Journal of Commodity Markets 31 (2023) 100346

Contents lists available at ScienceDirect

Journal of Commodity Markets

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

Regular article

Wheat price volatility regimes over 140 years: An analysis of daily price ranges☆

Marco Haase[1], Heinz Zimmermann[1,*], Matthias Huss[1]

University of Basel, Switzerland

ARTICLE INFO

JEL classification:
E30
C58
G13
N21
N51
Q02

Keywords:
Commodity futures volatility
Wheat futures
Historical price analysis
Structural volatility breaks

ABSTRACT

We analyze Chicago based daily wheat price volatility over more than 140 years using a novel data set of daily high and low futures prices starting in 1877. We identify five long-run regimes and find that volatility shifts between regimes are statistically more pronounced than fluctuations within regimes, even when conditioning on economic states. Historical volatility estimates derived from average commodity price data, a common practice in empirical studies, exhibit a regime-dependent upward bias between 0% and 22%. The magnitude of the bias and the importance of regimes potentially explain contradictory findings on volatility patterns in earlier studies.

1. Introduction

This paper takes an innovative look at commodity price volatility by exploring a new data set based on daily high and low price data of nearby and second-nearby wheat futures prices covering a time span of more than 140 years. The long data history allows us to compare the magnitude of typical short-run shifts in volatility within long-run volatility regimes to volatility shifts between regimes. Using Bai–Perron structural break tests applied to range based volatility estimates, a total of five volatility regimes can be identified.

Commodity price volatility is largely undiversifiable for many economic agents. It is directly related to aggregate welfare effects and economic costs as discussed by Gilbert and Morgan (2010) or Jacks et al. (2011). Is therefore not surprising that the study of commodity price volatility has a long historical tradition, with early papers dating back to the beginning of the 20th century (e.g., Mayo-Smith (1900)).

Further, it is well established that volatility is time-varying and clustered, meaning that occurrence of high volatility is likely to be followed by subsequent periods of above-average volatility. Commodity price volatility makes no exception from this observation. Extended periods of high food price volatility are particularly concerning given their immediate effects on the food security and livelihoods of large parts of the world's population. These periods recurringly spark public interest in the topic, and trigger political debates on the implementation of regulatory frameworks designed to counteract "abnormally high" volatilities. A prominent example

☆ We are grateful to the Swiss National Science Foundation (SNSF) for financial support under the project "Volatility in early commodity markets" (grant number 100018-172681). The comments by an anonymous referee and an associate editor have greatly improved the paper.

* Correspondence to: Department of Finance, Faculty of Business and Economics, University of Basel, Peter Merian Weg 6, CH-4002 Basel, Switzerland.
E-mail addresses: marco.haase@unibas.ch (M. Haase), heinz.zimmermann@unibas.ch (H. Zimmermann), matthias.huss@unibas.ch (M. Huss).

[1] Department of Finance, Faculty of Business and Economics, University of Basel, Peter Merian Weg 6, CH-4002 Basel, Switzerland.

https://doi.org/10.1016/j.jcomm.2023.100346
Received 31 December 2022; Received in revised form 9 June 2023; Accepted 26 June 2023
Available online 5 July 2023
2405-8513/

is the United Nations (UN) 2030 Agenda for Sustainable Development, which defines adopting "measures to ensure the proper functioning of food commodity markets and their derivatives [...] in order to help limit extreme food price volatility" (UN General Assembly, 2015, p. 16) as the central target (2.c).

Yet, little is known about the level and patterns of commodity volatilities over very long observation periods (e.g., decades or centuries), and how these possibly changed over time. One of the reasons for this knowledge gap is that the existing literature has placed surprisingly little effort in quality data. This is remarkable, since such data had been recorded – and thus was potentially available – already from the very first days of organized commodity exchanges (in the 19th century), although certainly not in readily machine readable form.

The typical empirical papers addressing commodity price volatility use either short observation frequencies (e.g., daily or even intraday price data) over relatively short sample periods, or longer spaced (e.g., annual, quarterly or monthly) price data for longer observation horizons, which impedes the comparability of recent to historical results across studies. These shortcomings are amplified by the fact that the historical commodity literature typically uses macroeconomic price data (e.g., published by the National Bureau of Economic Research, NBER, or the International Monetary Fund, IMF), where it is common, or even standard, to average prices over a particular month, quarter, or harvest year. While such data is instructive for the analysis of long-run price variability, it is an insufficient basis for the estimation of *volatility*, since the efficiency of volatility estimates substantially increases with the observation frequency of the underlying data.

This paper is the first to use consistent point estimates for daily wheat price volatility over an observation period of more than 140 years. We focus on wheat, as wheat always was, and is still one of the most important food grain sources for human consumption. Our estimates are derived from the daily high and low prices of wheat futures contracts, which have persistently been traded at the Chicago Board of Trade (CBOT), and its predecessors, since the start of our sample period in 1877. Wheat futures traded on the CBOT were the most important commodity contracts in the mid 19th century and remain the most important wheat contracts traded worldwide. The unique role of CBOT wheat futures prices, and their importance for global markets, already a hundred years ago, is reflected in the following quote from Dies (1925, p. 24):

> "Chicago is the clearing house of the world wheat trade. Accordingly vast quantities of wheat that never were intended for shipment to Chicago are hedged in that market. It is the price insurance center, used constantly for insurance purposes by dealers of Europe, Argentina, Australia, Canada, and a score of American interior and export cities".

The CBOT, today, forms a part of the Chicago Mercantile Exchange (CME), the world's largest commodity exchange, which is unparalleled in terms of liquidity and price discovery. The nature of our data allows for a coherent analysis of recent, and historical volatility patterns. The major contribution of our paper is twofold.

First, we endogenously estimate volatility regimes. We document significant heterogeneity in volatility across the five regimes we identify from our time series of daily wheat price volatility. Specifically, we find that the volatility differences between regimes are substantially higher than volatility differences within regimes, even when we condition our analysis according to state variables commonly used in the literature, such as the business cycle, the convenience yield (as a measure for scarcity), and inflation rates. Our results suggest that the identified volatility regimes coincide with monetary policy regimes.

Second, we find that historical volatility estimates derived from averaged price data – as is typically done in the earlier literature, often at a monthly or even lower frequency – are upward biased. The bias is approximately 11 percent over our full observation period, and is again time varying. While we find the highest bias (22 percent) in the regime with the lowest volatility level, the bias is negligible (even slightly negative) in the most recent regime.

Third, the volatility as measured from regime to regime does not appear to decrease over the most recent decades[2] as claimed in much quoted papers like Gilbert and Morgan (2010) or Jacks et al. (2011). Although the pattern is also evident in the averaged price data, thereby confirming the importance of regimes for these data as well, the picture is clearer for the range based data. The regime with by far the largest volatility – both in terms of volatility from daily range data and average daily volatility – is the most recent from 1996 to 2018.

Our results have important implications for inferences drawn from the comparison of recent to historical volatilities along several dimensions. Most importantly, our results suggest that a mere comparison of volatility estimates across studies (e.g., recently published estimates in comparison to results reported in earlier studies) can be severely misleading if the data sets used for the analysis are inconsistent in terms of observation frequency or filters applied to the underlying data (e.g., averaging, smoothing, detrending).

