Received: 19 January 2024 | Accepted: 27 January 2024

DOI: 10.1002/fut.22492

RESEARCH ARTICLE

FUTURES MARKETS WILEY

Risky times: Seasonality and event risk of commodities

Dominik Boos

ZHAW School of Management and Law, Institute of Wealth & Asset Management, Winterthur, Switzerland

Correspondence
Dominik Boos, ZHAW School of Management and Law, Institute of Wealth & Asset Management, Winterthur, Switzerland.
Email: dominik.boos@zhaw.ch

Funding information
(Publikationsförderung) of the ZHAW School of Management and Law.

Abstract

The seasonal risk of wheat, corn, and soybean is modeled by a novel seasonality filter based on a generalized ridge regression. Then, using a component GARCH model, seasonal risk is combined with event risk and a short-term risk dynamics. The resulting model is robust, generates seasonal patterns related to the crop cycle, and significantly outperforms the standard GARCH(1,1) in terms of out-of-sample risk prediction. Results are relevant for risk management and portfolio construction.

KEYWORDS

ARCH, commodities, event risk, ridge regression, seasonality

JEL CLASSIFICATION

C14, C51, G17, Q02

1 | INTRODUCTION

Commodity prices are largely determined by supply and demand. Accordingly, news about supply and demand must consequently be the key driver of commodity prices changes. This news flow and thus price fluctuations are, however, far from constant, and vary considerably over the course of the year. For instance, many agricultural commodities exhibit distinct seasonal patterns both in terms of risk and pricing (e.g., Hevia et al., 2018; Schneider & Tavin, 2021; Sørensen, 2002).

For the three main crops analyzed in this article—wheat, corn, and soybean—the seasonal pattern is strongly related to the crop cycle. During the main growing period and to a lesser extent during the planting season, crops are particularly vulnerable to weather risks such as drought, flood, frost, or hail—these are risky periods. In contrast, in the dormant period and when no crop is planted the weather risk remains limited.

Other important drivers of the news flow are large government reports that among other data provide information about crop conditions, projected crop yields, prospected planting areas, and inventories. These can lead to considerable price movements on the release date and possibly in the aftermath. Release days are risk days.

The aim of this article is to estimate risk patterns over the course of the year, combine this with information flows related to reports, and finally add a classical GARCH-type model to account for volatility unrelated to the first two components. It adds to the literature by introducing a novel, nonparametric seasonality filter based on a generalized ridge regression that is able to reproduce virtually any seasonal pattern. This flexibility is necessary because commodity volatility displays distinct patterns, such as longer periods of stable volatility during the dormant period, and also rapid rises or sharp declines. Typically, seasonal volatility is modeled with sinusoidal functions which impose rigid cycles and have problems capturing both stable periods and sharp level changes with a small number of parameters (e.g., Karali &

Thurman, 2010). They are, therefore, not flexible enough to reproduce the patterns specific to commodity seasonality unless a large number of variables are introduced. Also the use of monthly dummies as in Efimova and Serletis (2014) or an indicator for the growing season as in Simon (2002) are not satisfactory owing to the resulting discontinuities and because these approaches are not granular enough to model sharp (intra-month) changes.

The seasonality of commodity volatility has also been studied in Fackler and Tian (1999). Using Fourier transforms and cubic splines, they found robust patterns for option implied volatility, but the results for realized volatility remained puzzling. Other papers that analyzed option implied seasonal volatility are McKenzie et al. (2022), Arismendi et al. (2016), and Back et al. (2013).

This article also incorporates scheduled event risk into a volatility forecasting model—a rare exercise in the literature. One of the few examples for single stocks is Deng (2008). It uses dummies for earnings announcements, macroeconomic events, and the day of the week in an extended-GARCH model and detects a strong increase of volatility around earning announcements. Another example is Kim et al. (2023) who found that forecasts incorporating earnings announcement information can better predict short-run stock market volatility. Standalone analyses of the event risk that do not use a fully fledged risk model are, however, more frequent. For the United States Secretary of Agriculture (USDA) reports relevant here, for example, Adjemian (2012), study risk and returns around World Agricultural Supply and Demand Estimates (WASDE) reports, McKenzie and Ke (2022) found risk spillovers from the WASDE report to Chinese soybean and corn markets, Isengildina-Massa et al. (2021) analyzed how the information content of USDA reports varies from month to month, Dorfmann and Karali (2015) inferred which reports contain market-moving information. Moreover, McKenzie (2008) outlined why the August USDA Crop Production report induces price reaction despite the fact that previously released private reports are more accurate, Isengildina-Massa et al. (2008b) show that implied volatility drops immediately after the release of the WASDE report and that the report therefore resolves uncertainty, and Isengildina-Massa et al. (2008a) show that the impact of WASDE reports has increased over time. Dubinsky et al. (2018) studied equity price uncertainty surrounding information released on earnings announcement dates. This paper also provides a good overview of this strand of literature.

Note that the event risk methodology in this article notably differs from the extensive stochastic jump literature following Merton's (1976) seminal paper. Here, event days are predefined and known in advance, which greatly simplifies estimation as no hidden jump intensity process must be filtered. See for example, Bernard et al. (2008) and Diewald et al. (2015) for GARCH models with jumps in the commodity market.

Intraday volatility is also seasonal and can be combined with a GARCH model in the same way as proposed in this present paper (see Engle & Sokalska, 2012). Intrady patterns can, however, be estimated over several 100 days, whereas for annual seasonality, data can be collected at best for a few dozen cycles. This means that the requirements for the method of evaluating annual seasonality are substantially increased relative to the methods used for intraday volatility.

2 | METHODOLOGY

2.1 | GARCH models

The work horse of volatility forecasting and also the benchmark in this paper is the standard GARCH(1,1) model derived by Bollerslev (1986) and Engle (1982):

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2, \tag{1}$$

where $r_t = \sigma_t\varepsilon_t$ are daily log returns at day t and $\varepsilon_t \sim N(0, 1)$.

Following Engle and Sokalska (2012) and Engle and Rangel (2008), I propose a multiplicative component GARCH (CGARCH) model which specifies the conditional return variance σ_t^2 as a product of a calendar variance pattern c_t and a residual component h_t:

$$r_t = \sqrt{h_t c_t}\,\varepsilon_t, \tag{2}$$

where the error term ε_t follows as a standard normal distribution as in (1).

The calendar variance, in turn, is the sum of two components as in Ding and Granger (1996):

$$c_t = s_t + e_t, \tag{3}$$

where s_t reflects the seasonal trend and e_t is an event risk component that incorporates the risk markup associated with major announcements. If there is no event, $e_t = 0$.

The model can be fitted in two steps (see Engle & Rangel, 2008). First, the calendar variance is estimated following the procedure outlined in Section 2.2 below. Then, the daily returns are normalized by the calendar component, leading to an estimate of the residual component

$$z_t = r_t / \sqrt{c_t} = \sqrt{h_t}\varepsilon_t. \tag{4}$$

A GARCH(1,1) process is used to model the volatility of the residual component. Given its wide acceptance, this is the natural choice although any volatility process could be used to model z_t. However, the main goal of this article is to identify the improvement in forecasting risk achieved by the calendar component. Therefore, it is essential to use the benchmark model to fit the residual component since this procedure ensures any improvement in the model arises solely from the calendar component. A more sophisticated risk model is also not critically required because there are no conclusive deviations from GARCH(1,1) for commodities. For instance, in contrast to stock market indices where asymmetric response to positive and negative returns is a generally accepted fact (Glosten et al., 1993; Nelson, 1991), this leverage effect is typically inverted and significant only for some commodities (Chen & Mu, 2021).

