method is easy to implement since it only requires the p-values of the N time series tests. However, it does not answer the question which individual hypothesis can be rejected once the overall hypothesis is rejected. Hommel (1988) provides a general procedure to test arbitrary sub-hypotheses and particularly states how Simes' framework can be used to identify the individual hypothesis that can be rejected considering the multiple nature of the entire test procedure.

Following the identification of the time series exhibiting explosive behavior using the Hommel procedure, we could use the provided backwards SADF sequence to date-stamp the bubble periods. Figures 1 to 4 depict the backwards SADF sequence and the critical value sequence of our data. Note, however, that the critical values are not adjusted to the rolling window scheme, but must be understood as evaluations of the corresponding single test. The time-stamping should therefore be interpreted as descriptive, it does not need to be consistent with the statistically sound results of the overall test. This means that several SADF-sequences cross the line of critical values, but are not characterized as explosive by our procedure.[6]

We investigate daily settlement price data for the following ten US agricultural commodity futures contracts:

- grains: corn, cbot[7] wheat (cbwheat), kcbt[8] wheat (kwheat)
- livestock: live cattle, feeder cattle, hogs
- soft commodities: cocoa, coffee, sugar, cotton

Thereby, we cover the most commonly traded agricultural derivative markets in the world, relying on a relatively long sample period from January 3, 1980 to June 30, 2015.[9] All data are taken from Thomson Reuters Datastream.

We use perpetual series of commodity futures prices. They start at the nearest contract month, delivering the first price value for the continuous series

[5] Note, however, that we just used 2 000 realizations as a basis for the provided graphs in Figure 1-3.

[6] Fixing the critical value to a constant is standard in applied work. For more information on the time-stamping procedure see Caspi et al. (2014) section 3.1.

[7] Chicago Board of Trade.

[8] Kansas City Board of Trade.

[9] To ensure a long sample we exclude the markets for soybeans and soybean oil. These two commodities have only been traded in derivatives markets since April 2006.

until either the contract reaches its expiry date or until the first business day of the notional contract month, whichever is sooner. At this point, prices from the next trading contract month are taken. No adjustment for price differentials is made. We deflate the prices by dividing nominal prices by a CPI normed to 100 in August 1983.

Table 1 provides summary statistics for our futures prices, their respective trading volume and open interest. Due to differing trading days across markets as well as slightly deviating market characteristics, the number of usable observations varies across commodities. Measured by the number of outstanding contracts (open interest), the market for corn is the largest and the most liquid. The market for feeder cattle, in turn, is the smallest and exhibits the lowest mean trading volume compared to the others.

3 Empirical Results

Figures 1 to 4 summarize our empirical results. They depict three time series for each of the ten commodities under scrutiny. The green line illustrates real commodity prices called NORM*COMMODITYNAME*, respectively. Further, the blue line paints the ADF statistic sequence and the red line shows the corresponding critical value sequence.[10] As outlined before, the time stamping implied by our ADF statistic sequence should be understood as descriptive.

Figure 5 visualizes the joint decision about which commodity prices exhibit explosive behavior over time based on our estimation procedure. It outlines the logarithmized p-values corresponding to the GSADF-tests on the ten commodity price series. Following the procedure described in the preceding section, we can reject the null hypothesis for those commodities with p-values smaller than α/j, which is depicted by the dotted line.[11]

We find explosive behavior in the prices for wheat, cotton, feeder cattle, cocoa and coffee. Prices for corn, sugar, live cattle and hogs do not show explosive behavior. The results for corn prices are particularly interesting. Its corresponding p-value of 2.3% would clearly imply the conclusion of explosiveness in an isolated test. However, in our multiple setting, we are unable to deduce explosive behavior. As depicted in Figure 5 the p-value lies above the adjusted critical value implied by the Hommel (1988) procedure. Our panel data approach therefore avoids the statistical mistake of wrongly concluding that there exists explosive behavior in the corn market. This emphasizes the necessity of the applied procedure in comparison to conducting the GSADF-test on isolated time series.

[10] Cp. citeca/13 Section 5.

[11] Note that the p-value corresponding to cbwheat is numerically equal to zero and therefore termed $-INF$ after logarithmization. Obviously, its true value is larger than zero but extremely small.