Further, the comparison of volatilities estimated from rolling windows, by splitting samples in equidistant periods, such as half-point, or decades (as done in e.g., Gilbert and Morgan (2010)) necessarily results in the analysis of weighted averages of volatility levels observed in different regimes. The weight of each regime (i.e., the number of observations considered in a specific regime) can be considered a random variable itself.

While this may only command for particular caution in the interpretation of a descriptive analysis, the neglect of regimes can lead to a severe misinterpretation of the results when volatilities are correlated with exogenous variables: Since regimes appear to dominate the influence of other state variables on volatility by far, conclusions drawn from conditioning volatility on (or correlate volatility with) such factors can be heavily erroneous, if the analysis does not control for the related regimes.

[2] It should be noted that the temporal reference point of this study always refers to the end of the sample period, which is 2018.

Since the focus of our paper is exclusively on wheat futures contracts, external validity is a general limitation of our study. Whether the results presented in this paper can be generalized to other agricultural goods, or other commodities in general, remains a question for future research. This limitation notwithstanding, our paper raises a caution flag on (published) interpretations of volatility "trends", "shifts" and other "systematic" changes in volatility levels that stem from the comparison of inconsistent, or inefficient data.

2. Commodity price volatility: Some critical remarks about data used in the literature

At first glance it may seem surprising that the use of adequate data in empirical volatility estimation is worth emphasizing in the context of academic research. However, in the case of commodities, this is indeed an issue. Surprisingly, it is just in the field of commodity markets where the use of average prices or moving averages common in the analysis of volatilities. Sometimes, even both is done at the same time. For example, Sumner (2009) studies the variability of US corn and wheat prices since 1869 by using "the percentage deviations of the real price of corn from the three-year moving average of past prices" (p. 1252), where "prices are marketing year average prices" (p. 1253).

Unlike in asset markets, average commodity prices are often regarded as more "representative" than end-of-period prices, despite their undesirable statistical or even distorting properties, briefly addressed below. The use of long-run averages is also motivated by alleged short-run excess volatilities caused by small traders; averaging is supposed to give a clearer picture of the economic fundamentals that determine price changes.[3] While average price data may be sufficient for characterizing representative price levels or general price trends, they are an incomplete basis for studying price volatility, and how volatility changed over time.

The effects of time averaging of random time series on biases and misinterpretations of the resulting series or, more generally, the use of linear filters for smoothing purposes or detrending, have long been studied in statistics and economics. In this context, two effects are emphasized: the bias of dispersion measures, and a bias towards temporal dependency in the averaged or filtered series which does not exist in the original data. Among the prominent early papers in this field, Slutzky (1937) demonstrated that the "summation of random causes" can produce cyclical patterns indistinguishable from observed business cycles. This is closely related to the discussion of real business cycles in the 1980s. For example, Nelson and Kang (1981) noticed that inappropriate detrending can lead to spurious cycles in the detrended series. In the commodity literature, Working (1960) showed analytically that first differences of temporal averages of a random walk process exhibit positive autocorrelation, which is moreover directly related to the variance of the first difference of the basic, non-averaged series.

It is therefore surprising that these effects have not been sufficiently recognized in the commodity literature, especially with respect to the estimation of volatility and cyclical properties of prices and returns. Ironically, commodity exchanges in the US and Europe have recorded annual, quarterly, monthly, weekly and even daily high and low price quotes in their statistical publications since their earliest days, so that price ranges as indicators of volatility are available over a very long time horizon. Since this data has been publicly available – although not in readily accessible electronic format – it is surprising that empirical studies on volatility issues do not exploit this important information. More generally, it is surprising how little the existing studies are concerned with a careful description of their (and their peers') data.

This criticism can be illustrated by examining the few papers that analyze the behavior of commodity price volatility over really long-run time horizons. Calvo-Gonzalez et al. (2010) give a detailed review of recent papers on changes in commodity price volatility over time (pp. 3–6), but they provide no information what "monthly" prices, "annual" data means in the various studies. Their own empirical investigation is based on "a monthly unbalanced panel of observations from Global Financial Data (GFD) covering the period 1784–2009", but they leave nature of the data uncommented. Unfortunately, the publicly available information about these data does not provide the necessary information either.

More clarity is revealed in a study of Cashin and McDermott (2002) who find "small trends and big variability" in the Economist's index of industrial commodity prices from 1862. These are annual index data constructed as averages of monthly price observations.[4] They find substantial evidence of increasing price volatility over their sample period, with two major breaks at the beginning of the twentieth century and after 1971 (the breakdown of the Bretton Woods system). They attribute the rise in volatility in the early 1900s to greater amplitude of price movements, the further rise in 1971 to the increased frequency of large price movements, i.e. to a fall in the duration of large price cycles. Our results give a more detailed picture of these findings.

The duration and magnitude of cycles for 36 individual commodity prices is analyzed in Cashin et al. (2002), although for a shorter time period from 1957 to 1999. The authors use monthly data from IMF's International Financial Statistics (IFS) database. They find a pronounced asymmetry in commodity price cycles with price slumps lasting longer than price booms. A major drawback of their study, however, is that their data is a mixture of "1/average of daily quotations, 2/average of weekly quotations, 3/monthly quotations" (p. 296), which hardly allows an unbiased identification of cycles (averaging implies cyclical effects in time series unrelated to cycles in the underlying data), and limits a meaningful comparison between the individual commodities.

A further study which addresses changing commodity price volatility over several decades is Gilbert and Morgan (2010). The authors analyze whether the increase in food price volatility in 2008 should be regarded as permanent or not, given that "the historical experience has generally been that periods of high volatility have been relatively short and interspaced with longer periods

[3] See Irwin (1936) for an early statement of this common argument.

[4] Specifically, "for 1862–1910, the annual data are formed as an average of price observations in January and July of each year; from 1911 onward, the annual data are an average of observations for all months" (p. 183).

of market tranquility" (p. 3023). They find that opposite to the general claim of accelerating food price volatility over time, it is lower between 1990 and 2009 compared to the two preceding decades. The authors state that their analysis is based on "standard deviations of logarithmic changes in monthly average real US dollar prices at an annual rate" (p. 3025),[5] and is based on commodity price indices, not actual market prices, raising the question whether their conclusion also holds for non-averaged data.

Similarly Jacks et al. (2011) analyze monthly, quarterly and annual (log) price changes of a broad set of commodities, and conclude that commodity prices have not become more volatile over the period 1720 to 2008. Their analysis uses more than 500 individual time series drawn from several dozens of data sources. From this data universe, the authors construct one time series, to identify structural changes over time. While the focus on a large array of commodities is a major advantage of their study, it remains unclear how the conclusion is affected by the construction of the time series, which necessarily combines (i) prices in various locations and for diverse quality standards, (ii) prices before and after taxes, duties and delivery costs, or (iii) differences in price discovery across (centralized and decentralized) market places. Their volatility is "calculated as the unweighted standard deviation of period-over-period percentage changes in price across all commodities" (p. 805).