2.2 | Seasonality and event risk model

To model seasonality, an additional index $u\,(u = 1 \dots U)$ is introduced to indicate the season based on the (trading) day of the year. The event risk is modeled using an indicator function I_t^1 which is one if the first report is released on day t and zero otherwise. I_t^2 is respectively defined for report 2 and so on. With V events, Equation (3) can be rewritten as follows:

$$c_t = c_{t,u(t)} = s_u + I_t^1 e_1 + I_t^2 e_2 + \dots + I_t^V e_V, \tag{5}$$

where s_u represents the expected seasonality component of the variance at the uth day of the year and e_v represents the event risk (variance) related to report $v = 1 \dots V$. Note that heteroscedasticity, as defined by this equation, only depends on calendar information.

Several reports may be published on the same day. In this case, more than one indicator $I_t^{(\cdot)}$ is equal to one and the corresponding event risks add up on that day. This case is indeed relevant for the reports and commodities considered. In particular, during the growing season, additional information on crop production is included in the WASDE report. This additional information is modeled as a separate report. It induces more volatility on the joint reports relative to the WASDE reports without the crop production news. Similarly, at quarter ends, various reports are published on the same day.

As already outlined in Engle (1982), Equation (5) can be trained with an ordinary least square (OLS) estimator by plugging in squared returns, which leads to the following structure:

$$C = \begin{bmatrix} r_{1,1}^2 \\ r_{2,2}^2 \\ \vdots \\ r_{T,S}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 & I_1^1 & \dots & I_1^N \\ 0 & 1 & \dots & 0 & I_2^1 & \dots & I_2^N \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & I_T^1 & \dots & I_T^N \end{bmatrix} \begin{bmatrix} s_1 \\ \vdots \\ s_U \\ e_1 \\ \vdots \\ e_N \end{bmatrix} + \begin{bmatrix} \eta^1 \\ \eta^2 \\ \vdots \\ \eta^T \end{bmatrix} = [\Xi \quad I] \begin{bmatrix} S \\ E \end{bmatrix} + H, \tag{6}$$

where Ξ is a $T \times U$ matrix representing the seasonal dummies, I contains the event dummies, S and E are vectors of the seasonal components and event risk, and H summarizes the error terms. Note that the features matrix consists solely of seasonality and event dummies and the estimated variances are the loadings on these dummies. While OLS is feasible, estimates of S and possibly E will not be robust as we have up to 365 seasonality dummies as features. A regularization technique is needed to stabilize the parameter estimation.

The choice of the regularization method is motivated by two observations: First, there are no apparent breaks in the crop cycle. Even the beginning (planting) and the end (harvest) of the cycle are gradual processes that stretch over several weeks. Production is further smoothed out because crop cycles begin and end at different times in different regions. Second, there is very little theoretical reasoning about the exact form of seasonality, so we should allow for maximum flexibility. Consequently, we need a regularization method that allows for larger swings over intermediate time spans, but is steady and smooth over short time intervals.

For a smoothing problem with almost equivalent requirements, Boos and Grob (2023) implemented a generalized Tikhonov (1963) or ridge (Hoerl & Kennard, 1970) regression. I largely follow their approach but deviate from it in two ways: First, as in Tencate et al. (2016), I use a second-order difference matrix instead of a first-order difference matrix, making the fitted function smoother and more gently curved. Second, I use the day of the year instead of the lag and close the annual cycle by forcing a smoothing of the curve over the end of the year. Specifically, I minimize:

$$\min \sum_t \left(r_t^2 - s_u + I_t^1 e_1 + \cdots + I_t^V e_V \right)^2 + \gamma (s_U - 2s_1 + s_2)^2, \tag{7}$$

$$+ \sum_{j=2}^{U-1} \gamma (s_j - 2s_{j+1} + s_{j+2})^2 + \gamma (s_{U-1} - 2s_U + s_1)^2, \tag{8}$$

where the first term is the usual least squares loss function and the remaining terms are penalties for variation in the seasonal risk. The model accounts for cyclicality through the second and the last term which penalize differences between the last days of a year and the first days of the following year. Finally, the hyperparameter γ determines the amount of smoothing for each day of the year.

This approach is remarkably parsimonious since it is able to regularize the problem with a minimum set of restrictions and a single hyperparameter γ. Even more, γ is (in-sample) tuned by the data as described in Appendix A, so there is no discretion beyond the definitions in Equation (7).

As in the standard ridge case, (7) can be estimated with OLS by augmenting the data with pseudo observations that account for the regularization term. For this, we define the following circular Tikhonov matrix:

$$\Gamma = \begin{bmatrix} -2\gamma & \gamma & 0 & & 0 & \gamma \\ \gamma & -2\gamma & \gamma & 0 & & 0 \\ 0 & \gamma & -2\gamma & \gamma & 0 & \\ & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & & 0 & \gamma & -2\gamma & \gamma \\ \gamma & 0 & & 0 & \gamma & -2\gamma \end{bmatrix} \tag{9}$$

and then augment regression Equation (6) as follows:

$$\begin{bmatrix} C \\ 0 \end{bmatrix} = \begin{bmatrix} \Xi & I \\ \Gamma & 0 \end{bmatrix} \begin{bmatrix} S \\ E \end{bmatrix} + \widetilde{H}. \tag{10}$$

This stacked combination of observations and pseudo observations can be fitted by some regression method.

Fitting (7) respectively (10) using an OLS estimator has the drawback of not being efficient because Equation (5) is heteroscedastic by assumption. For typical GARCH models, the use of a weighted least squares (WLS) estimator largely suppresses this drawback, as pointed out in chapter 6.2 of Francq and Zakoian (2019). Therefore, throughout the paper, I use a multistage WLS procedure for parameter learning and tuning of the hyperparameter γ. See Appendix B for the details of this procedure.

3 | IN-SAMPLE RESULTS AND SEASONALITY

3.1 | Data

Daily returns and information on futures contracts are provided by Bloomberg Finance L.P. The selection of the commodities examined is driven by liquidity considerations. Corn and soybean are clearly the two most liquid grains contract at the Chicago Mercantile Exchange (CME). Chicago wheat is also very liquid, and in particular more liquid than the Kansas wheat contract. Therefore, I am using these three contracts in this paper.

To eliminates the Samuelson (1965) effect, namely, the effect that the return volatility of a futures contract decreases with time-to-maturity and to ensures that volatility does not jump when futures are rolled, I have constructed 3-month (92 days) constant maturity returns as follows: First, the term (maturity) is defined as the difference between the actual date and the first notice day. Then, for each trading day, of all the contracts with a remaining term of more than 92 days, the one with the lowest remaining term is determined. Similarly, the contract with the highest remaining maturity of less than 92 days is determined. The simple returns of the two contracts are then interpolated so that the weighted residual maturity is 92 days. Then, they are converted to log returns. If there is a contract with a maturity of exactly 92 days, the log return of that contract is used. Throughout the paper, I ignore cash or collateral returns and only consider excess returns.

Event days are the release dates of the two most highly regarded reports from the USDA, the WASDE, and the Grains Stock report. Over the course of the year, there are also three different planting surveys released together with the Grains Stock report. I differentiate between the four reports/group of reports:

- The WASDE report is released monthly, usually between the 8th and 12th of each month.[1] It contains annual forecasts for the supply and use of wheat, rice, coarse grains, oilseeds, and cotton in the United States and worldwide. The report also covers other agricultural commodities. Adjemian (2012) provides a very detailed description of the report. The WASDE website[2] has more details.
- The Crop Production reports are embedded in the WASDE report. They include information from planting reports (see below) and survey-based estimates of yield and production estimates for major crops consistent with their growth cycles: August through November for corn and soybeans, May through August for winter wheat, and July and August for spring wheat (Isengildina-Massa et al., 2021).
- The Grains Stock report is issued four times a year. For the first three quarters, publication is at the last trading day of each quarter while the fourth quarter is released together with the January WASDE report. It contains stocks (inventories) of all wheat, durum wheat, corn, soybeans, and many other agricultural commodities by individual States and the entire United States. See the Grains Stock website[3] for additional details.
- There is a series of planting reports which begins in January with the Winter Wheat Seedings[4] report (this includes data on the annual seeded acreage of winter wheat, durum wheat, and canola seedlings), followed by the Prospective Planting[5] report (this contains the expected plantings and previous year's harvest for principal crops) at the end of March and the Acreage[6] report (this presents acreage by planted and/or harvested areas) at the end of June. Following Dorfmann and Karali (2015), I treat these reports as a single report in most specifications.

