

New evidence for explosive behavior of commodity prices

Jeanne Diesteldorf, Sarah Meyer and Jan Voelzke[†]

50/2016

[†] Department of Economics, University of Münster, Germany

New evidence for explosive behaviour of commodity prices

Jeanne Diesteldorf · Sarah Meyer ·
Jan Voelzke

February 7, 2017

Abstract A number of papers have empirically investigated explosive behaviour of commodity futures prices by examining univariate time series. However, these individual tests ignore the multiple testing nature of the problem. To address this issue, we apply the Generalized Sup-ADF test by Phillips et al. (2015) on a panel data set of the ten most liquid commodity futures traded in the US over the past 35 years. We aggregate individual test results using a stagewise rejective multiple test procedure. Our results yield evidence for several periods of explosive behaviour in the markets for wheat, cattle, cocoa, coffee and cotton.

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

JEL Classification: C22 · C33 · G14 · Q14

J. Diesteldorf, Corresponding Author
University of Muenster
Department of Economics
Am Stadtgraben 9
D - 48143 Muenster
Telephone: +49 251 8325001
E-mail: jeanne.diesteldorf@wiwi.uni-muenster.de

S. Meyer
University of Muenster
Department of Economics
E-mail: mail@smeyer.org

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. There are several potential explanations based on changes in the fundamentals, such as growth in emerging markets, weather shocks, declining inventory or consumption growth. Nevertheless, since reliable data on fundamentals is hard to generate, neither a proof for the responsibility of these fundamental explanations, nor an unanimous quantification of their impact has successfully been done yet; see Adämmer & Bohl (2015).

When addressing this topic, one needs to carefully differentiate between empirical tests that examine speculative bubbles and those that focus on explosive behaviour of the underlying time series data; see Brooks et al. (2015) for a detailed discussion. The discourse is fostered by the theoretical debate on whether an econometric model needs a bubble component to explain empirically observed behaviour of price time series. Gürkaynak (2008) gives a vast review on the topic.

In particular, the test performed in the present paper investigates explosive behaviour. We do not model or measure the fundamental value which would be necessary for an examination of the presence of speculative bubbles.

A larger body of the literature uses the Sup-ADF test (Phillips et al. (2011)) to investigate explosive behaviour and/ or the existence of speculative bubbles in commodity futures markets.¹ Most papers focus on US commodity markets, reporting mixed evidence. However, all authors agree upon the fact that explosive behaviour is present in one way or another in some commodity futures markets under scrutiny; see Gutierrez (2013), Areal et al. (2014), Etienne et al. (2014), and Etienne et al. (2015).

However, this widely used approach entails an important econometric shortcoming so far ignored in the extant literature. Most empirical studies apply the same test on several individual time series and proceed to interpret the results and significance statement as if they were single tests. Yet, 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

¹ In contrast, Adämmer & Bohl (2015) employ the momentum threshold autoregressive (MTAR) approach to test for speculative bubbles in US corn, soybean and wheat prices. Paulson et al. (2013) apply a regime switching regression model to test for periodically collapsing bubbles in the US markets for wheat, corn, soybeans, rice, sugar and cotton.

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 the present investigation, we will rely on the structure and hypotheses of Simes (1986), with all their assumptions and implications, but also follow Hommel (1988) and use his advancement, allowing to build and test arbitrary subset-hypotheses.

Our paper conducts a panel test approach jointly investigating explosive behaviour in ten different agricultural commodity price time series. We thereby circumvent the drawback outlined above and employ a procedure that is statistically correct. We aim at detecting market exuberance in commodity futures prices, albeit remaining outside the discussion of the general necessity of a specific bubble model. To do so, we employ the Generalized Sup-Augmented Dickey Fuller test (GSADF) by Phillips et al. (2015). Already, the Sup-ADF test (Phillips et al. (2011)) proved to perform satisfactorily well.² Nevertheless, it has reduced power for long(er) time series, for those with multiple bubbles and/ or if employed on rapidly changing markets. The GSADF-test cures these shortcomings. Moreover, the procedure is robust concerning the sources of market exuberance, e.g. mildly explosive behaviour induced by changing fundamentals.³

The testing procedure can also be used to time-stamp phases of explosive behaviour, but, in the first place, delivers a classical test-statistic for the entire time series. This very value can then be compared to the corresponding critical value. Alternatively, 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 those in the existing literature. Finally, Section 4 concludes.

2 Method and Data

We examine explosive behaviour 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 behaviour. Hence, the underlying reduced form empirical model can be stated as:

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

with $\varepsilon_{i,t} \stackrel{iid}{\sim} N(0, \sigma_i^2)$. For each commodity i , 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

² Cp. e.g. Homm & Breitung (2012).