3. Data and descriptive statistics

The continuing important role of wheat futures and the constancy of the institutional setting, including the recording of price data, is an essential improvement to data used in the earlier studies, which makes our data set highly reliable and relevant for studying commodity price volatility.

Our analysis is based on a new, unique and hand-collected data set which consists of daily high and low prices for Chicago wheat futures contracts, starting in 1877. The data is manually extracted from the "Annual Reports of the Trade and Commerce of Chicago" (today, the Chicago Board of Trade, CBOT, which is part of the CME group). These reports have been published annually in print, in December, from 1877 until 1921. Specifically, the data set covers spot prices, as well as price quotations for futures contracts with different maturities. To fill the gap between our proprietary sample (1877 to 1921) and the recent data, we have purchased an additional data set from CBOT that contains raw data from 1922 through 1970. These data have been electronically recorded by CBOT. However, the data quality required additional work to ensure clean and high quality data. In particular, machine scripts are used to identify data errors, such as missing comma separators, decimal errors or abnormal values. The more recent data (1970 to 2018) are obtained from Datastream with conventional data quality. It must be noticed that trading in wheat futures at Chicago was suspended from August 25, 1917 to July 15, 1920.

Based on this raw data set, we construct time series of high and low prices based on daily and monthly frequencies. Data are available for two futures maturities, the nearby month (denoted by $F1$ in this paper) and the subsequent month (denoted by $F2$). In most of our tests, while displaying the results for both maturities, we focus on the nearby results and use those of the second maturity as robustness test. Returns on futures contracts are constructed using the conventional roll-over methodology which we have used in other papers.

Overall, our full sample period includes 37,165 vs. 37,388 daily price data for F1 and F2, respectively, from January 3, 1877 to February 28, 2018.

Volatility measures The focus of the subsequent analysis is on volatilities computed from our ranges, not the range data itself. We use the Parkinson (1980) estimator for estimating the variance from daily high and low data, which is given by

$$\hat{\sigma}_t^2 = \frac{1}{4\ln 2}[h_t - l_t]^2, \text{ where } \quad h_t = \ln(high_t), \text{ and } l_t = \ln(low_t) \tag{1}$$

for a single observation in the time interval t, and respectively, for T observations in the time interval t to $t+T$ by

$$\hat{\sigma}_{t,t+T}^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{\sigma}_t^2 \tag{2}$$

The estimator has a well-known bias if the underlying price data have a trend. This bias could be corrected if open and close data would be available, which is not the case for our data over the entire time period.[6] Since our major focus is on volatilities estimated from daily ranges the expected bias is small: on a daily basis, the price drift is small for reasonable annual drift rates. Moreover, we are analyzing futures prices for which the expected return is a pure risk premium, which is closer to zero than expected returns which include the rate of interest. Brandt and Kinlay (2005) find in simulations that, the overall efficiency of the Parkinson estimator is not substantially affected by the drift term compared to common alternatives. In empirical testing with stock index returns, the Parkinson estimators even outperforms all of the other estimators in terms of bias, MAD (mean absolute deviation) and MAPE (mean absolute percentage error).

Moreover, since the use of average price data is common practice in the existing literature, we also construct average monthly prices from daily mid-prices and compute their monthly returns by taking first differences. The derived volatility is referred to as "return volatility of monthly price averages". Daily mid-prices are defined by

$$p_t^{mid} = \frac{h_t + l_t}{2} \tag{3}$$

[5] The paper leaves open whether the prices reflect monthly averages of shorter term data, or annual data based on monthly averages.

[6] The respective corrected formulas rely on the Rogers and Satchell (1991) and Yang and Zhang (2000) estimators.

We use daily mid-prices throughout the entire time span for reasons of consistency, because a complete set of closing prices is only available since the 1920s.[7] For constructing mid prices, contract breaks are taken into account to avoid biases that otherwise occur if mid prices are not adjusted for the price gap between the nearby and the second nearby contracts in months where the contract expires or is rolled into the next maturity.[8] To obtain a continuous time series, contracts must be rolled from one maturity to the next. For consistency, both contracts, F1 and F2, are rolled into the next available maturity on the last trading day on which the shortest contract (F1) expires.

Finally, to improve comparability, all volatility estimates are displayed on a daily basis. The transformation of monthly volatilities is done with the effective number of trading days in each month using the square root of time.

A final remark is related to the number of weekly trading days which went from six to five following the dropping of Saturday trading sessions. The last Saturday on which trading took officially place was July 26, 1952.[9] Thus, the non-trading time from Friday to Monday extends over a longer calendar day span after that switch which could have a positive impact on the estimated volatility. If the point is valid, this should be reflected in an augmented range based volatility on Monday in the period after the switch. In order to check the magnitude of that effect, we compare the range-based volatilities on Mondays and Saturdays surrounding 100 weekends before that institutional switch with the volatilities on Mondays and Fridays covering the same number of weekends after. Prior to the change, the mean volatility for the nearby maturity (F1) on Saturday is 0.368% compared to 0.489% on Monday, an increase of 33.0%; the equivalent values after the switch are 0.545% for Friday and 0.689% for Monday, an increase of 26.5%. Thus, while the average volatility is larger on Monday than on Friday before and after the switch, the gap is even slightly smaller afterwards. These figures should be seen in the context of the overall mean daily volatility of 0.550% in that regime, and a standard deviation 0.379%; see Table 1, Panel A (columns 5 and 7). We therefore do not expect a bias in estimated volatilities from this institutional adjustment.

Summing up, two volatility measures are subsequently analyzed, which are based on

- daily log price ranges, abbreviated by $\sigma^{d,m}_{rng}$;
- monthly log returns based on log price averages, abbreviated by $\sigma^{m}_{ret(avg)}$.

Range based volatility: A first look Fig. 1 display the time series of daily volatilities extracted from daily log ranges, for the nearby (F1) and second nearby (F2) futures maturity. The observed pattern is similar for the two maturities. While visual inspection may not immediately allow for a conclusion whether overall futures price volatility increased over time (tested in the subsequent section), it is apparent that the number and duration of "spikes" (large temporary volatility jumps) was larger in the first part of the sample period. In this context, Markham (2015) for example states that

> "the decade of the 1880s was particularly infamous for manipulations. In 1881, there was, again, nearly a corner a month on the CBOT. (...) The wheat market was cornered four times in 1882, along with corners in lard, ribs, and corn". (pp. 4–5)

Descriptive statistics The distributional characteristics of the range based point estimates of daily volatility of the first (F1) and the second maturity (F2) are reported in Table 1. For an easier comparison, we report the first moment of return volatility of monthly price averages in column (1) of the table.

The following observations emerge from the Table: First, over the full sample period, volatility estimates from returns of monthly price averages are 1.32% (F1, Panel A) and 1.26% (F2, Panel B), and thus higher than the daily range based estimates 1.19% (F1) and 1.14% (F2).