The release dates for these reports were collected from the archives of their respective websites.

The WASDE report has been published monthly since 1985, which is why my data set starts on January 1, 1985. The Grain Stock report schedule changed to the current schedule in mid-1986, but even before that, four reports were published each year. Before 1994, both reports were published after the close of the domestic market. In these cases, following Adjemian (2012), the event dummy is set to one on the first trading day after the day of publication, but to zero on the day of publication itself. With a few exceptions at the beginning of the sample, the planting reports were published together with the Grains Stock report.

Seasonal dummies are derived from trading days. To reduce the number of features, I have combined two trading days into one dummy so that for the first and second trading day of the year, the first dummy is one, and for the third and fourth, the second dummy, and so on. I have used a total of 126 dummies. Some years have 253 trading days and 1996 even had 254. For these surplus days, the last dummy is set to one. Results are not sensitive to how the dummies are defined.

---

[1]Due to a government shutdown, the WASDE report for January 2019 was not published.

[2]https://www.usda.gov/oce/commodity/wasde

[3]https://usda.library.cornell.edu/concern/publications/xg94hp534

[4]https://usda.library.cornell.edu/concern/publications/z890rt24s; Winter wheat seedings have been published in different reports with different names; since 2017 in the Winter Wheat and Canola Seedings report.

[5]https://usda.library.cornell.edu/concern/publications/x633f100h

[6]https://usda.library.cornell.edu/concern/publications/j098zb09z

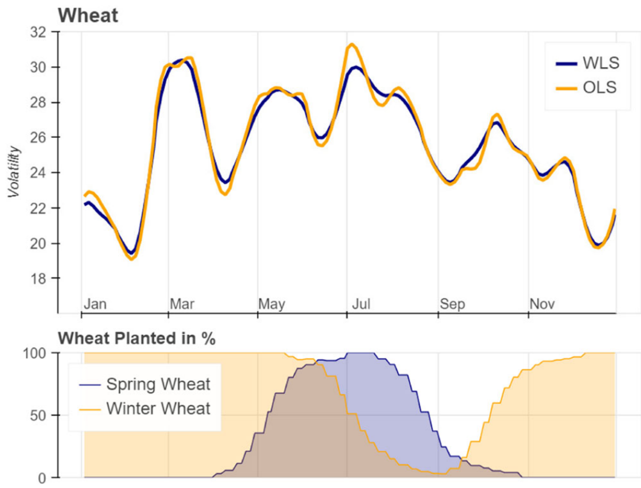

FIGURE 1 Seasonality pattern of wheat. The upper graph shows the fitted annualized volatility using weighted least square (WLS) and ordinary least square (OLS) estimates of (5). The data set consists of 9583 observation between 1985 and 2022. The lower graph depicts the crop cycle indicated as the percentage of wheat planted yet not harvested also using data between 1985 and 2022. [Color figure can be viewed at wileyonlinelibrary.com]

Finally, planting and harvesting progress data are taken from the USDA National Agricultural Statistics Service (NASS) through Quick Stats.[7] They are release in USDA's Crop Progress report[8] which is issued weekly during the growing season (April to November). I'm using the data items “progress, measured in pct planted” and “progress, measured in pct harvested” for winter and summer wheat, corn (grains) and soybean on national geographic level. They indicate the percentage of a crop that is planted respectively harvested for the entire United States.

3.2 | Seasonality pattern

In this section, I use the entire sample to estimate the model as outlined in Section 2, taking into account the WASDE report, the Crop Production report contained therein, and the Grains Stock report. Seasonality and the ARCH parameters are virtually the same for other variants of the model discussed in Section 4.6 below.

Figures 1 and 2 depict the seasonality patterns for the three commodities and a variables indicating average crop cycle over the course of the year measured as a percentage of the grain in the field. These data are form USDA Crop Progress report, see Section 3.1 above for the exact source. This variable is zero when no crop is planted, starts increasing when the first farmers begin planting (using the data item “percentage planted”), becomes one when the planting season is finished, and finally starts decreasing when the harvest begins (using 1—”percentage harvested”). It is based on averages of over the entire data period (1985–2022) and reflects crop production in the entire United States. For wheat, the crop cycle of both winter and summer wheat is depicted, but for corn and soybean there is only one cycle.

The volatility pattern for wheat is quite choppy and displays some larger short-term oscillations that look like an imperfect filtering of seasonality. These swings, however, coincide well with the vulnerability of the crop as outlined in Shroyer et al. (1995). In winter, before the wheat begins to grow, it is very resistant to low temperatures and hence volatility is lowest during this period. Then, when the crop starts growing in early spring volatility sharply increases as wheat has now little resistance and the risk of frost damage is particularly high. After a short decline once the risk of freezes has decreased, it remains elevated through the planting season of spring wheat and the growing period of both varieties. Risk falls to a lower level in late summer after the harvest but has a minor peak during the seeding of winter wheat. Minor swings might also be related to foreign production as (according to the WASDE report) the United States is only the fourth largest wheat producer globally. As production is concentrated in the northern hemisphere, foreign production must, however, follow a similar overall cycles.

[7]https://quickstats.nass.usda.gov/
[8]https://usda.library.cornell.edu/concern/publications/8336h188j?locale=en

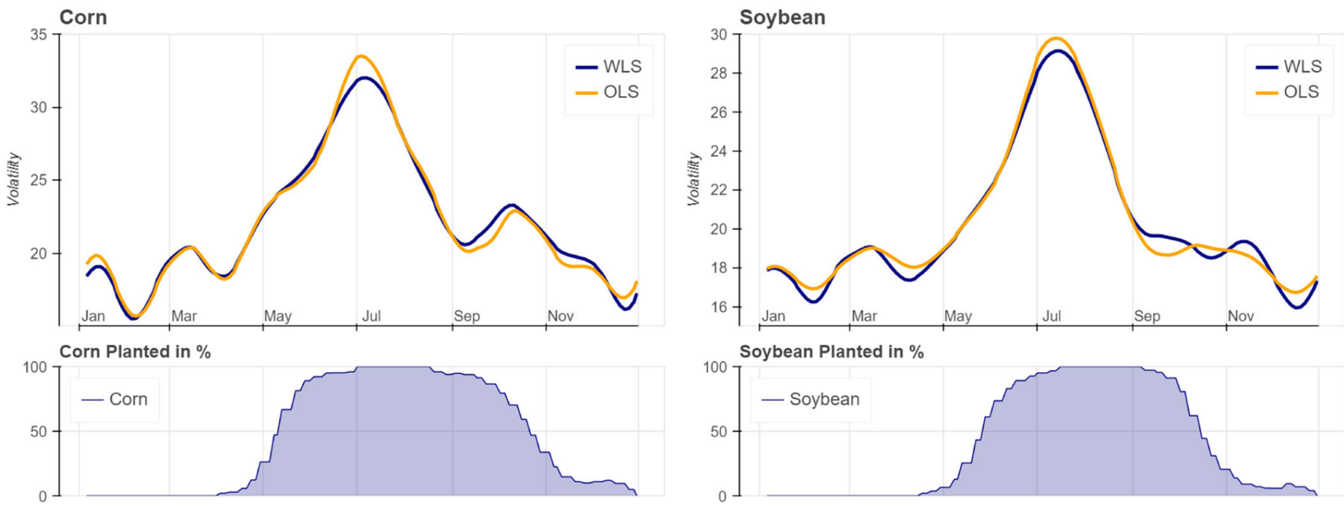

FIGURE 2 Seasonality pattern of corn and soybean. The upper graphs show the fitted annualized volatility using weighted least square (WLS) and ordinary least square (OLS) estimates of (5). The data sets consist of 9583 observation between 1985 and 2022. The lower graph depicts the crop cycle indicated as the percentage of corn/soybean planted yet not harvested also using data between 1985 and 2022. [Color figure can be viewed at wileyonlinelibrary.com]

Corn and soybean have similar patterns, reflecting the similarity of the two crop cycles and the fact that the United States is the dominant producer (according to the WASDE report). For both crops, volatility starts to increase in the planting season and reaches its peak in midsummer before rapidly declining. While the crops remain in the field until mid to late fall, they are less sensitive in the second part of the growth season as they have already reached a certain level of maturity and because the risk of hail and drought is lower once the heat has gone. Corn also shows a minor fluctuation during harvesting in October.

