Forecasting commodity markets volatility: HAR or Rough?

Mesias Alfeus[a], Christina Sklibosios Nikitopoulos[b,]

[a] *The African Institute for Financial Markets and Risk Management, Faculty of Commerce, University of Cape Town Rondebosch 7701, South Africa*
[b] *University of Technology Sydney, Finance Department, UTS Business School, PO Box 123 Broadway NSW 2007, Australia*

February 29, 2020

Abstract

Commodity is one of the most volatile markets and forecasting its volatility is an issue of paramount importance. We study the dynamics of the commodity markets volatility by employing fractional stochastic volatility and heterogeneous autoregressive (HAR) models. Based on a high-frequency futures price dataset of 22 commodities, we confirm that the volatility of commodity markets is rough and volatility components over different horizons are economically and statistically significant. Long memory with anti-persistence is evident across all commodities, with weekly volatility dominating in most commodity markets and daily volatility for oil and gold markets. HAR models display a clear advantage in forecasting performance compared to fractional volatility models.

Keywords: commodity markets, realized volatility, fractional Brownian motion, HAR, volatility forecast
JEL: C20, C53, C58, G13, Q02

Email addresses: `mesias.alfeus@uct.ac.za` (Mesias Alfeus), `christina.nikitopoulos@uts.edu.au` (Christina Sklibosios Nikitopoulos)

Electronic copy available at: https://ssrn.com/abstract=3520500

1. Introduction

Commodity markets exhibit the highest volatility of all asset classes and play a critical role in the economy (Chiang, Hughen, and Sagi (2015) and Kang, Nikitopoulos, and Prokopczuk (2019)). Bollerslev, Hood, Huss, and Pedersen (2018) documented strong similarities in realised volatility across asset classes and recent research predicts a strong correlation and substantial volatility spillovers between the equity and commodity markets (Chiang et al. (2015) and Basak and Pavlova (2016)). The long-memory property (slow decaying autocorrelation function which is slower than exponential) is a common stylised fact of financial markets volatility, a feature that cannot be captured by ordinary Brownian motion dynamics. There is strong empirical evidence in stock markets that long-memory volatility models such as fractional Brownian motion or heterogeneous autoregressing models can reproduce the empirical features of the volatility and improve its forecasting ability. While there is still no consensus on which is the most appropriate forecasting model, this paper aims to study the volatility of commodity markets with long-memory models and to shed some light on the commodity volatility forecasting, which is the most important and relevant domain of application.

To analyse the volatility of commodity markets, we propose two models for realized volatility that capture long-memory characteristics, a fractional stochastic volatility model and a Heterogeneous AutoRegressing (HAR) model. We find that the volatility in commodity markets features long-memory with anti-persistence and we confirm the empirically observed roughness of commodity markets. The Hurst parameter in commodity markets is lower than the estimate (of 0.15) obtained for equity markets and indexes, with energy markets have the highest Hurst parameter estimates (0.110 to 0.152) and metals have the lowest estimates (0.014 to 0.041). Furthermore, the informational content of this volatility is economically and statistically significant over different time horizons. We also compare the forecasting ability of these two long-memory models and find that HAR models have an advantage in volatility forecasting compared to rough fractional stochastic volatility models. After 2004,

Electronic copy available at: https://ssrn.com/abstract=3520500

the Hurst parameter has increased for most commodities and HAR models tend to perform better in term of volatility forecasting in the two sub-samples (separated in 2004).

This paper links two main strands of literature on long-memory volatility modelling and forecasting. The first one is based on modelling volatility with a fractional Brownian motion. Comte and Renault (1998) was the first to consider long memory in volatility with the sources of uncertainty following a fractional Brownian motion with Hurst parameter $H \in [\frac{1}{2}, 1)$. Recent empirical studies using high frequency data of realized volatility time series shows that the log-volatility have a behaviour which can be best explained by a fractional Ornstein-Uhlenbeck process yet of a Hurst parameter less than half and closer to zero (Bayer, Friz, and Gatheral (2016), Fukasawa (2017), El Euch, Fukasawa, and Rosenbaum (2018), Gatheral, Jaisson, and Rosenbaum (2018) and Fukasawa, Takabatake, and Westphal (2019)). Gatheral et al. (2018) showed that log-realized volatility computed from high-frequency data of major stock indices shows anti-persistence or roughness and revolutionalised the literature to understand the financial implications when $H < 1/2$. This model gives a general interpretation of the volatility dynamics from high-frequency behaviour in the market, and fit superfluously well at-the-money skew than any other classical stochastic volatility model, while it preserves long memory features of volatility. Following these seminal papers, there are many applications to various markets such as, to cryptocurrency (Takaishi (2019)), to electricity market (Bennedsen (2017)), to index volatility markets (Da Fonseca and Zhang (2019)). Livieri, Mouti, and Rosenbaum (2017) and Alfeus, Overbeck, and Schlögl (2019) conclude that at-the-money short term volatility from S&P 500 options is rough with a Hurst parameter of order $H = 0.3$. We add to this literature that lacks a comprehensive study on the volatility of commodity markets.

The second stand is based on HAR models and has been prominently used to forecast volatility using realized variance. From the seminal paper by Corsi (2009), the introduction of positive and negative realized semivariances by Barndorff-Nielsen, Kinnebrock, and Shephard (2010) and further the consideration of signed jumps by Patton and Sheppard (2015), the

Electronic copy available at: https://ssrn.com/abstract=3520500

volatility forecasting underwent significant advances, in particular in intraday equity and equity indexes markets. The literature on commodity markets is comparatively limited. According to Corsi (2009), volatility persistence can be predicted by past realized variances sampled over different time horizons. Extensions of this basic model incorporating leverage, semivariances and signed jumps have been employed in oil and agricultural markets and show that HAR models can mimic long memory and provide a comparatively good forecasting performance (Sévi (2014), Gong and Lin (2017) and Luo and Chen (2019)). Degiannakis and Filis (2018) demonstrate the importance of combining information from high-frequency data and oil market fundamentals to better forecast oil prices under the Mixed-Data Sampling (MIDAS) framework. In terms of the long memory feature in commodity markets, Elder and Serletis (2008) provide empirical evidence of anti-persistent commodity prices with a sample from 1994 to 2005. Based on rolling sample tests such as Lo (1991) modified R/S analysis and three Whittle methods, Wang and Wu (2012) show that long memory does not exist in energy markets, except times of extreme events such as the GFC (see also Cunado, Gil-Alana, and De Gracia (2010)). As the question of long-memory in commodity markets and the urge for more effective forecasting of commodity volatility still remain, we aim to make a contribution by assessing the long-memory feature of a large sample of commodity markets and by providing an insight on dominant forecasting predictors.

The remainder of the paper is structured as follows: Section 2 presents the characteristics of the realized volatility of commodity markets. Section 3 discusses the dynamics of released volatilities of commodity markets under fractional stochastic volatility and HAR model specifications. A detailed discussion of the impact of the roughness of the commodity markets on forecasting and comparison with HAR models is presented in Section 4. Section 5 presents robustness tests. Section 6 concludes.

Electronic copy available at: https://ssrn.com/abstract=3520500

2. Realized volatility of commodity markets

We denote as F_t the futures price of a commodity at time $t \geq 0$. A widely accepted definition of the daily realized volatility (RV), which is supported by quadratic variation arguments and in-fill asymptotic arguments (Andersen, Bollerslev, Christoffersen, and Diebold (2013)), is given by[1]

$$RV_t^{(d)} \equiv \sum_{i=1}^{M} [\log(F_{t-1+i\Delta}) - \log(F_{t-1+(i-1)\Delta})]^2, \tag{2.1}$$

where M is the number of intra-day observations expressed as $M = \frac{1}{\Delta}$ with Δ be the time between two observations. This estimate computes the true variation on day t, as intra-day returns are sampled within a time interval Δ (with $\Delta \to 0$). If σ_t denotes the spot volatility of these returns (which are not observable) then volatility can be approximated by using the observable realized volatility. When $\Delta \approx 0$, the realised volatility converges to the quadratic variation or integrated volatility, i.e.,

$$RV \to IV_t = \int_{t-1}^{t} \sigma_s^2 ds,$$

where the length of 1 day is normalised to $[0, 1]$. Realized volatility over multi-day horizons is estimated by normalized sums of the daily realized volatilities, e.g. weekly realized volatility is estimated by

$$RV_t^{(w)} = \frac{1}{5} \sum_{j=1}^{5} RV_{t-jd}^{(d)}. \tag{2.2}$$

To account for the market microstructure noise, we follow the approach of Hansen and Lunde (2005). They adopt a multiplicative factor that adjusts the realised volatility and the

[1]Equation (2.1) explicitly computes daily realized variance and consequently $(RV_t^{(d)})^{1/2}$ represents the daily realized volatility. As it is customary in the literature, we use the same term (realized volatility) interchangeably for both these two representations of released variation.

Electronic copy available at: https://ssrn.com/abstract=3520500

factor is given by

$$c \equiv \frac{\sum_{t=1}^{N} r_t^2}{\sum_{t=1}^{N} RV_t}, \quad t = 1, 2, \cdots, N, \tag{2.3}$$

where $r_{t,i} = \log(F_{t-1+i\Delta}) - \log(F_{t-1+(i-1)\Delta})$ reports the $i-$th return on day t and N is the number of days in the sample. Thus, the associated estimator of realised volatility on day t is given by

$$\hat{RV}_t = cRV_t^{(d)}. \tag{2.4}$$

2.1. Commodity markets data

We employ a large dataset of 5-minute[2] intra-day prices for the front-month commodity futures contracts in 22 commodities over the period from January 1996 to August 2019. The data are collected by Thompson & Reuters DataScope. We include six energy commodities (crude oil (WTI and Brent), natural gas, heating oil, gasoline), four precious metals (gold, silver, platinum and palladium), one industrial metal (copper), four grains (corn, soybean, rice, wheat), five softs (sugar, cocoa, coffee, cotton and orange juice) and three livestock (live cattle, feeder cattle and lean hogs).

In line with Andersen, Bollerslev, and Diebold (2007) and Bollerslev et al. (2018), data cleaning consists of two stages. In the first stage, we eliminate errors from the raw data by omitting data with zero trades, with settlement prices below the bid or above the ask and with bid quotes above the ask quotes. In the second stage, we filter realized volatility and omit data with zero realized volatility and any day with low total number of trades (using a threshold based on the total daily number of trades).

[2]Empirical studies on reliability of realized volatility estimates reveal that 5-min sampling frequency outperforms the majority of volatility estimators (Liu, Patton, and Sheppard (2015)).

Electronic copy available at: https://ssrn.com/abstract=3520500

2.2. *Realized volatility characteristics*

Table 1 reports summary statistics for the daily realized volatilities of the 22 commodities of the sample, where all statistics are reported in annualised terms (% per annum). Energy markets are the most volatile markets together with palladium from metals and coffee and orange juice from softs, while gold markets exhibit the least volatility. On average, the annualized volatility of energy markets is 33.52%, followed by 30.44% in softs, 24.43% in grains, 21.94% in metals and 17.43% in livestock. It is also evident that volatility persistence increases significantly with time across all commodity markets. The realized volatility series for all commodities are stationary.[3]

The realized volatility of commodity markets for a set of representative commodities is presented in Figure 2, which depicts the realized volatility in annualised terms for crude oil (energy), gold (precious metals), copper (industrial metals), soybean (grains), cocoa (softs) and feeder cattle (livestock). The series in generally display patterns idiosyncratic to the associated commodity market, which are though marked by key economic events, such as the Global Financial Crisis and the trend of an increased activity after 2004.