Second, the volatility pattern of the second maturity (F2) is not fundamentally different from the first maturity (F1), but the coefficients are consistently lower. For example, the range based volatility of F1 for the full sample period is 1.19% compared to 1.14% for F2, which is a difference of 4.4%; this is in line with the Samuelson (1965) maturity effect which suggests that futures price volatility should increase as contracts approach maturity. The same picture emerges for the average daily range based volatilities which are 0.97% for F1 and 0.93% for F2.

4. Structural breaks

Computing unconditional volatilities over a time span of more than 140 years is at best descriptive. Changing volatilities can be addressed by either identifying structural breaks or estimating conditional models. We select a combination: we identify structural breaks following the approach suggested by Bai and Perron (1998) and Bai and Perron (2003), and compare the distribution of range based volatilities (in Section 5) conditional on economic states for each of the identified regimes.

The Bai–Perron test methodology is the standard approach for testing the presence and dating of structural breaks, specifically multiple unknown breakpoints, in linear models and is routinely implemented in major statical software packages. Unlike other approaches, e.g. Bayesian models in the spirit of Pesaran et al. (2006) aimed at estimating and forecasting time series with multiple

7 While closing prices are recorded since February 1900, the data are incomplete and not suitable for empirical analysis.

8 Such bias clearly depends on the shape and steepness of the term structure, also known as the convenience yield, and is time varying.

9 See, e.g., the announcement of the vote of the Board of Trade in the New York Times on April 11, 1952 (p. 35), and an article covering the rollout of the decision in the New York Times on July 18, 1952, on the front page. Note that there were ten weeks without Saturday trading in the previous calendar year as well.

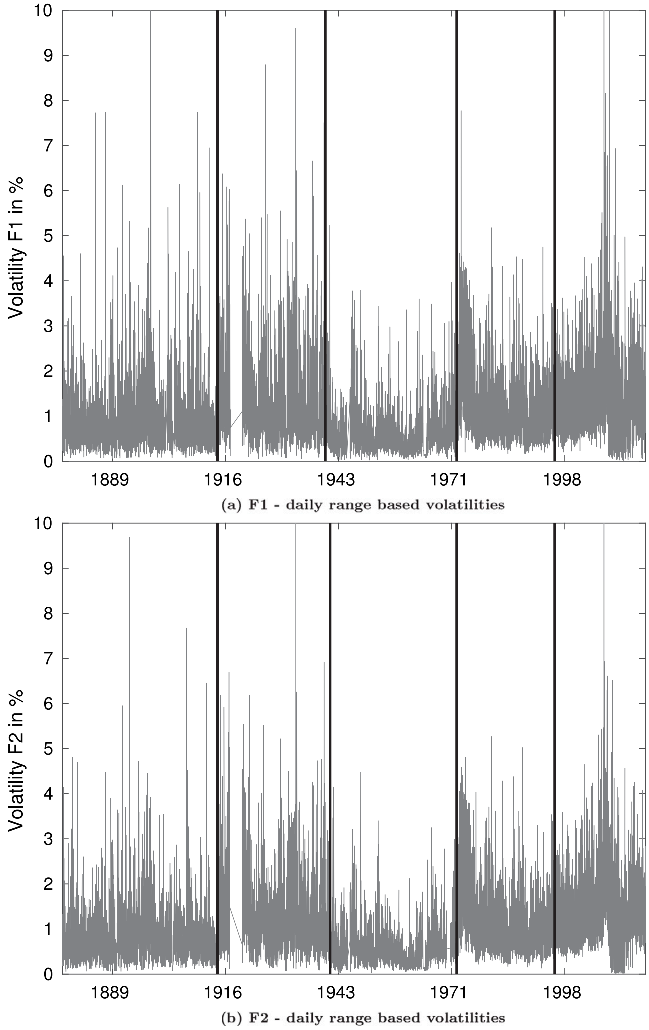


Fig. 1. Wheat futures volatility: Breakpoints. The figure displays the time series of daily volatilities extracted from daily and log ranges, for the nearby (F1)and second nearby (F2) futures maturity. To allow comparison, all volatility estimates are displayed on a daily basis. The time series includes 37,164 daily observations from January 3, 1877 to February 28, 2018. Break points are estimated using the methodology of Bai and Perron (2003). The number of break points is selected by the BIC criteria, where the maximum number of breaks is set to 5 (default). Structural breaks are displayed as vertical black lines.

structural breaks, the Bai–Perron test is based on a linear regression model with regime-dependent parameters estimated by least-squares.[10] The global minimum of the sum of squared residuals (SSR) is achieved by applying a dynamic programming algorithm and estimating the parameters sequentially. The test allows either all parameters, or a subset of parameters, to break simultaneously (pure or partial structural change). Since we only estimate a single parameter to break, the intercept, this distinction is not relevant in our subsequent tests. The model tests the null of k breaks against a fixed number of breaks (which can be zero), or k breaks against $k + 1$. We chose the second approach; as recommended by Bai and Perron (2003) the number of breaks is determined using the BIC criterion; we set the maximum number of breaks to five.[11] The approach also allows to estimate confidence intervals for break dates; they can be asymmetric around the break date if the underlying process is autoregressive, which is an important feature in our context due to the well-known volatility clustering.

[10] While the test is about structural breaks in linear regression parameters, it is based on minimizing the sum of squared residuals and can be easily applied to structural breaks in the average of a data series by fitting a constant to the data; see Zeileis et al. (2003), Section 4.1.

[11] Allowing for a higher number of breaks leads to the same major breaks but a finer subdivision of the here reported regimes. Confidence intervals for those sub-regimes are wide, however.

Table 1
Wheat futures volatility: Distribution of range based estimates.

Observation period	$\sigma^m_{ret(avg)}$ (1)	σ^d_{rng} (2)	min (3)	median (4)	mean (5)	max (6)	std (7)	skew (8)	kurt (9)	ρ_1 (10)	ρ_3 (11)	ρ_{20} (12)	bias (13)
Panel A: F1													
03.01.1877–28.02.2018	1.3213	1.1862	0.0178	0.8061	0.9696	11.3968	0.6833	2.4892	16.6647	0.55***	0.48***	0.38***	11.39%
03.01.1877–24.07.1914	1.1675	1.0082	0.0623	0.7254	0.8463	10.4710	0.5479	3.5673	32.8787	0.46***	0.35***	0.15***	15.80%
25.07.1914–06.09.1940	1.6568	1.4234	0.0527	1.0271	1.1996	9.5999	0.7661	2.1372	12.3787	0.47***	0.38***	0.27***	16.40%
07.09.1940–30.06.1972	0.8137	0.6678	0.0249	0.4614	0.5501	5.2351	0.3786	2.7675	17.3660	0.51***	0.41***	0.29***	21.85%
03.07.1972–28.03.1996	1.4977	1.2871	0.1666	0.9558	1.1178	7.7747	0.6382	1.9019	8.9370	0.47***	0.41***	0.34***	16.36%
29.03.1996–28.02.2018	1.5646	1.6014	0.0178	1.2286	1.3908	11.3968	0.7938	2.6012	18.7836	0.38***	0.33***	0.24***	−2.30%
Panel B: F2													
03.01.1877–28.02.2018	1.2623	1.1361	0.0165	0.7774	0.9331	12.8254	0.6481	2.3905	16.3105	0.58***	0.53***	0.43***	11.11%
03.01.1877–27.07.1914	1.1004	0.9265	0.0564	0.6916	0.7941	9.6892	0.4773	3.1899	30.7151	0.45***	0.37***	0.18***	18.77%
28.07.1914–31.10.1941	1.5648	1.3655	0.0729	0.9954	1.1613	11.1715	0.7183	2.3261	15.2647	0.50***	0.41***	0.29***	14.60%
01.11.1941–28.06.1972	0.7516	0.6155	0.0249	0.4286	0.5098	4.4795	0.3448	2.6799	16.2968	0.54***	0.46***	0.33***	22.11%
29.06.1972–22.03.1996	1.4444	1.2474	0.1330	0.9068	1.0743	5.2657	0.6339	1.8914	7.9839	0.52***	0.46***	0.39***	15.79%
25.03.1996–28.02.2018	1.4927	1.5369	0.0165	1.1998	1.3411	12.8254	0.7507	2.3409	18.3773	0.44***	0.40***	0.31***	−2.88%