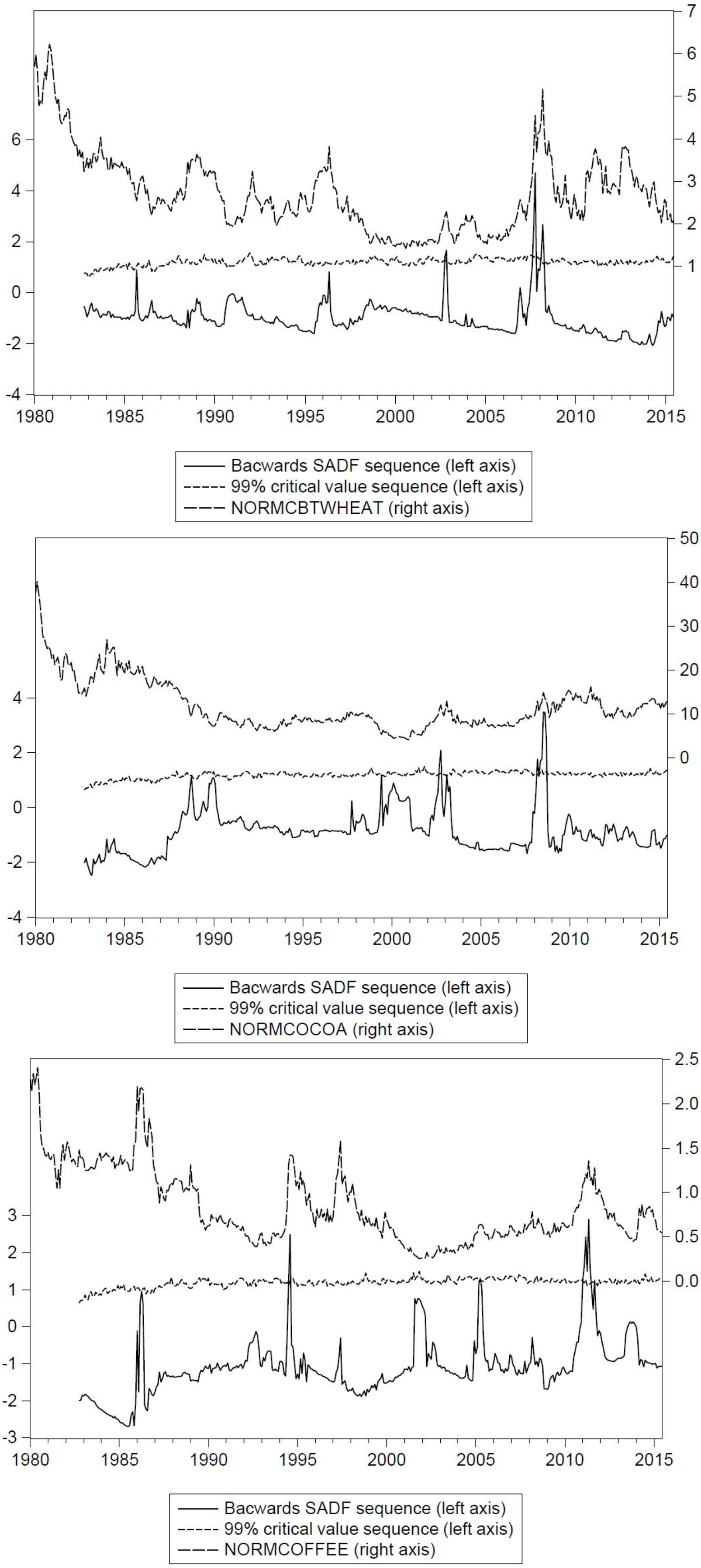

Fig. 1 Normalized prices, the backwards SADF sequence and the critical value sequence