Especially for corn and wheat, there are minor short-term swings that are difficult to interpret and may reflect imperfect fitting. It is important to bear in mind that there are only 38 annual observations and that the error terms are auto-correlated beyond seasonality by the GARCH component in (3). As the first step of the estimation procedure ignores ARCH effects, it must interpret any auto-correlation as seasonality. In such a small sample, volatility shocks unrelated to seasonality may lead to spurious swings in the seasonality pattern.

Analogous seasonality patterns have been estimated before. Specifically, Figure 1 of Fackler and Tian (1999) plots the volatility of soybean over the course of the year and strongly resembles the soybean pattern in Figure 2. Karali and Thurman (2010) Figure 2 plots seasonal volatility for the three crops presented here as well as for oats. They resemble those shown in Figures 1 and 2 although their wheat pattern is much smoother. Isengildina-Massa et al. (2021), Schneider and Tavin (2021), and Wang and Garcia (2011) calculate volatility for each month of the year and their findings are also inline with the seasonality pattern presented in Figures 1 and 2.

3.3 | ARCH persistence and event risk

The parameters of the GARCH models are very similar across commodities and across the models as seen in Table 1. The fact that the GARCH process changes so little is because two opposing effects are at work. On the one hand, the filtering out of seasonality leads to a lower persistence, while on the other hand, the event risk decays very quickly, resulting in a higher auto-correlation of the residuals. It is worth noting that the decay β is very close to the standard industry value for daily returns of 0.94 that goes back to RiskMetrics (1996).

Finally, events generate substantial risk on top of the seasonal risk as shown in Panel C of Table 1. Parameter values are transformed to daily volatility. The WASDE report is only marginally significant in the case of wheat but insignificant for corn and soybean. Conversely, Crop Production is insignificant for wheat but highly significant for the other two crops with point estimates of 1.84 and 1.85. As Crop Production is always released in conjunction with the standard WASDE report, daily variance on these days adds up. The induce variance is indicated in the last row of Table 1 it varies between 0.97 (Wheat) and 1.89 (Soybean). With t-statistics above 4, the Grains Stock report shows the

TABLE 1 Full sample parameters.

	Wheat		Corn		Soybean	
	Para	***t*-stat**	**Para**	***t*-stat**	**Para**	***t*-stat**
Panel A	GARCH(1, 1)					
ω	0.018	3.61	0.020	3.99	0.026	4.98
α	0.056	8.72	0.077	10.65	0.071	11.65
β	0.938	128.39	0.917	112.09	0.916	129.59
σ	28.57	–	28.11		22.21	–
Panel B	Residual GARCH					
ω	0.005	3.11	0.008	3.65	0.012	5.24
α	0.047	7.31	0.058	8.56	0.058	11.37
β	0.949	130.34	0.935	116.93	0.930	155.72
σ	1.072	–	1.053	–	1.008	–
Panel C	Event risk					
WASDE	0.75	1.63	0.10	0.05	0.35	0.63
G. stocks	2.31	4.08	2.72	4.57	2.15	4.48
Crop Prod.	0.61	0.50	1.84	3.58	1.85	3.91
WASDE cp	0.97		1.85		1.89	

Note: Panel A provides estimated parameters and the associated t-statistics of the standard GARCH(1,1) model defined in (1). ω is the constant term, α the first ARCH coefficient, β the decay rate, and σ the unconditional volatility implied by the model. Panel B provides the parameter of the residual GARCH model of z_t defined in (4). σ is the estimated unconditional volatility of z_t which is theoretically 1 by construction. Events include the WASDE report, Crop Production, and Grain Stocks. GARCH parameters and t-statistics are estimated in Python using the function arch_model of the arch package. Panel C contains the estimated event risk and WLS t-statistics (they are virtually the same as those based on Halawa & Bassiouni, 2000, which account for the ridge regularization). The last row gives the event risk for WASDE reports with embedded Crop Production reports which is derived from the two subreports. Event risk parameters are transformed into percentage points of daily volatility. The data set consists of 9583 observation between 1985 and 2022.

highest significance for all commodities and generates an event day volatility between 2.15 for soybean and 2.72 for corn. So, Grains Stock is the report that induces the highest volatility for all three crops.

Isengildina-Massa et al. (2021) also calculate event day volatility as well as monthly nonreport day volatility over a very similar time span (1985–2018). These values are not directly comparable to those reported in Panel C of Table 1 because they combine the seasonality component and the event risk. Using Equation (3) or (5) and their nonreport day volatility, the event risk implied by their results can be factored out. This procedure shows that their implied event risk is very similar to the one provided her. Specifically, as with the results presented here, the Grains Stock report shows the highest event risk for all three commodities (2.60 wheat, 2.97 corn, 2.04 soybean). For corn (1.95) and soybean (1.97) the Crop Production report and for wheat the WASDE report[9] (0.83) are also relevant sources of event risk. These values are in general marginally higher than those found in this paper because Isengildina-Massa et al. (2021) use the front month contract for their calculations whereas I use a constant maturity return which have a somewhat lower risk.

4 | FORECASTING ABILITY AND MODEL COMPARISON

4.1 | Design

The model's forecasting power is assessed by a backtest-style out-of-sample cross-validation exercise which starts in 2010 and hence spans 13 years or a third of the entire data sample. For the forecasts in 2010, the model is fitted using data up to the last trading day of 2009. Then, based on these parameters, for each day, a 1 day ahead forecast is

[9]For months where the risk on report days is lower than the risk on nonreport days the event risk is set to zero for this calculation.

calculated using return information up to day $t-1$. At the end of the year, the model is refitted using information before 2011. Based on these estimates, I have calculated 1 day ahead forecasts for 2011. This procedure is repeated until the end of the sample and generates 3276 daily forecasts using 13 model estimates. The forecast using the GARCH model is calculated using the same design.

4.2 | Evaluation measures

As in Boos and Grob (2023), I primarily quantified the model's goodness-of-fit using the coefficient of determination of the out-of-sample forecast:

$$R^2 = 1 - \frac{\text{MSE}}{\text{ASS}} = 1 - \frac{\frac{1}{T}\sum_{t\in W}\left(\varepsilon_t^2 - \widehat{\sigma_t}\right)^2}{\frac{1}{T}\sum_{t\in W}\left(\varepsilon_t^2 - \bar{\varepsilon}^2\right)^2}, \tag{11}$$

where MSE is the model's mean squared error, ASS is the average sum of squares of the observed data, W the set of days, where an out-of-sample forecast $\hat{y}$ has been calculated, $\bar{\varepsilon}^2$ the average variance over the out-of-sample period, and T the number of out-of-sample observations.

To detect biases and scaling issues, I ran the Mincer and Zarnowitz (1969) regression using the same set of data points:

$$\varepsilon_t^2 = a + b^*\widehat{\sigma_t}^2 + \eta_t. \tag{12}$$

A good forecast is unbiased and has an intercept of $a = 0$. The slope coefficient b should be close to one because otherwise a simple scaling of the forecast would improve its fit. Therefore, we tested the hypothesis that the intercept equals zero and the slope is equal to one.