Figure 1 depicts the average autocorrelation function of the daily realized volatilities for a representative commodity per class: crude oil, gold, copper, soybean, cocoa and feeder cattle. The daily realized volatility autocorrelations for all selected commodities display a hyperbolically decaying autocorrelation function of the daily volatility (except gold) implying long memory effects. While most of the commodities display similar patterns with strong persistence that lasts for a few months, crude oil and gold behave differently. In the short-term crude oil displays higher autocorrelation while gold displays much lower autocorrelation.[4]

[3]Using ets function in the forecast package of R, we find no seasonality in the daily realised volatility series.

[4]Hansen and Lunde (2014) address the impact of the measurement error in the realized volatility for different assets, that results in making the short-lagged autocorrelations non-comparable.

Electronic copy available at: https://ssrn.com/abstract=3520500

Table 1: **Commodities RV summary statistics**

Series	Mean	Stand. Dev.	Skewness	Excess Kurtosis	max	min	Ljung-Box 1-day	Ljung-Box 20-day	ADF	no. of days
Energy										
Crude oil (WTI)	32.1	14.3	2.2	7.9	148.3	5.0	3341.3	43887.4	-5.97^{**}	5,937
Crude oil (Brent)	29.2	12.9	2.1	6.8	125.4	6.6	3238.3	44395.4	-5.94^{**}	5,855
Natural gas	44.2	18.7	1.7	5.3	186.8	6.9	2358.6	26358.2	-8.96^{**}	5,554
Heating oil	31.1	12.9	1.3	3.0	111.2	5.5	2910.6	42307.0	-6.05^{**}	5,528
Gasoline	31.0	14.8	2.3	7.8	136.8	6.4	1996.9	31765.4	-3.94^{**}	3,161
Metals										
Gold	7.0	6.6	2.0	5.5	47.3	0.1	420.4	1255.8	-12.22^{**}	2,768
Silver	25.9	17.5	1.5	3.2	109.5	0.1	557.1	6337.3	-5.93^{**}	2,416
Platinum	21.5	14.0	1.7	4.4	95.0	0.1	88.7	600.5	-6.66^{**}	813
Palladium	34.0	25.9	2.0	4.9	155.8	1.1	181.5	1160.2	-5.39^{**}	641
Copper	21.3	13.3	2.3	7.8	110.7	0.6	663.9	7685.5	-7.45^{**}	3,855
Grains										
Corn	23.0	10.8	2.4	9.3	111.8	7.0	1354.3	16266.5	-6.76^{**}	4,830
Soybean	19.1	8.8	2.6	11.6	92.8	2.4	1596.5	19023.9	-6.97^{**}	5,146
Wheat	26.6	12.3	3.0	17.2	157.0	5.9	1241.3	14159.7	-7.31^{**}	4,842
Rough Rice	29.0	17.5	2.6	10.9	156.6	3.6	1002.2	7576.4	-7.88^{**}	4,700
Softs										
Cocoa	26.8	11.1	1.6	5.1	112.2	5.8	1199.7	10960.4	-8.22^{**}	4,299
Coffee	34.2	14.0	1.6	3.9	111.6	11.8	1104.0	11738.8	-7.65^{**}	3,223
Cotton	27.4	14.4	1.9	5.1	110.1	5.7	1194.8	13295.2	-6.80^{**}	4,241
Orange juice	35.3	16.5	1.9	6.1	151.7	8.6	513.9	4402.6	-7.36^{**}	3,013
Sugar	28.5	12.0	1.5	5.4	124.4	6.7	1594.3	19963.3	-5.87^{**}	4,300
Livestock										
Live cattle	15.1	5.9	1.6	4.2	49.2	4.2	850.4	8933.4	-6.29^{**}	3,437
Feeder cattle	14.2	8.4	3.9	31.6	129.8	2.3	812.8	8784.7	-8.60^{**}	5,284
Lean hogs	23.0	11.1	2.3	9.4	106.0	3.8	1223.2	13719.9	-8.11^{**}	5,391

The table presents summary statistics for daily realized volatilities of commodities. Volatility is reported in annualised terms (% per annum). The number of lagged changes p is selected based on the Akaike Information Criterion, testing up to a maximum of $p = 15$ lags. The symbols ***, **, and * indicate rejection of the unit-root null hypothesis at the 99%, 95%, and 90% confidence levels, respectively.(Dickey and Fuller (1979))

As also shown in Figure 3, the partial autocorrelation functions display a downward trend as recent data containing more information (being more relevant).

The empirical characteristics of the daily realized volatilities of commodity markets include volatility persistence, heteroscedasticity and long memory. Models capable to reproduce these stylized facts of commodity markets are employed next to analyse the volatility dynamics in these markets.

Electronic copy available at: https://ssrn.com/abstract=3520500

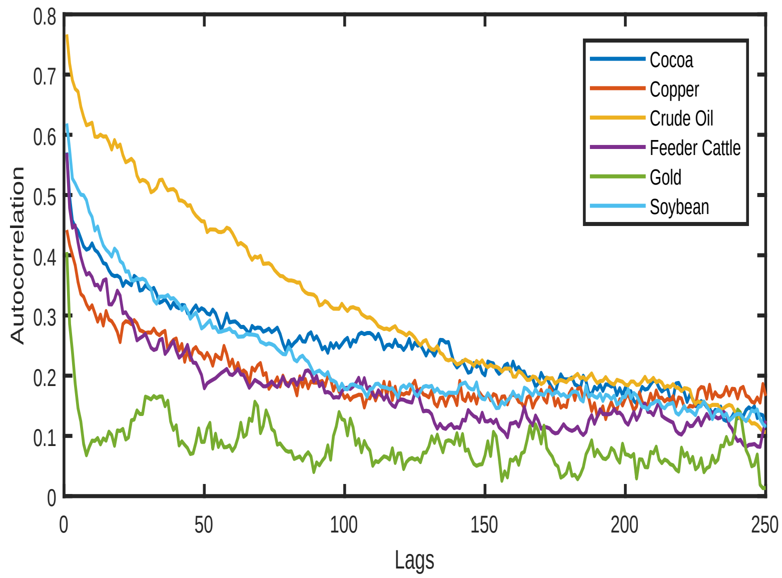


Figure 1: **Autocorrelation function of daily realized volatilities**
The figure plots the average autocorrelation function of the daily releazed volatilities for crude oil, gold, copper, corn, sugar and live cattle.

3. Volatility dynamics

We empirically study the realized volatility by using two well-known models of long-memory processes: a fractional stochastic volatility model (FRSV) (Gatheral et al. (2018)) and a Heterogeneous Autoregressing (HAR) model (Corsi (2009)).

3.1. Fractional stochastic volatility (FRSV)

The roughness (or smoothness) of the volatility is captured by a fractional Brownian motion. By definition, as introduced in Mandelbrot and Van Ness (1968), a fractional Brownian motion with Hurst parameter H, with $0 < H < 1$, is a centered Gausian process W_t^H with stationary increments, such that, for $t \in \mathbb{R}$,

$$\mathbb{E}[|W_{t+\Delta}^H - W_t^H|^q] = \mathbb{E}[|Z|^q]\Delta^{qH}, \quad Z \sim N(0,1), \tag{3.5}$$

where $\Delta \geq 0$, $q > 0$. The dynamics of the so-called rough stochastic volatility are modelled via a fractional Brownian motion with a Hurst parameter $H \in (0, 1/2)$. Then the process is anti-persistent, consequently it corresponds to a process with negative autocorrelation. An

Electronic copy available at: https://ssrn.com/abstract=3520500

ordinary Brownian motion is obtained, when $H = 1/2$. A fractional Brownian motion with a Hurst parameter $H \in (1/2, 1)$ represents a persistent process with positive autocorrelation.

An example of a rough volatility commodity model may follow the dynamics

$$\begin{aligned} d\log F_t =& \mu dt + \sigma_t dW_t, \\ d\log \sigma_t =& \alpha(m - \log \sigma_t)dt + \nu dW_t^H, \end{aligned} \tag{3.6}$$

where $m \in \mathbb{R}$, $\nu > 0$, $\alpha > 0$ and σ_t is represented by RV_t. An appealing feature of assessing the roughness of the volatility is that it is not tied up to model assumptions. Thus to estimate the Hurst parameter H, it is not required to use model specifications such as (3.6). By following Gatheral et al. (2018), the Hurst parameter H can be extracted from (3.5), by using the expectation

$$\mathbb{E}[|\log RV_{t+\Delta} - \log RV_t|^q], \tag{3.7}$$

for varied value of $q > 0$ and a small $\Delta > 0$. At times $t = 0, \Delta, 2\Delta, \ldots$ and for $N = T/\Delta$, the volatility values $RV_0, RV_\Delta, \ldots, RV_{k\Delta}, \ldots$ with $k \in \{0, 1, \ldots, t/\Delta\}$ are used to calculate

$$m(q, \Delta) = \frac{1}{N} \sum_{k=1}^{N} |\log RV_{k\Delta} - \log RV_{(k-1)\Delta}|^q. \tag{3.8}$$

The quantity $m(q, \Delta)$ cannot be computed exactly as the volatility values $RV_{k\Delta}$ are not directly observable. We rely to daily approximations of the unobservable $RV_{k\Delta}$. For the high-frequency data, we use the realized variance of 5-minute returns, while for the daily

Electronic copy available at: https://ssrn.com/abstract=3520500

data we use the realized variance of 10-day and 20-day returns.[5] Based on the approximation

$$RV_t \sim \sum_{i=1}^{M} (\log F_{t_i} - \log F_{t_{i-1}})^2, \tag{3.9}$$

where t_i for $i = 0, 1, \ldots, M$, represents the partition of interval $[0, \Delta]$ (e.g. 5-minutes for high frequency data, 1-day for daily data).

Under the assumption that the log-volatility process is stationary and that a law of large numbers holds, then from equation (3.5), qH has monofractual scaling poperties which implies that

$$m(q, \Delta) \propto \Delta^{qH},$$

with constant of proportionality K_q. Thus, we have

$$m(q, \Delta) \approx K_q \Delta^{qH}, \tag{3.10}$$

which reflects the empirical counterpart of (3.7). We further use the market data via the regression

$$\log m(q, \Delta) = \beta_1 + \beta_2 \log \Delta + \epsilon, \tag{3.11}$$

that provides an estimate for $H = \frac{\beta_2}{q}$ for several orders q, where $\beta_1 = \log K_q$. The regression of the slope in (3.11) against q reveals that the different orders of q lead to the same estimate of H as the one obtained when $q = 2$.

3.1.1. The roughness of commodity markets volatility

To assess the roughness of the commodity markets volatility, we estimate the Hurst parameter H. Under stationarity conditions, the scaling property (3.5) assumes that, inde-

[5] Several estimations methodologies of the daily variance of returns has been used in the literature including the Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009) and the Realized Library of Oxford-Man Institute. No significant differences have been reported using different estimation methodologies (Glaserman and He (2018)).

Electronic copy available at: https://ssrn.com/abstract=3520500

pendent of the value of q, there is a linear relation between $\log m(q, \Delta)$ and $\log \Delta$ that is associated to the Hurst parameter H.