The table shows daily return-based volatilities, in percent, calculated from monthly average prices ($\sigma^m_{ret(avg)}$), and daily range based volatilities, calculated from daily high and low prices (σ^d_{rng}), and the distributional characteristics of range based point estimates of daily volatility. All volatilities are displayed on a daily basis; the transformation of monthly volatilities is done with the effective number of trading days in each month using the square root of time. Observation periods include the full sample period and subsamples determined by statistical breakpoints. The coefficients ρ_1, ρ_3 and ρ_{20} denote serial correlation at lags 1, 3 and 20 days. The last column *bias* denotes the relative deviation of (1) to (2). Range based estimates refer to the Parkinson (1980) volatility estimator. F1 refers to the nearby futures maturity, F2 to the second maturity. The sample includes 37,164 daily resp. 1,653 monthly return observations from January 3, 1877 to February 28, 2018.

Our implementation follows broadly the recommendations in Bai and Perron (2003) and Zeileis et al. (2003). We estimate our models using the Matlab computer algorithm provided by Pierre Perron and applying the default parameters suggested there.[12]

Breakpoints In the context of testing for structural breaks, one specific questions is of interest: Is there a difference in the identification of regimes if daily and monthly absolute returns based on monthly average prices are used?

The results are displayed in Table 2. We identify four structural breaks for the first, and the second maturity using daily range based volatilities:

- F1: 24 Jul 1914, 06 Sep 1940, 30 Jun 1972, 28 Mar 1996
- F2: 27 Jul 1914, 31 Oct 1941, 28 Jun 1972, 22 Mar 1996

In general, the structural breaks are close across the two maturities with a difference of only a few days for 3 out of 4 break points. A larger difference occurs at the second breakpoint where the dates differs by more than a year!

If structural breaks are identified from absolute returns based on monthly price averages, only three breakpoint can be identified, and there is a substantial difference to the daily range data in the second breakpoint in the first maturity contract (F1):

- Monthly averages: Apr 1914 (Jun 1914), Aug 1941 (Aug 1941), May 1971 (May 1972)

While the structural breaks are identified without any economic reasoning, three of the breaks seem to be related to historical events that may be regarded as major institutional changes. Break one and three coincide with structural changes in the monetary system of the US:

- The FED opens for business in 1914: While the monetary flexibility changed drastically with abandonment of the gold standard, the effect is hardly visible in e.g. the level of US interest rates. However, Cowing (2015), p. 188, argues that it had a substantial impact on futures prices. This seems to be reflected in our data, although in futures return volatility, not futures price.
- Nixon shock and breakdown of Bretton Woods system (1971–1973): President Nixon temporally suspended the convertibility of the dollar into gold or other reserve assets in 1971, and the system of fixed exchange rates ended in 1973. This time interval includes our third breakpoint.

The second break point appears to coincide with the U.S. Agricultural Adjustment Act of 1941. This observation is substantiated by the fact that the breaks for F1 and F2 react differently (lagged) to the event and its announcement. Wheat prices have been largely controlled by the Commodity Credit Corporation (CCC) since the U.S. Agricultural Adjustment Act of 1938, until the Act was modified in 1941, see USDA (1984).

[12] The code is available through Perron's website https://blogs.bu.edu/perron/codes/ under "Computation and hypothesis testing in models with multiple structural changes"; a trimming parameter of eps1 = 0.15 is selected. Details of the estimation results are provided upon request.

Table 2
Wheat futures volatility: Breakpoints.

95%	90%	Break point	90%	95%
Panel A1: Daily range based volatility, Maturity F1				
12-Jan-1914 (−6.4)	09-Mar-1914 (−4.6)	24-Jul-1914	28-Sep-1914 (2.2)	27-Oct-1914 (3.2)
26-Aug-1940 (−0.4)	30-Aug-1940 (−0.2)	06-Sep-1940	22-Oct-1940 (1.5)	07-Nov-1940 (2.1)
02-May-1972 (−2)	18-May-1972 (−1.4)	30-Jun-1972	17-Jul-1972 (0.6)	21-Jul-1972 (0.7)
16-May-1995 (−10.6)	17-Aug-1995 (−7.5)	28-Mar-1996	09-Sep-1996 (5.5)	21-Nov-1996 (7.9)
Panel A2: Daily range based volatility, Maturity F2				
07-Feb-1914 (−5.7)	28-Mar-1914 (−4)	27-Jul-1914	05-Sep-1914 (1.3)	28-Sep-1914 (2.1)
22-Oct-1941 (−0.3)	25-Oct-1941 (−0.2)	31-Oct-1941	13-Dec-1941 (1.4)	30-Dec-1941 (2)
19-Apr-1972 (−2.3)	09-May-1972 (−1.7)	28-Jun-1972	11-Jul-1972 (0.4)	17-Jul-1972 (0.6)
06-Apr-1995 (−11.7)	20-Jul-1995 (−8.2)	22-Mar-1996	02-Oct-1996 (6.5)	26-Dec-1996 (9.3)
Panel B1: Monthly average price based absolute return, Maturity F1				
Jun 1902 (−142)	Nov 1905 (−101)	Apr 1914	Dec 1921 (92)	Feb 1924 (118)
Jan 1941 (−6)	Apr 1941 (−4)	Aug 1941	Jan 1944 (29)	Dec 1944 (40)
Aug 1965 (−69)	Mar 1967 (−50)	May 1971	May 1972 (12)	Dec 1972 (18)
Panel B2: Monthly average price based absolute return, Maturity F2				
Apr 1904 (−122)	Mar 1907 (−87)	Jun 1914	Sep 1921 (87)	Aug 1923 (110)
Oct 1940 (−9)	Jan 1941 (−6)	Aug 1941	Oct 1943 (26)	Jul 1944 (35)
Aug 1965 (−81)	Mar 1967 (−62)	May 1972	Aug 1973 (15)	Apr 1974 (23)

Break points are estimated using the methodology of Bai and Perron (2003). The number of break points is selected by BIC criteria, where the maximum number of breaks is set to 5 (default). The 90% and 95% confidence bounds are noted with 90% and 95%, respectively, and confidence bound in months are given in brackets. The sample includes for F1 37,165 and F2 37,388 daily ranges resp. for F1 and F2 1'694 monthly return observations from January 3, 1877 to February 28, 2018.