For comparison of the CGARCH with the benchmark model, as in Aït-Sahalia and Mancini (2008) I ran augmented Mincer and Zarnowitz (1969) regression that included the forecasts of both models:

$$\varepsilon_t^2 = a + b_1^*\widehat{\sigma_t^{mc}}^2 + b_2^*\widehat{\sigma_t^{bm}}^2 + \eta_t. \tag{13}$$

A superior model should squeeze out the other model in this regression. Assuming the component dominates, b_1 is significant and close to 1; b_2, on the other hand, is insignificantly different from zero.

Finally, I also assessed the forecasting ability using the following measures also used in Nguyen and Walther (2020) and discussed in Lopez (2001). The mean absolute error

$$\text{MAE} = \frac{1}{T}\sum_{t\in W}\left|\varepsilon_t^2 - \widehat{\sigma_t}^2\right|, \tag{14}$$

and the loss function implied in the Gaussian quasi-maximum likelihood function result in:

$$\text{MLE} = \frac{1}{T}\sum_{t\in W}\left[\ln(\widehat{\sigma_t}^2) + \frac{\varepsilon_t^2}{\widehat{\sigma_t}^2}\right]. \tag{15}$$

4.3 | Out-of-sample predictability and model compaison

The out-of-sample mean squared error, or its normalized version, the R^2, is used as a key criterion for assessing the predictive power of the model. As Panel A of Table 2 shows, it is higher for the CGARCH model compared to the benchmark model in all three cases. Corn shows the largest increase in R^2 (from 6.23 to 16.52); wheat the lowest, but still shows an increase of more than three percentage points, from 9.06 to 12.08. All other criteria in Panel A consistently favor the CGARCH model over the benchmark.

The Mincer Zarnowitz regression (12) confirms the results of the R^2 simulation and reveals that the quality of the CGARCH is indeed very high. For all three commodities, neither parameters significantly deviates from its theoretical

TABLE 2 Out-of-sample prediction.

	All Observations Wheat GARCH	CGARCH	Corn GARCH	CGARCH	Soybean GARCH	CGARCH
Panel A						
MSE	4.16	4.02	2.55	2.27	1.11	1.02
R^2	9.06	12.08	6.23	16.52	4.09	12.17
MAE	3.46	3.41	2.63	2.47	1.72	1.64
GMLE	−7.10	−7.14	−7.45	−7.56	−7.85	−7.92
Panel B	Nonevent days					
MSE	3.54	3.50	1.76	1.71	0.80	0.78
R^2	9.40	10.44	7.16	9.88	3.82	6.00
MAE	3.32	3.25	2.39	2.25	1.59	1.50
GMLE	−7.17	−7.17	−7.60	−7.62	−7.96	−7.97
Panel C	Event days					
MSE	1.40	1.23	1.52	1.13	0.60	0.47
R^2	−0.25	11.89	−13.22	15.99	−6.72	16.20
MAE	0.57	0.60	0.65	0.61	0.38	0.38
GMLE	−6.08	−6.57	−4.94	−6.55	−6.22	−7.14

Note: Evaluation criteria of out-of-sample forecasts: mean squared error (MSE), R^2, mean absolute error (MAE), and the quasi-maximum likelihood function (MLE) for all 3276 observations in Panel A, non event days (3081 observations) in Panel B, and event days (195) in Panel C. The out-of-sample period spans 2010 and 2022. Events include the WASDE report, Crop Production, and Grain Stocks. Models and commodities are indicated at the top of the table.

values as can bee seen in Table 3. The *b*s are between 0.93 and 1.03 and thus exceptionally close to their theoretical value of 1.

However, the GARCH models do not pass the Mincer Zarnowitz test. In all three cases, the *b*s are significantly lower than one. The forecasts are thus "overconfident" in the sense that shrinking them towards zero would improve the forecasting ability of the model.

Finally, the GARCH model is squeezed out in the augmented Minzer Zarnowitz regression (13). The loadings related to it (b_{GARCH}) are all negative, and for corn and soybean significantly negative. This means that in an optimal combination of the two models, the GARCH model B would enter with negative weight.

4.4 | Robustness

The model's robustness is demonstrated in three ways. First, Table 4 presents R^2 values for each year in the out-of-sample set. For wheat, the CGARCH model performs better than the GARCH model in 9 out of 13 years; for corn, it outperforms the GARCH model every year; and for soybeans, in 12 out of 13 years. This indicates that the CGARCH model's superior predictive power is robust across all commodities and over time.

Second, I split the out-of-sample period into event days and non event days and ran the out-of-sample analysis on each of the subsamples separately. The results are reported in Panel B and C of Table 2. While the R^2 of the GARCH model remains generally stable on non-event days relative to the full sample, the R^2 of the CGARCH model drops substantially. On event days, the CGARCH model has a similar R^2 than in the entire sample whereas the R^2 of the standard GARCH is even negative. This negative value occurs because the standard GARCH cannot foresee events and thus systematically underestimates risk on these days. These findings show that event risk is an important feature of the CGARCH model.

TABLE 3 Mincer zarnowitz regressions.

	Wheat		Corn		Soybean	
	para	***t*-stat**	**para**	***t*-stat**	**para**	***t*-stat**
	GARCH(1, 1)					
a	4.0e-5	2.02	5.3e-5	3.49	3.3e-5	2.87
b	0.89	−2.26	0.76	−4.83	0.75	−4.09
	CGARCH					
a	2.7e-5	1.47	9.1e-6	0.73	−8.5e-6	−0.89
b	0.93	−1.50	0.97	−0.80	1.03	0.66
	Augmented MZ-regression					
a	3.4e-5	1.73	5.4e	3.81	1.9e-5	1.70
β_{GARCH}	−0.11	−1.03	−0.49	−6.40	−0.44	−4.99
$\beta_{MCGARCH}$	1.02	10.52	1.29	20.57	1.32	17.58

Note: The GARCH parts displays estimates of the Mincer Zarnowitz regressions (12) which tests for bias and scaling. The regression coefficients are followed by corresponding t-statistics. It tests if the intercept a is different from zero and if the slope coefficient b is different from one. The augmented Mincer Zarnowitz regressions (13) in the last panel tests if parameters are different for zero. All tests use the out-of-sample variance forecasts and compare them with realized squared daily returns. The out-of-sample period consists of 3276 observation between 2010 and 2022. Commodities are indicated at the top of the table.

TABLE 4 R^2 by year.

	Wheat		Corn		Soybean	
	garch	**mcgarch**	**garch**	**mcgarch**	**garch**	**mcgarch**
2010	1.27	4.5	−1.93	10.26	1.95	11.89
2011	−2.28	6.34	−1.71	11.27	−0.56	6.81
2012	0.51	10.5	7.65	14.35	4.23	6.44
2013	1.31	10.67	1.75	9.48	0.52	10.91
2014	−0.88	−5.52	−2.88	16.41	1	4.28
2015	0.38	5.62	−0.61	16.25	0.3	15.66
2016	−1.45	−4.49	2.81	8.27	8.72	11
2017	3.56	9.5	3.98	10.49	0.35	14.44
2018	3.24	2.18	1.73	15.81	−0.61	3.85
2019	2.19	5.42	2.77	17.34	1.65	3.55
2020	−1.95	1.75	−3.08	4.64	−1.15	8.86
2021	−1.81	7.64	8.01	20.64	1.99	12.83
2022	8.53	6.64	7.05	9.06	4.49	10.13

Note: R^2 of the out-of-sample variance forecast relative to squared daily returns per year. All parameter estimates are based on the years preceding the year indicated on the left. Models and commodities are indicated at the top of the table.

However, seasonality also contributes to model performance, as the CGARCH model also performs better on nonevent days, and consistently for all three futures. The difference in the R^2 is smallest for wheat (10.43 vs. 9.40), which may be related to the somewhat unstable seasonality discussed in Section 3.2.

Finally, the results for the basic OLS estimator are presented in the second to the last line of Table 5. While the OLS estimator slightly (less than 0.1% in terms of R^2) underperforms the theoretically more efficient WLS estimator, it still substantially outperforms the standard GARCH model.

TABLE 5 Relevance of components.