Figure 4 plots the $\log m(q, \Delta)$ as a function of $\log \Delta$ for different values of q (q=0.5, 1, 1.5, 2, 3) for six commodities - crude oil, gold, copper, soybean, cocoa and feeder cattle. From Figure 4, we conclude that all commodities have the same scaling behaviour, which implies that the Hurst parameter H can be estimated. Table 2 reports the Hurst parameter estimates for the range of 22 commodity markets considered in the study (see first panel of Table 2). Hurst parameter estimates are well below 1/2 and consistently below the Hurst parameter estimates of 0.15 obtained for the realize volatility of major equity indexes and US equities (Gatheral et al. (2018) and Bennedsen, Lunde, and Pakkanen (2017)). The only commodities with Hurst parameter estimates ranging between 0.11 and 0.16 are the energy commodities, while most of the other commodities have a Hurst parameter that is less than half of 0.15. More precisely (over the whole sample period), metals have the lowest Hurst parameter estimates (ranging between 0.014 to 0.041), with higher Hurst parameter estimates for softs (0.045 to 0.69) and livestock (0.061 to 0.73) and lastly grains (0.038 to 0.080).

Even though, a form of persistence is evident in the log-volatility, this cannot be translated to a long memory in the volatility, in the classical power law sense. Gatheral et al. (2018) have demonstrated that the autocovariance of both the rough fractional stochastic volatility and index data do not behave (more precisely, do not decay) as a power law.[6]

3.2. HAR models

The Heterogeneous Autoregressing (HAR) model proposed by Corsi (2009) is a three-factor stochastic volatility model with each factor associated to realized volatilities of different

[6]As explained in Gatheral et al. (2018), if the model of the estimation procedure is misspecified, then in terms of the long memory of the volatility, spurious results may be attained, for instance as in Andersen, Bollerslev, Diebold, and Labys (2003).

Electronic copy available at: https://ssrn.com/abstract=3520500

Table 2: **Hurst parameter** H

Series	$1996-2019$	1996-2004	2005-20019
Energy			
Crude oil (WTI)	0.118	0.064	0.153
Crude oil (Brent)	0.110	0.064	0.146
Natural gas	0.148	0.125	0.159
Heating oil	0.111	0.090	0.122
Gasoline	0.152	-	0.152
Metals			
Gold	0.018	0.017	0.020
Silver	0.032	0.023	0.040
Platinum	0.014	0.043	0.001
Palladium	0.034	0.065	0.021
Copper	0.041	0.022	0.052
Grains			
Corn	0.080	0.062	0.088
Soybean	0.080	0.082	0.077
Wheat	0.061	0.034	0.077
Rough rice	0.038	0.035	0.042
Softs			
Cocoa	0.050	0.063	0.046
Coffee	0.069	0.060	0.075
Cotton	0.066	0.051	0.081
Sugar	0.062	0.035	0.082
Orange juice	0.045	0.029	0.052
Livestock			
Live cattle	0.073	0.070	0.075
Feeder cattle	0.061	0.062	0.059
Lean hogs	0.072	0.102	0.060

The table displays estimates of the Hurst parameter H for commodity prices.

frequencies, namely daily, weekly, and monthly. Precisely, the realized volatility is expressed as

$$\log RV_{t+h}^{h} = \beta_0 + \beta_D \log RV_t + \beta_W \log RV_t^W + \beta_M \log RV_t^M + \epsilon_t, \tag{3.12}$$

where RV_t^W and RV_t^M denotes the weekly (5-day) and monthly (22-day) realized volatilities, respectively. Our specifications differ from the original Corsi (2009) model as we consider a) logs of RV to ensure positiveness of the volatility process and that RV distribution is lognormal, and b) the realised variance instead of realized volatility, a choice that has been

Electronic copy available at: https://ssrn.com/abstract=3520500

tested and provides similar results, see Andersen et al. (2007), Sévi (2014) and Patton and Sheppard (2015). Further our lag structure avoids overlapping terms, see Sévi (2014). The model reproduces patterns of volatility persistence which is consistent with empirical observations, including the long-memory dynamic behavior featured in commodity futures markets.[7]

3.2.1. Model estimation

The parameters of model (3.12) are estimated by using standard OLS regression estimators with the Newey-West (Newely and West (1987)) covariance correction to account for serial correlation in the data.

Table 3 reports the estimated HAR parameter estimates for the range of commodity markets considered in the study. By considering results over the full sample (see first panel of Table 3), we find that all the coefficients are highly significant with energy markets displaying the highest explanatory power (ranging between 60% and 70%) followed by grains and softs (around 45% on average),and livestock (around 40%). Metals have the lowest explanatory power with the exceptions of silver, palladium and copper which are around 30%.

Supported by the Heterogeneous Market Hypothesis that was introduced by Müller, Dacorogna, Dav, Olsen, Pictet, and von Weizsacker (1997), the model assumes that markets are driven by distinct layers of investment horizons as "agents perceive, react and cause different volatility components". Accordingly, the model allows to assess the contribution of corresponding market components aggregated over different time horizons to the overall market activity. Weekly volatility dominates in most commodity markets, while daily volatility marginally leads in oil and gold markets. In silver, rice, cocoa and cotton markets, monthly volatility matters the most. Intuitively, commodity markets most likely exhibit

[7]As shown in Granger (1980) and LeBaron (2001) combining $AR(1)$ processes reproduced long memory effects. On than note, asymptotically short-memory models can be empirically indistinguishable with long-memory models, when the aggregation level is not much higher than the frequency component. HAR model is a short-memory process that when is aggregated at these three aggregation intervals (daily, weekly, monthly), it produces the long memory observed in empirical data acting as a pseudo long-memory process.

Electronic copy available at: https://ssrn.com/abstract=3520500

lower mean tick arrival frequency and more market microstructure noise compared to equity of FX markets (Bollerslev et al. (2018)). A daily frequency estimation sustains more noise (and consequently less information) compared to the averaged weekly and monthly realized volatilities thus they carry higher weights. However, crude oil and gold markets are the most liquid commodity markets, and thus more information is disseminated on daily basis.

4. Volatility forecasting

Long memory models have demonstrated superiority on forecasting realized volatility compared to standard volatility predictors (Bollerslev and Wright (2001)). We compare next the forecasting ability of the two long-memory models considered in this study, namely FRSV and HAR models.[8] We consider three forecast horizons, 1 day, 1 week (5-day) and 1 month (22-day) for out-of-sample analysis on a rolling window of 1000 days in which the models are re-estimated daily, weekly, and monthly accordignly. The forecasting performances are assessed using the Mean Square Error (MSE) and the P-ratio between the MSE of the corresponding predictor and the approximate variance of the log-volatility.[9]

4.1. Rough fractional stochastic volatility (RFSV) predictors

Given that for all commodities $H < 1/2$ (see Table 2), the fractional stochastic volatility models used in the analysis are denoted as Rough Fractional Stochastic Volatility models (RFSV). Accordingly, the RFSV predictors are constructed using the expression

$$\mathbb{E}[W_t^H|\mathcal{F}_t] = \frac{\cos(H\pi)}{\pi}\Delta^{H+1/2}\int_{-\infty}^{t}\frac{W_t^H}{(t-s+\Delta)(t-s)^{H+1/2}}ds, \tag{4.13}$$

[8]We compare only RFSV and HAR predictors. Other forecasting predictors such as Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models have been used in the literature, but they have shown no evidence of significantly outperforming compared to HAR and RFSV models, see Andersen, Bollerslev, Christoffersen, and Diebold (2006) and Gatheral et al. (2018).

[9]The $P-$ratio is evaluated by

$$P-ratio = \frac{\sum_{k=1000}^{N-\Delta}(\log RV_{k+\Delta} - \overline{\log RV_{k+\Delta}})^2}{\sum_{k=1000}^{N-\Delta}(\log RV_{k+\Delta} - \mathbb{E}[\log RV_{t+\Delta}])^2},$$

with $\mathbb{E}[\log RV_{t+\Delta}]$ be the empirical mean of the log-volatility (Gatheral et al. (2018)).

Electronic copy available at: https://ssrn.com/abstract=3520500

where $\mathcal{F}_t$ is the filtration generated by the fractional BM W_t^H with $H < 1/2$. For suitable constants ν and C, the log volatility can be approximated by $\log \sigma_t^2 \approx 2\nu W_t^H + C$ (Gatheral et al. (2018)), which reduces the prediction formula for log-variance to

$$\mathbb{E}[\log RV_{t+\Delta}|\mathcal{F}_t] = \frac{\cos(H\pi)}{\pi}\Delta^{H+1/2}\int_{-\infty}^{t}\frac{\log RV_s}{(t-s+\Delta)(t-s)^{H+1/2}}ds. \tag{4.14}$$

Note that the constants ν and C cancelled out in the derivations. The forecast frequency is controlled by Δ in equation (4.14).

4.2. HAR predictors

Accordingly, the HAR predictors are constructed as follows:[10]

$$\log(RV_{t+\Delta}) = \beta_0^{\Delta} + \beta_D^{\Delta}\log(RV_t) + \beta_W^{\Delta}\frac{1}{4}\sum_{i=1}^{4}\log(RV_{t-i}) + \beta_M^{\Delta}\frac{1}{17}\sum_{i=5}^{21}\log(RV_{t-i}). \tag{4.15}$$

4.3. Out-of-sample predictions

Table 4 presents the results of out-of-sample daily, weekly, and monthly predictions. Out-of-sample forecasting results are based on a rolling forecasting scheme with a $1,000$-day rolling window. We find that HAR models consistently outperform the RFSV models across all commodity classes and for all forecasting horizons. To further assess the statistical significance of the out-of-sample predictability, we employ the Diebold and Mariano (1995) (DM) tests and the results are reported in Table 5 (see first panel). A positive statistics indicates that HAR model outperforms the RFSV model and insignificant cases are reported in grey. Most of the commodities have positive and statistical significant DM statistics. In summary, HAR model forecast better in short and longer horizons. RFSV model do not provide evidence of comparable forecasting ability in relation to standard HAR models. Thus, for forecasting volatility in commodity markets, HAR and its extension which account

[10]Out-of-sample forecasting for HAR model is carried out using the available package in R for out of sample forecast (*HARForecast*). This package allows one to change the forecasting frequency (i.e. h=1,h=5, h=22; for daily, weekly and monthly respectively).

Electronic copy available at: https://ssrn.com/abstract=3520500

for the decomposition of positive and negative parts, as well as jumps, would be effective forecasting models (see for example Sévi (2014)).

5. Robustness tests

As a robustness test and to partially address the financialisation of commodity markets, we split our dataset in 2004 and conduct the analysis over two sample periods.[11]

5.1. Volatility dynamics

As Table A.1, which presents the RV summary statistics, reveals that there are notable differences between the two sub-periods: 1996-2004 and 2005-2019. After 2004 in the energy markets, the mean realized volatility has reduced but its standard deviation has increased and the market has exhibited more extreme values. In the second sub-period 2005-2019, most metals and softs markets experience lower mean and standard deviation of the realized volatility, while most grains and livestock markets have much higher realized volatility mean and standard deviation. Exemptions include silver, orange juice and sugar, where an increase in realized volatility has observed after 2004. Consistently, all commodity markets display a substantial increase in the volatility persistence in the second sub-period.

The last two panels of Table 2 display the Hurst parameters in the two sub-periods 1996-2004 and 2005-2019. Table 2 reveals an increase in the Hurst parameter estimates after 2004 for most of commodities, with the most substantial increase appearing in the energy markets and in particular crude oil. This is a shared feature with equity indexes markets. Indeed, after 2004, the Hurst parameter estimates for energy commodities has increased and ranges from 0.122 to 0.159, which is very close to the Hurst parameter estimates obtained for equity indices markets from the Oxford-Man Dataset in the time period which includes the global financial crisis (Gatheral et al. (2018)). Thus, since 2005 energy markets are becoming more

[11] There is empirical evidence of a structural break on 2004 (Irwin and Sanders (2011), Tang and Xiong (2012) and Hamilton and Wu (2014)), thus we choose to split the sample in December 2004 which marks the initiation of the financialisation of the commodity markets.