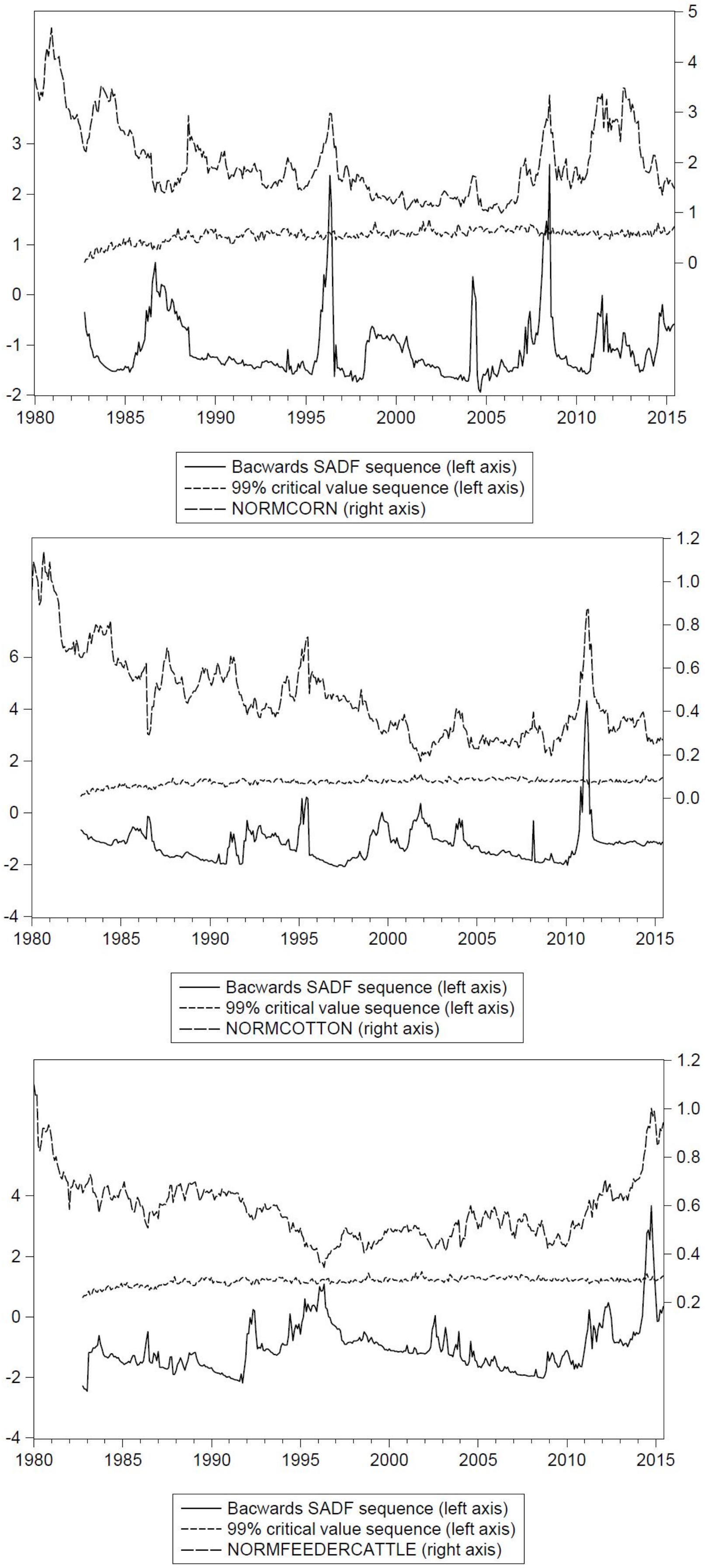

Fig. 2 Normalized prices, the backwards SADF sequence and the critical value sequence