	Components			R^2		
Method	seas	event	GARCH	Wheat	Corn	Soybean
WLS	x	x	x	12.08	16.52	12.17
WLS	x	x		1.47	10.75	6.51
WLS	x		x	9.61	9.05	5.65
WLS	x			−0.58	3.91	0.86
WLS		x	x	11.05	15.34	11.90
WLS		x		0.40	7.62	5.18
OLS	x	x	x	12.00	16.45	12.07
–			x	9.06	6.23	4.09

Note: The last three columns of the table display out-of-sample R^2 based on a subset of the three components. An x in the seas row indicates that the model has a seasonality component. Similarly for event and GARCH. Events include the WASDE report, Crop Production, and Grain Stocks. The out-of-sample period consists of 3276 observation between 2010 and 2022.

4.5 | Importance of the components

In this section, to assess the importance of the three components of the model, I calculated variants of the model that omit one or two components. The results are summarized in Table 5. First, the GARCH-part is skipped and out-of-sample forecast directly based on Equation (3) are calculated. For corn and soybean, they show considerable forecasting power that exceeds that of the pure GARCH model and is as high as 10.75 for corn. For wheat, the pure scheduled risk component shows only a weak performance that lags behind the GARCH model.

Next, the model was estimated without events but only seasonality dummies in (3). This model displays higher R^2 than the benchmark model for all three crops. The pure seasonality model, however, shows a negative R^2 for wheat. Finally, the model was trained without the seasonality filter. This model clearly improves the standard GARCH model and its performance is close to that of the complete model. The stand alone event model delivers R^2s between 0.40 for wheat and 7.62 for corn.

Overall, we can conclude that all three components add to the predictive power. Seasonality is the weakest predictor and the pure seasonality model demonstrates a negative performance for wheat. However, once combined with the other factors, seasonality consistently adds to performance (including for wheat). Event risk and the GARCH are roughly of equal importance and deliver positive contributions throughout.

4.6 | Variants and the importance of reports

In this section, I analyze the importance or the information content of different reports in an out-of-sample set up. I do this by comparing the predictive power when various reports are included or not or when sub-reports are differentiated. Specifically, I vary the features (dummies) in the first step of the risk forecasting process equation (3). For the WASDE reports, I separate reports that include a crop production report from those without. For the quarter-end reports, there are three variants. First, treat all as a single report (gs), then add a dummy when there is a planting report (pr), and finally differentiate between all four reports using four dummies (q4). Additionally, I created dummies 1 day prior (lead) to an event and 1 day after event days (lag) to capture increased volatility surrounding event days. The lead term can capture volatility induced by repositioning ahead of an event. Events might also change the auto-correlation structure of squared returns, an effect modeled through the dummy after the event. As an alternative specification of this effect, I also estimated a model that includes an AR(1) term in (3).

The results are summarized in Table 6. These are broadly in line with the findings of Isengildina-Massa et al. (2021) who analyzed the information content of USDA events in an insample set-up. The differentiation of the WASDE reports into crop production reports and normal reports adds to the performance of the model, especially with soybean where the R^2 increases from 10.51 to 12.17, as can be seen in the first two lines of Table 6. Skipping WASDE reports without crop production reduces the R^2 only for wheat (model 5).

TABLE 6 Event selection.

	Variables			R^2		
model	wasde	quarter	lead/lag	Wheat	Corn	Soybean
0	w/cp	gs		12.08	16.52	12.17
1	w	gs		12.12	15.9	10.51
2	w/cp	gs/pr		12.04	16.52	12.27
3	w	gs/pr		12.08	15.92	10.81
4	w/cp	q4		11.7	16.34	11.84
5	cp	gs		11.83	16.54	12.17
6	w/cp	gs	lag	11.98	16.72	12.13
7	w/cp	gs	lead	12.1	16.21	11.86
8	w/cp	gs	AR(1)	12.72	13.16	11.5

Note: The last three columns of the table display out-of-sample R^2 based on a subset of the three components. The second column indicates the reports included at the WASDE dates, the third column indicates the quarter end reports used, the fourth column indicates dummies surrounding the events (lag/lead: dummies one day after and one day before events, AR(1): a single AR term one day after events). The out-of-sample period consists of 3276 observations between 2010 and 2022.

Abbreviations: cp, crop production; gs, grains stock; pr, planting reports; q4, all four quarters are differentiated with its own dummy; w, WASDE.

The inclusion of the planting reports leaves the performance essentially unchanged (model 0/1 vs. 2/3). Differentiating all quarterly reports even reduces the performance of the model. Finally, the lead and lag variables do not improve the predictive power (models 6–8). Apparently, event volatility is concentrated on the event day and large event shocks fade at a very similar speed to other market shocks. The release procedure, however, changed twice over the sample period, so this matter needs further investigation.

5 | OUTLOOK

The seasonality filter and event risk methodology introduced in this paper has been used to fit a component volatility model where the residual risk was fitted with a standard GARCH(1,1) model. This novel method provides very accurate and robust out-of-sample forecasts that clearly dominate the basic GARCH model. I have also demonstrated that both seasonality and event risk contribute to a better risk forecast and that the Grains Stocks report and the Crop Production part of the WASDE report are the most important events.

While the use of a standard model to fit the second component of the CGARCH model helps isolate the core result (because it simplifies the exposition and reduces the danger of overfitting), it also leaves room for improvement. The calendar risk methodology can be combined with any model in the GARCH zoo without adding much complexity. Specifically, all generalizations proposed in Ding et al. (1993) provide some evidence regarding commodity risk. For instance, McKenzie et al. (2001) analyzed the base metal market and found results that suggest conditional standard deviation rather than conditional variances should be the focus of ARCH models as proposed in Taylor (1986) and Schwert (1989). These models are also known to be less sensitive to outliers and hence more robust, which are important properties for out-of-sample forecasts. Asymmetry in the commodity market was studied by Chen and Mu (2021), who found evidence of an inverse-leverage effect. This observation might also be explained by seasonal inventories. Stock-outs risks that can drive both prices and volatility up is highest just before the harvest and clearly seasonal. The framework presented in this paper can be extended to address whether or to what extent seasonality drives the inverse-leverage effect. Similarly, a good fit of seasonal volatility can also add to the long-memory literature in commodities (e.g., Alfeus & Nikitopoulos, 2022; Chang et al., 2012; Charfeddine, 2014) because any seasonality may give rise to spurious long memory as outlined in Gatheral et al. (2018).

The regression-based seasonality filter can also be used in conjunction with range-based data following Alizadeh et al. (2002) (see also Adjemian, 2012, for an application with USDA data) or with intraday data following Andersen et al. (2003), for example. It is also possible to include noncalendar variables in the first stage regression (5). For instance, Nguyen and Walther (2020) predicted commodity market volatility with long-term economic and financial variables. Similarly, inventory levels or the futures basis could be included as implied by the theory of storage (see e.g., Geman & Smith, 2013; Symeonidis et al., 2012).

Another improvement might result from a one-stage instead of two-stage estimation procedure presented here, for example, through a penalized maximum likelihood framework as in Fang et al. (2020). This might lead to a better estimate of parameters and seasonality pattern. Finally, the seasonality filter introduced in this paper is interesting in its own right and could have other applications in data science and econometrics.

ACKNOWLEDGMENTS

I thank Linus Grob for the joint development of programs used in this paper and for many helpful discussions, Eduardas Lazebnyj for proof reading and helpful comments, John Christian for proof reading as well as David Jaggi, Mark Thompson, Patrick Hauf, and seminar participants at the Zurich University of Applied Sciences for helpful comments. This work was financially supported by the Publikationsförderung of the ZHAW School of Management and Law. Open access funding provided by ZHAW Zurcher Hochschule fur Angewandte Wissenschaften.