Electronic copy available at: https://ssrn.com/abstract=3520500

integrated with financial markets (Chiang et al. (2015), Chiarella, Kang, Nikitopoulos, and Tô (2016) and Baur and Dimpfl (2018)) compared to other commodity markets. In general, Hurst parameters for all markets are less than 0.5 justifying rough volatility in commodity markets.

Table 3 presents the estimated parameters of the HAR model in the two sub-periods 1996-2004 and 2005-2019 (see the last two panels of this table). The general trend is that after 2004, there is an increase of the component weight across most of time horizons. Most specifically, in the energy market, the weekly volatility maintains the higher information share but with much higher weights since 2004, except oil markets where the daily volatility is more informative and with higher weight after 2004. Most of the softs and livestock have shifted their stronger weights from weekly (or monthly) to daily, while there are no changes (before and after 2004) for most grains. Silver maintains the highest weight in monthly volatility, while gains maintain the highest weight in weekly volatility (except rice where the monthly coefficient is higher after 2004).

In summary, weekly volatility tends to be more important and more informative than short-term volatility in most commodity markets, with the exception of crude oil, gold, coffee, cotton, sugar, lean hogs and feeder cattle markets where daily volatility carries more information in particular after 2004. Following the financialization, the increased liquidity in softs and livestock could have also been associated to commercial traders targeting protection in longer maturity hedges.

5.2. Volatility forecasting

Appendix B compares out-of-sample predictions in the two sub-periods 1996-2004 and 2005-2019, see Table B.2 and Table B.3 respectively. In general, HAR model outperforms the RFSV model in both sub-periods with further improvement after 2004. Table 5 displays the DM statistics in two sub-periods 1996-2004 and 2005-2019, see second and third panel, respectively. All commodities have positive and significant DM statistics for daily, weekly and monthly forecasts. Before 2004, the results for several monthly predictions are not

Electronic copy available at: https://ssrn.com/abstract=3520500

significant but, after 2004, predictions for all commodities and maturities are significant (expect palladium daily predictions). Thus, even in the two sample periods, HAR models consistently outperform the RFSV models, with stronger results after 2004.

6. Conclusion

This paper presents an empirical study on the realized volatility characteristics of commodity markets by using two long-memory models and comparing their forecasting ability. We use high-frequency intra-day futures prices for 22 commodities and we employ fractional stochastic volatility and HAR models. The analysis confirms that the volatility of commodity markets is rough and features long memory with anti-persistence. For energy commodities the Hurst parameter varies between 0.11 and 0.16, while for the other commodities is about half of these values.

Further, volatility components over different horizons are economically and statistically significant. Weekly volatility matters in most commodity markets, while daily volatility is more informative for oil and gold. Out-of-sample predictions of HAR models are consistently performing better compared to the fractional stochastic volatility specifications (over the whole sample and two sub-sample separated in 2004). More sophisticated HAR models decomposing negative and positive shocks, as well as continuous from discontinuous parts have demonstrated superiority in volatility forecasting. Even though fractional volatility models are successful in detecting the long memory characteristics of commodity markets, HAR models can forecast better its volatility on daily, weekly and monthly basis.

The financialisation of commodity futures markets impacted the markets in several ways. Trading volumes and open interest of commodity futures skyrocketed and a remarkable change in market participants, such as emerging passive index investors and institutional investors led to the integration of commodity and financial markets (Irwin and Sanders (2012), Basak and Pavlova (2016) and Kang et al. (2019)). However, the idiosyncratic features of each commodity market has also played an important role on altering the level

Electronic copy available at: https://ssrn.com/abstract=3520500

and variability of each market's volatility. Typically, weather patterns impact grains and softs markets, whereas other macro-economic effects such as Brazil's sugar based ethanol production and the recent trade tensions between US and China have plunged the prices of key commodities such as copper and soybean. For instance, corn's increasing use for fuel ethanol production and soybean oil's use for biodiesel led to tight supplies in 2010-2011.[12] The volatility of commodity markets is the product of the interactions between physical commodity and futures markets and disentangling the contribution of fundamentals from financial and macroeconomic factors merits further research which will provide useful insights to commodity volatility forecasting.

[12] See https://www.agmrc.org/renewable-energy/renewable-energy-climate-change-report/november-2010-newsletter/corn-and-soybean-availability-for-biofuels-in-2010-11

Electronic copy available at: https://ssrn.com/abstract=3520500

References

Alfeus, M., Overbeck, L., Schlögl, E., 2019. Regime switchin rough Heston model. Journal of Futures Markets 39 (5), 538–552.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., Diebold, F., 2006. Volatility and correlation forecasting. In: Elliot, G., Granger, C. W. J., Timmermann, A. (Eds.), Handbook of the Economic Forecasting. Elsevier.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., Diebold, F., 2013. Financial risk measurement for financial risk management. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. Elsevier: Amsterdam.

Andersen, T. G., Bollerslev, T., Diebold, F., 2007. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. Review of Economics and Statistics 89 (4), 701–720.

Andersen, T. G., Bollerslev, T., Diebold, F., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2009. Realized kernels in practice: Trades and quotes. Econometrics Journal 12 (3), C1–C32.

Barndorff-Nielsen, O. E., Kinnebrock, S., Shephard, N., 2010. Measuring downside risk – Realised semivariance. In: Bollerslev, T., Russell, J., Watson, M. (Eds.), Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle. New York:Oxford University Press.

Basak, S., Pavlova, A., 2016. A model of financialization of commodities. Journal of Finance LXXI (4), 1511–1555.

Baur, D. G., Dimpfl, T., 2018. The asymmetric return-volatility relationship of commodity prices. Energy Economics 76, 378–387.

Bayer, C., Friz, J., Gatheral, J., 2016. Pricing under rough volatility. Quantitative Finance 16 (6), 887–904.

Bennedsen, M., 2017. A rough multi-factor model of electricity spot prices. Energy Economics 63, 301–313.

Bennedsen, M., Lunde, A., Pakkanen, M. S., 2017. Decoupling the short- and long-term behavior of stochastic volatility. CREATES Research Papers 2017-26, department of Economics and Business Economics, Aarhus University.

Bollerslev, T., Hood, B., Huss, J., Pedersen, L. H., 2018. Risk everywhere: Modeling and managing volatility. Review of Financial Studies 31 (7), 2729–2773.

Electronic copy available at: https://ssrn.com/abstract=3520500

Bollerslev, T., Wright, J. H., 2001. High-frequency data, frequency domain inference and volatility forecasting. Review of Economics and Statistics 83, 596–602.

Chiang, I., Hughen, W. K., Sagi, J. S., 2015. Estimating oil risk factors using information from equity and derivatives markets. The Journal of Finance 70, 769–804.

Chiarella, C., Kang, B., Nikitopoulos, C. S., Tô, T., 2016. The return-volatility relation in commodity futures markets. Journal of Futures Markets 36 (2), 127–152.

Comte, F., Renault, E., 1998. Long memory in continuous-time stochastic volatility models. Mathematical Finance 8 (4), 291–323.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics 7 (2), 174–196.

Cunado, J., Gil-Alana, L., De Gracia, F. P., 2010. Persistence in some energy futures markets. The Journal of Futures Markets 30 (5), 490–507.

Da Fonseca, J., Zhang, W., 2019. A rough multi-factor model of electricity spot prices. The Journal of Futures Markets 39 (5), 600–611.

Degiannakis, S., Filis, G., 2018. Forecasting oil prices: High-frequency financial data are indeed useful. Energy Economics 76, 388–402.

Dickey, D. A., Fuller, W. A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74 (366a), 427–431.

Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. Journal of Business & Economic Statistics 13, 253–263.

El Euch, O., Fukasawa, M., Rosenbaum, M., 2018. The microstructural foundations of leverage effect and rough volatility. Finance and Stochastics 22 (2), 241–280.

Elder, J., Serletis, A., 2008. Long memory in energy futures prices. Review of Financial Economics 17 (2), 146–155.

Fukasawa, M., 2017. Short-time at-the-money skew and rough fractional volatility. Quantitative Finance 17 (2), 189–198.

Fukasawa, M., Takabatake, T., Westphal, R., 2019. Is Volatility Rough ? arXiv preprint arxiv.org/abs/1905.04852.

Gatheral, J., Jaisson, T., Rosenbaum, M., 2018. Volatility is rough. Quantitative Finance 18 (6), 933–949.

Glaserman, P., He, P., 2018. Buy rough, sell smooth. $https://papers.ssrn.com/sol3/papers.cfm?abstract_id = 3301669$.

Electronic copy available at: https://ssrn.com/abstract=3520500

Gong, X., Lin, B., 2017. Forecasting the good and bad uncertainties of crude oil prices using HAR framework. Energy Economics 67, 315–327.

Granger, C., 1980. Long memory relationships and the aggregation of dynamic models. Journal of Econometrics 14, 227–238.

Hamilton, J. D., Wu, J. C., 2014. Risk premia in crude oil futures prices. Journal of International Money and Finance 42, 9–37.

Hansen, P., Lunde, A., 2005. A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? Journal of Applied Econometrics 20 (7), 873–889.

Hansen, P., Lunde, A., 2014. Estimating the persistence and the autocorrelation function of a time series that is measured with error. Econometric Theory 30, 60–93.

Irwin, S., Sanders, D., 2011. Index funds, financialization, and commodity futures markets. Applied Economic Perspective and Policy 33, 1–31.

Irwin, S., Sanders, D., 2012. Financialization and stuctural change in commodity futures markets. Journal of Agricultural and Aplied Economics 44 (3), 371–396.

Kang, B., Nikitopoulos, C., Prokopczuk, M., 2019. Oil futures volatility and the economy. $https://papers.ssrn.com/sol3/papers.cfm?abstract_id = 3417706$.

LeBaron, B., 2001. Stochastic volatility as a simple generator of apparent financial power laes and long memory. Quantitative Finance 1 (6), 621–631.

Liu, L. Y., Patton, A. J., Sheppard, K., 2015. Does anything beat 5-minute RV? a comparison of realised measures across multiple asset classes. Journal of Econometrics 187 (1), 293–311.

Livieri, G., Mouti, S., Rosenbaum, M., 2017. Rough volatility: Evidence from option prices. arxiv preprint arXiv:1702.02777.

Lo, A. W., 1991. Long-term memory in stoch market prices. Econometrica 59 (5), 1279–1313.

Luo, J., Chen, L., 2019. Multivariate realized volatility forecasts of agricultural commodity futures. Journal of Futures Markets 39 (12), 1565–1586.

Mandelbrot, B. B., Van Ness, J. W., 1968. Fractional brownian motions, fractional noises and applications. SIAM Review 10 (4), 422–437.

Müller, U., Dacorogna, M., Dav, R., Olsen, R., Pictet, O., von Weizsacker, J., 1997. Volatilities of different resolutions - Analysing the dynamics of market components. Journal of Empirical Finance 4, 213–239.

Newely, W. K., West, K. D., 1987. A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Electronic copy available at: https://ssrn.com/abstract=3520500

Patton, A. J., Sheppard, K., 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. Review of Economics and Statistics 97, 683–697.

Sévi, B., 2014. Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models. European Journal of Operational Research 235, 643–659.

Takaishi, T., 2019. Rough volatility of Bitcoin. arXiv preprint arxiv.org/pdf/1904.12346.

Tang, K., Xiong, W., 2012. Index investment and the financialisation of commodities. Financial Analysts Journal 68, 54–74.