³ Cp. Phillips et al. (2015), p. 1045.

calculates the supremum over all ADF test statistics for all possible starting points and interval lengths that lie within the time period examined.⁴

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.⁵ For the calculation of the p -values, we simulate 20 000 realizations of the test statistic under the null hypothesis.⁶

Simply rejecting the null hypotheses for all 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 $H_{i,0}$, $i = 1, \dots, N$ are true. The main advantage, making this method very feasible, is that it only needs the p -values of the N time series tests as input. 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.

Subsequent to the identification of the time series exhibiting explosive behaviour using the Hommel procedure, we could use the provided backwards SADF sequence to date-stamp the periods of explosiveness. 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.⁷

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

⁴ For a detailed description see Phillips et al. (2015).

⁵ Further parameters $d = 1$ and $\eta = 1$ are set in accordance to the literature.

⁶ Note, however, that we just used 2 000 realizations as a basis for the provided graphs in Figure 1-3.

⁷ 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.

- grains: corn, cbot⁸ wheat (cbwheat), kcbt⁹ 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, 2015.¹⁰ All data are taken from Thomson Reuters Datastream.

We investigate perpetual series of commodity futures prices. They start at the nearest contract month, delivering the first price value for the continuous series 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 dashed line illustrates real commodity prices called `NORMCOMMODITYNAME`, respectively. Further, the solid line paints the ADF statistic sequence and the dotted line shows the corresponding critical value sequence.¹¹ 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 behaviour 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 , depicted by the dotted line.¹²

We find explosive behaviour in the prices for wheat, cotton, feeder cattle, cocoa and coffee. Prices for corn, sugar, live cattle and hogs do not show explosive behaviour. The results for corn prices are particularly interesting. Its

⁸ Chicago Board of Trade.

⁹ Kansas City Board of Trade.

¹⁰ 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.

¹¹ Cp. Caspi et al. (2014), Section 5.

¹² 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.

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 behaviour. 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 error of wrongly concluding that there exists explosive behaviour in the corn market. This emphasizes the necessity of the applied procedure in comparison to conducting the GSADF-test on individual time series and thereby neglecting its multiple testing nature.

While Gutierrez (2013) supports our empirical evidence for explosive prices in the wheat market, he also finds explosive behaviour 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 behaviour.

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

Apparently, our results are in line with what others have found before. Moreover, our empirical evidence underlines the fact that explosive behaviour 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 behaviour in 1995 and 1996; the market for cocoa shows explosive behaviour between 2002 and 2004.

It jumps to the eye that the markets for cocoa, coffee, corn and kwheat exhibit two or even three pronounced periods of explosiveness. The other markets, such as feeder cattle and cotton, show explosive behaviour towards the end of the sample period only. Feeder cattle is the only market characterized by explosive behaviour as late as 2014/2015. It is also the only livestock futures market with explosive price; 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 explosiveness in 2011, but not during the period of increasing financialization beginning in the middle of the last decade. Some markets exhibit explosive behaviour long before: wheat and corn in 1996 (both grains), coffee a little earlier in later 1994.