During this period, the CCC's regulation implies a "quasi price floor", which triggers whenever the post harvest price was below a certain level of a defined parity price[13] at the beginning of the marketing year, implying an asymmetric effect on the volatility level, as prices are restricted to fluctuate below, but not above the floor.

This price floor was tightened in autumn 1941, when the congress passed Public Law 74 (Steagall Amandement), see Shepherd (1944). However, according to BIS (1941, p. 80) the intended floor-price increase was public a year before but the definitive level was uncertain until the beginning of the next marketing year, i.e. autumn 1941.

It is further interesting to notice that the major regulatory measures on grain futures trading introduced in the 1920s and 1930s are not reflected in volatility breaks,[14] but that two of the breaks seem to be related to major monetary regime shifts, while the launch of the Bretton Woods system in 1944 cannot be detected. However, this major monetary event is superimposed by the U.S. agricultural price support program.

Precision of structural breaks Besides the mentioned differences in detected breakpoints between the two data sets, they crucially differ in terms of precision. The 90% and 95% confidence time interval for all break points are displayed in Table 2. Consider, for example, the first breakpoint and the nearby futures maturity (F1). With average price data, the 95% confidence interval includes plus/minus 118/142 = 260 months, with daily range data, the interval collapses to plus/minus 3.2/6.4 = 9.6 months. This is a remarkable improvement and should eventually prove helpful in the search for the economic causes of breaks. From a statistical perspective it improves the identification of clear-cut regimes.

Volatility pattern between regimes The breakpoints can be used to get a precise picture about the changing volatility pattern over time, specifically, *between* and *within* the identified regimes. The results are displayed in Column (13) of Table 1.

The general pattern of the results is quite homogeneous and indicates that, with the exception of the 1940–1972 (respectively 1941–1972) time interval, volatility has increased over the five subperiods. Interestingly, the most volatile period is the most recent, since 1996, followed by the 1914–1941 time interval subsequent to the creation of the FED. It is not, as often claimed, the "speculative" area in the late 19th century which experienced above-average volatility.

Our previous observations can be confirmed by looking at these figures: monthly volatilities estimated from average daily price data are uniformly above the daily range based estimates, except in the most recent regime. This implies that by comparing range based volatility estimates for recent time periods (the "best" actual volatility measure) to return-based volatilities from price averages in the old days (the "only" measure mostly available for historical periods) makes volatility differences appear to be *less* pronounced (e.g. 1.60 vs. 1.32 between the last regime and the entire observation period, for F1) than they actually are (1.60 vs. 1.19). This is against the intuition which would suggest that using average-based proxies leads to downward-biased results.

[13] See Montgomery (1953).

[14] The Grain Futures Act of 1922 required that futures trading must only be done at licensed exchanges, and was replaced by the Commodity Exchange Act in 1936.

Notice that our result is in stark contrast to the findings of Jacks et al. (2011) who claim that price volatility of commodities has "unambiguously" not risen over time, and questions the statement of Gilbert and Morgan (2010) who find that the reverse of "the general tendency for commentators to assert that food price volatility has increased over time [...]" appears to be true (p. 3033). However, the identified bias in volatility of averaged price estimated across different regimes might explain the contradicting results.

Volatility differences between and within regimes The remaining columns of Table 1 providing a more detailed picture about the structural shifts of the volatilities, with several interesting insights:

The median daily volatility *within* each regime (column 4) is in line with the daily range based volatility (column 2), i.e. the latter show the typical picture of the volatility within each regime. It is not surprising to observe that the point estimates of daily volatilities (i.e. range based volatilities for a single day based on formula 1) exhibits a high dispersion, with standard deviations ranging from 0.38% to 0.79% for the nearby maturity (column 7). It is however interesting to notice that the differences *between* the volatilities of subsequent regimes (column 2) – each estimated from thousands of data points – are in the same order of magnitude: 0.41%, −0.75%, 0.62% and 0.32% for the nearby maturity. This means that volatility differences *between* the identified regimes are at least as important as volatility fluctuations *within* regimes. This finding is reinforced by our analysis in Section 5 where we are unable, with a single exception, to discriminate between high and low volatility states within the regimes using well-known conditioning instrumental variables.

This implies that research efforts aimed at explaining volatility patterns in commodity markets should not only address short-run volatility fluctuations, but also the determinants of the structural shifts between volatility *regimes*.

Finally, as stated before, the "speculative" area in the late 19th century does not experience above-average volatility, but that period reveals an exceptionally high skewness and kurtosis, at least in the nearby maturity. This indicates the erratic behavior of daily volatility and confirms the visual impression from Fig. 1 in the previous Section. Also note that the kurtosis is substantially smaller between 1972 and 1996 — which is however not the period when volatility was small. Therefore, volatility and kurtosis seem to be two separate characteristics in wheat futures volatility.

5. Conditioning state variables

Volatility shifts can either occur at specific points in time, as analyzed in Section 4, or across economic states. The importance of good and bad states (e.g. expansion and recession, low and high inflation) for explaining commodity returns in the long run is analyzed by Levine et al. (2018). The authors find that economic states are important drivers of commodity returns, even after conditioning on whether commodity markets are in backwardation or contango.

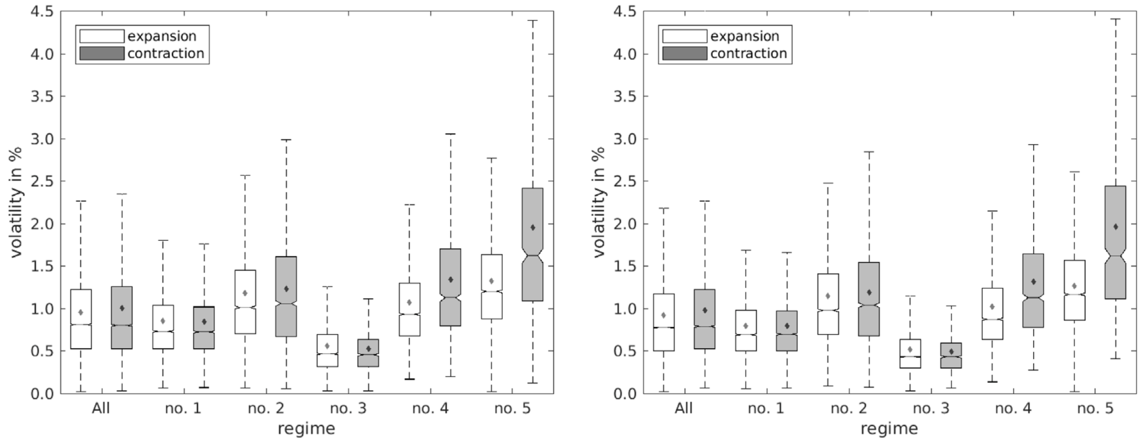
In this section, we analyze three possible conditioning state variables and their impact on daily range based volatilities across and within the five regimes identified in the previous Section. The following state variables are specified: (a) NBER business cycles, (b) the slope of the term structure of futures prices (i.e. backwardation and contango), and (c) CPI inflation. The volatility effects are analyzed for the entire time period and the five subperiods separately. For an easier interpretation, the findings are presented as Boxplots. The two center boxes indicate the interquartile range of the data, and the horizontal bars represent the multiple of 1.5 of that range. The diamond indicates the arithmetic mean.