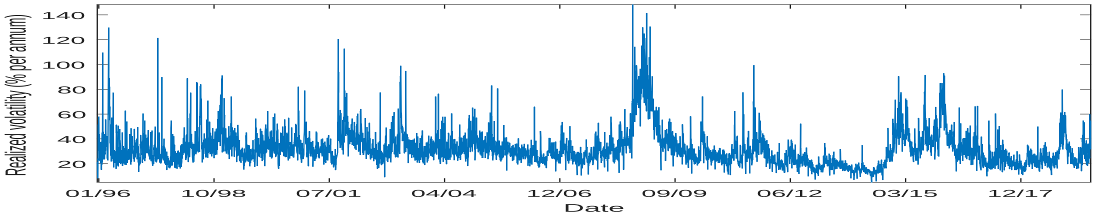
Wang, Y., Wu, C., 2012. Long memory in energy futures prices: Further evidence. Resources Policy 37 (3), 261–272.

Appendix A. Summary Statistics: 1996-2004 and 2005-2019

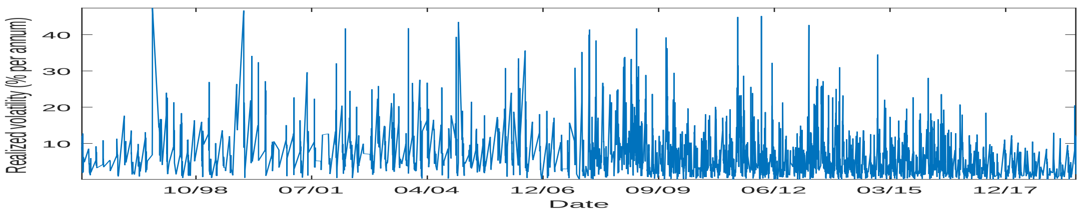
Appendix B. Out-of-sample predictions 1996-2004

Tables B.2 and Tables B.3 report the out-of-sample performance of the two volatility forecasting models, HAR and FRSV, over two sample periods 1996–2004 and 2005–2019.

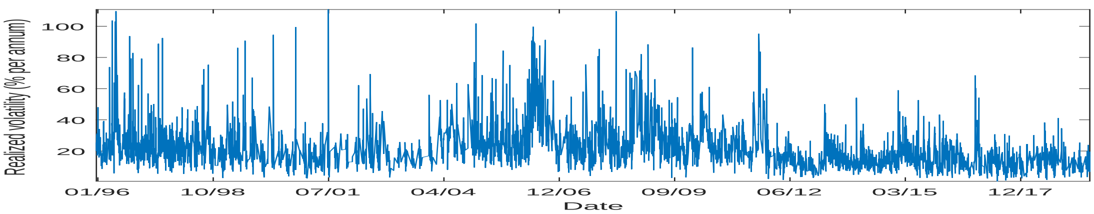
Electronic copy available at: https://ssrn.com/abstract=3520500



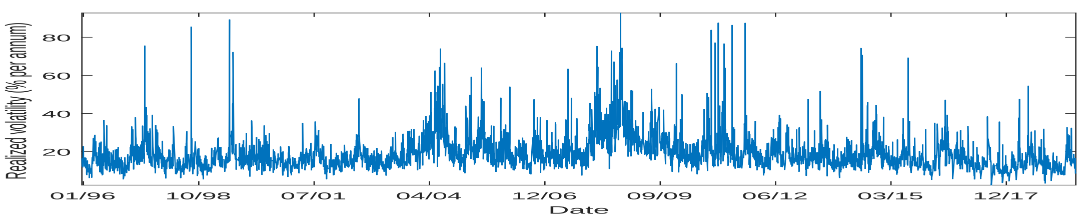
2.a: Crude oil (energy)



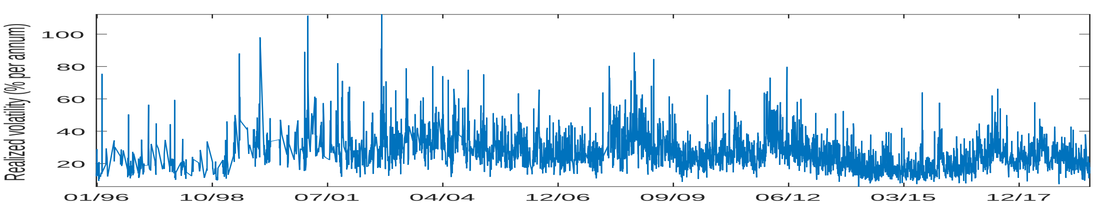
2.b: Gold (precious metals)



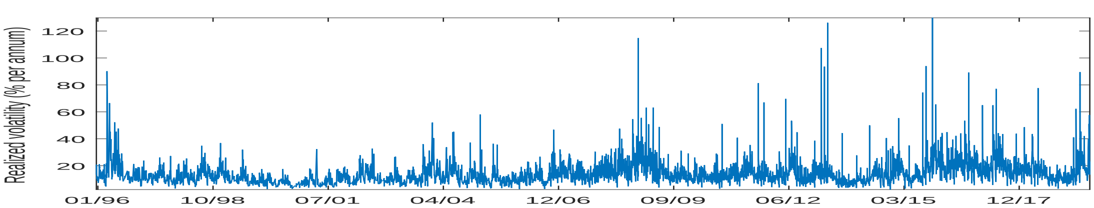
2.c: Copper (industrial metals)



2.d: Soybean (grains)



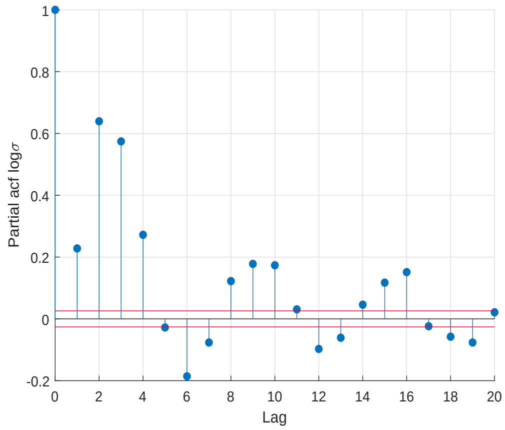
2.e: Cocoa (softs)



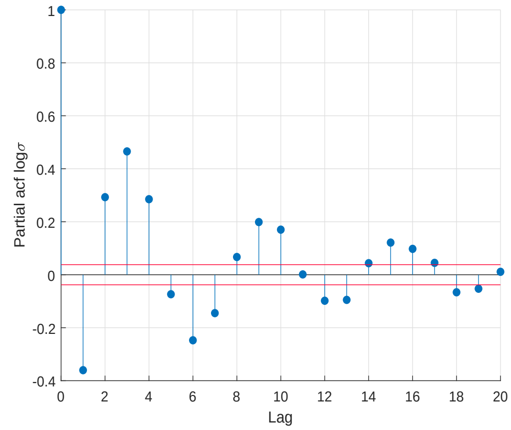
2.f: Feeder cattle (livestock)

Figure 2: **Released volatility of commodity markets**
The figure plots the time series of the realized volatilities in annualised terms for the representative commodity per class: crude oil (energy), gold (precious metals), copper (industrial metals), soybean (grains), cocoa (softs) and feeder cattle (livestock).

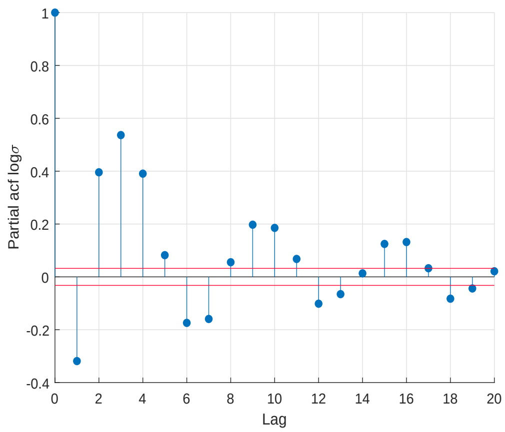
Electronic copy available at: https://ssrn.com/abstract=3520500



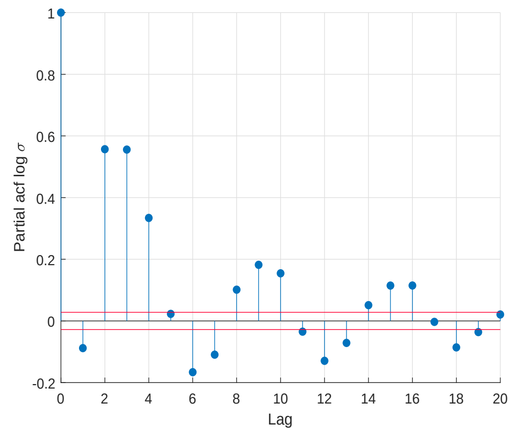
3.a: Crude oil (energy)



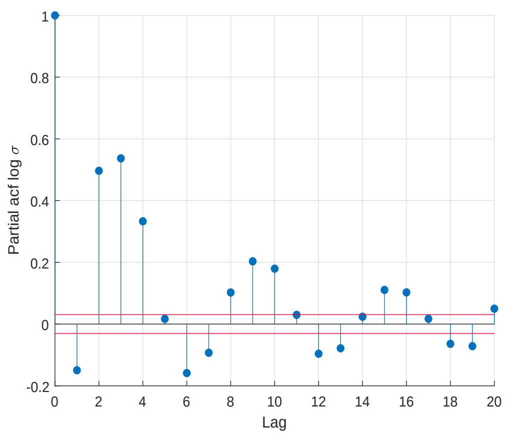
3.b: Gold (precious metals)



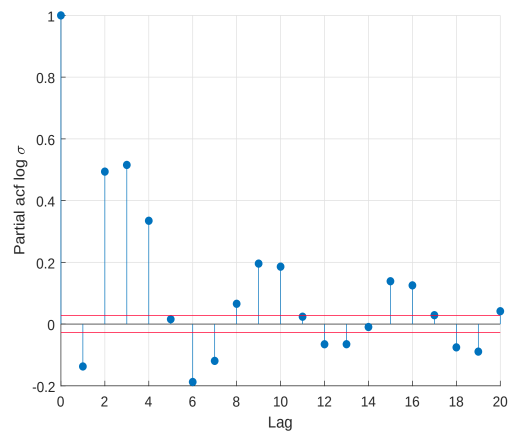
3.c: Copper (industrial metals)



3.d: Soybean (grains)



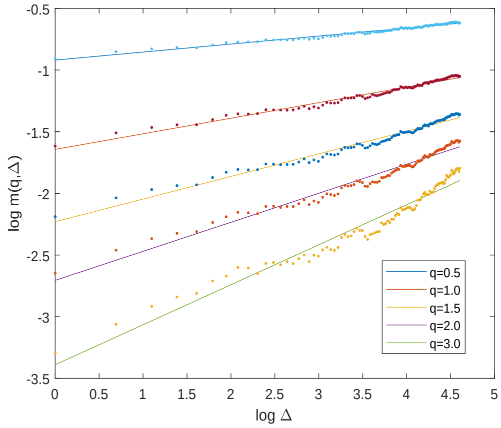
3.e: Cocoa (softs)



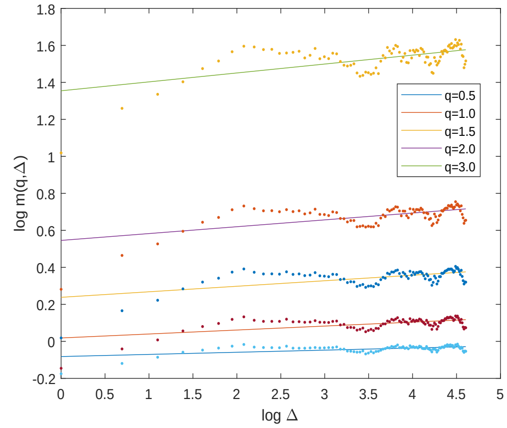
3.f: Feeder cattle (livestock)

Figure 3: Partial Autocorrelation Function
The figure depicts the partial autocorrelation function of the released volatilities for crude oil, gold, copper, soybean, cocoa and feeder cattle.