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

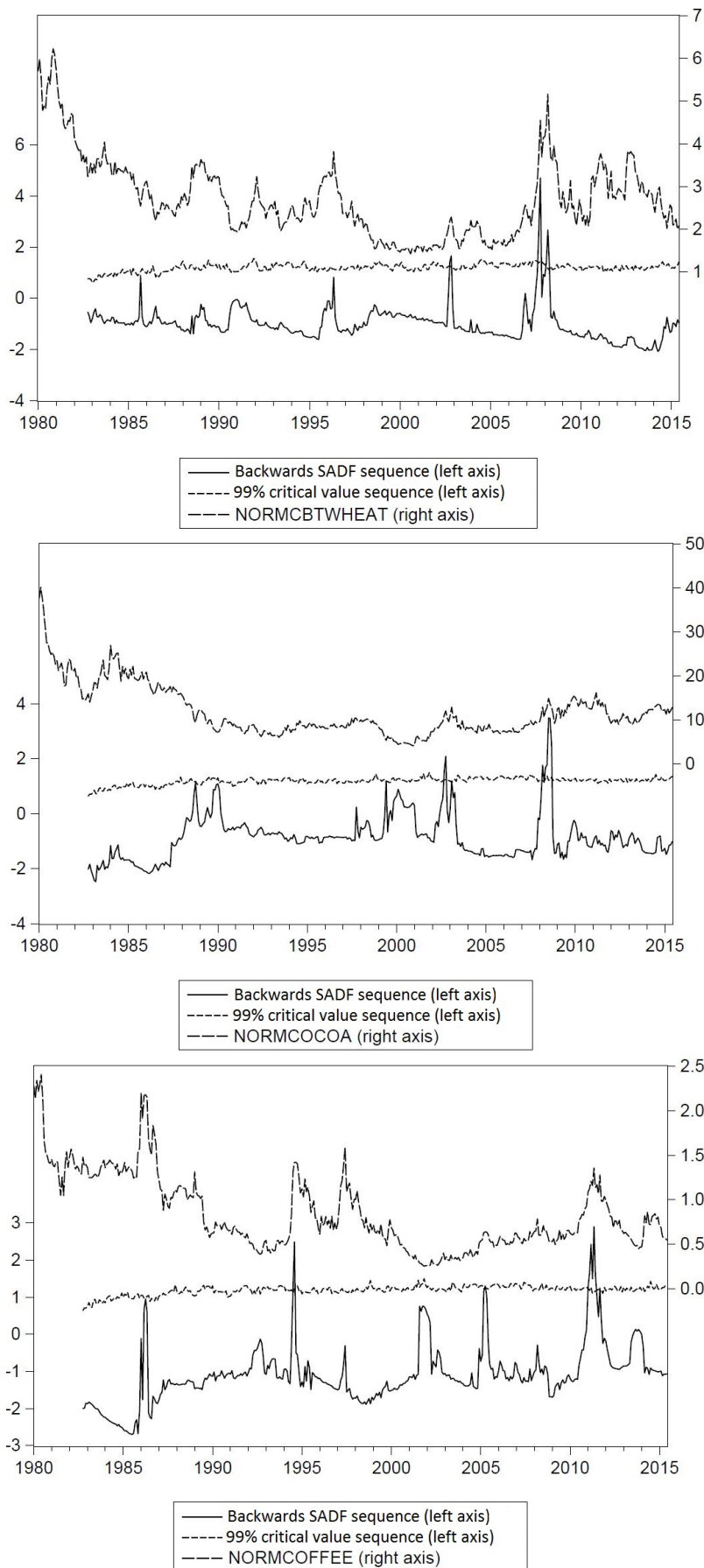


Fig. 1 Normalized prices, the backwards SADF sequence and the critical value sequence