US business cycles data are taken from the NBER. The results are displayed in Fig. 2a. For the full time period (All), there is no noticeable difference between volatilities in booms and recessions, neither with respect to the median nor with respect to the typical quantiles. The subperiods offer a differentiated picture: two periods which include severe recessions (fourth: Oil Crisis 1970s; fifth: Dotcom Bubble and Financial Crisis 2000s) exhibit elevated volatilities in recession states compared to booms.[15] However, the effect is rather small in the fourth subperiod, in contrast to the most recent subperiod (1996–2016) where the interquartile range of daily volatilities increases from 0.9%–1.6% in booms to 1.1% to 2.5% in recessions.

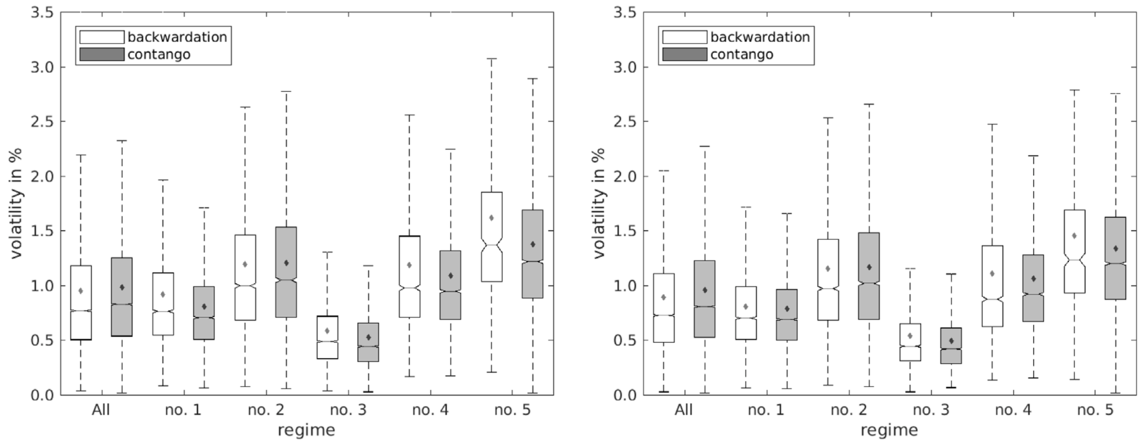
The slope of the term structure of commodity price, which can be measured just at the shortest end with our data, is a natural conditioning state variable in commodity markets. Backwardation (negative slope) and contango (positive slope) are implicit measures of the benefits from holding commodity in physical stock, the so called "convenience yield", and are thus regarded as indicators of scarcity and excess. Based on Working's "theory of storage", most theories predict a positive relationship between scarcity and volatility; see e.g. Pindyck (2001). Empirical evidence for the posited positive relationship covering 21 commodities from 1993–2011 is provided by Symeonidis et al. (2012). Our results are presented in Fig. 2b. The general picture does not support a substantial difference in median volatilities and their distribution between the two states. In the full sample period, opposite to the hypothesis, volatilities in backwardation are even smaller than in contango. In the individual subperiods, only the most recent regime (1996–2016) reveals some difference in the distribution of volatilities between the two states, with slightly higher values in backwardated markets. In the other regimes, the volatility differences are almost negligible and exhibit no systematic pattern.

One possible explanation of this result is that the slope between the first and second maturity futures price is just an insufficient statistic for measuring backwardation and contango properly; however, no other historical data are available for the full historical time period. An analysis using the full set of longer contract durations to measure backwardation and contango periods would be an interesting route to follow. We leave this question for a separate study, as this data is not available for the full historical time period.

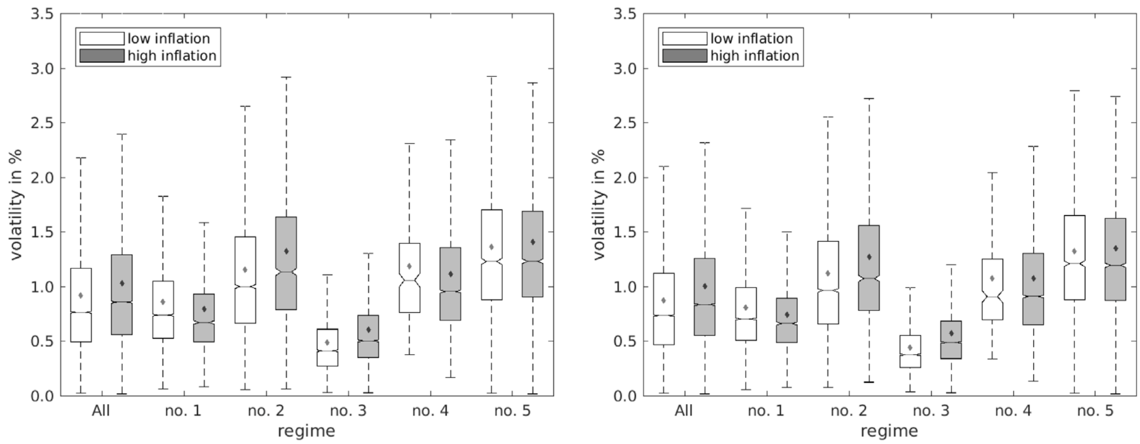
[15] However, the same observation is virtually absent in the second subperiod which includes the Great Depression.



(a) Business cycle: NBER expansion and contraction periods (left: F1, right: F2)



(b) Term structure: Backwardation and Contango (left: F1, right: F2)



(c) Inflation: CPI below and above 2% (left: F1, right: F2)

Fig. 2. Wheat futures volatility: Conditioning economic states. The figure shows the distributional properties of daily range based volatilities for the full time period and the five regimes, conditional on economic states, according to state variables indicated in the subcaptions. The diamond indicates the arithmetic mean. The sample includes 37,164 daily observations from January 3, 1877 to February 28, 2018.

Inflation is a potentially important state variable in long-term price studies. There are several periods of extraordinary inflation in the US which have a potential effect on the volatility of nominal commodity prices. Many empirical studies using monthly data use real prices to control for that effect. Since the focus on this paper is on daily volatility, a natural way to address the impact of inflation is by separating high and low inflation regimes. We chose the long-run inflation target of 2% as threshold for separating the two regimes. We use CPI-inflation which is compiled from several sources by the Bureau of Labor Statistics and available throughout the 19th century. The results are displayed in Fig. 2c. Apparently, volatility differences are essentially non-existent between low and high inflation states — with a certain exception in the second regime, which is insofar an interesting observation because that period (1914–1940) includes the most extreme annual inflation rates (from +20% to −10%).