DATA AVAILABILITY STATEMENT

Daily returns and information on futures contracts are provided by Bloomberg Finance L.P. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the authors with the permission of Bloomberg Finance L.P. Event dates are openly available from USDA using the following links: WASDE reports (https://www.usda.gov/oce/commodity/wasde), Grains Stock reports (https://usda.library.cornell.edu/concern/publications/xg94hp534), Winter Wheat Seedings (https://usda.library.cornell.edu/concern/publications/z890rt24s), Prospective Planting reports (https://usda.library.cornell.edu/concern/publications/x633f100h), and Acreage reports (https://usda.library.cornell.edu/concern/publications/j098zb09z). A collection of the event days used in this study can also be found in Boos (2023). The Crop Progress reports are available at (https://usda.library.cornell.edu/concern/publications/8336h188j) but can more conveniently be downloaded using USDA's Quick Stats (https://quickstats.nass.usda.gov/).

REFERENCES

Adjemian, M. K. (2012). Quantifying the wasde announcement effect. *American Journal of Agricultural Economics*, *94*, 238–256. https://doi.org/10.1093/ajae/aar131

Alfeus, M., & Nikitopoulos, C. S. (2022). Forecasting volatility in commodity markets with long-memory models. *Journal of Commodity Markets*, *28*, 100248. https://doi.org/10.1016/j.jcomm.2022.100248

Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance*, *57*, 1047–1091. https://doi.org/10.1111/jofi.2002.57.issue-3

Allen, D. M. (1974). The relationship between variable selection and data augmentation and a method for prediction. *Technometrics*, *16*, 125–127.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*, 579–625. https://doi.org/10.1111/ecta.2003.71.issue-2

Arismendi, J. C., Back, J., Prokopczuk, M., Paschke, R., & Rudolf, M. (2016). Seasonal stochastic volatility: Implications for the pricing of commodity options. *Journal of Banking & Finance*, *66*, 53–65. https://doi.org/10.1016/j.jbankfin.2016.02.001

Aït-Sahalia, Y., & Mancini, L. (2008). Out of sample forecasts of quadratic variation. *Journal of Econometrics, Special Issue on Econometric Modelling in Finance and Risk Management: An Overview*, *147*, 17–33. https://doi.org/10.1016/j.jeconom.2008.09.015

Back, J., Prokopczuk, M., & Rudolf, M. (2013). Seasonality and the valuation of commodity options. *Journal of Banking & Finance*, *37*, 273–290. https://doi.org/10.1016/j.jbankfin.2012.08.025

Bernard, J.-T., Khalaf, L., Kichian, M., & Mcmahon, S. (2008). Forecasting commodity prices: Garch, jumps, and mean reversion. *Journal of Forecasting*, *27*, 279–291. https://doi.org/10.1002/for.v27:4

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*, 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Boos, D. (2023). *Event days for "Risky times: Seasonality and event risk of Commodities"*. Zenodo. https://zenodo.org/record/8374198

Boos, D., & Grob, L. (2023). Tracking speculative trading. *Journal of Financial Markets*, *64*, 100774. https://doi.org/10.1016/j.finmar.2022.100774

Chang, C.-L., McAleer, M., & Tansuchat, R. (2012). Modelling long memory volatility in agricultural commodity futures returns. *Annals of Financial Economics*, *7*, 1250010. https://doi.org/10.1142/S2010495212500108

Charfeddine, L. (2014). True or spurious long memory in volatility: Further evidence on the energy futures markets. *Energy Policy*, *71*, 76–93. https://doi.org/10.1016/j.enpol.2014.04.027

Chen, Y.-F., & Mu, X. (2021). Asymmetric volatility in commodity markets. *Journal of Commodity Markets*, *22*, 100139. https://doi.org/10.1016/j.jcomm.2020.100139

Deng, Q. (2008). Predictable components of individual stock volatility. Available at SSRN: https://doi.org/10.2139/ssrn.1156621

Diewald, L., Prokopczuk, M., & Wese Simen, C. (2015). Time-variations in commodity price jumps. *Journal of Empirical Finance*, *31*, 72–84. https://doi.org/10.1016/j.jempfin.2015.02.004

Ding, Z., & Granger, C. (1996). Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics*, *73*, 185–215. https://EconPapers.repec.org/RePEc:eee:econom:v:73:y:1996:i:1:p:185-215

Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, *1*, 83–106. https://doi.org/10.1016/0927-5398(93)90006-D

Dorfmann, J., & Karali, B. (2015). A nonparametric search for information effects from usda reports. *Journal of Agricultural and Resource Economics*, *40*, 124–143.

Dubinsky, A., Johannes, M., Kaeck, A., & Seeger, N. J. (2018). Option pricing of earnings announcement risks. *The Review of Financial Studies*, *32*, 646–687. https://doi.org/10.1093/rfs/hhy060

Efimova, O., & Serletis, A. (2014). Energy markets volatility modelling using garch. *Energy Economics*, *43*, 264–273. https://doi.org/10.1016/j.eneco.2014.02.018

Engle, R., & Rangel, J. (2008). The spline-garch model for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies*, *21*, 1187–1222. https://EconPapers.repec.org/RePEc:oup:rfinst:v:21:y:2008:i:3:p:1187-1222

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*, 987–1007.

Engle, R. F., & Sokalska, M. E. (2012). Forecasting intraday volatility in the us equity market. multiplicative component garch. *Journal of Financial Econometrics*, *10*, 54–83.

Fackler, P. L., & Tian, Y. (1999). Volatility models for commodity markets. In *Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management*.

Fang, T., Lee, T.-H., & Su, Z. (2020). Predicting the long-term stock market volatility: A garch-midas model with variable selection. *Journal of Empirical Finance*, *58*, 36–49. https://doi.org/10.1016/j.jempfin.2020.05.007

Francq, C., & Zakoian, J.-M. (2019). *GARCH models: Structure, statistical inference and financial applications* (2nd ed.). Wiley. https://doi.org/10.1002/9781119313472

Gatheral, J., Jaisson, T., & Rosenbaum, M. (2018). Volatility is rough. *Quantitative Finance*, *18*, 933–949. https://doi.org/10.1080/14697688.2017.1393551

Geman, H., & Smith, W. O. (2013). Theory of storage, inventory and volatility in the lme base metals. *Resources Policy*, *38*, 18–28. https://doi.org/10.1016/j.resourpol.2012.06.014

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, *48*, 1779–1801.

Golub, G. H., Heath, M., & Wahba, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, *21*, 215–223.

Halawa, A., & Bassiouni, M. E. (2000). Tests of regression coefficients under ridge regression models. *Journal of Statistical Computation and Simulation*, *65*, 341–356. https://doi.org/10.1080/00949650008812006

Hevia, C., Petrella, I., & Sola, M. (2018). Risk premia and seasonality in commodity futures. *Journal of Applied Econometrics*, *33*, 853–873. https://doi.org/10.1002/jae.v33.6

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *42*, 80–86.

Isengildina-Massa, O., Cao, X., Karali, B., Irwin, S. H., Adjemian, M., & Johansson, R. C. (2021). When does usda information have the most impact on crop and livestock markets? *Journal of Commodity Markets*, *22*, 100137. https://doi.org/10.1016/j.jcomm.2020.100137

Isengildina-Massa, O., Irwin, S. H., Good, D. L., & Gomez, J. K. (2008a). The impact of situation and outlook information in corn and soybean futures markets: Evidence from wasde reports. *Journal of Agricultural and Applied Economics*, *40*, 89–103.

Isengildina-Massa, O., Irwin, S. H., Good, D. L., & Gomez, J. K. (2008b). Impact of wasde reports on implied volatility in corn and soybean markets. *Agribusiness*, *24*, 473–490. https://doi.org/10.1002/agr.v24:4

Karali, B., & Thurman, W. N. (2010). Components of grain futures price volatility. *Journal of Agricultural and Resource Economics*, *35*, 167–182.