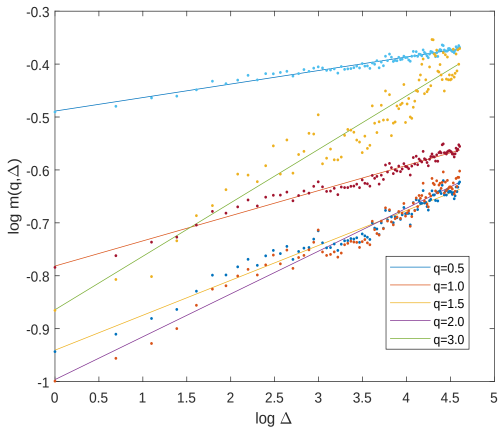
Electronic copy available at: https://ssrn.com/abstract=3520500



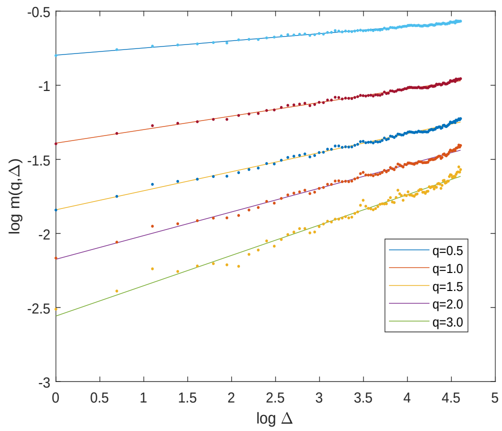
4.a: Crude oil (energy)



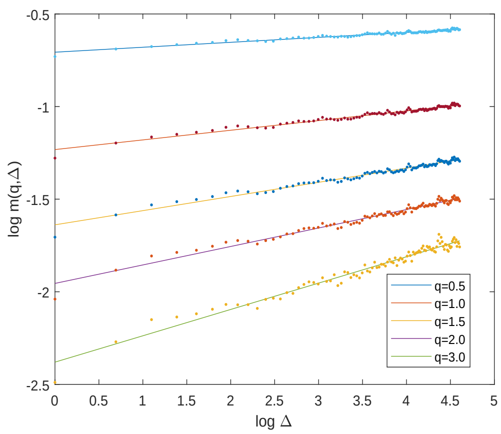
4.b: Gold (precious metals)



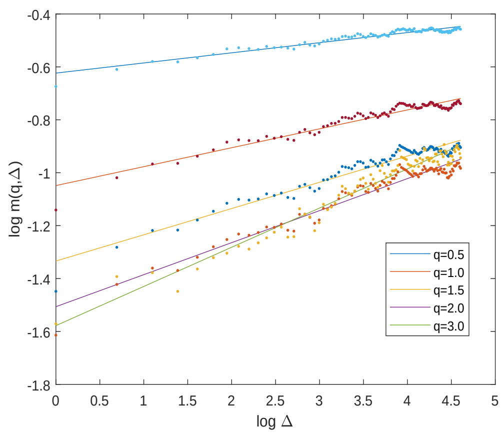
4.c: Copper (industrial metals)



4.d: Soybean (grains)



4.e: Cocoa (softs)



4.f: Feeder cattle (livestock)

Figure 4: $\log m(q, \Delta)$ **as a function of** $\log \Delta$
The figure depicts the $\log m(q, \Delta)$ as a function of $\log \Delta$ including a linear fit. The variance proxies are computed using 5-minute released variance estimates for crude oil, gold, copper, soybean, cocoa and feeder cattle.

Electronic copy available at: https://ssrn.com/abstract=3520500

Table 3: **HAR estimation**

	1996-2019					1996-2004					2005-20019				
Series	β_0	β_D	β_W	β_M	R^2	β_0	β_D	β_W	β_M	R^2	β_0	β_D	β_W	β_M	R^2
Energy															
Crude Oil (WTI)	-0.401***	0.375***	0.359***	0.216***	0.660	-1.333***	0.336***	0.321***	0.171***	0.333	-0.292***	0.422***	0.354***	0.188***	0.754
	(-5.4)	(29.5)	(20.1)	(12.5)		(-5.7)	(15.9)	(10.6)	(4.9)		(-3.8)	(26.7)	(16.1)	(9.3)	
Crude Oil (Brent)	-0.423***	0.371***	0.338***	0.239***	0.646	-1.276***	0.310***	0.318***	0.213***	0.333	-0.276***	0.433***	0.321***	0.213***	0.759
	(-5.4)	(29.0)	(18.5)	(13.3)		(-5.3)	(14.4)	(9.9)	(5.7)		(-3.6)	(27.4)	(14.7)	(10.6)	
Natural gas	-0.536***	0.276***	0.510***	0.141***	0.589	-1.024***	0.289***	0.486***	0.080**	0.436	-0.441***	0.285***	0.503***	0.153***	0.627
	(-6.6)	(20.9)	(26.2)	(7.4)		(-5.6)	(12.5)	(14.4)	(2.4)		(-4.7)	(17.8)	(21.3)	(6.6)	
Heating Oil	-0.299***	0.321***	0.371***	0.270***	0.708	-1.086***	0.276***	0.353***	0.228***	0.342	-0.280***	0.345***	0.367***	0.255***	0.739
	(-4.3)	(24.2)	(19.1)	(14.3)		(-4.6)	(11.8)	(10.1)	(5.8)		(-3.5)	(21.3)	(15.6)	(11.4)	
Gasoline	-0.334***	0.303***	0.440***	0.215***	0.704						-0.334***	0.303***	0.440***	0.215***	0.704
	(-3.6)	(17.1)	(17.1)	(8.9)							(-3.6)	(17.1)	(17.1)	(8.9)	
Metals															
Gold	-3.548***	0.308***	0.220***	0.170***	0.181	-6.364***	0.287***	-0.036	0.168*	0.0849	-3.987***	0.331***	0.219***	0.116***	0.187
	(-7.7)	(16.4)	(7.9)	(4.3)		(-4.1)	(6.8)	(-0.5)	(1.3)		(-7.4)	(15.8)	(7.1)	(2.7)	
Silver	-0.915***	0.186***	0.210***	0.499***	0.319	-1.542**	0.247***	0.065	0.515***	0.161	-0.805***	0.174***	0.291***	0.442***	0.378
	(-3.8)	(9.1)	(6.3)	(12.4)		(-2.0)	(6.3)	(1.0)	(5.3)		(-3.3)	(7.2)	(7.4)	(9.8)	
Platinum	-2.123***	0.157***	0.353***	0.254***	0.188	-4.310***	0.125***	0.376***	-0.020	0.0937	-4.000***	0.133***	0.279***	0.165*	0.081
	(-3.8)	(4.4)	(6.0)	(3.5)		(-4.1)	(2.1)	(3.9)	(-0.2)		(-3.9)	(2.9)	(3.6)	(1.5)	
Palladium	-1.653***	0.307***	0.352***	0.138**	0.304	-3.114***	0.462***	0.252***	-0.127*	0.341	-1.744***	0.190***	0.282***	0.326***	0.213
	(-3.5)	(7.7)	(6.0)	(2.1)		(-3.7)	(6.9)	(2.9)	(-1.3)		(-2.5)	(3.8)	(3.4)	(3.3)	
Copper	-1.179***	0.190***	0.378***	0.301***	0.316	-2.466***	0.161***	0.255***	0.305***	0.121	-0.969***	0.214***	0.409***	0.270***	0.393
	(-6.0)	(11.8)	(14.9)	(10.3)		(-4.4)	(5.6)	(5.4)	(4.5)		(-4.6)	(10.9)	(13.5)	(8.2)	
Grains															
Corn	-0.697***	0.284***	0.356***	0.280***	0.497	-1.971***	0.285***	0.335***	0.162***	0.266	-0.723***	0.307***	0.342***	0.266***	0.498
	(-5.7)	(19.9)	(16.6)	(12.5)		(-5.3)	(10.9)	(8.7)	(3.5)		(-5.1)	(18.0)	(13.4)	(10.0)	
Soybean	-0.819***	0.312***	0.359***	0.238***	0.487	-0.996***	0.297***	0.375***	0.219***	0.443	-0.783***	0.323***	0.347***	0.243***	0.496
	(-6.5)	(22.6)	(17.7)	(11.3)		(-4.3)	(12.9)	(11.0)	(6.1)		(-5.1)	(18.7)	(13.7)	(9.2)	
Rough rice	-0.884***	0.290***	0.233***	0.371***	0.363	-1.514***	0.340***	0.192***	0.283***	0.265	-0.654***	0.263***	0.241***	0.418***	0.426
	(-5.8)	(20.1)	(10.5)	(14.4)		(-4.7)	(14.4)	(5.4)	(6.3)		(-3.8)	(14.5)	(8.5)	(13.0)	
Wheat	-0.780***	0.279***	0.348***	0.280***	0.449	-2.192***	0.217***	0.326***	0.203***	0.189	-0.702***	0.305***	0.342***	0.267***	0.488
	(-6.1)	(19.6)	(16.2)	(12.1)		(-5.9)	(9.0)	(8.7)	(4.3)		(-4.8)	(17.2)	(12.9)	(9.6)	
Softs															
Cocoa	-0.782***	0.296***	0.277***	0.333***	0.424	-0.952***	0.330***	0.315v	0.237***	0.415	-0.792***	0.297***	0.244***	0.365***	0.411
	(-5.5)	(19.7)	(12.1)	(13.1)		(-3.2)	(10.6)	(6.9)	(4.8)		(-4.7)	(17.3)	(9.3)	(12.2)	
Coffee	-0.728***	0.301***	0.352***	0.254***	0.477	-1.182***	0.301***	0.366***	0.174***	0.355	-0.942v	0.332***	0.299***	0.252***	0.408
	(-5.2)	(17.2)	(13.6)	(9.4)		(-4.0)	(9.8)	(8.1)	(3.5)		(-4.6)	(15.6)	(9.6)	(7.4)	
Cotton	-0.737***	0.312***	0.282***	0.317***	0.451	-0.843***	0.284***	0.237***	0.379***	0.390	-0.696***	0.350***	0.315***	0.250***	0.501
	(-5.4)	(20.7)	(12.3)	(12.7)		(-3.4)	(12.0)	(6.4)	(9.0)		(-4.3)	(17.9)	(10.9)	(8.3)	
Sugar	-0.465***	0.257***	0.347***	0.339***	0.566	-0.538***	0.227***	0.289***	0.420***	0.497	-0.470***	0.336***	0.330***	0.276***	0.591
	(-4.3)	(16.9)	(14.9)	(14.0)		(-2.4)	(8.3)	(6.6)	(8.8)		(-3.8)	(18.5)	(12.4)	(10.3)	
Orange juice	-1.233***	0.239***	0.308***	0.295***	0.288	-3.218***	0.260***	0.233***	0.101	0.126	-0.974***	0.258***	0.308***	0.309***	0.348
	(-6.0)	(13.2)	(10.8)	(8.8)		(-4.6)	(6.7)	(3.9)	(1.2)		(-4.7)	(12.5)	(9.6)	(8.4)	
Livestock															
Feeder cattle	-1.112***	0.325***	0.309***	0.251***	0.405	-1.364***	0.365***	0.373***	0.125***	0.441	-1.072***	0.316***	0.258***	0.313***	0.376
	(-7.05)	(24.0)	(15.4)	(11.3)		(-5.0)	(15.2)	(11.0)	(3.6)		(-5.4)	(19.2)	(10.4)	(11.0)	
Live cattle	-0.973***	0.238***	0.391***	0.268***	0.411	-2.039***	0.277***	0.397***	0.111*	0.282	-0.866***	0.238***	0.390***	0.281***	0.440
	(-5.3)	(14.0)	(15.0)	(9.3)		(-2.9)	(5.9)	(5.8)	(1.4)		(-4.6)	(13.0)	(13.7)	(9.1)	
Lean hogs	-1.133***	0.246***	0.373***	0.250***	0.356	-0.782***	0.255***	0.487***	0.167***	0.506	-1.176***	0.320***	0.284***	0.260***	0.351
	(-7.5)	(18.0)	(18.1)	(11.0)		(-4.1)	(11.2)	(14.3)	(4.9)		(-6.0)	(19.1)	(11.5)	(9.1)	

The table reports the HAR model estimates, where $\log(RV^h_{t+h}) = \beta_0 + \beta_D \log(RV_t) + \beta_W \log(RV^W_t) + \beta_M \log(RV^M_t) + \epsilon_t$. We report t-statistics in parentheses and *, **, *** indicate significance at the 10%, 5% and 1% significance levels, respectively. Newely and West (1987) corrected standard errors with 12 lags have been used in the estimations.