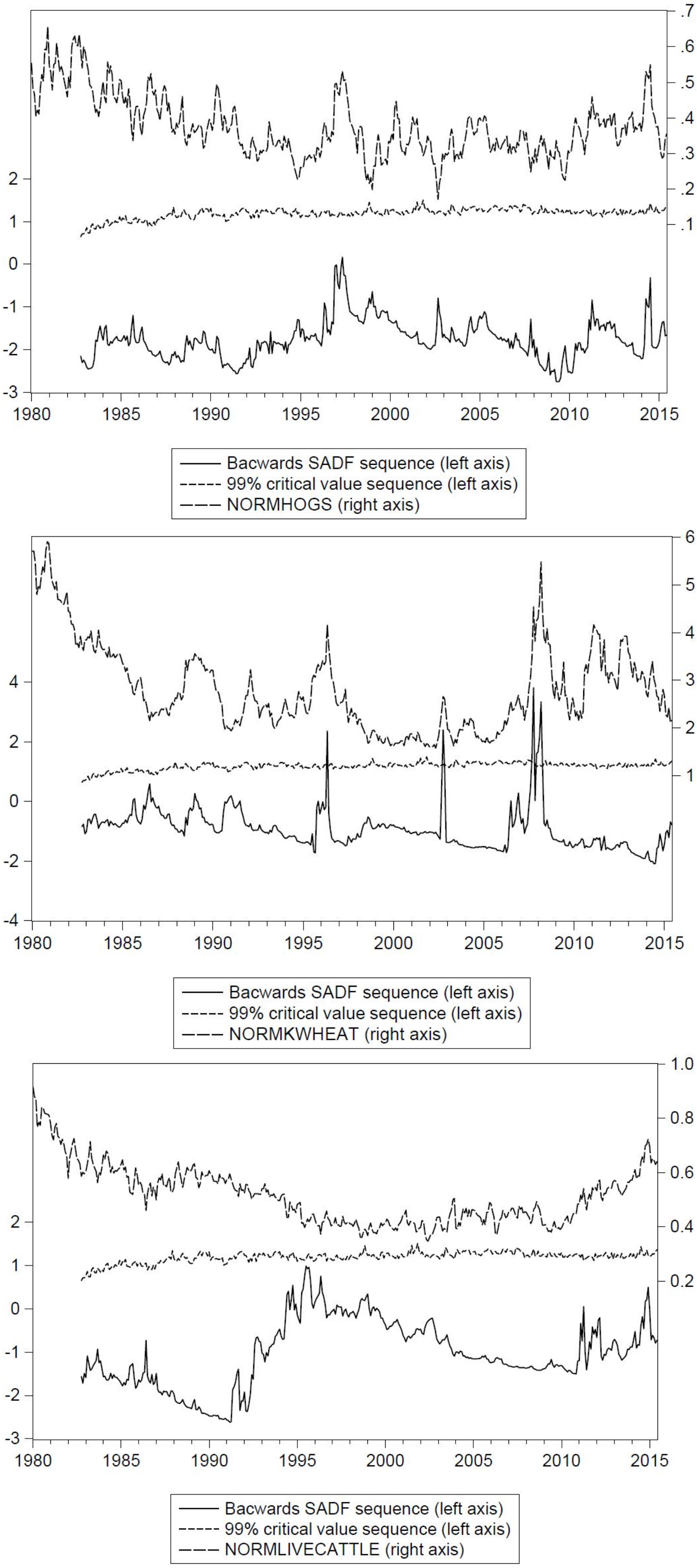

Fig. 3 Normalized prices, the backwards SADF sequence and the critical value sequence

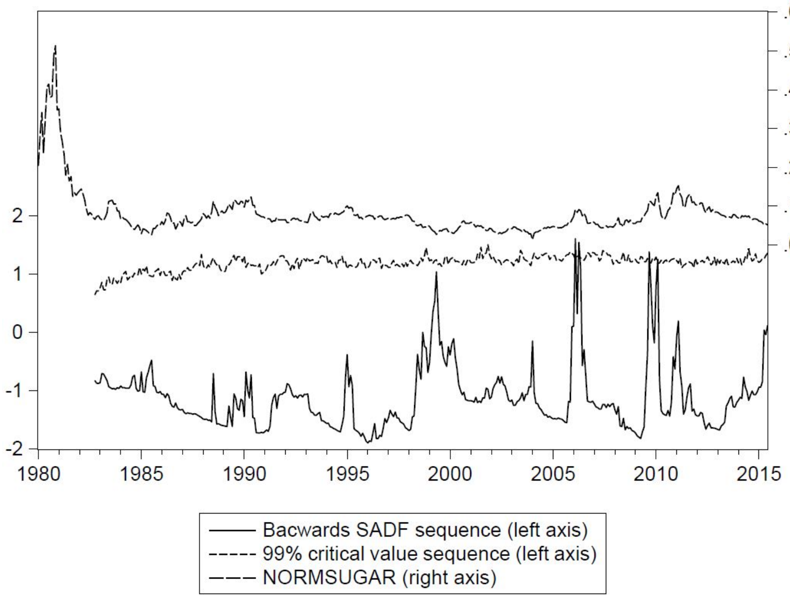

Fig. 4 Normalized prices, the backwards SADF sequence and the critical value sequence