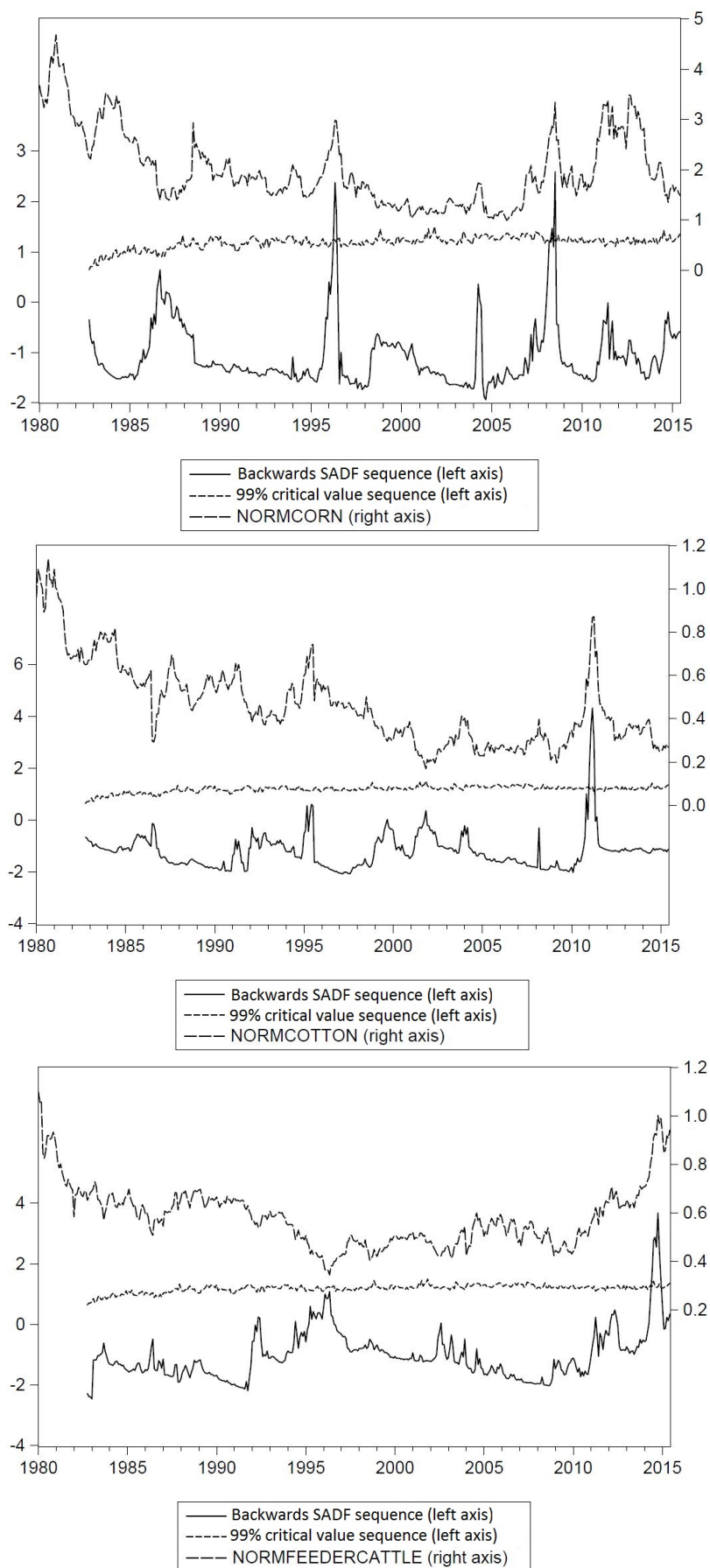


Fig. 2 Normalized prices, the backwards SADF sequence and the critical value sequence

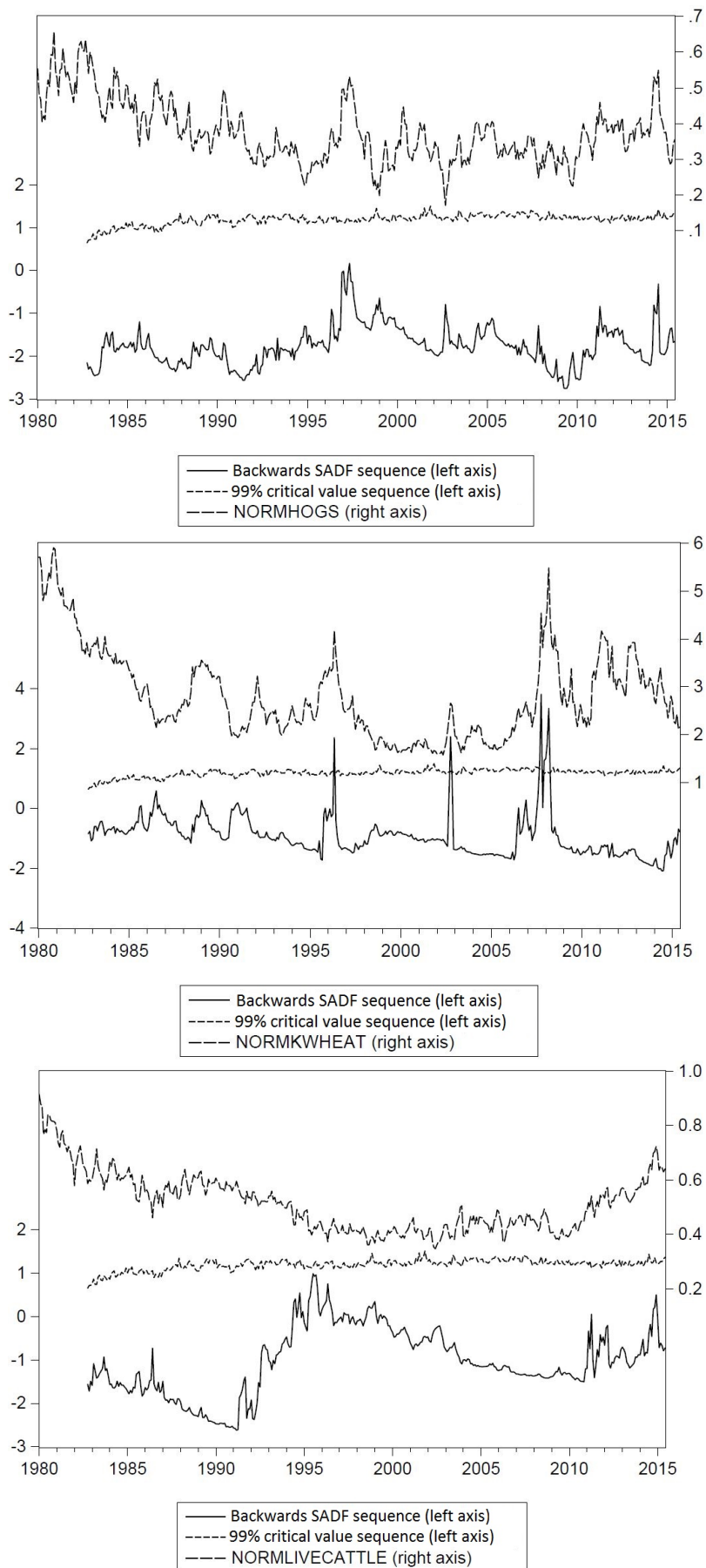


Fig. 3 Normalized prices, the backwards SADF sequence and the critical value sequence