Electronic copy available at: https://ssrn.com/abstract=3520500

Table 4: **Out-of-sample predictions: 1996 − 2019**

	HAR						*FRSV*					
	$1-day$		$5-day$		$22-day$		$1-day$		$5-day$		$22-day$	
Series	$P-ratio$	MSE	$P-ratio$	MSE	$P-ratio$	MSE	$P-ratio$	MSE	$P-ratio$	MSE	$P-ratio$	MSE
Energy												
Crude Oil (WTI)	**0.309**	**0.198**	**0.411**	**0.264**	**0.554**	**0.357**	0.339	0.220	0.436	0.287	0.591	0.401
Crude Oil (Brent)	**0.625**	**0.585**	**0.752**	**0.702**	**0.874**	**0.813**	0.694	0.694	0.796	0.816	0.891	0.933
Natural gas	**0.402**	**0.242**	**0.477**	**0.287**	**0.693**	**0.418**	0.484	0.296	0.502	0.316	0.762	0.509
Heating Oil	**0.268**	**0.178**	**0.337**	**0.223**	**0.446**	**0.296**	0.302	0.202	0.361	0.244	0.478	0.331
Gasoline	**0.357**	**0.199**	**0.461**	**0.257**	**0.640**	**0.358**	0.394	0.222	0.494	0.283	0.662	0.392
Metals												
Gold	**0.863**	**4.097**	**0.992**	**4.777**	**0.949**	**4.562**	0.962	7.656	1.056	9.006	1.005	9.303
Silver	**0.674**	**1.499**	**0.693**	**1.540**	**0.744**	**1.670**	0.818	2.224	0.793	2.282	0.819	2.594
Platinum	**0.948**	**2.174**	**0.989**	**2.313**	**1.045**	**2.527**	1.087	3.190	1.018	3.136	1.063	3.496
Palladium	**1.044**	**1.294**	**1.017**	**1.313**	**1.048**	**1.533**	1.167	1.974	1.066	2.026	1.036	2.446
Copper	**0.628**	**0.868**	**0.701**	**0.969**	**0.752**	**1.041**	0.745	1.159	0.761	1.226	0.800	1.347
Grains												
Corn	**0.506**	**0.318**	**0.593**	**0.373**	**0.721**	**0.453**	0.564	0.365	0.633	0.419	0.744	0.505
Soybean	**0.497**	**0.305**	**0.595**	**0.364**	**0.742**	**0.455**	0.552	0.349	0.634	0.410	0.766	0.511
Rough rice	**0.614**	**0.633**	**0.680**	**0.703**	**0.741**	**0.766**	0.706	0.804	0.745	0.872	0.772	0.937
Wheat	**0.546**	**0.304**	**0.629**	**0.350**	**0.737**	**0.411**	0.608	0.350	0.659	0.385	0.756	0.450
Softs												
Cocoa	**0.599**	**0.329**	**0.678**	**0.373**	**0.750**	**0.414**	0.670	0.388	0.726	0.429	0.778	0.476
Coffee	**0.619**	**0.251**	**0.739**	**0.301**	**0.868**	**0.352**	0.683	0.288	0.781	0.339	0.891	0.400
Cotton	**0.574**	**0.452**	**0.663**	**0.523**	**0.801**	**0.627**	0.647	0.540	0.710	0.610	0.826	0.730
Sugar	**0.482**	**0.254**	**0.561**	**0.296**	**0.640**	**0.338**	0.542	0.296	0.598	0.331	0.667	0.378
Orange juice	**0.717**	**0.443**	**0.799**	**0.493**	**0.908**	**0.563**	0.821	0.554	0.853	0.596	0.925	0.680
Livestock												
Feeder cattle	**0.625**	**0.585**	**0.752**	**0.702**	**0.874**	**0.813**	0.694	0.694	0.796	0.816	0.891	0.933
Live cattle	**0.574**	**0.299**	**0.655**	**0.342**	**0.795**	**0.418**	0.659	0.355	0.687	0.379	0.809	0.464
Lean hogs	**0.694**	**0.476**	**0.818**	**0.562**	**0.915**	**0.630**	0.771	0.557	0.862	0.641	0.941	0.727

The table displays the out-of-sample predictions results for 1-day, 5-day and 22-day realized volatility predictions for two models; HAR and RFSV.

Electronic copy available at: https://ssrn.com/abstract=3520500

Table 5: **DM statistics**

	1996 -2019			1996-2004			2005-2019		
	1-day	5-day	22-day	1-day	5-day	22-day	1-day	5-day	22-day
Energy									
Crude Oil (WTI)	8.918	6.036	3.860	6.197	3.093	3.054	6.535	4.552	2.674
Crude Oil (Brent)	8.400	5.793	3.932	5.717	3.552	2.289	6.146	4.831	2.611
Natural gas	11.065	6.061	5.105	6.407	3.282	2.587	9.542	3.455	4.284
Heating Oil	8.909	6.933	4.003	5.318	3.958	2.196	5.850	4.826	3.087
Gasoline	4.454	4.256	2.230				4.454	4.256	2.230
Metals									
Gold	23.116	14.988	18.338	6.926	4.594	2.993	20.223	13.369	15.091
Silver	13.083	11.241	9.051	1.743	0.745	1.565	8.554	8.453	6.971
Platinum	6.963	5.972	3.368	3.021	2.397	21.918	4.044	3.272	3.445
Palladium	5.020	5.564	3.326	2.109	3.488	0.977	1.220	8.670	3.678
Copper	14.169	9.460	6.078	5.743	5.159	2.866	11.527	8.642	6.845
Grains									
Corn	8.344	6.622	3.022	4.338	3.283	1.548	7.922	6.374	4.729
Soybean	9.865	7.786	3.434	5.401	3.819	1.376	7.424	7.276	5.154
Rough rice	13.898	9.890	5.989	7.331	5.079	4.008	9.633	7.854	4.666
Wheat	8.991	5.094	2.654	6.537	3.693	2.253	7.495	5.640	3.564
Softs									
Cocoa	10.574	8.105	4.644	3.747	2.732	1.972	7.922	6.469	3.384
Coffee	6.546	5.998	3.668	3.748	3.581	5.357	3.446	3.234	2.029
Cotton	9.917	7.561	3.525	6.696	4.960	2.028	5.801	5.327	2.804
Sugar	7.155	5.841	2.983	4.829	3.309	0.517	7.425	5.519	2.110
Orange juice	10.114	7.423	4.259	4.291	3.964	2.371	8.011	7.263	4.346
Livestock									
Feeder cattle	12.797	8.030	3.580	5.780	2.826	1.180	11.719	8.463	4.049
Live cattle	8.627	4.842	2.524	2.602	0.482	0.701	7.828	3.576	2.067
Lean hogs	9.752	7.645	4.341	7.086	3.006	1.393	8.620	5.785	3.405

The table displays the Diebold and Mariano (1995) statistic with three forecast windows, 1-day, 5-day and 22-day for commodity markets. A positive statistic indicates that HAR outperforms the FRSV model. The statistic is computed using Newely and West (1987). No significant statistics are in grey.