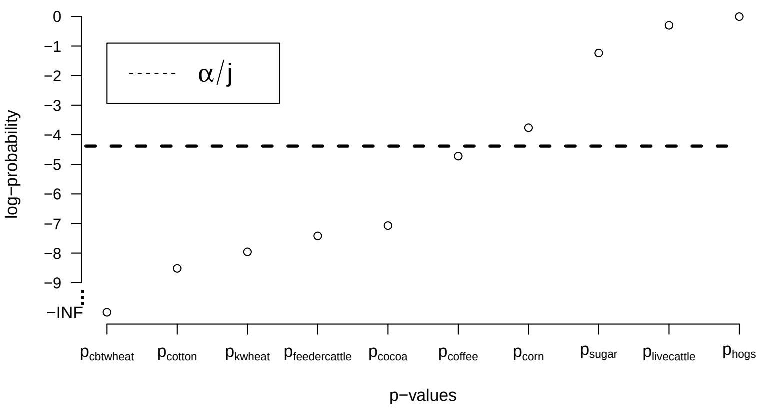

Fig. 5 Logarithmized p-values and the cut-off calculated as in Hommel (1988). Commodities with a p-value under the line show significant explosive behaviour.

Compared to the findings in the extant literature, our results partly underline what other research has shown before. While Gutierrez (2013) supports our empirical evidence for speculative bubbles in the wheat market, he also finds explosive behavior in the corn market, which, in turn, we do not. Etienne et al. (2014) report evidence for bubbles in all of the ten markets we examine. However, they stress that these bubble episodes represent a small portion of total price behavior.

Based on an econometrically different approach, Paulson et al. (2013) find no bubbles in the markets we consider here, while Adämmer & Bohl (2015) report evidence in favor of speculative bubbles in wheat prices between 2003 and 2013, supporting our evidence. Further, their empirical findings yield inconclusive results for the market for corn futures, where we find no signs of explosive behavior.

Apparently, our results are in line with what others have found before. Moreover, our empirical evidence underlines the fact that explosive behavior in commodity futures markets has existed prior to the price spikes induced by increasing financialization. Figures 1 to 4 outline that the price times series for coffee, corn and kwheat exhibit explosive behavior in 1995 and 1996; the market for cocoa shows explosive behavior between 2002 and 2004.

Moreover, it jumps to the eye that the markets cocoa, coffee, corn and kwheat exhibit two or even three pronounced bubble periods. The other markets, such as feeder cattle and cotton, show explosive behavior towards the end of the sample period only. Feeder cattle is the only futures market characterized by explosive behavior as late as 2014/ 2015; it is also the only livestock futures market with price bubbles; live cattle and hogs prices show no signs of explosiveness.

The markets for wheat, cocoa and corn show signs of explosiveness between 2006 and 2011; the markets for cotton and coffee (both soft commodities) show bubbles in 2011, but not during the period of increasing financialization beginning in the middle of the last decade. Some markets exhibit bubbles long before: wheat and corn in 1996 (both grains), coffee a little earlier in later 1994.

Our results underline that explosive behavior exists in some commodity futures markets regardless of the markets' trading volume, open interest or the characteristics, such as market structure or institutional setting. In these terms, the markets under scrutiny are very similar. Also, we detect bubbles across commodity categories in the market for grains, livestock and soft commodities, respectively.

4 Conclusion

We examine explosive behavior in a panel dataset of ten US agricultural commodity futures markets. To this end, we estimate the generalized sup-ADF test and observe explosive behavior in the market for wheat, feeder cattle, cocoa, coffee and cotton. No bubble-like behavior is found in the markets for sugar, live cattle, hogs and corn.

Our results confirm some of the findings of the extant literature on the matter. However, our estimations identify several periods of bubble-like behavior in the markets for coffee, corn and wheat futures long before the price spikes caused by an increasing financialization of agricultural derivative markets.