Kim, J., Schonberger, B., Wasley, C., & Yang, Y. J. (2023). Forecasting market volatility: The role of earnings announcements. *The Accounting Review*. 1–29. In press. https://doi.org/10.2308/TAR-2021-0351

Lopez, J. A. (2001). Evaluating the predictive accuracy of volatility models. *Journal of Forecasting*, *20*, 87–109. https://doi.org/10.1002/1099-131X(200103)20:2<87::AID-FOR782>3.0.CO;2-7

McKenzie, A. M. (2008). Pre-harvest price expectations for corn: The information content of usda reports and new crop futures. *American Journal of Agricultural Economics*, *90*, 351–366. https://doi.org/10.1111/ajae.v90.2

McKenzie, A. M., & Ke, Y. (2022). How do usda announcements affect international commodity prices? *Journal of Commodity Markets*, *28*, 100239. https://doi.org/10.1016/j.jcomm.2021.100239

McKenzie, A. M., Thomsen, M. R., & Adjemian, M. K. (2022). Characterizing implied volatility functions from agricultural options markets. *American Journal of Agricultural Economics*, *104*, 1605–1624. https://doi.org/10.1111/ajae.12288

McKenzie, M. D., Mitchell, H., Brooks, R. D., & Faff, R. W. (2001). Power arch modelling of commodity futures data on the London metal exchange. *The European Journal of Finance*, *7*, 22–38. https://doi.org/10.1080/13518470123011

Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*, 125–144. https://EconPapers.repec.org/RePEc:eee:jfinec:v:3:y:1976:i:1-2:p:125-144

Mincer, J., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In J. Mincer (Ed.), *Economic forecasts and expectations: Analysis of forecasting behavior and performance* (pp. 3–46). National Bureau of Economic Research, Inc. https://EconPapers.repec.org/RePEc:nbr:nberch:1214

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, *59*, 347–370. https://EconPapers.repec.org/RePEc:ecm:emetrp:v:59:y:1991:i:2:p:347-70

Nguyen, D. K., & Walther, T. (2020). Modeling and forecasting commodity market volatility with long-term economic and financial variables. *Journal of Forecasting*, *39*, 126–142. https://doi.org/10.1002/for.v39.2

RiskMetrics. (1996). *RiskMetrics—technical document*. Technical Report fourth edition, Morgan Guaranty Trust Company of New York.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, *6*, 41–50.

Schneider, L., & Tavin, B. (2021). Seasonal volatility in agricultural markets: Modelling and empirical investigations. *Annals of Operations Research*. In press. https://doi.org/10.1007/s10479-021-04241-7

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance*, *44*, 1115–1153. https://doi.org/10.1111/jofi.1989.44.issue-5

Shroyer, J. P., Mikesell, M. E., & Paulsen, G. M. (1995). *Spring freeze injury to Kansas wheat*. Technical Report C-646, Kansas State University.

Simon, D. P. (2002). Implied volatility forecasts in the grains complex. *Journal of Futures Markets*, *22*, 959–981.

Sørensen, C. (2002). Modeling seasonality in agricultural commodity futures. *Journal of Futures Markets*, *22*, 393–426. https://doi.org/10.1002/fut.v22:5

Symeonidis, L., Prokopczuk, M., Brooks, C., & Lazar, E. (2012). Futures basis, inventory and commodity price volatility: An empirical analysis. *Economic Modelling*, *29*, 2651–2663. https://doi.org/10.1016/j.econmod.2012.07.016

Taylor, S. (1986). *Modelling financial time series*. John Wiley and Sons.

Tencate, A., Kalivas, J. H., & Andries, E. (2016). Penalty processes for combining roughness and smoothness in spectral multivariate calibration. *Journal of Chemometrics*, *30*, 144–152. https://doi.org/10.1002/cem.v30.4

Tikhonov, A. N. (1963). Solution of incorrectly formulated problems and the regularization method. *Soviet Mathematics*, *4* 1035–1038.

van Wieringen, W. N. (2023). Lecture notes on ridge regression. arXiv:1509.09169. https://doi.org/10.48550/arXiv.1509.09169

Wang, X., & Garcia, P. (2011). Forecasting corn futures volatility in the presence of long memory, seasonality and structural change. In *2011 Annual Meeting, July 24-26, 2011, Pittsburgh, Pennsylvania*. Vol. 103749, Agricultural and Applied Economics Association. https://EconPapers.repec.org/RePEc:ags:aaea11:103749

How to cite this article: Boos, D. (2024). Risky Times: Seasonality and Event Risk of Commodities. *The Journal of Futures Markets*, 1–17. https://doi.org/10.1002/fut.22492

APPENDIX A: HYPERPARAMETER TUNING

The general Tikhonov regularization problem can be stated as

$$\begin{bmatrix} Y \\ 0 \end{bmatrix} = \begin{bmatrix} X \\ \Gamma \end{bmatrix} B + \begin{bmatrix} G \\ H \end{bmatrix}, \tag{A1}$$

where X and Y are the data (features and labels), Γ is the regularization matrix that depends on a set of hyperparameters $\gamma_1 \dots \gamma_N$, and G and H are vectors of error terms. Define the hat matrix as:

$$A(\Gamma) = X(X^T X + \Gamma^T \Gamma)^{-1} X^T. \tag{A2}$$

Following Allen (1974), a the hyperparameters can be efficiently determined using leave-one-out cross-validation (LOOCV). This is done by estimating Equation (A1) with all but the τth observation. This yields estimated B denoted as $\hat{B}_{i,-\tau}(\Gamma)$ and an out-of-sample forecast:

$$\widetilde{Y}(\Gamma) = X\hat{B}_{i,-\tau}(\Gamma) \tag{A3}$$

Allen's PRESS (prediction sum of squares) statistics $P(\Gamma)$ is the average mean squared error over all observations:

$$P(\Gamma) = \frac{1}{T}\|\widetilde{Y}(\Gamma) - Y\|^2. \tag{A4}$$

Using the Sherman–Morrison–Woodbury formula, it can be shown that:

$$P(\Gamma) = \frac{1}{T}\|B(\Gamma)(I - A(\Gamma))Y\|^2 \tag{A5}$$

where $B(\Gamma)$ is a diagonal matrix with jjth entry $1/(1 - a_{jj}(\Gamma)$, a_{jj} being the jjth entry of $A(\Gamma)$ and $\|\cdot\|$ is the Euclidean norm. See van Wieringen (2023) and Golub et al. (1979) for further details.

Golub et al. (1979) show that the numerically far more efficient generalized cross-validation (GCV) criteria

$$V(\Gamma) = \frac{1}{T}\|(I - A(\Gamma))Y\|^2 \Big/ \left[\frac{1}{T}\operatorname{Trace}(I - A(\Gamma))\right]^2, \tag{A6}$$

a rotation-invariant version of PRESS, is a similarly appropriate criteria. Throughout the paper, I choose $\Gamma = \Gamma(\gamma)$ as the value that minimizes GCV. Note that with this procedure, the hyperparameters are tuned by the data and therefore there is no discretion in their choice.

APPENDIX B: DETAILS ON THE WLS ESTIMATION PROCEDURE

I estimate the model in four steps using the tuning function described in the previous section (A). Note, that following chap. 6.2 in Francq and Zakoian (2019) the standard deviation of ε_t is proportional to $\tilde{c}_t^1$.

1. Tune λ and fit (10) with OLS.
2. Define a weighting matrix W with the inverse of the normalized variances $\tilde{c}_t^1/\hat{c}^1$ in the diagonal, where $\hat{c}^1$ is the mean of $\widetilde{C}^1$. Using λ from step 1 define a weighted version of (10) as

$$\begin{bmatrix} WC \\ 0 \end{bmatrix} = \begin{bmatrix} W\Xi & WI \\ \Lambda & 0 \end{bmatrix}\begin{bmatrix} S \\ E \end{bmatrix}. \tag{B1}$$

   and estimate it with OLS.
3. Now we can update W using the the improves estimate of C from the first WLS run in step 2. Estimate (B1) with updated W retuned λ using OLS.
4. repeat step 3.

Further iterations hardly alter the estimates.