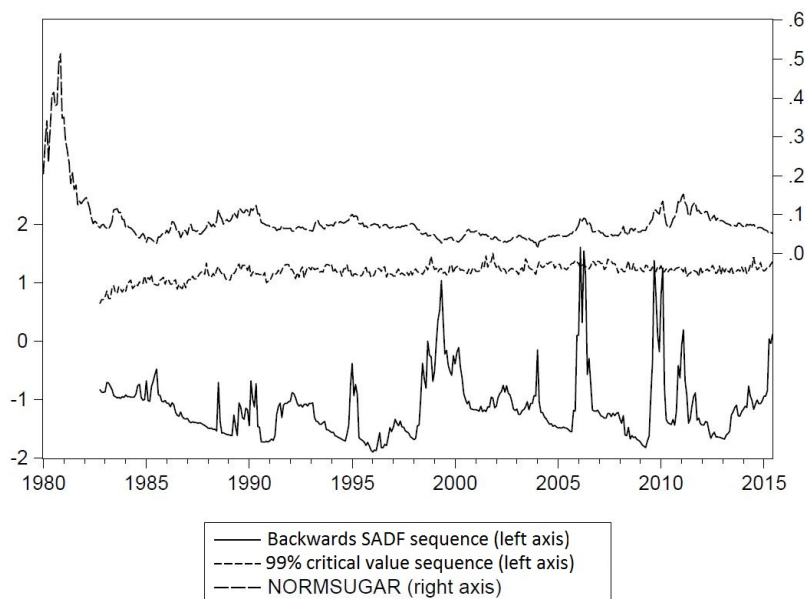


Fig. 4 Normalized prices, the backwards SADF sequence and the critical value sequence

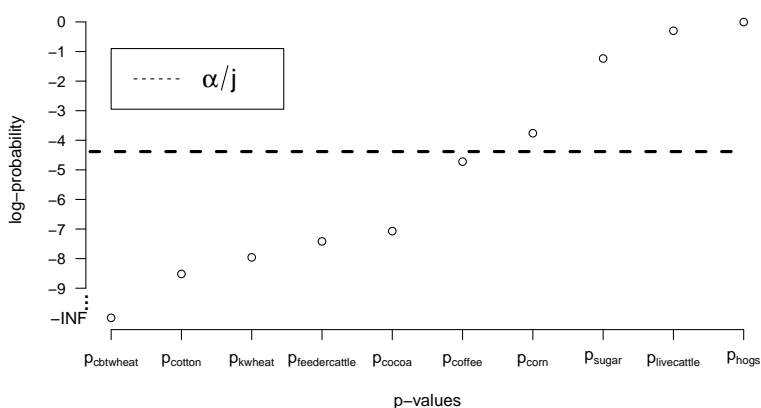


Fig. 5 Logarithmized p -values and the cut-off calculated as in Hommel (1988). Commodities with a p -value beneath the line show significant explosive behaviour.

4 Conclusion

The present paper examines explosive behaviour in US agricultural commodity futures markets. To this end, we employ a panel data set of the ten most liquid markets to conduct the GSADF test. We aggregate test results for individual

commodities using the Hommel (1988) procedure to control the Familywise Error Rate. Thereby, we get a statistically precise overall statement, instead of several separate statements about single commodities.

Our results yield empirical evidence for explosive behaviour in the market for wheat, feeder cattle, cocoa, coffee and cotton. No explosive behaviour 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 explosive behaviour in the markets for coffee, corn and wheat futures long before the extraordinary price spikes caused by an increasing financialization of agricultural derivative markets.

Most importantly, we appreciate the multiple testing nature of the empirical problem by employing a panel approach and thereby address a methodological shortcoming so far ignored in the literature. Note, however, that we are unable to draw any conclusion about possible causes of explosive behaviour exhibited by the data. 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 possible causes of 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 Paper* **14**.
- Brooks, C., Prokopczuk, M. & Wu, Y. (2015), 'Booms and busts in commodity markets: Bubbles or fundamentals?', *Journal of Futures Markets* **35**(10), 916–938.
- Caspi, I. et al. (2014), Rtaadf: 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.