By employing a panel approach, we appreciate the multiple testing nature of the problem. Nonetheless, we are unable to draw a conclusion about possible causes of explosive behavior exhibited. In future research, we may do so in setting up structural models for the time-stamp data provided by our applied procedure. Alternatively, entirely theoretical (DSGE-)models might explain the empirical evidence outlined here.

Table 1 Summary Statistics

Price

	cwheat	cotton	kwheat	feeder cattle	cocoa	coffee
Mean	417.82	68.44	440.25	92.74	1854.61	125.96
Max	1282.50	214.14	1337.00	242.93	3822.00	314.80
Min	230.75	28.52	235.75	47.65	682.00	42.00
Obs.	9781	9732	9524	9525	9732	9732

	corn	sugar	live cattle	hogs
Mean	312.40	11.98	81.94	58.48
Max	831.25	44.80	171.00	132.65
Min	142.75	2.35	50.12	27.95
Obs.	9521	9732	9471	9523

Trading Volume

	cwheat	cotton	kwheat	feeder cattle	cocoa	coffee
Mean	37521.81	12006.96	12736.48	3336.275	9742.634	11487.16
Max	351063	120809	69786	28848	78252	105303
Min	0	0	0	0	0	0
Obs.	9454	9365	9212	9218	9357	9352

	corn	sugar	live cattle	hogs
Mean	102077.9	42728.25	24780.59	15632.83
Max	845770	410033	147566	120692
Min	0	0	0	0
Obs.	9202	9352	9159	9212

Open Interest

	cwheat	cotton	kwheat	feeder cattle	cocoa	coffee
Mean	165680.1	86561.73	70354.3	19546.53	89278.7	66497.29
Max	562198	302683	250506	56556	726500	207077
Min	10	0	353	4440	5853	907
Obs.	9781	9731	9524	9525	9731	9731

	corn	sugar	live cattle	hogs
Mean	526257.7	296582	140778.9	82285.16
Max	1745258	1113863	398748	335916
Min	71072	25797	0	14455
Obs.	9521	9731	9470	9523

Sample period: January 3 1980 to June 30 2015. All data are taken from Thomson Reuters Datastream.

References

Adämmer, P. & Bohl, M. T. (2015), 'Speculative bubbles in agricultural prices', *The Quarterly Review of Economics and Finance* **55**, 67–76.

Areal, F. J., Balcombe, K. G. & Rapsomanikis, G. (2014), 'Testing for bubbles in agricultural commodity markets', *ESA Working Paperp* **14**.

Caspi, I. et al. (2014), Rtadf: Testing for bubbles with eviews, Technical report, University Library of Munich, Germany.

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

Etienne, X. L., Irwin, S. H. & Garcia, P. (2015), 'Price explosiveness, speculation, and grain futures prices', *American Journal of Agricultural Economics* **97**(1), 65–87.

Gürkaynak, R. S. (2008), 'Econometric tests of asset price bubbles: taking stock', *Journal of Economic Surveys* **22**(1), 166–186.

Gutierrez, L. (2013), 'Speculative bubbles in agricultural commodity markets', *European Review of Agricultural Economics* **40**(2), 217–238.

Holm, S. (1979), 'A simple sequentially rejective multiple test procedure', *Scandinavian journal of statistics* pp. 65–70.

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

Hommel, G. (1988), 'A stagewise rejective multiple test procedure based on a modified bonferroni test', *Biometrika* **75**(2), 383–386.

Meyer, S. & Trede, M. (2016), Explosive earnings dynamics: Whoever has will be given more, Technical report, Center for Quantitative Economics (CQE), University of Muenster.

Paulson, N., Liu, X., Filler, G. & Odening, M. (2013), 'Testing for speculative bubbles in agricultural commodity prices: a regime switching approach', *Agricultural Finance Review* **73**(1), 179–200.

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

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

Rüger, B. (1978), 'Das maximale signifikanzniveau des tests: lehne h_0 ab, wenn k unter n gegebenen tests zur ablehnung führen', *Metrika* **25**(1), 171–178.

Simes, R. J. (1986), 'An improved bonferroni procedure for multiple tests of significance', *Biometrika* **73**(3), 751–754.