The only consistent observation across all three Figures is that the volatility differences between the five volatility regimes is considerably more important than those across the conditioning states. This observation reinforces the conclusions drawn from Table 1, namely the importance of volatility differences *between* the identified regimes compared to those within the regimes. Applying widely used conditioning instruments does not seem to discriminate between high and low volatility states within the five regimes. As noticed, the only exception is the business cycle effect in the most recent volatility regime (1996–2016). This implies that empirical studies focusing on recent data tend to overstate the relevance of the business cycle for explaining commodity price volatility.

6. Summary

We analyze Chicago based daily wheat price volatility from 1877 to 2018 using a novel data set of daily high and low futures prices starting in 1877. These data allow us to perform consistent point estimates for daily wheat price volatility over an observation period of more than 140 years. Our main findings can be summarized as follows:

First, the analysis of structural breaks reveals that volatility differences *between* the identified regimes are substantially higher than volatility fluctuations within regimes. This is particularly true for volatility differences between economic states *within* the regimes: Applying widely used conditioning instruments (e.g. business cycle, inflation and convenience yield) does not seem to discriminate between high and low volatility states within regimes. Therefore, research efforts aimed at explaining volatility patterns in commodity markets should not only address short-run volatility fluctuations, but also the determinants of the structural shifts between volatility *regimes*. Two identified break points appear to coincide with shifts in the monetary policy.

Second, return volatility of monthly price averages is substantially higher than the daily range based estimate. The identified bias is on average 11 percent but time variant. Surprisingly, the highest bias of 22 percent is found in the regime with the lowest volatility level and disappears in the latest but highest volatility regime from 1996 to 2018.

Third, the volatility measured by regimes does not reveal a decreasing trend over the most recent decades as found by several papers using data covering several decades such as Gilbert and Morgan (2010) or Jacks et al. (2011). The identified bias in volatility of averaged price estimated across different regimes might explain the contradicting results. The regime with by far the largest volatility – both in terms of volatility from daily range data and average daily volatility – is the most recent.

Future research has to show whether the results presented in this paper can be generalized to other commodity futures prices.

CRediT authorship contribution statement

Marco Haase: Data curation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Heinz Zimmermann:** Conceptualization, Data curation, Funding acquisition, Methodology, Writing – original draft, Writing – review & editing. **Matthias Huss:** Data curation, Methodology, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

This research has been partially funded by the Swiss National Science Foundation (SNSF) under the project "Volatility in early commodity markets" (grant number 100018-172681). No other external financial sources have been provided. No conflicts of interest must be declared.

Data availability

Data will be made available on request.

References

Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66 (1), 47–78.
Bai, J., Perron, P., 2003. Computation and analysis of multiple structural change models. J. Appl. Econometrics 18 (1), 1–22.
BIS, 1941. Eleventh annual report. Tech. rep., Bank for International Settlements, Basel, 9th June 1941.
Brandt, M.W., Kinlay, J., 2005. Estimating historical volatility. Tech. rep., Investment Analytics.
Calvo-Gonzalez, O., Shankar, R., Trezzi, R., 2010. Are Commodity Prices more Volatile now? a Long-Run Perspective. The World Bank.
Cashin, P., McDermott, C.J., 2002. The long-run behavior of commodity prices: small trends and big variability. IMF Staff Pap. 49 (2), 175–199.
Cashin, P., McDermott, C.J., Scott, A., 2002. Booms and slumps in world commodity prices. J. Dev. Econ. 69 (1), 277–296.
Cowing, C.B., 2015. Populists, Plungers, and Progressives: A Social History of Stock and Commodity Speculation, 1868-1932. Princeton University Press.
Dies, E.J., 1925. The Wheat Pit. The Argyle Press.

Gilbert, C.L., Morgan, C.W., 2010. Food price volatility. Philos. Trans. R. Soc. B 365 (1554), 3023–3034.
Irwin, H., 1936. Technical conditions are important factors in short-time movements of wheat prices. J. Farm Econ. 18 (4), 736–742.
Jacks, D.S., O'Rourke, K.H., Williamson, J.G., 2011. Commodity price volatility and world market integration since 1700. Rev. Econ. Stat. 93 (3), 800–813.
Levine, A., Ooi, Y.H., Richardson, M., Sasseville, C., 2018. Commodities for the long run. Financ. Anal. J. 74 (2), 55–68.
Markham, J.W., 2015. Law Enforcement and the History of Financial Market Manipulation. Routledge.
Mayo-Smith, R., 1900. Price movements and individual welfare. Political Sci. Q. 15 (1), 14–36.
Montgomery, G., 1953. Wheat Price Policy in the United States. Tech. rep..
Nelson, C.R., Kang, H., 1981. Spurious periodicity in inappropriately detrended time series. Econometrica 49 (3), 741–751.
Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. J. Bus. 53, 61–65.
Pesaran, M.H., Pettenuzzo, D., Timmermann, A., 2006. Forecasting time series subject to multiple structural breaks. Rev. Econom. Stud. 73 (4), 1057–1084.
Pindyck, R.S., 2001. The dynamics of commodity spot and futures markets: A primer. Energy J. 22, 1–29.
Rogers, L.C.G., Satchell, S.E., 1991. Estimating variance from high, low and closing prices. Ann. Appl. Probab. 1, 504–512.
Samuelson, P.A., 1965. Proof that properly anticipated prices fluctuate randomly. Ind. Manag. Rev. 6 (2), 41–49.
Shepherd, G., 1944. The coordination of wheat and corn price controls. Iowa Agric. Home Econ. Exp. Station Res. Bull. 27 (330), 1.
Slutzky, E., 1937. The summation of random causes as the source of cyclic processes. Econometrica 5 (2), 105–146.
Sumner, D.A., 2009. Recent commodity price movements in historical perspective. Am. J. Agric. Econ. 91 (5), 1250–1256.
Symeonidis, L., Prokopczuk, M., Brooks, C., Lazar, E., 2012. Futures basis, inventory and commodity price volatility: An empirical analysis. Econ. Model. 29 (6), 2651–2663.
UN General Assembly, 2015. Transforming our world: The 2030 Agenda for Sustainable Development. Resolution adopted by the General Assembly on 25 September 2015, available at: http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1{&}Lang=E.
USDA, 1984. History of Agricultural Price-Support and Adjustment Programs, 1933-84: Background for 1985 Farm Legislation. US Department of Agriculture.
Working, H., 1960. Note on the correlation of first differences of averages in a random chain. Econometrica 28 (4), 916–918.
Yang, D., Zhang, Q., 2000. Drift independent volatility estimation based on high, low, open and close prices. J. Bus. 73, 477–491.
Zeileis, A., Kleiber, C., Krämer, W., Hornik, K., 2003. Testing and dating of structural changes in practice. Comput. Statist. Data Anal. 44 (1–2), 109–123.