Electronic copy available at: https://ssrn.com/abstract=3520500

Table A.1: **Commodities RV summary statistics**

Series	Mean	Stand. Dev.	Skewness	Excess Kurtosis	max	min	Ljung-Box 1-day	Ljung-Box 20-day	ADF	no. of days
				1996-2004						
Energy										
Crude oil (WTI)	34.3	12.0	2.2	9.0	129.7	8.3	536.9	2837.7	-7.86^{**}	2,180
Crude oil (Brent)	29.9	11.1	2.1	8.1	115.8	6.6	486.1	3559.6	-7.18^{**}	2,087
Natural gas	50.4	20.2	1.8	5.3	186.8	17.4	543.1	4473.3	-6.59^{**}	1,847
Heating oil	36.9	11.2	1.6	5.2	111.2	16.5	444.2	3121.1	-6.36^{**}	1,827
Metals										
Gold	8.8	7.0	2.1	6.3	47.3	0.4	22.5	58.6	-7.73^{**}	556
Silver	22.3	15.2	1.6	3.8	103.1	1.5	81.3	483.6	-5.77^{**}	658
Platinum	27.2	15.1	1.5	3.0	93.6	1.9	15.4	90.8	-4.82^{**}	308
Palladium	43.3	31.0	1.4	1.7	155.8	4.9	80.0	400.8	-3.11^{**}	221
Copper	22.4	13.3	2.6	10.7	110.7	2.6	35.5	275.4	-8.14^{**}	1,223
Grains										
Corn	18.4	6.6	2.5	15.0	91.7	7.8	241.3	1331.9	-6.83^{**}	1,436
Soybean	17.6	7.7	2.9	15.9	89.3	5.5	562.6	6171.8	-4.95^{**}	1,871
Wheat	22.5	8.5	3.7	35.6	149.2	8.0	142.0	922.6	-7.70^{**}	1,739
Rough Rice	30.1	18.6	2.9	13.5	156.6	4.1	328.17	916.5	-9.31^{**}	1,718
Softs										
Cocoa	30.5	13.7	1.5	4.6	112.2	9.4	250.3	1515.2	-6.65^{**}	992
Coffee	42.3	16.5	1.2	1.9	111.6	12.8	271.2	2235.9	-5.21^{**}	1,067
Cotton	25.8	14.5	1.9	5.4	110.1	5.7	284.8	2762.6	-7.30^{**}	1,740
Orange juice	33.2	17.6	2.6	10.0	151.7	8.6	52.8	148.2	-6.79^{**}	676
Sugar	25.1	12.1	1.8	7.2	121.0	6.7	334.1	4171.1	-5.19^{**}	1,344
Livestock										
Live cattle	14.6	5.7	2.1	7.2	48.3	4.8	85.5	538.1	-4.70^{**}	468
Feeder cattle	12.2	6.3	3.0	21.0	90.1	2.8	482.9	3790.0	-6.12^{**}	1,700
Lean hogs	23.6	11.4	2.2	8.2	105.9	6.5	679.5	8671.6	-5.18^{**}	1,937
				2005-2019						
Energy										
Crude oil (WTI)	30.8	15.3	2.3	7.78	148.3	5.0	2630.4	40188.8	-4.14^{**}	3,756
Crude oil (Brent)	28.8	13.8	2.1	6.2	125.4	6.7	2641.8	40506.7	-4.05^{**}	3,767
Natural gas	41.1	17.0	1.6	4.8	181.6	6.9	1785.3	21656.2	-6.43^{**}	3,706
Heating oil	28.2	12.6	1.6	3.8	105.0	5.5	2207.7	34882.6	-4.14^{**}	3,700
Gasoline	31.0	14.8	2.3	7.8	136.8	6.4	1996.9	31765.4	-3.94^{**}	3,161
Metals										
Gold	6.6	6.4	1.97	5.2	45.1	0.1	412.4	1194.3	-11.70^{**}	2,211
Silver	27.2	18.2	1.5	2.9	109.5	1.0	445.3	5486.9	-5.93^{**}	1,757
Platinum	18.1	12.1	2.0	7.0	95.0	0.1	40.9	159.7	-6.54^{**}	504
Palladium	29.0	21.2	2.5	8.9	154.4	1.1	61.4	535.2	-5.46^{**}	419
Copper	21.0	13.3	2.1	6.5	109.5	0.6	729.4	8852.5	-7.43^{**}	2,631
Grains										
Corn	25.0	11.6	2.2	7.8	111.8	7.0	854.1	9967.4	-6.93^{**}	3,393
Soybean	19.9	9.2	2.4	10.1	92.7	2.4	994.1	11965.1	-7.03^{**}	3,274
Wheat	28.8	13.2	2.8	14.0	157.0	5.9	828.3	9184.7	-6.31^{**}	3,102
Rough Rice	28.3	16.8	2.3	8.14	154.1	3.6	680.1	8349.2	-5.62^{**}	2,981
Softs										
Cocoa	25.7	9.9	1.3	3.4	88.6	5.8	884.0	9051.9	-6.89^{**}	3,306
Coffee	30.1	10.6	1.7	5.5	104.4	11.8	545.0	4191.8	-6.49^{**}	3,223
Cotton	28.4	14.2	1.9	5.0	110.1	6.7	948.8	11124.3	-5.53^{**}	2,500
Orange juice	35.8	16.1	1.7	4.7	151.7	8.9	486.3	5112.3	-6.35^{**}	2,336
Sugar	30.0	11.6	1.5	5.3	124.4	8.6	1215.6	14563.0	-5.81^{**}	2,955
Livestock										
Live cattle	15.1	5.9	1.5	3.9	49.2	4.2	762.2	8547.6	-5.60^{**}	2,968
Feeder cattle	15.2	9.1	3.9	30.6	129.7	2.3	409.1	4643.0	-7.43^{**}	3,583
Lean hogs	22.7	10.9	2.4	10.1	104.1	3.8	565.3	5613.8	-7.37^{**}	3,453

The table presents summary statistics for daily realized volatilities of commodities. Volatility is reported as an annualised percentage. The number of lagged changes p is selected based on the Akaike Information Criterion, testing up to a maximum of $p = 15$ lags. The symbols ***, **, and * indicate rejection of the unit-root null hypothesis at the 99%, 95%, and 90% confidence levels, respectively.

Electronic copy available at: https://ssrn.com/abstract=3520500

Table B.2: **Out-of-sample predictions: 1996–2004**

	HAR						RSFV-RV					
	1-day		5-day		22-day		1-day		5-day		22-day	
Series	Pratio	MSE	Pratio	MSE	Pratio	MSE	Pratio	MSE	Pratio	MSE	Pratio	MSE
Energy												
Crude oil (WTI)	**0.659**	**0.210**	**0.824**	**0.263**	**0.969**	**0.302**	0.715	0.238	0.842	0.286	1.009	0.350
Crude oil (Brent)	**0.697**	**0.294**	**0.830**	**0.349**	**0.985**	**0.409**	0.760	0.334	0.855	0.382	1.015	0.458
Natural gas	**0.598**	**0.293**	**0.709**	**0.358**	**0.968**	**0.476**	0.699	0.353	0.732	0.395	1.039	0.586
Heating oil	**0.660**	**0.200**	**0.859**	**0.241**	**0.979**	**0.277**	0.732	0.228	0.904	0.263	1.020	0.309
Metals												
Gold	**0.946**	**2.425**	**0.992**	**2.697**	**1.003**	**4.030**	1.023	4.198	1.165	5.411	1.001	6.748
Silver	**1.022**	**1.238**	**1.149**	**1.440**	**1.151**	**1.595**	1.036	1.774	1.096	1.835	1.055	2.074
Platinum	**1.011**	**0.794**	**1.026**	**0.790**	**1.041**	**0.803**	1.190	1.165	1.122	1.081	1.110	1.140
Palladium	**1.034**	**1.231**	**0.952**	**1.165**	**1.000**	**0.234**	1.095	2.021	1.012	2.656	1.000	1.195
Copper	**0.920**	**1.087**	**0.975**	**1.158**	**0.984**	**1.192**	1.017	1.401	1.041	1.462	0.980	1.432
Grains												
Corn	**0.710**	**0.256**	**0.845**	**0.306**	**0.912**	**0.347**	0.774	0.293	0.890	0.346	0.956	0.396
Soybean	**0.527**	**0.264**	**0.620**	**0.308**	**0.785**	**0.391**	0.591	0.305	0.664	0.344	0.813	0.428
Rough rice	**0.812**	**0.807**	**0.950**	**0.944**	**0.953**	**0.923**	0.877	0.988	1.003	1.158	0.987	1.137
Wheat	**0.845**	**0.312**	**0.896**	**0.332**	**0.937**	**0.349**	0.972	0.382	0.943	0.376	0.968	0.391
Softs												
Cocoa	**0.870**	**0.331**	**1.078**	**0.404**	**1.059**	**0.367**	0.938	0.391	1.116	0.482	1.086	0.466
Coffee	**0.698**	**0.450**	**0.788**	**0.514**	**0.973**	**0.545**	0.776	0.521	0.826	0.574	0.986	0.628
Cotton	**0.766**	**0.596**	**0.848**	**0.663**	**0.926**	**0.758**	0.852	0.724	0.919	0.801	0.962	0.889
Sugar	**0.843**	**0.392**	**0.928**	**0.434**	**1.078**	**0.492**	0.922	0.459	0.966	0.486	1.043	0.508
Orange juice	**0.866**	**0.631**	**0.960**	**0.620**	**1.033**	**0.761**	0.935	0.783	1.005	0.797	1.081	0.952
Livestock												
Feeder cattle	**0.601**	**0.443**	**0.776**	**0.572**	**0.980**	**0.758**	0.654	0.517	0.789	0.642	0.961	0.839
Live cattle	**0.814**	**0.382**	**1.099**	**0.555**	**0.944**	**0.467**	0.891	0.439	1.185	0.587	0.957	0.515
Lean hogs	**0.482**	**0.348**	**0.609**	**0.422**	**0.880**	**0.612**	0.553	0.411	0.628	0.459	0.884	0.675

The table displays the out-of-sample predictions results for 1-day, 5-day and 22-day realized volatility predictions for two models; HAR and RFSV.

Electronic copy available at: https://ssrn.com/abstract=3520500

Table B.3: **Out-of-sample predictions: 2005–2019**

	HAR						RSFV-RV					
	1-day		5-day		22-day		1-day		5-day		22-day	
Series	Pratio	MSE	Pratio	MSE	Pratio	MSE	Pratio	MSE	Pratio	MSE	Pratio	MSE
Energy												
Crude oil (WTI)	**0.223**	**0.178**	**0.325**	**0.257**	**0.513**	**0.390**	0.242	0.194	0.344	0.278	0.541	0.436
Crude oil (Brent)	**0.213**	**0.164**	**0.313**	**0.239**	**0.474**	**0.348**	0.229	0.177	0.332	0.258	0.499	0.386
Natural gas	**0.376**	**0.218**	**0.453**	**0.263**	**0.691**	**0.4010**	0.4572	0.269	0.466	0.281	0.761	0.491
Heating oil	**0.253**	**0.167**	**0.340**	**0.222**	**0.479**	**0.302**	0.278	0.184	0.360	0.24040	0.509	0.339
Gasoline	**0.357**	**0.199**	**0.461**	**0.258**	**0.640**	**0.358**	0.394	0.222	0.494	0.28341	0.6623	0.3922
Metals												
Gold	**0.878**	**4.240**	**1.006**	**4.915**	**0.969**	**4.741**	0.974	8.407	1.056	9.744	1.004	10.152
Silver	**0.771**	**1.823**	**0.795**	**1.895**	**0.818**	**2.024**	0.912	2.512	0.877	2.587	0.876	2.919
Platinum	**1.019**	**3.260**	**1.060**	**3.519**	**1.034**	**2.834**	1.139	4.847	1.058	4.641	1.063	4.288
Palladium	**1.194**	**1.898**	**1.123**	**1.618**	**1.025**	**1.544**	1.354	2.578	1.282	2.319	1.046	2.341
Copper	**0.763**	**0.895**	**0.864**	**1.014**	**0.916**	**1.093**	0.886	1.177	0.915	1.289	0.946	1.466
Grains												
Corn	**0.541**	**0.324**	**0.633**	**0.380**	**0.742**	**0.449**	0.613	0.379	0.682	0.434	0.784	0.526
Soybean	**0.588**	**0.320**	**0.708**	**0.385**	**0.837**	**0.453**	0.648	0.367	0.754	0.440	0.881	0.540
Rough rice	**0.609**	**0.559**	**0.686**	**0.632**	**0.744**	**0.685**	0.687	0.690	0.749	0.778	0.773	0.846
Wheat	**0.550**	**0.301**	**0.627**	**0.344**	**0.735**	**0.402**	0.619	0.351	0.672	0.389	0.769	0.464
Softs												
Cocoa	**0.613**	**0.322**	**0.700**	**0.368**	**0.774**	**0.409**	0.677	0.372	0.747	0.419	0.794	0.461
Coffee	**0.507**	**0.215**	**0.652**	**0.276**	**0.791**	**0.335**	0.549	0.239	0.684	0.305	0.804	0.372
Cotton	**0.479**	**0.375**	**0.591**	**0.463**	**0.703**	**0.551**	0.527	0.428	0.634	0.530	0.743	0.649
Sugar	**0.556**	**0.208**	**0.675**	**0.253**	**0.787**	**0.296**	0.616	0.237	0.713	0.281	0.790	0.328
Orange juice	**0.749**	**0.424**	**0.844**	**0.474**	**0.916**	**0.516**	0.836	0.510	0.907	0.571	0.945	0.632
Livestock												
Feeder cattle	**0.749**	**0.424**	**0.844**	**0.474**	**0.916**	**0.516**	0.836	0.510	0.907	0.571	0.945	0.632
Live cattle	**0.522**	**0.300**	**0.611**	**0.351**	**0.754**	**0.433**	0.600	0.354	0.632	0.381	0.768	0.478
Lean hogs	**0.638**	**0.509**	**0.787**	**0.627**	**0.884**	**0.702**	0.700	0.593	0.823	0.715	0.914	0.822

The table displays the out-of-sample predictions results for 1-day, 5-day and 22-day realized volatility predictions for two models; HAR and RFSV.

Electronic copy available at: https://ssrn.com/abstract=3520500